Proceedings

# of the

11th European Conference on Information Warfare and Security

The Institute Ecole Supérieure en Informatique Electronique et Automatique, Laval, France

5-6 July 2012

Edited by

Eric Filiol and Robert Erra

ESIEA, Laval

France

Papers have been double-blind peer reviewed before final submission to the conference. Initially, paper abstracts were read and selected by the conference panel for submission as possible papers for the conference.

Many thanks to the reviewers who helped ensure the quality of the full papers.

These Conference Proceedings have been submitted to Thomson ISI for indexing.

Further copies of this book and previous year’s proceedings can be purchased from http://academic-bookshop.com

CD version ISBN: 978-1-908272-56-0
CD version ISSN: 2048-9897
Book version ISBN: 978-1-908272-55-3
Book Version ISSN: 2048-9870

Published by Academic Publishing International Limited
Reading
UK
44-118-972-4148
www.academic-publishing.org

Contents

Preface

This year sees the 11th European Conference on Information Warfare and Security (ECIW 2012), which is hosted by the Institute Ecole Supérieure en Informatique, Electronique et Automatique, Laval, France The Conference Chair is Eric Filiol from ESIEA, Laval, France and I am pleased to be the Programme Chair along with Laurent Beaudoin.

The Conference continues to bring together individuals working in the area of Information Warfare and Information Security in order to share knowledge and develop new ideas with their peers. The range of papers presented at the Conference will ensure two days of interesting discussions. The topics covered this year illustrate the depth of the information operations' research area, with the subject matter ranging from the highly technical to the more strategic visions of the use and influence of information.

The opening keynote is given by Rainer Fahs, Chairman of the European Institute of Computer Antivirus Research (EICAR) on the topic of "Cyber warfare: Prospective aspects from the EICAR perspective". The second day will be opened by Lieutenant-colonel Eric Freyssinet, head of the cybercrime division, Gendarmerie nationale in France. Eric will address the issue of the necessary continuum between fighting cybercrime and cyberdefense.

With an initial submission of 90 abstracts, after the double blind, peer review process there are 42 papers published in these Conference Proceedings. These papers come from around the world including Australia, Canada, Czech Republic, Estonia, Finland, France, Iran, Japan, Malaysia, Portugal, South Africa, The Netherlands, Turkey, United Kingdom and the United States of America.

I wish you a most interesting conference and an enjoyable stay in France.

Robert Erra and Eric Filiol

ESIEA, Laval, France

July 2012

Conference Executive

Eric Filiol, ESIEA, Laval, France
Robert Erra ESIEA, Paris, France
Laurent Beaudoin ESIEA, Laval, France

Conference Committee

Nasser Abouzakhar (University of Hertfordshire, UK); Kari Alenius (University of Oulu, Finland); Colin Armstrong (Curtin University, Australia); Debi Ashenden (Cranfield University, Shrivenham, UK); Maumita Bhattacharya (Charles Sturt University, Australia); John Biggam (Glasgow Caledonian University, UK); Andrew Blyth (University of Glamorgan, UK); Martin Botha (South African Police, South Africa, South Africa); Svet Braynov (University of Illinois at Springfield, USA); Bill Buchanen (Napier University, UK); Catharina Candolin (Defence Command Finland, Finland); Joobin Choobineh (Texas A & M University, Texas, USA); Maura Conway (Dublin City University, Ireland); Michael Corcoran (DSTL, UK); Paul Crocker (Universidade de Beira Interior, Portugal); Josef Demergis (University of Macedonia, Greece); Moses Dlamini (SAP Research, South Africa); Geoffrey Darnton (Bournemouth University, UK); Paul Dowland (University of Plymouth, UK); Marios Efthymiopoulos (Political Science Department University of Cyprus, Cyprus); Ramzi El-Haddadeh (Brunel University, UK); Robert Erra (ESIEA PARIS, France); John Fawcett (University of Cambridge, UK); Eric Filiol (Ecole Supérieure en Informatique, Electronique et Automatique, France); Chris Flaherty (University of New South Wales, Australia); Steve Furnell (University of Plymouth, UK); Javier Garci'a Villalba (Universidad Complutense de Madrid, Spain); Kevin Gleason (KMG Consulting, MA, USA); Stefanos Gritzalis (University of the Aegean, Greece); Julio Cesar Hernandez Castro (Portsmouth University, UK); Ulrike Hugl (University of Innsbruck, Austria); Aki Huhtinen (National Defence College, Finland); Bill Hutchinson (Edith Cowan University, Australia); Berg Hyacinthe (Assas School of Law-CERSA-CNRS, La Sorbonne, France); Abhaya Induruwa (Canterbury Christ Church University,); Ioannis Mavridis (University of Macedonia, Greece); Hamid Jahankhani (University of East London, UK); Amit Jain (BenefitFocus Inc, USA); Helge Janicke (De Montfort University, UK); Andy Jones (BT, UK); James Joshi (University of Pittsburgh, USA); Nor Badrul Anuar Jumaat (University of Malaya, Malaysia); Maria Karyda (University of the Aegean, Greece); Vasilis Katos (Democritus University of Thrace, Greece); Auli Keskinen (National Defence College, Finland); Jyri Kivimaa (Cooperative Cyber Defence and Centre of Excellence, Tallinn, Estonia); Spyros Kokolakis (University of the Aegean, Greece); Ahmet Koltuksuz (Yasar University, Turkey); Theodoros Kostis (University of the Aegean, Greece); Prashant Krishnamurthy (University of Pittsburgh, USA); Dan Kuehl (National Defense University, Washington DC, USA); Peter Kunz (DiamlerChysler, Germany); Pertti Kuokkanen (University of Helsinki, Finland); Takakazu Kurokawa (National Defence Acadamy, Japan); Rauno Kuusisto (Finish Defence Force, Finland); Tuija Kuusisto (Internal Security ICT Agency HALTIK, Finland); Michael Lavine (John Hopkins University's Information Security Institute, USA); Martti Lehto (National Defence University, Finland); Tara Leweling (Naval Postgraduate School, Pacific Grove, USA); Paul Lewis (technology strategy board, UK); Sharman Lichtenstein (Deakin University, Australia); David Llamas (University of St Andrews, UK); Hossein Malekinezhad, (Islamic Azad University, Naragh Branch, Iran); Mario Marques Freire (University of Beira Interior, Covilhã, Portugal); Rob McCusker (Teeside University, Middlesborough, UK); Durgesh Mishra (Acropolis Institute of Technology and Research, India); Yonathan Mizrachi (University of Haifa, Israel, Israel); Edmundo Monteiro (University of Coimbra, Portugal); Evangelos Moustakas (Middlesex University, London, UK); Kara Nance (University of Alaska Fairbanks, USA); Muhammad Naveed (IQRA University Peshawar, Pakistan); Daniel Ng (C-PISA/HTCIA, China); Rain Ottis (Cooperative Cyber Defence Centre of Excellence, Estonia); Tim Parsons (Selex Communications, UK); Andrea Perego (Università degli Studi dell’Insubria, Varese, Italy); Michael Pilgermann (University of Glamorgan, UK); Fred Piper (Royal Holloway, University of London, UK); Engur Pisirici (govermental - independent, Turkey); Jari Rantapelkonen (National defense University, Finland); Andrea Rigoni (for Booz & Company,, USA); Raphael Rues (DigiComp Academy, Switzerland); Filipe Sa Soares (University of Minho, Portugal); Henrique Santos (University of Minho, Portugal); Damien Sauveron (Mathematics and Computer Sciences, University of Limoges, France); Richard Sethmann (University of Applied Sciences Bremen, Germany); Paulo Simoes (University of Coimbra, Portugal); Jill Slay (University of South Australia, Australia); Anna Squicciarini (University of Milano, Italy); Iain Sutherland (University of Glamorgan, Wales, UK); Jonas Svava Iversen (Danish Broadcast

Corporation, Denmark); Sérgio Tenreiro de Magalhães (Universidade Católica Portuguesa, Portugal); Peter Trommler (Georg Simon Ohm University Nuremberg, Germany); Theodore Tryfonas (University of Bristol, UK); Craig Valli (Edith Cowan UniversitY, Australia); Rudi Vansnick (Internet Society, Belgium); Richard Vaughan (General Dynamics UK Ltd, United Kingdom); Stilianos Vidalis (Newport Business School, Newport, UK); Paulo Viegas Nunes (Military Academy, Lisbon, Portugal); Natarajan Vijayarangan (Tata Consultancy Services Ltd, India); Teemupekka Virtanen (Helsinki University of Technology, Finland); Marja Vuorinen (University of Helsinki, Finland); Michael Walker (Vodafone, UK); Mat Warren (Deakin University, Australia, Australia); Kenneth Webb (Edith Cowan University , Australia); Trish Williams (Edith Cowan University, Australia); Simos Xenitellis (Royal Holloway University, London, UK); Omar Zaafrany (Ben-Gurion University of the Negev, Israel); Omar Zakaria (National Defence University of Malaysia,

Biographies

Conference Chair

Eric Adrien Filiol is the head of the Operational Cryptology and Virology at ESIEA a French Engineer School in Computer Science, Electronics and Control Science. He has spent 21 years in the French Army mainly as a ICT security expert (cryptanalysis, computer virology, cyberwarfare). He is also senior officer reservist in the French DoD. He holds a Engineer diploma in Cryptology, a PhD in applied mathematics and computer science and a Habilitation Thesis in Computer Science. His main research interest are Symmetric Cryptosystems analysis (especially from a combinatorial point of view), Computer virology (theoretical and experimental study of new form of malware and anti-malware technologies), Computer warfare techniques. He is also the Scientific Director of the European Institute in Computer Antivirus Research (EICAR) in Germany and the Editor-in-chief of the Journal in Computer Virology. He likes playing Bass Guitar (Jazz), running (marathon and half marathon) and good wine/food.

Programme Chairs

Robert Erra holds a Phd in Computer Science from the University of Rennes I and is currently Professor of CS Scientific Director of the Masters in Information & System Security at ESIEA Paris opened in 2004. He is interested in developments and analysis of algorithms for information security, from cryptanalysis of asymmetric cryptography to malware analysis and in secure programming.

Laurent Beaudoin received a PhD from Télécom Paristech in image processing and remote sensing. He has worked in Ecole Supérieure d'Informatique d'Electronique et d'Automatique (ESIEA), a french engineering school, since 2001. He founded in 2004 the Image and Signal Processing R&D department (ATIS laboratory). His main research activities concern Defence and Security, exploring robots (UAS, AUV), remote sensing and ICTs for persons with disabilities. With his students, he regularly participates to national or international challenges (minidrone DGA-ONERA challenge, SAUC-E NATO.)

Keynote Speakers

Rainer Fahs is currently employed as Senior Information Systems Security Engineer at the NATO Air Command and Control Management Agency (NACMA) where he is responsible for the security architecture of the newly developed NATO Air Command & Control System (ACCS). In this capacity he is also the Chairman of the ACCS Security Accreditation Board which is responsible for the security accreditation of the ACCS system. In 2003 Rainer retired from the German Air Force where he spent most of his time in flight safety and flying operations inclusive many hours of F104 and Alpha Jet flying. His last active job was in HQ Air Force Ramstein (Germany) where he spent four years in tactical evaluation for offensive flying. Rainer has been with NATO in Brussels as NATO civil employee where he started as system/network administrator in an intelligence project and developed back in 1991 the first NATO Secret network based on PCs before he was hired by NACMA in 1995. In this new job he had to start over again and get back to school to learn the secrets of cryptography, formalization and the inner architectures of computers and networking. Part of his job is also to represent his agency at NATO Committees dealing with

NATO Security Policy and in particular the INFOSEC Policy. Rainer joined the European Institute for Computer Anti-Virus Research (EICAR) in 1992 where he participated in the working group looking into the AV technology. In 1994 he was elected Board member and became the Director of EICAR Working Groups. In 1996 he was first time elected Chairman of the Board and has been in that position until today.

Lieutenant-colonel Eric Freyssinet, head of the cybercrime division, Gendarmerie nationale. Chairman of the Expert group on IT Crime - Europe of Interpol. Education: Ecole Polytechnique (general engineering, X1992), Mastère spécialisé in Network and IT security from Telecom ParisTech (2000), and currently PhD student at University Paris 6 on the subject of the fight against botnets. Pour les horaires de train, dès que je les aurai fixés.

__Mini Track Chairs__

Prof Kevin M. Gleason has over 30 years of experience combining computers and education. A long-time student of computer hacks and information breaches his lectures introduced disaster recovery analysis and preparation. An author of technical textbooks and a 2001 NASA/ASEE Summer Faculty Research Fellow at the NASA—Goddard Space Flight Center in Greenbelt MD. In the aftermath of the 2001 terrorist attacks, was the principle investigator of 'psycho-metrics' (a method of identifying the author through written text). He is currently semi-retired working as an adjunct professor to several colleges and a consultant to business in the Greater Boston Massachusetts area.

Professor Aki Huhtinen LTC(G.S) PhD. is Docent of practical philosophy in the University of Helsinki and Docent of social consequences of media and information technology in the University of Lapland. The author is also Docent of information security and information operations in the University of technology in Tampere. Huhtinen works at the Department of Leadership and Military Pedagogy at the Finnish National Defence University.

Amit Jain is currently working in the R&D at BenefitFocus Inc, a Charleston, South Carolina based enterprise providing health benefits management platform for employers and insurance carriers. His current position involves researching and developing semantics enabled health care systems providing efficient and secure data management. Earlier he was a part of BeliefNetworks Inc, a startup company that sought to create tools for knowledge generation from structured and unstructured data. He holds a Masters and a PhD from University of South Carolina, Columbia. His dissertation focused on using semantics for authorizations on ontologies and syntax independent data. His research interests include security policies, ontology based enterprise systems, identity management, digital rights management and information warfare

Dr. Helge Janicke is currently working as a Senior Lecturer in Computer Security at De Montfort University, Leicester (UK). He is leading the research theme on Computer Security and Trust within the Software Technology Research Laboratory and is working with De Montfort's Centre for Secure Computing. His research interests are in the area of computer security and formal methods for modelling security systems.

Jari Rantapelkonen LTC, D.Sc(mil) is a researcher and teacher in Finnish National Defence University. He works at the Department of Operational Art and Tactics

Henrique Santos received his first degree in Electric and Electronic Engineering, by the University of Coimbra, Portugal, in 1984. In 1996 he got his PhD in Computer Engineering, at the University of the Minho, Portugal. Currently he is an Associate Professor at the Information Systems Department, University of Minho, being responsible for several projects and the supervision of several dissertations, mainly in the Information Security and Computer Architecture areas. He is the president of a national Technical Committee (CT 136) for information system security standards. In 1990, he was teaching at the University of Bristol, United Kingdom.

Dr Tim Watson is the Head of Department of Computer Technology at De Montfort University and the Director of its Centre for Secure Computing. Tim is a regular media commentator on information security and digital forensics and a member of various advisory groups, including DSTL's Cyber and Situational Awareness Information Exchange, the CESG Academic Advisory Group, the National Information Assurance Forum, the IAAC Academic Liaison Panel and the UK ISO 27000 series standards body.

Peter Norris is a teacher fellow and principal lecturer at De Montfort University where he leads the security strand of the Center for Secure Computing. Originally trained as an engineer, he spent nine years in industry, before joining academia. He has advised local government on identity management, supervised research on network security, helped develop the UK input to ISO standards on both digital forensics and network security, and is currently researching the security of motor vehicle digital systems. His overarching interest is in the security consequences of the practical interaction of heterogeneous systems.

Biographies of Presenting Authors

Andrew Adams is Professor of Information Ethics in the Graduate School of Business Administration and Deputy Director of the Centre for Business Information Ethics at Meiji University in Tokyo. He is the chair of ACM SIGCAS. He holds a BSc, MSC and PhD in Computing subjects and an LLM in Advanced Legal Studies.

Adetunji Adebiyi is a Doctoral student with the University of East London UK. His research focuses on integrating security into software design during SDLC. He is a member of the British Computer Society. His research has led him to give talks and presentations in conferences and seminars he has attended.

Kari Alenius is Associate Professor in the Department of History at the University of Oulu, Finland, since 1998. He also has Adjunct Professorship at the University of Oulu (1997). His research interests include the history of propaganda and mental images, the history of Eastern Europe between the World Wars, and the history of ethnic minorities.

Olga Angelopoulou is a lecturer in Digital Forensics at the University of Derby. She obtained a doctorate in Computing with the title: ‘Analysis of Digital Evidence in Identity Theft Investigations’. Her research interests include Computer Forensics, Digital Evidence, Identity theft, Online fraud, Computer crime investigations and the Online Social Networking.

Rabia Aslanoglu got her B.S. degree from Izmir University of Economics, Department of Software Engineering in 2009. Currently, she is studying for her M.S. degree in Izmir Institute of Technology, Department of Computer Engineering, Izmir, Turkey. Her research interests are public-key cryptography and information warfare.

Melanie Bernier is a Defence Scientist with the Defence Research and Development Canada’s Center for Operational Research and Analysis. She has a Masters Degree in Electrical Engineering and experience in modelling and simulation, concept development and experimentation, joint C4ISR, and computer networks. She is currently leading studies in force development for the cyber environment.

Matt Bishop received his Ph.D. in computer science from Purdue University in 1984. His main research areas are the analysis of vulnerabilities in computer systems, security of electronic voting systems and election processes, and data sanitization, network security, and malware analysis. His textbook, *Computer Security: Art and Science*, was published in 2002 by Addison-Wesley Professional.

Abílio Fernando Costa Cardoso, Occupation or position held: Researcher in Cloud Computing, IT governance and computer networks, professor at Department of Innovation, Science and Technology, manager of IT area at Portucalense University. M Sc in Computer Science, 1995, B Sc in Applied Mathematics, 1987, Class teaching: Project planning and network management, Java advanced programming

Christian Czosseck is scientist at the NATO CCD COE in Tallinn, Estonia. Serving in the German military for more than 14 years, he held several information assurance positions. Christian holds a M.Sc. equivalent in computer scienceand is currently PhD candidate at the Estonian BusinessSchool, Tallinn looking into cyber security and botnet related issues.

Michel Dubois is teacher and researcher at the Cryptology and Operational Virology (C+V)° laboratory. Currently he is a PhD student in cryptography and works on a new approach of the cryptanalysis of the AES.

Tim Hartog graduated in 2005 at the Technical University of Twente, he has been active in the field of Information Security. During his work at TNO, the Dutch Organization for Applied Scientific Research, Tim has been working on topics like Trusted Computing, Trusted Operating Systems and Cross Domain Solutions.

Major Arto Hirvelä is instructor (leadership) in Research Group at the Finnish National Defence University. His research interests are information environment, strategic communication and information operations.

Dr. Berg P. Hyacinthe (Ph.D., Florida State University; LLD Candidate, Assas School of Law, CERSA-CNRS, en Sorbonne) is Director of the “Centre de Recherche Scientifique et d’Études Cybernétiques” (CRESEC) at “l’Université d’État d’Haiti”. His research agenda is set at the intersection of law and technology, exploring legality/illegality of advanced digital technologies in armed conflicts.

Anna-Marie Jansen van Vuuren began her career as a journalist in 2004. She holds a Masters degree in Journalism and Media Management from the University of Stellenbosch and currently she is pursuing a Ph.D. at the University of Pretoria. Anna-Marie is a lecturer at the University of the Witwatersrand and freelances as a broadcast journalist and producer at the South African Broadcast Corporation.

Joey Jansen van Vuuren is the Research Group Leader for Cyber Defence at the CSIR South Africa. , mainly involved in research for the Defence and Government sectors.. She obtained her Masters from UNISA and her research is focussed on National Cybersecurity and the analysis of Cyber threads. She is also involved in Cyber awareness programs in South Africa

Saara Jantunen has studied English language and culture in the University of Groningen, and English philology in the University of Helsinki. Currently she is writing her doctoral dissertation in the Finnish National Defence University, where she majors in leadership. Her research interests include language and identity, ideology in discourse, strategic communication and multimodal discourse. Jantunen currently works in education.

Roman Jašek is the head of Department of Informatics and Artificial Intelligence, Faculty of Applied Informatics, Tomas Bata University in Zlín. His habilitation thesis focused on implementing information security paradigm into commercial organizations as well as tertiary education institutions. Professional interests include computer security auditing, knowledge protection, information systems, and informatics.

Eli Jellenc leads analysis of geopolitical cyber security and directs research operations in the EU and Japan for VeriSign-iDefense, where he previously created the International Cyber Intelligence unit. He holds an MA in International Security from Georgetown University, and has authored 3 cyber risk country-profiles in the “Cybercrime and Security” journal (Oxford University Press).

Jeremy Julien is an Engineering student at the ESIEA school in France. Jeremy has been responsible for the plotting of our keywords and related cables. Jeremy also assisted Nils by searching for related cables, keywords and other relevant sources.

Harry Kantola is working as a teacher at the department of tactics and operational arts in Finnish National Defense University (FNDU), teaching senior staff officers. Kantola has studied (Masters of arts) in Swedish National Defense College (SNDC) and to General Staff Officer at FNDU. His thesis at these schools dealt with CNO in general.

Anssi Kärkkäinen Captain, M.Sc. (Eng.) graduated from the Finnish National Defence University in 2000. He also graduated a Master of Science (Engineering) degree from Helsinki University of Technology (currently Aalto University) in 2005. Currently he is carrying out doctoral studies at the same university. His current assignment is a Staff Engineer for Defence Command Finland.

William Aubrey Labuschagne lectured programming, networking, and security subjects at Tshwane University of Technology (TUT) 2002-2009. Obtained Red Hat Certified Technician (RHCT - 2006) and is qualified SCRUM Master (Scrum Alliance). Qualifications consist of NDip: Computer System Engineering and

BTech: Information Technology at Pretoria Technikon Currently completing MTech at UNISA in field of security awareness in rural areas and is technologist at CSIR.

Pierre Leandre is a student in his second year at ESIEA where hestudy informatics and electronics. At school he is part of the OCV lab (Operational Cryptology and Virology Laboratory).

Sylvain (Sly) Leblanc is an Assistant Professor at the Royal Military College of Canada. Sly was a Canadian Army Signals Officer for over 20 years, where he developed his interest in computer network operations. His research interests are in computer security and computer network operations.

Andrew Nicholson received his Masters degree in Computer Security in 2010 at De Montfort, University in England. In 2010 Andrew began PhD research looking into attribution of cyber attacks. He is a member of the De Montfort University Cyber Security Centre (CSC) research group. Andrews main interests are network security, cyber-conflict and online anonymity.

Mzukisi Niven Njotini is a currently working as a lecturer in College of Law at the University of South Africa (South Africa). Njotini has an LLB and LLM (Cum Laude) degree and is currently enrolled for an LLD or Phd degree. Njotini's area of speciality is information technology law.

Azah Norman BA(Hons.) in Information Science from National University of Malaysia (UKM) (2000)and Masters in Secure e-Commerce from Royal Holloway University of London, United Kingdom (2003). Research focuses on information systems security and e-commerce. Worked as a consultant in a Malaysian premier security company under Malaysia's national information & communication technology corporation (over three years). Research medals for research in Halal-RFID verification and currently pursuing PhD in information systems security management.

Karlis Podins has graduated University of Latvia with master degree in Computer Science, and is working for Cooperative Cyber Defence Centre of Excellence since 2008. Apart from dormant interest in Finite Automata and Natural Language Processing, his research has touched several areas related to security.

Dr Graeme Pye is a Lecturer with the School of Information Systems, Deakin University, Australia. His research is continuing to focus on investigating the security and resilience aspects of Australian critical infrastructure and the relationships between associated infrastructures. Furthermore, he is also interested in modelling emergency management and disaster response systems, including resilience and security management aspects.

Neil Rowe is Professor of Computer Science at the U.S. Naval Postgraduate School where he has been since 1983. He has a Ph.D. in Computer Science from Stanford University (1983). His main research interests are the modeling of deception, information security, surveillance systems, image processing, and data mining.

Mirva Salminen (M.Soc.Sc.) is a doctoral candidate at the University of Lapland, Finland, researching on the outsourcing of the state's security related functions. She has studied International Relations and Political Science at the University of Tampere, Finland; Military History and Strategy at the Finnish National Defence University; and Security Studies at Aberystwyth University, the UK.

Libor Sarga is a doctoral worker at the Department of Statistics and Quantitative Methods, Faculty of Management and Economics, Tomas Bata University in Zlín. His dissertation will be focused on computer and data security in the presence of unreliable human element as an exploitable attack vector. His personal interests include technology, hardware and software architectures.

Emad Shafie PhD student at The Software Technology Research Laboratory, in De Montfort University, I have achieved a master degree in Information technology 2009 from De Montfort University and bachelor degree in computer science 2000 from King Abdul-Aziz University. I have worked as leader of programming department for five years at Umu Alqura University in Saudi Arabia.

Torsti Sirén Lieutenant Colonel, General Staff, Ph.D. (Pol. Sc.), B.A. (Slavonic philology) Torsti Sirén is Head of Research Group in the Department of Leadership and Military Pedagogy at the Finnish National Defence University, Helsinki.

Ignus Swart obtained his masters degree in computer science at Tshwane University of Technology where he studied with a full scholarship. After several years working as both a software developer and security professional, he is currently in the employ of the CSIR, Cyber Defence department and pursuing his PhD studies at Rhodes University.

Selma Tekir is working at Izmir Institute of Technology Department of Computer Engineering. In 2009, she worked as a visiting researcher at Faculty of Computer Science Chair for Databases, Data Analysis and Visualization at University of Konstanz-Germany. Tekir received a PhD in computer engineering from Ege University, Turkey in 2010.

Solomon Uwagbole is currently a full-time research student in the Centre for Distributed Computing, Networks, and Security at School of Computing, Edinburgh Napier University, Scotland. He is also a freelance Microsoft Certified Trainer (MCT). He holds B.Sc. in Zoology (Hons) from the University of Delhi, India and M.Sc. in Distributed Computing from Brunel University, UK.

Professor Matt Warren is the Head of School at the School of Information System, Deakin University, Australia. He has gained international recognition for his scholarly work in the areas of Information Security, Risk Analysis, Electronic Commerce and Information Warfare. He has authored/co-authored over 180 books, book chapters, journal and conference papers.

Stuart Weinstein is Associate Head, University of Hertfordshire School of Law. Stuart earned his BA from Williams College, Massachusetts, his JD from Columbia University School of Law, and an MBA from University of Hertfordshire Business School. He is a member of the Bars of California, New York and DC and a Solicitor of Senior Courts of England & Wales.

A Non-Militarised Approach to Cyber-Security

Andrew Adams[1], Pauline Reich[2] and Stuart Weinstein[3]
[1]Meiji University, Tokyo, Japan
[2]Waseda University, Tokyo, Japan
[3]University of Hertfordshire, Hatfield, UK
aaa@meiji.ac.jp
pcreich@yahoo.com
weinstein_stuart@yahoo.com

Abstract: In 2011 cyberspace came under highly visible military threat. This threat was not cyber-attack by governments or terrorists, but the threat of a militaristic approach to cyber-security. The US and UK military establishments (among others) made strong arguments about the need to expand their online presence from use of the Internet for their own information transmission and into cyber-attack capabilities. Responding to claims of the Russian and Chinese governments sponsoring cracking attacks against Estonia, Georgia and Google, cyberspace in 2011 became the fifth arena of warfare (land, (under)sea, air, space and now cyberspace). Although development of the basic concept and protocols of the Internet was funded by DARPA, a military research agency, the military and civilian uses of Internet systems rapidly diverged in the early days. This separation allowed the development of a free, generative and borderless Internet whose base flexibility and civilian orientation made it one of the core technologies of modern life by 2011. Just as it has become an essential platform for legitimate activity, illegitimate activity has also flourished online. The very automation which makes computers and the Internet so valuable can also be utilised for negative purposes such as Denial of Service Attacks, malware distribution and fraud. There are claims that some governments are sponsoring attacks and cyber-espionage against their enemies (other states or large corporations), and claims about the rise and dangers of cyber-terrorism. Military forces, faced with a diminishing role in preparations for large scale physical conflicts, have begun claiming that civilian cyberspace needs to be (re-)militarised and that the armed forces should be given both the technical tools and the legal rights to conduct not just cyber-defence activities, but offensive cyber-attacks. In this paper we argue from both philosophical and practical standpoints that a pacifist approach to cyber-security is more appropriate. Based on the constitutional pacifism of Germany and Japan, we argue that investment in cyber-defence would be better targetted at improving the physical and electronic infrastructure of the Internet in general (for example, by funding the free distribution of malware signatures to all users or research and development of better technological security tools). This would provide better cyber-security for the citizens of the world than an arms race to develop military cyber-attack capabilities. The borderless and non-geographic topology of the Internet provide little capacity for avoiding collateral damage which, we argue, is likely to prove more costly than the original dangers identified or forecast. Technological measures used within the parameter of laws protecting the privacy, civil rights and civil liberties of citizens and utilized for defensive purposes, along with further research on thwarting cyber-attacks on critical information infrastructures, would be more beneficial and are evaluated in this pacifist context.

Keywords: militarisation of cyberspace, cyberattack, cyberdefence, pacifism

1. Introduction: The military threat to cyberspace

In 2011 the Internet came under a highly visible threat from the military. We do not refer to military forces attacking either the physical or informational infrastructure of the Internet, but of growing claims by the US, UK and other military forces that they should be funded and authorised to conduct cyber-attacks to counter apparent threats to national security or national interests. While the concept of using communications infrastructure for military activity dates back to at least the early 90s, it is only very recently that the militaries of democratic regimes began proposing a clear doctrine of legitimate cyber-offense. In this paper we analyse the validity of these proposals by the military and find them lacking in justification both philosophically and practically. Based upon the constitutional pacifism of Germany and Japan we propose that military assets be focussed on improving cyber-defence capabilities and not authorised to develop or deploy cyber-attack capabilities.

The Internet clearly owes a great deal to the military, and the US military in particular. While other networking and internetworking systems and approaches were developed in the seventies and eighties such as the JANET Coloured Book protocols (Cooper, 2010), it was the open, generative (Zittrain, 2008) nature of the Internet protocols which led to its dominance over walled garden systems (such as GEnie, AOL and CIX) for business and home computers by 2000 and with a similar trajectory appearing for mobile devices by 2010. This open generative nature allows bad actors as well as good to operate online. When the citizens, governments, commercial and non-commercial information systems of a country are under attack, is the military the correct place for them to turn for assistance? Is making the military capable of (counter-)attacking a useful tool for online security, or would focussing solely on cyber-defence be a more useful role for the military in the information age?

1.1 Cyber-attacks

The military case for expanding its sphere of activities to cyberspace is based upon the reality and potential for cyber-attacks. Although in the twentieth century there was a clearly defined principle (not always followed) that military personnel should be deployed primarily against other military personnel in attack and only in defence against civilian targets, this principle has come under pressure in many quarters in the last twenty years. It has been asserted that government-sponsored, or even covertly condoned, cyber-attacks constitute military action (Gorman and Barnes, 2011).

1.1.1 Government sponsored

The evidence so far for cyber-attacks which have been carried out by government operatives or those supported in some way by governments is limited. There were allegations that the government of the Russian Federation was behind the cyber-attacks experienced by Estonian financial institutions (and other targets in Estonia) in 2007. However, Ottis (2008) found that while that government did nothing to help block the attacks, many of which originated in Russia and others of which were probably directed from Russia through third countries, there is no evidence that government agencies were involved in encouraging, aiding or sponsoring such attacks. Attacks on Google which appeared to target Chinese dissidents (Jacobs and Helft, 2010) have never been clearly linked to the Chinese government, although both the apparent targets and the general level of control exerted by the Chinese government on Internet activity imply that these perpetrators had at least some covert backing from the Chinese government, even if they were not directly employed. The Stuxnet worm (Falliere, Murchu and Chien, 2011) is credited by some including (Chen, 2010) as the real start of cyber warfare. Although the origins of the worm are still unknown, the target (Iranian control systems for uranium centrifuges) and the sophistication of the system, including use of inside information from the manufacturer of the target system (Siemens) lead many to infer that only a government organisation such as the US' NSA or the UK's GCHQ would have the capability and *desire to carry out this highly targetted attack.*

1.1.2 Organised crime and hacktivism

Although early cyber attacks were primarily committed by individuals with the goal of personal aggrandisement or simply for the personal (perverse) pleasure of having an impact on the world, more recently a significant proportion of attacks have been linked to organised crime (Gandhi et al, 2011). Hacktivism, cracking attacks aimed at achieving political goals, have also been on the rise, in particular in relation to the group Anonymous.

1.1.3 Retaliation

Cyber-attacks include targetted cracking aimed at entering and copying and/or changing/deleting information on a target site, distributed denial of service attacks and distribution of malware. Retaliatory cyber attacks on the apparent origin of an attack by targetted cracking might cause significant problems for the attacker, but such attacks frequently use intermediate already-cracked systems (often virtual networks of home or ordinary office computers) as proxies to guard their own systems against retaliation. Such proxy sites usually remain in general use for their intended purposes and taking them down to prevent a successful crack could potentially cause more harm than the crack being attempted. In general, if a cracking attack is noticed while in progress, it is easier to shut down or protect the target system than take down the source. If an attack is only noticed after the fact, the chances of tracking the attack back to its real source and not just to an innocent intermediate victim, are relatively small. Distributed denial of service attacks take one of two forms. Botnets of compromised computers under the control of a single perpetrator, or a small group of perpetrators (acting on their own behalf or as agents of a third party) may be used. The constituent parts of these botnets are most often compromised home computers. Successful botnet operators will only use a small fraction of the computing power and network connectivity of their compromised bots to avoid alerting the machine owner to the compromised status of the owner's machine. A counter-attack would require overloading or cracking into a large number of attackers, most of whom are, it must be emphasized, innocent victims themselves. In the case of groups like Anonymous, the machines are acting under the control of their owners/legitimate operators and only in this case does a counter-attack actually affect the perpetrators. Malware infections are very difficult to track back to source and in the interim, attacks against infected machines attempting to spread the infection seems like fighting fire by trying to quell the flames by smothering it with gasoline.

1.2 The military-sponsored case for the right to conduct military cyber-(counter-) attacks

Owens, Dam and Lin (2009) claim that the US military has been gradually building up its cyber-attack capabilities quietly for many years, hiding the budget appropriations for the work in numerous small (in US military terms – millions of dollars) projects. Only in 2011 did the armed forces of the US and other democracies bring forward solid public claims for the resources to develop and the right to deploy cyber attack capabilities. The claims are familiar and mirror other elements of the security-industrial complex that has grown up since 2001. Modern threats are reportedly less likely to come from large conventional forces such as those of the Soviet Union, the People's Republic of China or the Democratic People's Republic of Korea. Instead, the threats are generally from distributed groups such as Al Qaeda, coordinated by and recruiting through electronic communications media, although also featuring in-person small scale training camps. Their attack vectors include guerilla warfare, information operations (propaganda) and cyber-attacks. The military sector claim is that cyber-defence is not adequate, and that only a responsive cyber-attack capability will allow it to ensure the security of each home nation (Lynn, 2010).

2. Pacifism and security

2.1 German constitutional pacifism and cyber-security

The constitution of the federal government of Germany includes the following pacifist elements (Tomuschat, Currie and Kommers, 2010):

> *Article 24 [International organizations]*
>
> *...... (2) With a view to maintaining peace, the Federation may enter into a system of mutual collective security; in doing so it shall consent to such limitations upon its sovereign powers as will bring about and secure a lasting peace in Europe and among the nations of the world.*
>
> *Article 26 [Ban on preparations for war of aggression]*
>
> *(1) Acts tending to and undertaken with intent to disturb the peaceful relations between nations, especially to prepare for a war of aggression, shall be unconstitutional. They shall be made a criminal offense.*

Daalgard-Nielsen (2005) points out that the defeat of Germany in World War II gave rise to two *interpretations by German politicians:*

> *Two major camps representing competing interpretations of the defeat eventually emerged. On the left, a pacifist interpretation took hold and 'never again war' became the rallying cry for a disparate coalition of intellectuals, educators, unionists, politicians and Protestant clergy. In spite of differences in focus and agenda, most of these groups held self-seeking nationalism, excessive militarism, Prussian authoritarianism or crass capitalist materialism responsible for the two world wars. They converged on a principled objection to the use of force in international affairs and argued that Germany, because of its history, carried a special responsibility to work for peace and peaceful conflict resolution.*

> *On the centre-right, the conclusion was different. 'Never again alone' was the precept for Germany's democratization, rehabilitation and reconstruction … it was necessary to pursue a policy that contained no new sources of mistrust, and a firm commitment to the West had to replace the perennial wavering of the past. The notion of Germany as part of a greater Western community had to take the place of excessive nationalism.*

According to Daalgard-Nielsen (2005), the US focus since 9/11 on the dual danger of terrorism and proliferation of weapons of mass destruction has created a different operating environment for traditional German views concerning military activity. She notes that an emphasis on pre-emptive action challenges the composite German working consensus between 'never again go alone' or 'never again at all'. As a result, German politicians have been thrust into a long battle over Germany's military role in the post-Cold War world and forced to choose between Germany's multilateral and diplomatically-oriented strategic culture and the US' military action-oriented strategic culture.

Does the German pacifism policy translate into the cyber-security arena? To a certain extent it does. (Federal Ministry of the Interior, 2011) released a comprehensive cyber-security strategy to create two high-level government agencies devoted exclusively to cyber-security issues: the National Cyber Response Centre and the National Cyber Security Council, which issued Basic Principles of Cyber Security:

> *Cyber security must be based on a comprehensive approach. This requires even more intensive information sharing and coordination. The Cyber Security Strategy mainly focuses on civilian approaches and measures. They are complemented by measures taken by the Armed Forces to protect its capabilities and measures based on mandates to make cyber security a part of Germany's preventive security strategy. Given the global nature of information and communications technology, international coordination and appropriate networks focusing on foreign and security policy aspects are indispensable. This includes cooperation not only in the United Nations, but also in the EU, the Council of Europe, NATO, the G8, the OSCE and other multinational organizations. The aim is to ensure the coherence and capabilities of the international community to protect cyberspace.*

At the heart of the Cyber Security Strategy is the recognition that it "requires the enforcement of international rules of conduct, standards and norms. Only a mix of domestic and external policy measures will be appropriate for the dimension of the problem. Cyber security can be improved by enhancing the framework conditions for drawing up common minimum standards (code of conduct) with allies and partners." It also strongly calls for cooperative activity:

> *In global cyberspace security can be achieved only through coordinated tools at national and international level. At EU level we support appropriate measures based on the action plan for the protection of critical information infrastructures, the extension and moderate enlargement of the mandate of the European Network and Information Security Agency (ENISA) in view of the changed threat situation in ICT ...*
>
> *We will shape our external cyber policy in such a way that German interests and ideas concerning cyber security are coordinated and pursued in international organizations, such as the United Nations, the OSCE, the Council of Europe, the OECD and NATO. An increasingly multilateral approach must be brought in line with the necessity of sovereign evaluation and decision-making powers. In this context, a code for state conduct in cyberspace (cyber code) should be established, which is signed by as many countries as possible and includes confidence building security measures ... NATO must take cyber security appropriately into account in its entire range of responsibilities. We are in favour of the alliance's commitment to establishing uniform security standards, which Member States may also use for civilian critical infrastructures on a voluntary basis, as foreseen in NATO's new Strategic Concept.*

The German Cyber Security Strategy can be seen clearly as fitting well-within the parameters of Articles 24 and 26. In compliance with Article 24, it is based upon the NATO system of mutual collective security. Article 26 appears to be followed in the Cyber Security Strategy in that nowhere does it suggest that Germany will act alone nor make or support “pre-emptive” strikes in the cyber environment. It can be clearly seen that the German Cyber Security Strategy owes its roots to the position of ‘never again alone’ emphasizing above-all the importance of collective security and international cooperation as the keys to combat cyber-attacks.

2.2 Japanese constitutional pacifism

The Japanese constitution is even clearer than the German Federal constitution in explicitly denying Japan the right to make war, even supposedly denying it the right to maintain a military:

> *Article 9.*
>
> *1. Aspiring sincerely to an international peace based on justice and order, the Japanese people forever renounce war as a sovereign right of the nation and the threat or use of force as means of settling international disputes.*
>
> *2. In order to accomplish the aim of the preceding paragraph, land, sea, and air forces, as well as other war potential, will never be maintained. The right of belligerency of the state will not be recognized.*

Despite these elements of its constitution, Japan maintains a significant “self-defence force” with land, sea and air divisions which not only undertake manoeuvres in Japan and on exercise in other countries, but have engaged in peacekeeping operations overseas. The Japanese government has been accused of engaging in “linguistic gymnastics to claim that it technically complies with its pacifist constitution” by (Kaufman, 2008). Despite this, Japanese citizens appear wedded to at least lip-service to a pacifist constitutional stance: “more than 50% of the population has been against revision of Article 9.” (Eyfells, 2010)

Japan’s Defense Ministry Budget Request for FY2012 requests 6 billion yen in funding for “Projects contributing to ensuring information infrastructure, such as improving information security”, specifically improvement of the information security of the Ministry of Defense/Self Defense force, e.g. development of computer protection system; improvement of information and communications infrastructure of MOD/SDF.1

The 2011 Defense of Japan White Paper Part I: Security Environment Surrounding Japan notes that “cyber attacks on the information and communications networks of governments and militaries as well as on important infrastructure significantly affect national security. Japan must continue to pay attention to developments in cyberspace threats.” The Information Strategy for Protecting the Nation of the Ministry of Defense/Japan Self Defense Force effective from 2010-2013 includes the following concepts (Sasaki, 2011):

> *Reinforcing policies and capabilities for responding to cyber-attack incidents*
>
> *Establishing policies to address the new environment*
>
> *-Promotion of policy from the standpoints of nation/users*
>
> *-Reinforcement of international alliances*
>
> *Encouraging organizations to take active rather than passive measure.*

2.2.1 Cyber-attacks and responses in Japan

In September 2011, there were reports (Tabuchi, 2011)of online breaches at “defense contractors including Mitsubishi Heavy Industries, which builds F-15 fighter jets and other American-designed weapons for Japan’s Self-Defense Forces…” which had started in August but were only reported to the Government of Japan in September. Another news source reported that “The IHI Corporation, a military contractor that supplies engine parts for fighter jets, may have also been a target”. It was unknown whether any classified information had been compromised, “but an investigation by a security company has revealed that connections were made to14 overseas sites, including at least 20 servers in China, Hong Kong, the United States and India… China, especially, has vehemently denied that the attack originated from within its borders". Tabuchi (2011) also notes that "Mitsubishi Heavy Industries won 215 contracts worth $3.4 billion from Japan's Defense Ministry in the year ending last March, or a quarter of the ministry's spending that year".

3. Pacifist cyber-defence versus militarized cyber-attack

Historically, military activity operated primarily under a "might makes right" justification, with the divine right of kings providing an appeal to a sovereign decision to deploy military forces as a backup. In the modern (post-WWII, UN) era various international legal norms have emerged surrounding rights to self-defence against attack and, more recently, humanitarian arguments in favour of intervention in places such as Libya. As described in the Introduction, the arguments for permitting military cyber-attack capability development *and deployment* have been made with increasing volume and regularity over the last few years, and the beginnings of such capabilities on a broad scale are being brought into being in both the UK and US at least, with claims that more secretive regimes from Russia to China to the Middle East have already developed and deployed such groups within their military structures. Is the militarized cyber-attack approach justified on either or both philosophical or practical grounds?

3.1 Philosophical arguments

In modern times, in liberal democratic countries whose military is under civilian ultimate authority, military forces are generally deployed primarily to deal with external threats. When deployed internally they are deployed under very strict rules of engagement and almost never on an ongoing basis. In those cases where they have been deployed on home soil on a continuing basis, it has usually become as much a part of the problem as a part of the solution, such as the deployment of UK forces in Northern Ireland from 1969-2007 (Chief of the General Staff, 2006). Short term deployments of home or overseas military personnel are usually undertaken in very limited circumstances of supporting or protecting vital infrastructure[1], relief and recovery from disasters[2] or to support law enforcement in restoring order. Deployments of military personnel for any purpose other than disaster recovery and humanitarian aid are likely to be criticised. Even humanitarian deployments can cause problems as the military mindset, targetted towards security issues, may over-react to minor incidents or include a disregard for collateral damage such as numerous instances in Afghanistan including the killing of multiple civilians in April 2010 as reported by AFP (2009)

As the Internet has become a core element of life for a significant portion of humanity, the concept of cyberspace as a pseudo-geographic new place has been put forward many times (Berry, Kim and Spiegal, 2010)(Kitchin and Dodge, 2011). This same argument is being used by US and other militaries as a justification for funding and authorising the development of offensive capabilities in this new arena, to complement land, sea, undersea, air and outer space capabilities. Unlike the most recent other addition to these geographies of military engagement (outer space) cyberspace has already been massively colonised by ordinary citizens. This colonisation has included some of the best and some of the worst aspects of human activity, ranging from the massive collaborative efforts to improve access to knowledge represented by projects such as Wikipedia, to the deluge of spam and fraud that has rendered certain forms of communication useless and threatened to overwhelm others.

Just as the Outer Space Treaty[3] places severe restrictions on the use of celestial bodies, earth orbit or further into space for conducting aggressive military actions, so should the principle that cyber-defence should not be permitted to morph into militarized cyber-attacks be adopted. There are already too many forces conducting cyber-attacks on the Internet, poisoning the general information environment in terms of both volume and specific protocol misuse. Legitimising attacks by the military in democratic nations risks generating far more heat than light, and distracts from a potentially useful

[1] Such as the use of army personnel and equipment to replace civilian firefighting crews on strike in the UK, most recently in 2002 (Rayment, 2002).

[2] Such as the deployment of US Marines and Japanese Land Self-Defence Force after the Great North Eastern Earthquake in 2011.

[3] Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies www.oosa.unvienna.org/oosa/SpaceLaw/outerspt.html

role for the military in developing and distributing cyber-defence benefits to the civilian (individual and organisational, commercial, governmental and non-profit). If military activity in cyberspace is targeted at improving the defence mechanisms in the information infrastructure, then military developments should be shared and used by all. If the focus of the military is on, or even just includes, mainly or exclusively offensive capabilities, then not only will defensive capabilities be less well developed, but the military will seek to limit dispersal of its defensive developments until and unless they have offensive capabilities which can overcome those same defences. We need only consider the responses of the US and UK governments to cryptography (Levy, 2001) to see how such attitudes can play out, when the offensive (code-breaking) capabilities of state actors are prioritised over defensive (code-making) capabilities of citizens. Without the eventual capitulation of governments to the dissemination of strong encryption software, online commerce, online banking and many other elements of the Information Age would not exist or would be severely hampered.

3.2 Practical arguments

The purpose of the military in a liberal democratic state is to provide a structure within which individuals sublimate their individual well-being and autonomy to the state. In extreme cases military personnel are ordered to take actions which will almost certainly result in their deaths in order to achieve the (hopefully legitimate) goals of the government in deploying military forces. The equipment used by military forces is also such that, while necessary for applications of force in pursuit of those policies, it would be unwise to allow access to such equipment by ordinary citizens not under strong oversight and control. In considering the claims of senior military commanders that they should be authorised and funded to develop offensive cyberspace capabilities, the question needs to be asked as to whether either of these justifications holds for cyber attack capabilities. Such a case has not been clearly made, although some such as Kesan & Hayes (2011) have presented some supportive arguments. Even they, however, admit that the dangers of collateral damage are significant and may “be viewed as violations of the Law of War’s principles of distinction and proportionality”.

Cyber-attacks, like cyber-defence, are a matter either of mass computing power and bandwidth or of high levels of skill combined with dedicated action. Are there any benefits to locating cyber-attack or cyber-defensive capabilities within military organisations rather than within civilian agencies, ranging from the national security/intelligence services (the US' NSA and the UK's GCHQ are already major players in this field) to academic groups such as CMU's original CERT team (still in operation and still a major international player in worldwide Internet security despite the creation in 2003 of the Department of Homeland Security's US-CERT group under license from CMU's CERT for the use of the name). The existing expertise of such groups, and their record of cooperative working with the appropriate industries and the academic and commercial research communities, such as the development and release under a free software license of the SELinux (security enhanced GNU/Linux) system by the NSA, suggest that increased funding for such non-military bodies would be more effective overall in improving cyber-security. There is a justifiable role for the military in providing a quickly deployable set of experts to defend security and infrastructure interests against sustained cyber-attack, and to provide highly hardened hosting infrastructure to which systems under attack might be moved. Such capabilities can be regarded as the information systems equivalent of military forces providing logistical expertise and assets for disaster recovery.

4. Conclusions

It is clear that many non-democratic states regard cyberspace as a venue in which offensive military action is useful and “legitimate” (Liang & Xiangsui, 1999). There has been a slow covert development of offensive capabilities as a response by the military in some democratic states, which has now been made public; legitimation has been sought for current and future development and deployment. We have discussed above whether cyber attacks are at all legitimate in general in democratic states and, if so, whether they should be under the control of military or national security agencies, or both, as is the case in the United States. In our opinion, cyber attack is rarely useful and even when it is useful, its control by military agencies (given their mindset and immunity from most repercussions for collateral damage) is less useful than the more restricted usage likely to be made by national security/intelligence agencies. As we have seen, Japan’s responses to cyber attacks up to now have shown it behaving with self-restraint in light of offensive attacks. The published stance of the German federal government stands in agreement with this approach, in marked contrast to the belligerent statements of the US and UK governments and their militaries. The emphasis in a recent announcement by METI (2011) called for such measures as precautions to be taken by individuals to combat targeted attacks and “society-wide measures to prevent spread of damage.” It also places responsibility on corporations for taking information security measures and calls for “strict implementation of information security measures”. Further research is necessary to determine whether Japan can lead the rest of the world in more peaceful ways to deal with cyber-attacks of varying levels of severity, and whether the German and Japanese examples of constitutionally-inspired restraint can hold back the military threat of constant cyber-war, with the ordinary user's activities and machines the likely collateral damage.

References

AFP 2009. Anger after Afghan family killed in US raid. http://www.google.com/hostednews/afp/article/ALeqM5gXoRYVRJh9SXLK7hn3e1gST3ogOQ

Berry, Chris & Kim, Soyoung & Spigel, Lynn (Eds), 2010. Electronic Elsewheres. University of Minnesota Press, Minneapolis, MN.

Chen, Thomas M., 2010. Stuxnet, the Real Start of Cyber Warfare?, IEEE Network November/December, 2-3.

Chief of the General Staff, 2006. Operation Banner, An Analysis of Military Operations in Northern Ireland. Army Code 71842.

Cooper, Christopher S., 2010. JANET The First 25 Years. The INT Association, Didcot.

Daalgard-Nielsen, Anja, 2005. Germany, Pacifism and Pre-emptive Strikes, Security Dialogue 36(3), 339-359.

Eyfells, Eyjolfur, 2010. Japan's Security Dilemma, MA Thesis, Haskoli Islands. http://skemman.is/stream/get/1946/6120/17507/1/Eyjolfur_Eyfells_MA_Thesis_fixed.pdf

Falliere, Nicholas & Murchu, Liam O. & Chien, Eric, 2011. W32.Stuxnet Dossier, Symantec White Paper.

Gandhi, R. & Sharma, A. & Mahoney, W. & Sousan, W. & Zhu, Q. & Laplante, P., 2011. Dimensions of Cyber-Attacks: Cultural, Social, Economic, and Political, IEEE Technology and Society Magazine 30(1). http://www.rogerclarke.com/EC/05725605.pdf

German Federal Ministry of the Interior, 2011. Cyber Security Strategy for Germany. http://www.bmi.bund.de/SharedDocs/Downloads/DE/Themen/OED_Verwaltung/Informationsgesellschaft/cyber_eng.pdf

Gorman, Siobhan & Barnes, Julian. E., 2011. Cyber Combat: Act of War, Wall Street Journal, 31st May. http://online.wsj.com/article/SB10001424052702304563104576355623135782718.html

Jacobs, Andrew & Helft, Miguel, 2010. Google, Citing Attacks, Threatens to Exit China, New York Times, 13th January.

Kaufman, Zachary D., 2008. No Right to Fight: The Modern Implications of Japan’s Pacifist Postwar Constitution, Yale Journal of International Law 33(1), 266-273.

Kesan, Jay P. & Hayes, Carol M. 2011. Self Defense in Cyberspace: Law and Policy. Telecommunications Policy Research Conference 2011. http://ssrn.com/abstract=1979857.

Kitchin, Rob & Dodge, Martin, 2011. Code/Space: Software and Everyday Life. MIT Press, Cambridge, MA.

Levy, Steven, 2001. Crypto: Secrecy and Privacy in the New Code War. Allen Lane, London.

Liang, Q. & Xiangsui, W., 1999. Unrestricted Warfare. Foreign Broadcast Information Service, Langley, VA. http://www.cryptome.org/cuw.htm

Lynn, WIlliam J. III, 2010. Defending a New Domain: The Pentagon's Cyberstrategy, Foreign Affairs Sept/Oct, 97-108. http://www.foreignaffairs.com/articles/66552/william-j-lynn-iii/defending-a-new-domain

METI, 2011. Announcement and implementation of information security measures based on recent trends. http://www.meti.go.jp/english/press/2011/0527_02.html

Ottis, R., 2008. Analysis of the 2007 cyber attacks against Estonia from the information warfare perspective, Proceedings of the 7th European Conference on Information Warfare and Security, 163-168.

Owens, William A. & Dam, Kennet W. & Lin, Herbert S. (eds), 2009. Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities. National Academies Press, Washington, DC.

Rayment, Sean, 2002. 30,000 soldiers on alert to step in for striking firemen, The Telegraph, 4th August. http://www.telegraph.co.uk/news/uknews/1403508/30000-soldiers-on-alert-to-step-in-for-striking-firemen.html

Sasaki, Takahiro, 2011. Cyber Security of MOD/JSDF and Regional Cyber Security from Japanese Perspective, Regional Collaboration in Cyber Security and Cyber Terrorism Conference, 4-5.

Tabuchi, Hiroko, 2011. U.S. Expresses Concern about New Cyberattacks in Japan, The New York TImes, 22nd September. http://www.nytimes.com/2011/09/22/world/asia/us-expresses-concern-over-cyberattacks-in-japan

Tomuschat, Christian & Currie, Daivd P. & Kommers, Donald P., 2010. Basic Law for the Federal Republic of Germany, https://www.btg-bestellservice.de/pdf/80201000.pdf.

Zittrain, Jonathan, 2008. The Future of the Internet (and How to Stop It). Allen Lane, London.

Matching Attack Patterns to Security Patterns Using Neural Networks

Adetunji Adebiyi, Johnnes Arreymbi and Chris Imafidon
School of Architecture, Computing and Engineering, University of East London, London, UK
adetunjib@hotmail.com
J.Arreymbi@uel.ac.uk
C.O.Imafidon@uel.ac.uk

Abstract: The issue of information systems security and its consequences have raised so much concern in many public and private domains, and as hackers attacks on software continues to increase, the demand for secure software has also increased significantly. The industry has been looking at better ways of integrating security into every phase of software development lifecycle (SDLC). The use of security pattern is one of the ways that have been proposed in this research area to help software developers to integrate security into software application during development. However due to different types of security patterns and their taxonomy that have been developed, software developers are faced with the challenge of finding and selecting appropriate security patterns that addresses the security risks in their design. One of the solutions addressing this problem as proposed by Wiesauer, and Sametinger (2009), involves matching attack patterns identified in the software design to security patterns. This research investigated this area by matching attack patterns to security patterns using neural networks and finding how the outcome could be used to enhance software systems security at the design and development stages. The result of performance of the neural network and the advantage of this approach is presented in this paper. This study found that attack patterns can be matched to their corresponding security patterns using a neural network that has been trained for this purpose. Therefore, software developers using the trained neural network as a tool can easily be guided into selecting the appropriate security patterns meeting the security requirements of their software application.

Keywords: security pattern, attack pattern, neural network

1. Introduction

In today's hostile computing environment, the role software plays in ensuring the security of critical data and resources is very important. With the increase of attacks aimed directly at software applications the need for software to defend itself and continue functioning normally is of much interest to the software industry. However, when software applications are developed without security in mind, software attackers take advantage of the security flaws found in them to mount multiple attacks when they are deployed. As reliance on network security and other host based security measures for protecting software no longer provide adequate security, a new research field called software security emerged in the last decade with the aim of building security into software application during development. This approach views security as an emergent property of the software and much effort is dedicated into weaving security into the software all through software development lifecycle (SDLC) (McGraw, 2003)

Reportedly, 50% of security problems in software products today have been found to be design flaws (Halkidis et.al, 2006, McGraw, 2006). In this view, many authors argue that it is much better to find and fix flaws during the early phase of software development because it is more costly to fix the problem at a late stage of development and much more costly when the software has been deployed (Spampinato et.al, 2008, Mockel and Abdallah, 2011, Gegick and Williams, 2007).

However, as software developers often lack the knowledge of security, this necessitates the involvement of security experts who will help in identifying threats in the software during development. This poses a challenge because of the existing gap between security professionals and software developers. The disconnection between the two has led to lack of critical understanding of current technical security risks in software development efforts (Pemmaraju and McGraw, 2000). This is because security related activities such as threat modelling and risk analysis has historically been in the domain of security experts who have over the years observed system intrusions, dealt with malicious attackers and have studied software vulnerabilities (Kenneth, Wyk and McGraw, 2006) On the other hand, software developers trained to think of functions and features of their product and its on-schedule delivery often become frustrated with the added burden of security which is not their only or primary concern (Kienzle and Elder, 2003).

To reduce this knowledge gap, various techniques has been suggested that would enable software developers who are not necessarily experts in security to integrate security into software during the early phase of SDLC. One of these techniques is the use of security patterns that captures security expertise solutions to recurring security problems which help software developers to make informed decisions between security and other goals (Blakley, et.al 2004). In a similar way to design patterns which were developed and used for identifying and presenting solutions to reoccurring problems in object oriented programming, security patterns were developed to enhance the elimination of security flaws inserted into software during development and also provides mitigation for known vulnerabilities(Halkidis, et.al, 2006). As other security best practices approaches often burden the software developer with a list of what to do and what not to do during software development, the advantage of security patterns is that they provide constructive assistance with their inherent solutions and guidance on how they are to be applied (Kienzle and Elder, 2003).

While there are many security patterns that have been developed, software developers still have to find out which security patterns adequately address the security problem in their software design. This paper address this problem by proposing a tool based on neural networks that will match possible attack patterns identified in software design to corresponding security patterns that can help in mitigating the security risk.

2. Security patterns

Several works have been done based on security patterns since the pioneer work of Yoder et.al, 1997 who applied design patterns to specific security issues. Different security patterns have been developed by many authors in different context and with different definitions on security pattern (Halkidis, et.al, 2006). However, most authors define security patterns as patterns describing particular recurring security problem in specific context and presents a well-proven solution to it (Sametinger and Wiesauer, 2009, Laverdiere et.al, 2006, Kiiski, 2007, Kienzle and Elder, 2003, Blakley, et.al, 2004).

One of the benefits of security patterns is that it provides an effective way for software developers who are not expert in security to learn from security experts. Since security patterns documents proven solutions to recurring problems in a well-structured manner that is familiar and easily understood by software developers it also enhances reusability of the patterns (Hafiz and Johnson, 2006). Therefore by using security pattern, software developers who are not expert in security are able to expand their security focus from low level implementation to high level architectures (Schumacher, et.al, 2006)

In previous research, many authors describe security patterns for different purposes. This includes security patterns for web applications (Steel, et.al, 2005, Kienzle and Elder, 2003), security patterns for mobile Java code (Mahmoud, 2000), security patterns for agents systems (Mouratidis, et.al, 2003), Security patterns for Voice over IP (VoIP) (Fernandez, et.al, 2007) and security pattern for capturing encryption-based access control to sensor data (Cuevas, et.al, 2008). Also, to enable software developers to choose the appropriate security pattern addressing the security risks in their designs, several authors have proposed different classification schemes for security patterns. This include classification based on applicability (Blakley, et.al, 2004), classification based on product and process (Kinezle and Elder, 2003) classification based on logical tiers (Steel, et.al, 2005), classification based on application domain (Bunke, et.al, 2011) classification based on security concepts (Hafiz and Johnson, 2006), classification based on system viewpoints and interrogatives (Zachman, 1987), classification based on confidentiality, integrity and availability (CIA) model (Hafiz and Johnson, 2006) and classification based on attack patterns (Wiesauer and Sametinger, 2009)

3. Research approach

In this paper, we propose a new approach in which neural network is used to suggest possible solutions to possible attack patterns identified during the evaluation of a software design. This approach builds on the proposed selection criterion by Wiesauer and Sametinger (2009) which matches attack patterns to security design patterns. The authors argue that there is a need for a selection criterion because software developers who are not experienced in security are unable to select and apply security pattern in a correct and effective manner. Therefore by using their proposed taxonomy the authors state that software developers can match identified attack patterns in their software designs to corresponding security design patterns that would provide the appropriate mitigation. In our proposed approach, we achieve this by abstracting data from the 51 regularly expressed attack patterns by William and Gegick (2006) for training the neural network to match the attack patterns to security patterns that can provide mitigation to the threats in the attack patterns. Microsoft threat classification scheme (STRIDE) was also used to classify the attack patterns into six groups in order to align the attack patterns to their corresponding threat category. The data abstracted from the attack patterns formed the attributes of the attack patterns that were used in training the neural network. The attributes consists of:

The Attack ID: This is the unique ID that identifies the attack

Resource Attacked: This is the resource that is attacked in the attack pattern.

Attack Vector: This is method through which the attacker uses to attack the resource

Attack Type: This state whether the attack is an attack against confidentiality, integrity or availability

3.1 The neural network architecture

A feed-forward back-propagation neural network is used to analyse the attack patterns and generate possible solutions from the security design patterns that can be help in mitigating the threat identified in the attack patterns. The standard three layer neural network architecture consisting of the input layer, the hidden layer and the outer layer was adopted in training the neural network. To optimize the performance of the neural network, Resilient Back-propagation (RP) training optimization algorithm was applied. With respect to the transfer functions, a tan-sigmoid transfer function was applied to the various connection weights in the hidden nodes and output nodes. Since a supervised learning architecture (i.e. back-propagation) was adopted for training the neural network, the data discussed in section 3.2 was used for its training.

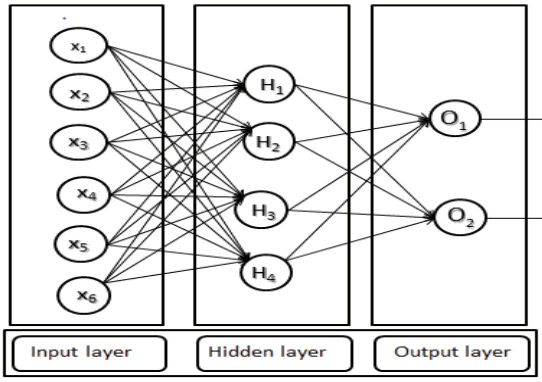


Figure 1: Neural network architecture

3.2 Data collection

The regularly expressed attack patterns by William and Gegick (2006) was used as the data source for training the neural network. To align the regularly expressed attack patterns to the threats that can exploit the security flaws they represent, Microsoft threat classification scheme (STRIDE) was used to classify them into six groups according to their corresponding threat category. Figure 2 shows that out of the 52 regularly expressed attack patterns, 1 of them was classified under spoofing identity attack category, 2 was classified under tamper with data attacks, none was classified under repudiation attacks, 6 was classified under the information disclosure attacks, 21 was classified under the denial of service attacks and 27 was classified under the elevation of privilege attacks. No attack was classified under repudiation attacks because none of the regularly expressed attack patterns demonstrated this type of attack. However, it was assumed that this attack was covered under the elevation of privilege attack because the attacker must have escalated his privileges before been able to cover his tracks in a multi-stage attack scenario.

Data was also collected from the security design patterns defined by Steel, et.al (2005), Blakley, et.al (2004) and Kinezle and Elder (2003). A total of 23 security design patterns were defined by Steel, et.al (2005). These were classified into four logical tiers consisting of the web tier, the business tier, web services and identity tier. A total of 13 security design patterns were defined by Blakley, et.al

(2004) and these were classified into two groups. This consisted of the available security design patterns and the protected security design patterns. The security design patterns defined by Kinezle and Elder (2003) were also classified into two categories. These include the structural patterns and procedural patterns. The structural patterns consist of 13 main security design patterns and 3 mini-patterns. The mini-patterns are less formal and shorter discussion that were included as a supplement to the main security design patterns. The procedural patterns consist of 13 security design patterns.

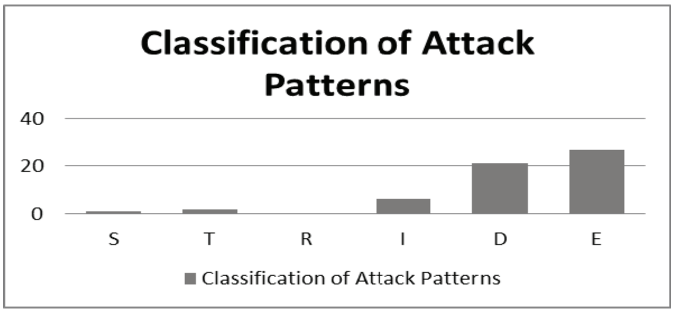


Figure 2: Number of attack patterns classified according to STRIDE

There are other security design patterns that have been defined by other authors different from the ones highlighted above For this reason, a decision had to be made on which security design patterns to be analysed for abstracting the data needed for training the neural network. The decision to use the security design patterns defined by Steel, et.al (2005), Blakley, et.al (2004) and Kinezle and Elder (2003) base on the following reasons:

- Security design pattern by Blakley, et.al (2004) was initiated by the Open Group Security Forum in a coordinated effort to resolve the problem of lack of clear definition of security design patterns. The security design patterns were defined based on a comprehensive list of existing security design pattern to be used as a guide by software developers.
- There is an existing research work by Halkidis, S.T. et al. (2006) in which the security design by Blakley, et.al (2004) was analysed qualitatively which provided insight into this research work
- Since security design patterns have been defined for different purposes, the security design patterns by Steel, et.al (2005) and Kinezle and Elder (2003) were chosen because they both address web related security issues.

3.3 Data encoding

In order to encode the data needed for training the neural network, the collected data were initially analysed. The information from the analysis on the attack components in the regularly expressed attack patterns was used to encode the input data. Table 1 shows the attributes of the regularly expressed attack patterns used in encoding the input data to the neural network

Table 1: Attributes of regularly expressed attack patterns

s\no	Attribute	Observable	Value
1	Attack ID	Attack Pattern	Attack ID
2	Resource Attacked	Attack Component	Attack Component ID
3	Attack Vector	Attack Component	Attack Component ID
4	Attack Type	Availability	1
		Integrity	2
		Confidentiality	3

The taxonomy of security design patterns by Wiesauer and Sametinger (2009) was based on the description of the attack patterns in Common Attack Pattern Enumeration and Classification (CAPEC) catalogue and the intent and purpose of the security design patterns. The authors stated that since the classification of the attack patterns in CAPEC catalogue is based on STRIDE, their proposed taxonomy on security design patterns could be considered as classification based on STRIDE as well. Building on this approach, the security design patterns defined by Steel, et.al (2005), Blakley, et.al (2004) and Kinezle and Elder (2003) were analysed. From previous research by Halkidis, S.T. et al. (2006), it was observed that security design pattern by Blakley, et.al (2004) was analysed qualitatively using Microsoft threat classification (STRIDE) to find out the security design pattern that provides protection on each of the threat category. The result of the analysis is shown on Table 2 and was used as part of the data needed for training the neural network.

Table 2: Classification of Security Design Pattern by Blakley, et.al (2004)

s\no	Security Pattern	S	T	R	I	D	E
1	Check pointed System					X	
2	Standby					X	
3	Comparator- Check Fault – Tolerant System					X	
4	Replicated System					X	
5	Error/ Detection/Correction					X	
6	Protected System	X	X		X		X
7	Policy	X	X		X		X
8	Authenticator	X	X		X		X
9	Subject Descriptor						
10	Secure Communication	X	X		X	X	X
11	Security Context		X		X		X
12	Security Association	X	X		X		X
13	Secure Proxy	X	X		X		X

In a similar manner, security design patterns defined by Steel, et.al (2005), and Kinezle and Elder (2003) were analysed. During the analysis of security design patterns by Kinezle and Elder (2003), the procedural patterns were not analysed because they consisted security patterns which were not implemented in the software application. Procedural patterns were defined for the purpose of improving the development process of mission critical software applications. They can impact the management of a software development project when adopted by software developers. Table 3 and Table 4 show the threat category that the security design patterns were classified after the analysis. It would be noticed that the secure assertion and dynamic service management on Table 3 and Table 4 respectively were not classified under any category. These security design patterns are related to each other and could not be classified under any threat category because they only provide monitoring and reporting of the system events but do not offer protection against STRIDE attacks

Table 3: Classification of security design pattern by Kinezle and Elder (2003)

s\no	Structural Patterns	**S**	**T**	**R**	**I**	**D**	**E**
1	Account Lockout	X	X		X		X
2	Authenticated Session	X	X		X		X
3	Client Data Storage				X		X
s\no	Structural Patterns	**S**	**T**	**R**	**I**	**D**	**E**
4	Client Input filters		X		X	X	X
5	Directed Session (M)		X				
6	Hidden Implementation (M)				X		
7	Encrypted Storage				X		X
8	Minefield	X		X			
9	Network Address Blacklist	X				X	
10	Partitioned Application						X
11	Password Authentication	X	X		X		X
12	Password Propagation	X	X		X		X
13	Secure Assertion						
14	Server Sandbox		X		X		X
15	Trusted Proxy	X	X		X		X
16	Validated Transaction		X				

Table 4: Classification of security design pattern by Steel, et.al (2005)

s\no	**Security Pattern**	**S**	**T**	**R**	**I**	**D**	**E**
1	Authentication Enforcer	X	X		X		X
2	Authorization Enforcer	X	X		X		X
3	Intercepting Validator		X		X		X
4	Secure Base Action	X	X		X	X	X
5	Secure Pipe	X	X		X	X	X
6	Secure Service Proxy	X	X		X		X
7	Secure Session Manager	X	X		X		X
8	Intercepting Web Agent	X	X		X		X
9	Secure logger		X	X	X		
10	Audit Interceptor		X	X	X		
11	Container Managed Security	X	X		X		X

s\no	**Security Pattern**	**S**	**T**	**R**	**I**	**D**	**E**
12	Dynamic Service Management						
13	Obfuscated Transfer Object				X		
14	Policy Delegate				X	X	
15	Secure Service Façade	X	X		X		X
16	Secure Session Object	X			X		X
17	Message Inspector Gateway	X	X	X	X		X
18	Secure Message Router	X			X		
19	Message Inspector	X	X	X	X		
20	Assertion Builder	X					
21	Credential Tokenizer	X		X			
22	Single Sign On (SSO) Delegator	X			X		
23	Password Synchronizer	X			X		

Following the analysis of the data collected on the regularly expressed attack patterns and the security design patterns, the data needed for training the neural network was encoded. A total of 226 training data samples were abstracted from the regularly expressed attack patterns using the attributes in Table 1. To encode the data, the corresponding value for the information abstracted by each attribute in the Table was used in encoding the data. For instance, regularly expressed attack pattern 1 is represented as:

(User^{+})(Server^{+})(Log^{+})(HardDrive^{+})

Based on the analysis of the data collected on this attack pattern, the information on Table 5 was abstracted from attack pattern 1 using the regularly expressed attack pattern attributes.

Table 5: Sample of Pre-processed training data from attack pattern

Attack ID	Resource Attacked	Attack Vector	Attack Type
1	Hard Drive	Log	Availability

In order to encode the information abstracted in the Table 5, the attack component ID for Hard Drive and Log was used for their encoding. The corresponding value for *Availability* in Table 1 was also used for its encoding. Table 6, shows the training data for the example above after it has been encoded.

Table 6: Sample of training data after encoding

Attack ID	Resource Attacked	Attack Vector	Attack Type
1	42	58	1

The next stage involves converting the encoded data into ASCII comma delimited format which can be used to train the neural network as shown below

1, 42, 58, 1

The data is then loaded into the neural network for training as shown in the following Table.

Table 7: Sample of input data into neural network

Input 1	Input 2	Input 3	Input 4
1	42	58	1

For the expected output, security design patterns by Blakley, et.al (2004), Steel, et.al (2005), and Kinezle and Elder (2003) were grouped into six groups with respect to STRIDE. Each group provides possible solutions to the threats identified under each threat category of STRIDE. A unique ID is assigned each group so that the neural network can match them to the corresponding attack patterns. Based on this encoding, the neural network is expected to identify the possible solution for the attack pattern by giving the following output:

1, 0, 0, 0, 0, 0

3.4 Neural network training

To train the neural network the training data set is divided into two sets. The first set of data is the training data sets (201 Samples) that were presented to the neural network during training. The second set (26 Samples) is the data that were used to test the performance of the neural network after it had been trained. The training performance is measured by Mean Squared Error (MSE) and the training stops when the generalization stops improving or when the 1000th iteration is reached.

Mat lab Neural Network tool box was used to perform the training. The training parameters also include the learning rate which is set to 0.01 with a goal of 0; maximum fail set to 6 and a minimum gradient of 0.000001

4. Result, analysis and discussion

The training was executed in five simulations to obtain the average results of its performance because the neural network is initiated with random weights during its training and this gives different results. Therefore, the average results on the training time, MSE and number of epoch were also used in analysis of the performance of the neural network.

Table 8: Actual and expected output of neural network

s\n	Test Data Sample	Actual Output	Expected Output
1	Sample 1	6.0000	6
2	Sample 2	5.9999	6
3	Sample 3	5.0000	5
4	Sample 4	4.9998	5
5	Sample 5	5.0000	5
6	Sample 6	5.0000	5
7	Sample 7	5.9999	6
8	Sample 8	6.0000	6
9	Sample 9	5.0000	5
10	Sample 10	2.6441	6
11	Sample 11	5.0000	5
12	Sample 12	5.0000	5
13	Sample 13	6.0000	6
14	Sample 14	4.9183	5
15	Sample 15	6.0000	6
16	Sample 16	5.0000	5
17	Sample 17	6.0000	2
18	Sample 18	1.7707	2
19	Sample 19	5.6890	6
20	Sample 20	4.0000	4

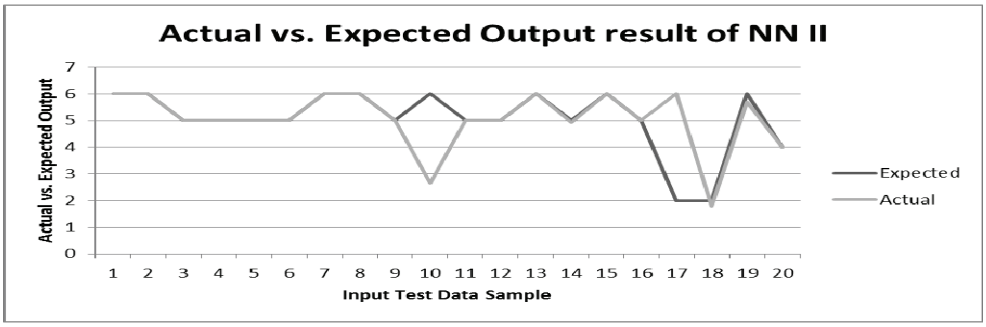


Figure 3: Actual vs. expected output of NN II

The performance of neural network was also tested using the test data. Table 8 shows the actual and expected output of the neural network. By comparing the expected and actual output, it would be seen that the network was able to match most the attack patterns to correct the group that will provide mitigation to vulnerabilities in the attack pattern. The output of neural network matched the expected output after approximation (i.e. -0.5 <X> 0.5: where X is the group ID) There were two instances in which the network failed to match the attack patterns to the correct group. This was when the network was used to evaluate test data sample 10 and 17. For test data sample 10, the network produce an output of 2.6441 when the expected output is 6 and for test data 17, the network produced an output of 6 when the expected output is 2. By looking at the data used in training the network for matching the attack patterns to their corresponding security pattern, it was seen that for these attack patterns, the attacker had multiple ways in which the attack could be carried out. This explains why the network failed to match the attack patterns. With a larger data sample for training the neural network, a better performance can be achieved. Figure 3 is a graph showing the difference between the actual and expected output

5. Future work

The regularly expressed attack pattern used in training the neural network is a generic classification of attack patterns. Therefore, any unknown attack pattern introduced to the neural network will be classified to the closet group of security design pattern that can provide mitigation. As the neural network has been trained to match attack patterns to three security design patterns, further work is required to train the network to match attack patterns to other security design patterns. In addition, the neural network needs to be thoroughly tested before it can gain acceptance as a tool for matching attack patterns to security design patterns.

6. Conclusion

This paper has demonstrated how neural network can be used has a tool to match attack patterns to security patterns. The main advantage of this approach is that software engineers who are novices in security can be aided by using neural network in this capacity to get expert solutions to security holes in their software design before it is coded. Previous research works have shown that the cost of fixing security flaws in software applications when they are deployed is 4–8 times more than when they are discovered early in the SDLC and fixed. The result of the evaluation shows that the neural network was able to match the identified attack patterns to the group of security design patterns that can provide mitigation for the attacks. Based on the information given by the proposed neural network tool, a software developer can make informed decision on what to do in order to integrity security into his software design.

References

Blackley, B. et.al (2004) 'Technical Guide Security Design Patterns', The *Open Group,* Available at: http://users.uom.gr/~achat/articles/sec_patterns.pdf (Last Accessed: December 2011)

Bunke, M., et.al, (2011) 'Application-Domain Classification of Security Patterns, In: *The Third International Conferences on Pervasive Patterns and Applications,* Rome, Italy, pp 138-143

Cuevas, A. et.al, (2008) 'Security Pattern for Capturing Encryption-Based Access Control to Sensor Data', In: *Proceedings of The Second International Conference on Emerging Security Information, Systems and Technologies,* Cap Estere, pp62-67

Fernandez, E.B et.al, (2007) 'Security patterns for Voice over IP' *Journal of Software,* Vol.**2**(2), 19-29.

Gegick, M. and Williams, L. (2006) 'On the design of more secure software-intensive systems by use of attack patterns', *Information and Software Technology*, Vol. **49**, pp381-397

Halkidis, S.T. et.al (2006) A qualitative analysis of Software security patterns, *Computer & Security*, Vol. **25**, pp379-392

Hafiz, M. and Johnson, E.(2006) 'Security Patterns and their Classification Schemes' *Technical Report for Microsoft's Patterns and Practices Group,* Available at: http://munawarhafiz.com/research/patterns/secpatclassify.pdf (Last Accessed: December, 2011)

Kenneth, R., Wyk, V., and McGraw, G. (2005) 'Bridging the Gap Software Development and Information Security', *IEEE Security & Privacy,* Vol. **3**(5), pp. 75-79,

Kienzle, M. D and Elder, M.C. (2002) 'Final Technical Report: Security Patterns for Web Application Development', *Defense Advanced Research Project Agency (DARPA),* Available at: http://www.scrypt.net/~celer/securitypatterns/final%20report.pdf (Last Accessed: November 2011)

Kiiski, L (2007) 'Security Patterns in Web Applications', *Publications in Telecommunications Software and Multimedia Laboratory,* Available at: http://www.tml.tkk.fi/Publications/C/25/papers/Kiiski_final.pdf (Last Accessed: November 2011)

Laverdiere, M. A, et.al. (2006) 'Security Design Patterns: Survey and Evaluation', In*: Proceedings of Canadian Conference on Electrical and Computer Engineering (CCECE) 06,* Ottawa, Ont, pp1605- 1608

Mahmoud, Q. (2000) 'Security Policy: A Design Pattern for Mobile Java Code' In: *Proceedings of the Seventh Conference on Pattern Languages of Programming (PLoP' 00)*, Illinois, USA.

McGraw, G. (2003), 'Building Secure Software. A difficult but critical step in protecting your business'*, Citigal, Inc*, Available at: http://www.cigital.com/whitepapers/dl/Building_Secure_Software.pdf (Last Accessed: November 2011)

McGraw, G. (2006)'The Role of Architectural Risk in Software', *Inform IT Network*, Available at: http://www.informit.com/articles/article.aspx?p=446451(Last Accessed: November 2011)

Mockel C and Abdallah, A.E (2011) 'Threat Modelling Approaches and Tools for Securing Architectural Designs of E-Banking Application', Journal of Information Assurance and Security', Vol. **6**(5), pp 346-356

Mouratidis, H. and Giorgini, P (2007) 'Security Attack Testing (SAT)- testing the security of information systems at design time', Information Systems, Vol. **32**, p1166- p1183

Pemmaraju, K., Lord, E. and McGraw, G.(2000) 'Software Risk Management. The importance of building quality and reliability into the full development lifecycle', Available at: http://www.cigital.com/whitepapers/dl/wp-qandr.pdf, (Last Accessed: July 2011)

Schumacher, et.al, (2006) 'Security Patterns: Integrating Security and System Engineering' *John Wiley & Sons, Ltd,* Chichester UK

Spampinato, D. G. (2008), ‘SeaMonster: Providing Tool Support for Security Modelling’, NISK Conference, Available at: http://www.shieldsproject.eu/files/docs/seamonster_nisk2008.pdf (Last Accessed: November 2011)

Steel, C., et.al (2005) ‘Core Security Patterns: Best Practices and Strategies for J2EE, Web Services and Identity Management’ *Pearson Education, Inc.,* Massachusetts, USA.

Wiesauer, A and Sametinger, J (2009) ‘A Security Design pattern Taxonomy Based on Attack Patterns - Findings of a Systematic Literature Review’, In Proceedings of SECRYPT'2009, pp387-394

Yoder, J., et.al (1997) ‘Architectural Patterns for Enabling Application Security’ In: *Fourth Conference on Patterns Languages of Programs (PLoP ‘97),* Monticello, Illinois.

Zachman, J.A. (1987) ‘A Framework for Information System Architecture’ *IBM System Journal,* Vol.**26**(3), pp276-292

An Exceptional war That Ended in Victory for Estonia or an Ordinary e-Disturbance? Estonian Narratives of the Cyber-Attacks in 2007

Kari Alenius
Department of History, University of Oulu, Finland
kari.alenius@oulu.fi

Abstract: In the spring of 2007 Estonia became the victim of a large-scale cyber-attack. Estimates of the significance of these events vary both in and outside of Estonia. For those who regard the events as being exceptionally important, the cyber-attacks launched against Estonia are seen as a milestone of modern warfare. Sometimes the term "Web War One" has even been used. At the other extreme, the events have been underestimated and their distinctiveness has been disputed. This study does not attempt to answer the question of which perspective is "right" and which is "wrong", especially when it is particularly difficult to provide an objective answer to this type of question. Instead, this study analyses Estonian interpretations of what occurred. The central elements of the Estonian main narrative crystallized during the summer and fall 2007. The narrative came to be composed of a few key elements describing the entire conflict in general and in a stereotypical way.

Keywords: rhetoric, narratives, cyber-attack, Estonia

1. Introduction

For this study, the internet resources of Estonia's leading media outlets *Eesti Päevaleht* (EPL), *Postimees* (PM), *Õhtuleht* (OL), *Eesti Rahvusringhääling* (ERR) have been examined from the end of April until the end of June. In addition, a Google search of published material on the internet has been done using keywords *cyber-attack, Estonia, 2007*, and their Estonian equivalents. In this way, individual published speeches have been found from among other publications and from the home pages of other quarters. In the case of the four aforementioned media outlets it is apparent that the Google search yielded almost exactly the same results as a systematic review of internet newspaper archives. Thus, it can be concluded that key Estonian internet data has been analyzed for this study.

The aim is to find out what kinds of narratives of these events were created in Estonia and why these narratives were a certain kind. As an alternative to narrative we can speak of discourses or mental images. Regardless of what the selected term is, in question is the examination of the processes that essentially guide human activity. Multiple sciences have convincingly demonstrated that above all, people act on the basis of their mental impressions, and not on the basis of "objective facts" that are empirically observable, although of course the former are built upon the latter. In many ways, mental images are stereotypical, in other words, simplified and coloured models of reality, and for the most part they arise as a result of largely unconscious and to a more minor degree, conscious psychological processes (Fält, 2002, 8-10; Ratz, 2007, 189-195).

Along with empirical findings, mental images are influenced by an individual's beliefs, fears, hopes, and all of the factors behind these – in short, the whole experiential history of an individual and their perception of the world. If a group of people have sufficiently similar images regarding a subject, then we can speak of collective images. Narratives and discourses partly reflect already existing mental impressions. They are also partly used for constructing images, clarifying images for oneself and spreading them to other people (Fält, 2002, 9-11; Ratz, 2007, 199-213). In any case, the importance of mental images in interaction between people and throughout the course of history justifies why, in the case of Estonia's cyber-attacks of 2007, it makes sense to analyze mental images and narratives created by events in Estonia.

2. The cyber-attacks in 2007 and their contexts

To understand the narratives generated there is first reason to briefly explain actual events and their associated contexts. The cyber-attacks were related to Estonia's so-called Bronze Soldier uproar. In 1947 the Soviet Union had set up a military statue in the center of Estonia's capital, Tallinn, the official name of which was "a monument to the liberators of Tallinn". After Estonian independence (1991) the fate of all Soviet monuments came into question. The Bronze Soldier was left in place but became a memorial for all those that had fallen during the WWII. However, these changes did not prevent the statue from becoming the focus of disputes. Some Estonian Russians organized celebrations annually near the statue on May 9 on Russia's so-called Victory Day, as well as on September 22, the anniversary of Tallinn’s “liberation”. In the minds of many Estonians, these kinds of celebrations were hostile actions towards Estonia, as on the Estonian side the statue was often regarded as a symbol of Soviet occupation. From the Estonian perspective, the open show of Russian and Soviet symbols during the celebrations was a glorification of the occupation and a distortion of history (Kaasik, 2006, 1893-1916).

On May 9, 2006 there was a confrontation at the statue in which Russian celebrators attacked protesters carrying the Estonian flag. After the conflict, demands that the statue be removed from Tallinn’s center and placed elsewhere strengthened. At the beginning of 2007 the Estonian parliament adopted two laws on the basis of which the Bronze Soldier and other similar monuments, as well as any dead buried in connection with them could be moved to a more suitable location. Preparations to move the Bronze Soldier and the Soviet soldiers buried nearby began on April 26, 2007. The statue and its surroundings were isolated with fences and unauthorized access to the site was prevented. The same evening, the Russians opposed to the operation were involved in large-scale rioting and sabotage in the center of Tallinn, and the unrest continued the following night. The Bronze Soldier was moved as planned to Tallinn’s military cemetery and opened to the public on April 30, and the situation in Tallinn calmed down (RKK, 2007, 1-3).

At the same time as the riots, targeted cyber-attacks against Estonia began on April 27, which mainly targeted the websites of Estonian state institutions. The attacks mainly consisted of massive spamming and DDOS attacks. On the last day of April the extent and technical level of the attacks rose sharply and the main focus became the DNS system of Estonian servers. The number of sites expanded to include Estonian Internet service providers and the Estonian media. The attacks continued in varying intensity on a daily basis until mid-May, after which the situation almost returned to normal. Most of the cyber-attacks came from Russia, and based on the technical factors and large-scale resource requirements for the attacks, this suggests that the Russian government was involved in the attacks. The Russian state naturally denied involvement (RKK, 2007, 2-4; Saarlane, 2007-05-17).

Russia has also denied responsibility for other aggressive actions against Estonia. However, as early as the beginning of 2007, Russia’s state leadership warned Estonia about moving the Bronze Soldier, and on April 23, left Estonia a formal diplomatic note concerning the issue (ERR, 2007-04-27). Already before the relocation of the statue there had been intensifying anti-Estonian verbal attacks in the state-controlled Russian media, and during the riots, the Russian Embassy in Tallinn, at the very least, kept close ties with the leaders of the riots. In Moscow, anti-Estonian protesters surrounded the Estonian Embassy for a week, apparently with the consent of the government, and prevented it from operating normally. In practice, Russia also undertook economic sanctions against Estonia and began a boycott of Estonian products in Russia (RKK, 2007, 3-4).

3. The main narrative

When we examine the reaction of the Estonian public to these cyber-attacks, it is apparent that quite soon after the onset of the attacks public debate began to develop in two rival narratives. On one hand, there was the mainstream public debate, which can be called the main narrative, and on the other hand, there was a side narrative that received less publicity, which can also be called a counter narrative. The main narrative was represented by Estonia’s state leadership as well as the majority of journalists and IT professionals that publicly commented on the issue. The creators of the counter-narrative consisted of a few individual commentators who belonged to the two latter groups.

During the first three days of the cyber-attacks there was uncertainty and confusion among the Estonian public, which did not yet provide sufficient conditions for the birth of a narrative. In the first few days the main focus was on the riot and its aftermath, which is understandable as this had never been seen before in Estonia and in its drama, it had a high news value. Then and in the next couple of days, the cyber-attacks were also relatively few, and no clear information was available regarding their nature and origin. The issue was also new and unexpected: it was not anticipated, and there were no precedents in Estonia or elsewhere in the world on the basis of which an image could immediately be built. Thus, public uncertainty and confusion was apparent in that the media took a neutral tone in news regarding the matter. For example, these reported that the websites of Estonian state institutions had been attacked or were being harassed, but other evaluations of these events were not presented (ERR, 2007-04-28; ERR 2007-04-29).

The birth of a main narrative can be considered April 30, the date on which the first indicative commentaries appeared (OL, 2007-04-30; Virumaa Nädalaleht, 2007-04-30). Over the next two weeks, the mainstream image presented and the narrative of the incident broadened and took its essential form. During the second half of May, a few additional elements related to the end of the attacks were added to this. Then, there were more time and better conditions for drawing conclusions and forming an overall picture. However, the most active phase of public debate occurred in mid-May, and from the perspective of the media, the actuality of the topic began to wane after this. During the summer and fall of 2007 the topic was rarely returned to in the Estonian public, but on the other hand, the main narrative only took its final form during this time.

The declaration of Estonia's Justice Minister Rein Lang to Estonian television on the evening of April 30 acted as the initiator of the main narrative. Lang stated that investigations into the IP addresses of the attackers had revealed that the attacks originated in Russia, among others, from government institutions in Moscow (OL, 2007-04-30). During the next couple of days, one of the basic elements crystallized among the Estonian public discussion: Russia was the attacker (Delfi, 2007-05-01; OL, 2007-05-03). Although a few news reports stated that the majority of the attacks came from other addresses besides those under the direct control of the Russian government, and additionally the so-called botnet technique hampered clarification regarding the origin of the attacks, the guilty party was now known (ERR, 2007-05-04).

If the Russian state did not itself organize the attacks, it was responsible for them, as it could have chosen to prevent them. Additionally, "Russia" was guilty as, in any case, the attacks came from there, whether they were implemented privately or by the government (Delfi, 2007-05-01; OL, 2007-08-09). The conclusion that Russia was guilty was probably affirmed by other circumstantial evidence such as the earlier threats and verbal attacks against Estonia by Russian government leaders and the media, as well as Russia's suspected involvement in the rioting, at the very least as an instigator, and where necessary, as an advisor.

The identification of an enemy was a relief to Estonians, as afterwards, it was easier to interpret the situation and more possible to design countermeasures – at least at the level of beliefs. A vague and faceless enemy is always experienced as more fearsome (Zur 1991, 345-347). In question was a general human psychological reaction, which one commentator descriptively put into words at the beginning of May: "...The issue also has a good side. Events on the streets of Tallinn illustrate who our enemies are, and how many of them there are. There are no more illusions. Enemies cannot be integrated" (Delfi, 2007-05-01).

When an enemy had been found for the main narrative, the narrative could be begun and almost inevitably, one began to build using the logic and structure of the general image of the enemy. Since this was a new type of situation which did not fully recall traditional war (for instance, an official declaration of war and the conventional use of military force were lacking), all of the typical elements in the image of the enemy could not be used. Nevertheless, a few main elements were included in the Estonian main narrative. Firstly, the terminology used portrayed a war and an enemy. There was an enemy that attacked and one's own country, which repelled these attacks. A clear polarization between "us" and "others" is necessary in perceiving the existence of an enemy or another party (Zur, 1991, 345-346).

Secondly, the perception of the enemy is related to clear valuations of "good" and "bad". One's own side represents good and acts properly, while the other party represents bad and acts incorrectly, both on a theoretical–moral and practical level (Zur, 1991, 345-353). According to this polarization of values there was no understanding shown for the acts of the enemy in the Estonian public, but they were categorically condemned. Usually, there is no room for pondering the actions of the enemy in the sense that consideration would be given to why the enemy views the situation of conflict in a different way, and could the enemy have any "reasonable" or even "legitimate" grounds for its actions, from its perspective. In the mainstream Estonian public debate, those responsible for these cyber-attacks were explicitly in the wrong, malicious, and criminal (Delfi, 2007-05-01; ERR, 2007-05-05; Arvutikaitse, 2007-05-09).

The third characteristic adopted in the Estonian main narrative can be considered the typical manner in which the strengths and weaknesses of the enemy were brought to light. An appropriate balance between these two characteristics is always sought in portraying the enemy. The enemy must be sufficiently strong so that the threat to one's own side is taken seriously and that there is sufficient readiness to fight against the enemy, and if necessary, to make sacrifices in order to achieve victory. At the same time, victory over a strong enemy emphasizes the heroism and ability of one's own side and acts as a mental factor in strengthening the community. However, in emphasizing the strength of the enemy one should not go too far, as if it is portrayed as being too strong, this can result in hopelessness and defeatism among one's own community (Zur, 1991, 346-360).

In mainstream Estonian debate, the strength of the enemy was emphasized by explaining openly and in detail how wide-ranging and how many types of cyber-attacks had been made against Estonia. At the same time however, it was remembered to note that these attacks had been repelled. If in some cases the enemy had gained an advantage, then this advantage at least was temporary and limited. The counter-measures of one's own side had already gained control of the situation and no vital area had truly been in danger (ERR, 2007-05-01; OL, 2007-05-03). Thus, both the listeners and perhaps also the narrators of this narrative were able to feel safe in relation to the overall developments and final result of the war.

One could also feel safe in regard to the fact that in spite of its strength, in the case of Estonia, the enemy was weak and inferior, according to classic models. These characteristics can occur, for example, in the ridiculousness of the enemy. A comparison that is directly or apparently made to one's own side strengthens the opposing characteristics of one's own side. In Estonian mainstream public debate, this element was reflected in the good-natured comments of a few IT professionals regarding the simplicity of some cyber-attacks and the fact that they were easily deflected. The inability of the attackers to understand that their IP addresses were also easily found and that they revealed themselves was also wondered at publicly. No setbacks in these contexts were mentioned (ERR, 2007-05-04).

4. The competing side narrative

Within the competing side narrative, these aforementioned issues were mostly denied or put in perspective. The common basic premise was that Estonia had ended up in difficulties. The critique was not so great that it would have questioned belonging to the same community (Estonia/Estonians) or that the existence of the conflict would have been in dispute. However, the description of details and their interpretations differed.

This counter model of interpretation did not interpret the Russian state as the main opponent. This did not actually mean that it took a positive stance towards Russia, but it emphasized the difficulty of tracing the nature of the attacks as well as the role of individual Russian entities. In practice, the data available was the same as that which supporters of the main narrative had, but supporters of the counter narrative thought that it justified lesser conclusions: the Russian state was not merely responsible. At the same time the intermittent success of cyber-attacks and at least one's own temporary insufficiency to respond to them was also highlighted (EPL, 2005-05-17; Elamugrupp, 2007-06-30).

In the minds of those supporting a counter narrative, the question was also perhaps of a war, but according to their interpretation, the issue was not only about a simple series of successful defences, as was presented in the mainstream debate. At the most extreme, Estonian defenders of cyber-attacks were accused of overreacting and even unknowingly playing into the hand of the enemy: if the goal of the enemy was to isolate Estonia from the rest of the world, then the Estonians had ultimately done this themselves by blocking the access of foreign Internet addresses to Estonian websites (EPL, 2007-05-17).

There were relatively few statements that built on a side narrative in Estonian public debate, about one tenth of all the news and commentary. There was roughly the same number of statements belonging to a "gray zone", and it is difficult to classify these as belonging to either group. Therefore, about eighty percent of all the statements belonged to the group that built on the main narrative. The differences were not between different publications; for example, there were no significant differences between Estonia's leading Internet publications. Individual supporters of a side narrative could be found in different publications without that they would have been concentrated in any of them.

Possibly, these journalists were practicing a culture that was characteristic within conditions wherein which there was a free exchange of information, particularly of Western democracies, which a few researchers have referred to as a symbolic shadow-boxing. In conditions where there is freedom of information, it is considered the duty of the media to provide the general public with an image that is not too uniform, regardless of the issue and situation. If differing interpretations of the “facts” are not otherwise born, self-respecting journalists must create them if necessary, in the name of criticality and pluralism. This can lead directly to the aforementioned “shadow-boxing”. In this case, the basic configuration of the crisis and the justification for defending one’s own side is not put into question, but it is considered necessary to also search for mistakes and failures in one’s own actions (Carruthers, 2000, 157-158).

5. Additional characteristics and further development

The key characteristics of the above-mentioned main and side narratives were created and established by mid-May. In the side narrative there was no apparent formation of additional characteristics after this, which is partly due to the fact that there were very few statements belonging to this side narrative after mid-May. Based on these few late comments, it is not possible to make any broader conclusions regarding possible changes in position (Vikerkaar 2008). In the case of the main narrative however, it is possible to continue an analysis of developments. By the second half of May two additional characteristics had joined the picture, and during the summer and fall of 2007 the image crystallized to take the shape of a few general interpretations.

The first additional characteristic was that by the end of May, it was already ventured to declare that Estonia was the victor in the war. No significant cyber-attacks had occurred for a week, and in perspective of the daily attacks that had occurred by the end of April and beginning of May, this seemed sufficient evidence to end the war. On the other hand, as logic regarding the image of the enemy entails, one’s own side has no reason to lull themselves with a false sense of security. Once the enemy has been found he continues to be a potential enemy, and it is unrealistic to hope for a world without an enemy. In the main narrative it was remembered to emphasize that the attacks against Estonia and elsewhere were possible in future, even probable. For this reason, Estonians had to remain vigilant and develop their capacity to combat future attacks (OL, 2007-05-17; EPL, 2007-05-25).

The second additional feature was closely related to the previous one. In principle, it contained two conflicting sides of the same issue. On one hand it was reported that Estonia was an object of admiration for NATO allies, and the statements of allied representatives visiting Estonia were quoted frequently in public. Accordingly, the guests came to learn from the Estonians (Äripäev, 2007-05-18; PM, 2007-05-25). Praise received from others is always pleasant and helps to construct a positive image of oneself or of one’s own group; to this extent, the quotation of statements had a clear general psychological background. On the other hand however, several Estonian statements which otherwise clearly belonged to the main narrative emphasized the need to gain support in rebuffing the cyber-attacks. Speeding up the construction of NATO’s cyber defense center in Estonia was met with joy, and additionally, there was a desire for international reform regarding the definition of cyber-attacks. It was considered that the existing international agreements were outdated: they did not take the matter seriously enough, take into account the technological development in the field, nor allow for a sufficiently effective legal and practical response (PM, 2007-05-14; Virumaa, 2007-05-25).

It is clear that the achievement of international agreements and definitions that would obligate other countries to assist states that have become the target of cyber-attacks would have been to Estonia’s advantage. For this reason, it was reasonable for Estonia’s representatives to demand this in both Estonia’s internal debate as well as abroad (EP, 2007-05-10). At the same time, assistance, in particular from other NATO countries, would have been welcome. In principle, however, it was questionable what other assistance would have been available from others if Estonia was already the world’s leading expert in cyber defence. This paradox was not mentioned in Estonian statements. According to the principles of propagandistic communication, contradictory elements can be used in communication, as long as they are not presented at the same time (Zur, 1991, 350-351). Thus, in the Estonian main narrative these things – the Estonia in need of assistance and Estonia as the world’s most skilled – never appeared in the same statements.

On the international scene, defining the cyber-attacks as a war was driven especially by the Estonian president Toomas Hendrik Ilves, as well as by Estonian members of the European Parliament (EP, 2007-05-10; President, 2007-06-18). For them, based on their positions and contacts, driving the national interests of Estonia abroad naturally fit well. At the same time, their positions were also heard by the Estonian public. The same message was forwarded particularly to the domestic audience by the speaker of the Estonian parliament Ene Ergma (EPL, 2007-05-25). When commander of the Estonian armed forces Ants Laaneots is added as a constructor of the main narrative, it can be said that Estonia's highest state leadership was more or less in favor of the main narrative. Out of all Estonia's influential public personas, Laaneots (PM, 2007-06-20), as well as former Prime Minister and Chairman of the leading right-wing party (IRL) Mart Laar (OL, 2007-08-09), most clearly stated that the Russian state was responsible for the cyber-attacks. Ilves and Ergma expressed the matter a little but more diplomatically, but even in their statements there was no doubt regarding the main opponent in the war.

6. Conclusions

Thus, the central elements of the main narrative crystallized during the summer and fall, so that over time the details and the exact course of events became side issues. These kinds of components fell or were dropped from the narrative and the image increasingly came to be composed of a few key elements describing the entire conflict in general and in a stereotypical way. All in all, it can be concluded that according to the Estonian main narrative, the cyber conflict consisted of the following components: 1) it was a war; 2) the Russian state was either directly or indirectly responsible for the attacks; 3) in question was a new, unprecedented kind of war; 4) the war ended in victory for Estonia.

Of the Estonian public debate, approximately eighty percent built on or supported the narrative described. In this sense, it can be regarded as a strong national narrative. The fact that in its contents it had a strong nationalistic emphasis and that its most visible supporters were individuals that belonged to a conservative right-wing also makes it a national narrative. As in many cases the same individuals also belonged to Estonia's state and political elite, the main narrative can also be described as Estonia's official narrative to a large extent. However, it was not a case of the Estonian government forcing members of the elite or the Estonian media to comply with this explanation. It was sufficient that the situation and conditions were such that the majority of those who participated in public debate came to similar conclusions on their own initiative. The experience of Estonian society coming under an unfair attack from the outside gave birth to a very large, uniform reaction of defence that was explained to oneself and to others in the form of this main narrative.

A side narrative was perhaps formed as a conscious counter-reaction to a main narrative that was experienced as being too uniform and thus propagandistic or implausible. It may also have been a case of differences in interpretation regarding other events, without any initial purpose of criticizing the mainstream narrative. The counter narrative questioned the nature of events as a war and preferred to support the interpretation of Internet harassment. At the same time, the interpretation that regarded the Russian state as opponent in the crisis was viewed as being too simplistic, and the private or unclear background of the attackers was referred to. The exceptionality of the events was also questioned and they were compared to known, large-scale operations of harassment and damage initiated by private parties. Similarly, the fourth main characteristic of the main narrative, victory in the war, was not seen as a legitimate interpretation: if there was no war, there was also no victory. In addition, the success in combating the operations varied according to this view.

In principle, both of these narratives were based on the same information regarding the events. Each narrative was also partly built on the basis of general psychological models, in which the role of apparent "facts" in shaping the narrative lessened. These narratives used empirical construction materials, but their development also partly followed models guided by a stereotypical, human way of thinking. Thus, for instance, images of the enemy are generally similar to a great extent, regardless of the circumstances. The appropriate selection and appropriate interpretation of the information available was essential so that they supported the perceived best, simple enough explanation of the model – a narrative. It is not possible to determine the relative weight of conscious and unconscious activity, but both have undoubtedly played a significant part in the birth of these models. To uncover the specific characteristics of these narratives, the content of the media in neighboring countries (for instance, Baltic and Scandinavian countries) could be examined for this same time period, but this requires a separate study in the future.

References

Äripäev (2007-05-18) 'Eestist saab NATO kübersüda ja IT-polügoon', http://leht.aripaev.ee/?PublicationId=464dc490-fb94-4024-9b75-258ddc8543a9&articleid=12282&paperid=A4DE138A-6A0D-4C2A-A1B6-6613E673D67A

Arvutikaitse (2007-05-09), ’9. maid tähistati küberrünnakuga’, http://www.arvutikaitse.ee/9-maid-tahistati-kuberrunnakuga/

Carruthers, S. (2000) *The Media at War: Communication and Conflict in the Twentieth Century*, Basingstoke: Macmillan

Delfi (2007-05-01) ‘Küberrünnak Eesti riigiasutustele’, http://www.delfi.ee/archive/kuberrunnak-eesti-riigiasutustele.d?id=15733528

Elamugrupp (2007-06-30), ’Vaenlane kasutas kübersõjas müstilisi e-pomme’, http://www.elamugrupp.ee/modules.php?op=modload&name=News&file=article&sid=1067&mode=thread&order=0&thold=0

EP (2007-05-10), ’EP palub Euroopa Liidul näidata üles solidaarsust Eestiga’, http://www.europarl.europa.eu/sides/getDoc.do?type=IM-PRESS&reference=20070507IPR06398&language=ET

EPL (2007-05-17) ’Sõja versioon 2.0 (beeta)’, http://www.epl.ee/news/arvamus/article.php?id=51087289

EPL (2007-05-25) ’Ergma: Eesti vastu suunatud küberrünnak ei jää EL-is viimaseks’, http://www.epl.ee/news/eesti/ergma-eesti-vastu-suunatud-kuberrunnak-ei-jaa-el-is-viimaseks.d?id=51088437

ERR (2007-04-27) ‘Venemaa andis Eesti suursaadikule pronksmehega seoses noodi’, http://uudised.err.ee/index.php?0573764

ERR (2007-04-28) ’Valitsuse kommunikatsioonibüroo hoiatas valeteabe eest’, http://uudised.err.ee/index.php?0573960

ERR (2007-04-29) ’Välismaised rünnakud häirivad valitsusasutuste veebilehti’, http://uudised.err.ee/index.php?0574001

ERR (2007-05-01) ’Rünnakud Eesti küberruumi vastu on sagenenud’, http://uudised.err.ee/index.php?0574069

ERR (2007-05-04) ‘IT-ekspert: Vene rünnakud serveritele on oskamatult tehtud’, http://uudised.err.ee/index.php?0574202

ERR (2007-05-05) ’Politsei pidas kinni esimese küberrünnakus osaleja’, http://uudised.err.ee/index.php?0574259

Fält, O. (2002) ‘Introduction’, in Alenius K., Fält O. and Jalagin S. (eds.) *Looking at the Other. Historical Study of images in theory and practice*, Oulu: Oulu University Press.

Kaasik, P. (2006) ‘Tallinnas Tõnismäel asuv punaarmeelaste ühishaud ja mälestusmärk’, *Akadeemia*, no. 4.

OL (2007-04-30) ’Rein Lang: küberründed Venemaalt tulevad riiklikelt aadressidelt’, http://www.ohtuleht.ee/227560

OL (2007-05-03) ’Venemaa küberrünnak Eesti pihta on Euroopa kohta tavatu’, http://www.ohtuleht.ee/227851

OL (2007-05-17) ’Kübersõda karmistub’, http://www.ohtuleht.ee/230007

OL (2007-08-09), ’Uurimise takistamine tõestab, et küberrünnak lähtus Venemaalt’, http://www.ohtuleht.ee/241417

PM (2007-05-14) ’Aaviksoo rääkis NATO juhiga küberrünnakutest’, http://blog.postimees.ee/170507/esileht/siseuudised/260640.php

PM (2007-05-25) ’USA eksperdid: küberrünnak Eesti vastu äratas meid’, http://rooma.postimees.ee/040607/esileht/siseuudised/262679.php

PM (2007-06-20) ’Venemaa muutub Eestile üha ohtlikumaks’, http://rooma.postimees.ee/210607/esileht/siseuudised/267665.php

President (2007-06-18), ‘Toomas Hendrik Ilves: "Kas küberrünnak on hädaolukord?’, http://www.president.ee/et/meediakajastus/intervjuud/3150-vabariigi-president-ajalehele-frankfurter-allgemeine-zeitung-18-juunil-2007/index.html

Ratz, D. (2007) ‘The Study of Historical Images’, *Faravid*, vol. 31.

RKK (2007) ‘Moskva käsi Tallinna rahutustes’, Rahvusvaheline Kaitseuuringute Keskus, http://www.icds.ee/index.php?id=73&tx_ttnews%5Btt_news%5D=179&tx_ttnews%5BbackPid%5D=99&cHash=a1145105e4

Saarlane (2007-05-17), ’Kreml eitas osalust Eesti küberrünnakutes’, http://www.saarlane.ee/uudised/uudis.asp?newsid=29727&kat=3

Vikerkaar (2008), ’Küberrünnakute moos aprillirahutuste kibedal pudrul’, http://www.vikerkaar.ee/?page=Arhiiv&a_act=article&a_number=4732

Virumaa (2007-05-25), ’VE: küberrünnakud’, http://www.virumaa.ee/2007/05/ve-kuberrunnakud/

Virumaa nädalaleht (2007-04-30), ’Minister Rein Lang: küberründed tulevad Venemaa riiklikelt IP-aadressidelt’, http://www.vnl.ee/artikkel.php?id=6804

Zur, O. (1991), ’The love of hating: the psychology of enmity’, *History of European Ideas*, vol. 13, no. 4.

Who are you Today? Profiling the ID Theft Fraudster

Olga Angelopoulou[1], Stilianos Vidalis[2] and Ian Robinson[2]
[1]School of Computing and Mathematics, Faculty of Business Computing and Law, University of Derby, Derby, UK
[2]School of Design, Engineering, Fashion and Technology, Faculty of Arts and Business, University of Wales, Newport, Newport, UK
o.angelopoulou@derby.ac.uk
stilianos.vidalis@newport.ac.uk
ian.robinson@students.newport.ac.uk

Abstract: Online Identity Theft (ID theft) is a significant problem in our modern knowledge-based and social-driven computing era. This type of cybercrime can be achieved in a number of different ways; and more of the point, various statistical figures suggest it is on the increase. The target is individual privacy and self-assurance, while efforts and measures for increased security and protection appear inadequate to prevent it. While personal identities are increasingly being stored and shared on digital media in virtualised environments, the threat of personal and private information that is used fraudulently cannot be eliminated. This trend in crime can result in complex investigations that involve virtualised information technologies, both as a medium for analysis and as evidence at the same time. Fraudsters are obtaining more sophisticated technological ways and increase their capability not only for committing but also for concealing their crimes. It is believed that fraudsters of this kind of crime are not acting individually, but rather they operate in an organised and well-structured manner. Indeed ID theft is nowadays directly linked to drug trafficking, money laundering and terrorism. ID theft, like almost all different types of crime, involves two parts, at least one victim and at least one fraudster. We argue that the differentiation of the investigation procedure between the victim's and the fraudster's side, depends on the ownership and control of the digital media involved in the crime, and can provide results on a more crime-focused basis. In addition it provides information gathering, understanding and knowledge about the way the fraudster acts and could potentially assist in future investigations. Different pieces of evidence can be discovered on each side (victim-fraudster) concerning the techniques that have been used to perpetrate the crime. The online ID theft techniques can leave evidence on both the victim's and the fraudster's system. However, the evidence tends to contain different elements on each side that can reveal information about the fraudster and eventually profile him in relation to the attack. There is an approach of profiling the ID theft fraudster based on the findings thatarise during the forensic investigation process in this paper. We discuss the extent of ID theft as a problem and the role of the fraudster in different ID theft techniques. We aim to demonstrate processes that could assist the profiling of the fraudster under the forensic investigation of ID theft.

Keywords: identity theft, computer crime, fraudster profiling

1. Introduction to ID theft

Identity is defined as:

[noun (pl. identities)] the characteristics determining who or what a person or thing is (The Oxford Dictionary of English)

Theft is defined as:

[mass noun] the action or crime of stealing: he was convicted of theft (The Oxford Dictionary of English)

The growth in identity related fraud resulted in the need for the establishment of specific terminology in order to identify this particular type of crime. Identity Theft is defined by the Home Office (2009) as:

Criminals can find out your personal details and use them to open bank accounts and get credit cards, loans, state benefits and documents such as passports and driving licenses in your name.

A paraphrased definition could be:

Identity theft is the use of your ***personal identity*** *in the form of personal information by another individual for their* ***financial gain.***

However, it may be argued that the gain is not necessarily always financial. Identity Theft may be aimed at satisfying some other objective; espionage, terrorism, revenge, illegal immigration or assuming a new identity to avoid criminal charges (Newman and McNally, 2005). However it is generally accepted that the end objective of Identity Theft is usually some form of financial gain as suggested by Gerard G. J. et al. (2004), who defined it as:

Identity theft is the criminal act of assuming the identity of another person with the expectation of gain. The gain is normally financial as a result of improperly extending credit, allowing banking transactions, establishing cellular telephone or other utility service, or gaining governmental benefits.

Therefore, based on the above ID theft can be defined as someone's action of using any sort of distinct personal private information with fraudulent intention; mainly for financial gain.

2. Computer crime and cybercrime

Based on Mohay et al. (2003) while the use of computers and the Internet become even more popular, at the same time fraudsters' take advantage and increase their ways of attacking systems. The computer can be used in three different ways in order to assist a crime. It can be used as the 'tool' that the fraudster uses for performing the crime, the 'target' that the fraudster manages to attack and penetrate and the 'storage area' that he can use in order to save information related to the crimes. (Shinder and Titel, 2002). For this paper the following definitions will be used.

A **computer** is defined as an electronic, magnetic, optical, electrochemical or other high speed data processing device performing logical, arithmetic, or storage functions, and includes any data storage facility or communications facility directly related to or operating with conjunction of such device. (Press, 2006)

Computer Crime is defined as the activities that target data and/or low level data operations using computers as the means for conducting those activities.

Cyber Crime is defined as the activities that target information and/or knowledge or the high level data operations that manage information and/or knowledge of legal bodies, including software programs, using computers as the means for conducting those activities.

E-Crime or economic crime is defined as those computer crime or cyber-crime activities where the main purpose is financial gain.

3. Issues on detecting and investigating ID theft cases

Based on the aforementioned definitions we will reason that ID theft is a cyber-crime. It is accepted by the academic and business world that having appropriate incident management procedures that will allow you to maintain an appropriate forensic readiness state for collecting potential evidence of crime, is essential.

Evidence *can be defined as items that can be used to assist a party or parties in either proving or disproving the fact of a matter and allow the burden of proof to be reached.*

During legal proceedings it is necessary to prove a case to a level that surpasses a burden of proof.For criminal cases "the prosecution bears the legal burden of proving every element of the offence charged, and disproving any defences raised, beyond reasonable doubt" (Hopkins, 2006: 9).

Therefore, evidence should be used to support each element of an allegation. Failure to adequately support an allegation with sufficient evidence would be likely to result in the Court removing those items from the case or the acquittal of the defending party on those points as the evidence to support those allegations could be deemed inadmissible as given in McAlhone& Stockdale (2002: 94) where it is stated that "the evidence will only be excluded if the judge...decides that the admission of the evidence would, in the circumstances, have such an adverse effect on the fairness of the proceedings that it ought not to be admitted".

The investigator of an ID theft incident has first of all to deal with and understand the motivation of the fraudster. A thief who targets unsecure systems is a relatively different case than someone who decides to attack an organisation in order to steal specific information. The impulse of the attacker can prove his knowledge, skills and intention (Casey, 2003). Fraudsters are cautious not to leave traces behind, mislead the investigator with anti-forensic techniques (Harris (2006), Smith (2006), Forte and Power (2007)) and remain undetected for a long period of time.

The ID theft fraudster does not always need to have advanced technical knowledge. Persuading and deceptive skills are always an advantage. For instance, a shoulder surfer needs to be quick and observant, while someone who performs phishing (Gajek and Sadeghi(2008), Lininger and Vines (2005)) or pharming (Jakobsson and Myers, 2006) attacks needs to know how to clone a URL, programming and web designing in order to design a professional web-site and write malicious code.

Phishing emails can often be difficult to identify, although simple spelling mistakes usually give them away. A user can identify the unusual email if the email only contains a large image instead of text, which is linked to the spoofed web page. In some cases e-mails contain a virus or have Trojan infected files attached. Spoofed web pages can be identified from spelling mistakes and omissions. For example, when a login web page is spoofed and the original has been updated. The spoofed pages are not maintained and consequently the changes are not conducted on the spoofed webpage.

It is also important to understand how the fraudsters choose a target. Financial profit is their main objective. When targeting individuals the profit may be less and when the target is corporate a large turnover may come at once. However, if someone targets and steals a large amount of money from a company or its clients, the loss will be easier to identify and more intensively investigated. For the individual on the other hand it can also take a longer period of time to find out that he/she is an ID theft victim. Meanwhile, the fraudster is free to hit other targets.

In ID theft, a person takes control of and abuses someone else's personal and private authentication information; it belongs to criminal offences. There are people who unfortunately act unethically and against the law in order to gain financial profit. In the past it used to be loss of goods and properties, but now it has also come to loss of personal and private data. (Pfleeger,2000).

In addition, individuals often do not adopt any measures to protect themselves. They can provide personal identity information to anyone, carry their credit and debit cards in their wallets that might get stolen, have their PINs and passwords written on pieces of paper that they carry in their wallet or save them as records in their mobile phones. In addition, they visit unreliable web sites and purchase goods by entering their bank account details without considering online security.

Regardless if there is a team of thieves, acting with the same identity on multiple locations, or one person that abuses more than one identity, an ID theft investigation will reveal more information about the illegal activities. During such an investigation we have to take into consideration a number of problems that deter the preservation of evidence trail. The most important is that fraudsters tend to discover more and more sophisticated behaviours and manage to hide evidence traces that can prove their guilt. Due to the complexity of the crime we believe that the majority of ID thieves do not work on their own, but in groups. This means that even when an ID thief is revealed, members of his group will still have the time to act undetected for a period of time. Furthermore, information that involves personal data might be refused to the investigator in the first place, putting a hold to certain evidence leads. There are also technical difficulties relating to the type of electronic devices and computer technologies used for the crime and how/if information can be retrieved from them. For example, virtualised environments present a number of operational, legal and technical difficulties to the investigator that is likely not to be adequately equipped to handle such a scenario. This is the subject of another paper that the authors are developing.

Even though not all ID theft incidents result in prosecution, when this happens the investigator needs to present his findings in an appropriate manner. Evidence in the court of law needs to be factual and detailed. The information presented needs to be complete and unbiased. The law today requires that all original electronic media related to the case must be preserved and most legal systems require to record any relevant information related with the electronic source the data was created (Pierce, 2003).The evidence from a computer or network system presented in the court needs to have the following attributes:

- *Authenticity*, the evidence can be related with the events of the incident;
- *Demonstrational*, they can be presentedin a form that can be submitted to the court;
- *Best Evidentially*, they represent the evidence in the more complete form;

- *Probative*, the information can be presented practically (Stephenson, 2002)

No matter how detailed theanalysis of the investigator is,during the Court hearing the defence will try to dispute both the investigator and his evidence, e.g. Trojan defence (see Haagman and Ghavalas (2005)). For this reason, the evidence provided in the Court should leave no doubt of its authenticity and completeness. It will enhance the professional image of the investigator when the case is represented without deficiency (Michaud, 2001). Generally evidence can be produced in three main forms, these being: real, hearsay, and testimonial.

Real Evidence*is an object that is involved in a case or may have played a role in the action in question. The object may also be able to be examined by a Court or an expert witness who is suitably qualified to provide guidance to the Court regarding the item.*

Hearsay Evidence*is less common, however, is information given to a Court by someone who is not a direct witness of an event or action and, therefore, does not have first-hand knowledge of something but has heard it from another source.*

Testimonial Evidence*is provided by a witness. This may be a witness to the alleged offence itself or a professional witness who has examined the item of real evidence and are providing findings based on that examination.*

We argue that it is important to capture features of past ID theftsas they can aid in the investigation of future crimes.There are two main evidence categories that are provided within Hopkins (2006): circumstantial and direct.
Circumstantial Evidence*can be used to draw an inference of guilt sufficiently to prove a case beyond reasonable doubt; however, it is not drawn from fact and can be open to interpretation.*

Direct Evidence*comprises of evidence of the allegation itself and if that evidence is accepted then the allegation itself will be proven.*

Through our profiling methodology we attempt to collect and analyse real and testimonial direct evidence for supporting the case.

4. Victim and fraudster

The Oxford Dictionary of English defines the victim as a person who is tricked or duped and the fraudster as a person who is intending to deceive.

The victim is the dupe, the prey, the target. The specific word has been selected in order to describe the person that has been affected emotionally or financially by ID theft.

The word fraudster has been chosen for this work among its synonyms (criminal, deceiver, perpetrator), as is approaches more accurately the intention of the person to commit fraud by acting deceitfully.

There is an approach of distinguishing and discriminating the investigation process between the victim and the fraudster in this research work. Different pieces of evidence can be discovered on each side (victim-fraudster) concerning the technique that has used to perpetrate the crime. While the online ID theft techniques can leave evidence on both the victim’s and the fraudster’s system, different elements are contained on each side. For example, a phishing e-mail could leave evidence on the victim’s internet logs, while on the fraudster’s system, information about the building of the scam could be found. A malware might leave unknown running processes on the victim’s machine and a source code library on the fraudster’s machine.

This is going to assist the investigator as it aims to focus the examination of the digital media on the side of interest. Therefore, the investigator will work on a structured ground under a procedure that includes only these elements that he needs to search for. A structured approach can reduce the duration that is required for the examination and reduce the chance of evidential material being overlooked.

5. The profile of the ID thief

We argue that in most cases, the investigator is aware whether the media belongs to the victim or the fraudster. No matter whether the investigation concentrates on the victim's or the fraudster's perspective, it is of great importance to identify the intention of the fraudster under a structured and justified basis. It will add expertise and intelligence on the way future investigations will take place.

Marcella and Greenfield (2002), discuss the concept of profiling the fraudster. The forensic examination of digital media will provide profiling elements both from the victim's and the fraudster's side. The profiling of the fraudster should be retrieved and drawn from a different aspect when it comes from the victim's media and different from the fraudster's media.

Rogers (2003) argues that computer based investigation procedures need to develop in a similar way to procedures in the classic forensic science and include the profiling of the fraudster. Numerous studies refer to the importance of profiling the fraudster (for example, O' Block et. al., 1991; Turvey, 2002; Kocsis, 2006). However, it is mainly based on the sex, race, age etc.

We separate the evidential data that assist in the fraudster profiling in three processes listed underneath. These are the result of a large research project regarding ID theft and its online investigation.

- The identification of the fraudster
- The profiling and analysis
- The profile structure

Information that assists in the profiling of the fraudster can be found during the investigation of the digital media amongst the evidential data. Figure 1 illustrates the processes that are required for the profiling and how they are linked with residual information. The appropriate handling and analysis of this information can transform it to evidential data that can be later used for creating the fraudster's profile. The procedure is described in the following sections.

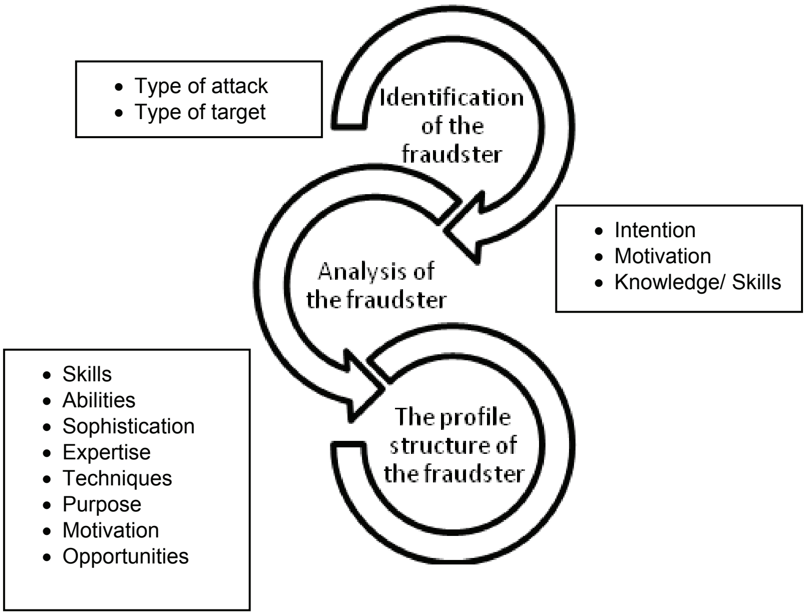


Figure 1: The profiling process

5.1 Identification of the fraudster

Concerning the discovered ID theft evidential data, the incident should be considered either as a direct insider attack or as an external attack as argued by Vidalis and Jones (2005). Their argument is based on the way the system has been accessed and consequently on the intention of the fraudster to perform an internal or external attack. The type of the attack will also determine the course of the investigation.

The investigator is then expected to identify the reason the machine and consequently the individual became a target. It is beneficial both for the outcome of an ID theft investigation and the profiling of the fraudster to be able to determine the source of the problem by recognising significant causes. Such data can be collected from either the victim's or the fraudster's side. The type of the target will interpret information gathered from the type of the attack, such as avulnerable system or published information available about an individual. A vulnerable system is more likely to become a target of an attack than a protected system, e.g. an antivirus is not installed or evidence of information being publicly available on social networking web sites.

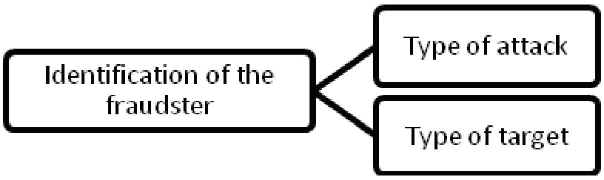


Figure 2: The elements for the identification of the fraudster

5.2 Analysis of the fraudster

The identification and analysis of the activities taken by the fraudsterassist in understanding his methods and techniques. Information that is gathered based on the complexity of the attack could reveal information such as the intention of the theft, e.g.financial or identity ID theft. The collection of evidence is able to provide such identification and offer additional valuable indications about the objective of the attack and the skills of the fraudster.

The collection of information and the analysis of the data provided concerning the fraudster is independentfrom the owner of the media (victim or fraudster). It is going to supply invaluable elements that could assist in structuring his profile.

Intention

The investigator should also determine whether the fraudster retrieved information from an individual or from a corporate system (Vidalis and Jones, 2005). This can reveal further information about the intelligence and the intention of the fraudster.

Based on the residual data in the digital media the investigator should be able to put in the picture the intention of the fraudster. The intention is based on the different forms of ID theft (Angelopoulou et al., 2007). For example, in the case where only identification data is stolen, then the intention of the fraudster appears to be identity ID theft. In some cases there may not be enough detail to prove this, but there should be indications in relation to the findings that support the fraudster's intention.
The fraudster's intention is divided in two categories:

The financial intention; the purpose of the ID thief is to gain access to financial information for financial gain.

The identity intention; the purpose of the ID thief is to gain access to someone's identification information or impersonate an individual for acquiring a new identity.

Motivation

Vidalis et al. (2004) present a threat agent list, where each different type of threat agent is motivated by his beliefs. In the case of investigating an ID theft incident, the threat agent is interpreted as the fraudster.

Jones (2002) refers to the components of the threat agent's motivation. According to this categorization, an ID thief has personal gain as a motive. However, the evidential data can disclose more than just this. Information that concerns a limited number of attacked systems for example indicates that the threat agent works in a focused target group.

The motivation of the fraudster can be revealed by examining the selected target and comparing the information with the treat agent list to identify the objective of the attack. For example, different information is obtained for a fraudster targeting a vulnerable system and different information for an attack motivated by organised

Knowledge/ Skills

The investigator gets an insight about the background knowledge and the expertise of the fraudster, the group that the fraudster belongs (Jones and Ashenden D., 2005). The complexity of the attack will provide the investigator with information about the fraudster's skills. For example, a complex attack involving the distribution of custom malicious code targeting financial data could reveal conspiracy intentions. Consequently, the fraudster should have advanced technical skills to manage the attack.

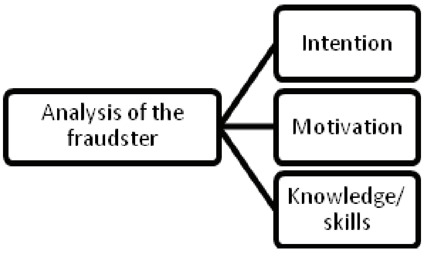


Figure 3: The elements for the analysis of the fraudster

5.3 The profile structure

The key issue is to determine the findings based on the victim's and the fraudster's side. The research work of Vidalis and Jones (2005) fits the needs of profile structure for the threat agent and the research work of Angelopoulou (2007) assists on the identification of specific information concerning the investigation of ID theft. The combination of these allows the attempt to structure a profile based on the detailed examination of the digital media.

Hence the profile structure needs to be distinguished between the victim's and the fraudster's side at this point.

5.3.1 The victim's digital media

The victim's digital media provides a vast amount of information about the penetration of the system. The handling of specific evidential elements can also provide information that will assist on the profiling of the fraudster.

Reveal technical skills

The information discovered after the analysis of the victim's side is obviously going to reveal clues concerning the technical skills of the fraudster. Brute force attacks and IP Spoofing are good examples for low and high technical level accordingly for the threat agent's skills. In addition, the security measures considered by the owner of the system can reveal the required by the threat agent technical skills to manage the attack.

Reveal programming skills

The abilities of the fraudster will provide information based on the method of the attack. The evidential data identified can show whether the intruder has programming skills, e.g. use of personal written scripts or an opportunist by using already written code.

Ability to convince someone

The fraudster's social engineering skills that are possibly identified after the media analysis. A phishing attack for example, could show that the threat agent has programming skills to develop the phishing and social engineering skills to conceal the fraud and convince the victim.

Ability to keep stealth action

The complexity of the attack is going to show the ability of the attacker to keep activities stealthy. The vulnerabilities of the system and the technical skills of the fraudster combined could possibly leave minimum traces for the investigator to identify further information about his actions.

5.3.2 The fraudster's digital media

When the object of the investigation is the digital media that belongs to the fraudster the information providing elements that assist the fraudster profiling is richer. This information if treated appropriately will deliver the profile of the fraudster.

Sophistication of tools

The collection of tools – programming, hacking, and security - identified in the fraudster's side can provide insight about his actions. Such tools can show the complexity of his capabilities.

Level of expertise

As the sophistication of the tools that are installed and probably also used by the attacker are able to define his level of expertise. The more advanced and complex the tools, the more experienced and advanced the fraudster.

Use of defensive techniques

There is no sophisticated intruder acting without considering the use of defensive techniques. These techniques would let him act unattended. For example, someone may accomplish a man-in-the-middle attack and introduce a third party involvement to the investigation.

Identify purpose of attacking

It is whether the method of the attack reveals that the fraudster acts based on his ego or his curiosity. The investigator should be able to draw a picture regarding the purpose of the attack.

Identify opportunities

The opportunities of attacking a system differ according to the technical skills, the motive, the flexibility and the background of the fraudster, as well as with his capabilities. However, an attack to an unsecure system provides different aspect to his profiling than an attack to a protected, secure, or even a corporate system.

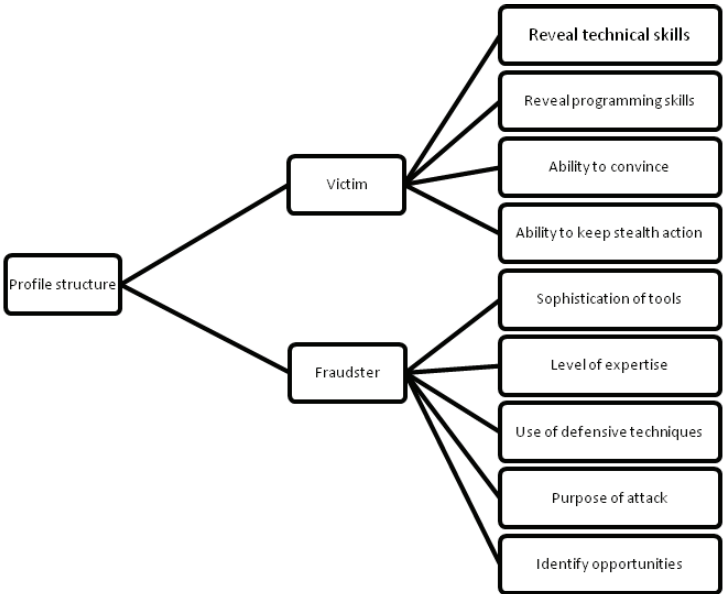


Figure 4: The profile structure

6. Conclusion

Technological innovations contributed in the rise of cybercrimes. The extensive increase of ID theft incidents, and the complexities of ID theft in a digital environment, both suggest a need for better understanding the fraudsters and the ways they act. This approach can support the investigator, and produce reliable, repeatable results.

The investigator is required to identify and analyse the digital media that constituted to ID theft. The analysis of the evidence aims to assist him identify the evidential data able to construct a profile for the fraudster concerning the story behind a specific incident. Throughout the digital media examination, information that concerns the target of the crime (victim) and the fraudster is gathered. This information can be of critical value for the investigator if analysed and treated with the aim to create a potential profile of the fraudster. The profiling can be performed by analysing information that is already available from the investigation and contains evidence that refers to the fraudster. We suggest three consecutive processes that can assist the investigator: the identification, analysis and profile structure.

Even though there have been efforts to identify the threats, eliminate the risks, inform and educate the users, attacks against someone's good name cannot be totally omitted. There lies the need for understanding the source of the problem that is people and for ID theft, fraudsters.

References

Angelopoulou O., Thomas P.,Xynos K., Tryfonas T., 2007, Online ID theft techniques, investigation and response. *Int. J. Electron. Secur.Digit. Forensic* 1, 1, May 2007, pp. 76-88

Casey, E., 2003, Determining Intent — Opportunistic vs Targeted Attacks, Computer Fraud & Security, Volume 2003, Issue 4, pp. 8-11

Casey E., 2004, Digital Evidence and Computer Crime: Forensics Science, Computers and the Internet, Amsterdam, Academic Press, 2nd ed., ISBN: 0121631044

Forte D., Power R., 2007, A tour through the realm of anti-forensics, Computer Fraud & Security, Vol. 2007, Issue 6, June 2007, pp. 18-20

Gajek S. and Sadeghi A.R., A Forensic Framework for Tracing Phishers, In: Fischer-HübnerS., Duquenoy P., Zuccato A., MartucciL., The Future of Identity in the Information Society, IFIP International Federation for Information Processing, 2008, Volume 262, 23-35, DOI: 10.1007/978-0-387-79026-8_2

Gerard G.J., Hillison W., Pacini C., January 2004, Identify Theft: An Organization's Responsibilities. pdf, downloaded from http://ruby.fgcu.edu/courses/cpacini/ courses, Accessed on: 26/05/2012

Haagman D. and Ghavalas B., 2005, Trojan defence: A forensic view, Digital Investigation, Vol. 2, Issue 1, February 2005, pp. 23-30

Harris R., 2006, Arriving at an anti-forensics consensus: Examining how to define and control the anti-forensics problem, Digital investigation, Vol. 3, Supplement 1, pp. 44-49, The Proceedings of the 6th Annual Digital Forensic Research Workshop (DFRWS '06)

Home Office, Identity Fraud Steering Committee, 2012, http://www.identity-theft.org.uk. Accessed on 27/05/2012

Hopkins B. 2006. Evidence: Key Fact. 2nd Edition. London: Hodder Education.

Jakobsson M., and Myers S., 2006, Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft, Wiley-Interscience.

Jones A., 2002, Protecting the Critical National Infrastructure – Developing a Method for the Measurement of Threat Agents in an Information Environment Information Security Technical Report, Vol 7, No. 2, p.22-36

Jones A. and Ashenden D., 2005, Risk Management for Computer Security: Protecting Your Network & Information Assets, Butterworth-Heinemann, ISBN: 0750677953

Kocsis R. N., 2006, Criminal Profiling: Principles and Practice, Humana Press Inc.,U.S., ISBN: 1588296393

Kornblum J., 2002, Preservation of Fragile Digital Evidence by First Responders, in proc. of Digital Forensic Research Workshop 2002, Syracuse, New York, available from: www.dfrws.org/2002/papers/Papers/Jesse_Kornblum.pdf

Lininger R. and Vines R.D, 2005, Phishing: Cutting the Identity Theft Line, John Wiley & Sons.

Marcella A.J. and Greenfield R.S., 2002, Cyber Forensics, A field manual for collecting, examining and preserving evidence of computer crimes, CRC Press LLC, ISBN: 0849309557

McAlhone, C. and Stockdale, M. 2002. Nutshells Evidence.3rd Edition. London: Sweet & Maxwell Limited.

Michaud D. J., 2001, Adventures in Computer Forensics .pdf, downloaded from: http://www.sans.org/reading_room/whitepapers/incident/adventures-computer-forensics_638, Accessed on: 26/05/2012

Mohay G., Anderson A., Collie B., De Vel O., Mckemmish R., 2003, Computer and Intrusion Forensics, Artech House inc., Computer security series, ISBN: 1-58053-369-8

Newman G.R. and McNally M.M., 2005, Identity Theft Literature Review, Presented at the National Institute of Justice Focus Group Meeting, available from http://www.ncjrs.gov/pdffiles1/nij/grants/210459.pdf. Accessed on: 26/05/2012

O' Block R. L., Donnermeyer J.F., Doeren S.E., 1991, Security and Crime Prevention, Butterworth-Heinemann, 2nd ed., ISBN: 0750690070

Palmer G.L., 2002, Forensic Analysis in the Digital World, International Journal of Digital Evidence (IJDE), Volume 1, Issue 1

Pfleeger C. P., 2000, Security in Computing, Prentice Hall PTR, USA, 2nd edition, ISBN: 0-13-337486-6

Pierce M., 2003, Detailed Forensic Procedure for Laptop computers. pdf, downloaded from: http://www.sans.org/rr/whitepapers/casestudies/1141.php. Accessed on05/01/2012

Smith A., 2006, Describing and Categorizing Disk-Avoiding Anti-Forensics Tools, Journal of Digital Forensic Practice, 1:4, pp. 309 – 313

Schweitzer D., 2003, Incident Response: Computer Forensics Toolkit, Wiley Publishing, Inc., Indianapolis, Indiana, ISBN: 0-7645-2636-7

Shinder D.L., Tittel E., 2002, Scene of the Cybercrime: Computer Forensics Handbook, Syngress Publishing, Inc., USA, ISBN: 1-931836-65-5

Stephenson, P., 2002, Collecting Evidence of a Computer Crime, Computer Fraud & Security, Volume 2002, Issue 11, Pages 17-19

The Oxford Dictionary of English (revised edition). Ed. Catherine Soanes and Angus Stevenson.Oxford University Press, 2012.Oxford Reference Online. Oxford University Press, http://www.oxfordreference.com/

Turvey B. E., 2002, Criminal Profiling: An Introduction to Behavioral Evidence Analysis, Academic Press, 2nd ed., ISBN: 0127050418

Vidalis S., Jones A. and Blyth A., 2004, Assessing cyber-threats in the information environment, Network Security, Volume 2004, Issue 11, Pages 10 - 16

Vidalis S. and Jones A., 2005, Analyzing Threat Agents & Their Attributes, in Proceedings of the 5th European Conference on i-Warfare and Security

The Islamic Republic of Iran's Strategy Against Soft Warfare

Ebrahim Anoosheh
Islamic Azad University, Rafsanjan, Iran
a_anooshin@yahoo.com

Abstract: New information and communication technologies in the emerging post-industrial society have led to new rules and concepts for politics and international relations. The notion of *soft warfare* is related to phenomena now in evidence in this changing environment. Today, various countries in the world – especially those challenging the present international order – are deeply involved deeply with this new concept. The Islamic Republic of Iran is among those countries which considered themselves as targets of *soft warfare*. Concepts such as *cultural invasion*, *cultural incursion* and *soft subversion* are commonly held by authorities of the Islamic Republic of Iran, indicating that the country believes itself to be involved in cyber and soft warfare. Iran's authorities believe that the (imperialistic) West, and especially the USA, is targeting soft warfare against the cultural integrity, national identity and security of Iran. Thus, the Islamic Republic of Iran has attempted to defend itself on the cyber and virtual battlefield by installing a number of negative policies, such as censoring and filtering, but also through some positive measures including improvements in media and satellite infrastructures. This paper investigates positions, strategies and solutions the Islamic Republic of Iran is deploying against new communication and information technologies and, especially, the concept and subject of *soft warfare* and media war.

Keywords: The Islamic Republic of Iran; USA; soft warfare; media war; national security

1. Soft warfare

The concept of *soft* warfare is in conceptual juxtaposition with that of *hard* warfare. There is little consensus on the precise definition of this concept, and thus different governments and movements generally ascribe their own interpretations. According to John Collins, of the USA's National War College, *soft warfare* is the, "…planned use of propaganda and its related tools to penetrate in the mind of the enemy by methods that to result in (cause) the goals of national security of the administrator" (Collins, 1991: pg. 487). The United States Army employs the following definition: "Soft warfare is exact and planned use of propaganda and other actions having the primary purpose of influencing the opinions, emotions, desires and behavior of enemy, impartial groups or friend groups in such a way as to support the achievement of national objectives" (www.tabnak.ir/fa/news/69658)

oft warfare therefore has strong linkages with many terms found in the political and military sciences. In military science terms such as *psychological warfare* or psychological operations are often employed, and in political science terms such as *soft subversion*, *soft threat*, *velvet revolution* and *color revolution* are felicitous examples. With all such terms the common goal lies in imposing one group's decisions on another group without using military methods. (Persianblog.ir)

The new practical and theoretical characteristics of soft warfare were designed and became operational after the onset of Cold War. These achieved a peak in the 1970's through the collaboration of prominent academics in the political and communication sciences. Joseph Nye, Harold Lasswell, John Collins and members of the Central Intelligence Agency, together with top Pentagon commanders, established the "Foreign Hazard Committee" during this period. The main aim of this committee was propaganda bombardment on Eastern Bloc countries, and especially the USSR. After the collapse of the Soviet system the group changed its name to the "Current Peace Committee" which has been active to the present date. (www.tabnak.ir/fa/news/68658)

The most important characteristics of soft warfare can be viewed as less bloodshed and fewer human casualties. Soft warfare reduces the duration of war, the number of physical confrontations and, therefore, the financial costs borne by an army. The winner of such a war is the protagonist that has made the most out of soft warfare tools and related methods. Nowadays, new information and communication technologies provide governments with new capabilities, and it is possible to discuss soft power based on technological advantage. Joseph Nye has stated,

> *"…soft power has a special focus on occupying mental space of other country through providing attraction, and a country will achieve soft power if it can apply information and knowledge to resolve disputes and display disputes in a way to take advantage from*

> *them. Soft power is not political propaganda and it includes rational issues and public values and its target is public opinion of people outside the country and then inside it. Today mass media doesn't think about transferring the realities but rather it tries to make new realities" (Farahani, 2010: pg. 198).*

Thus, *soft power* and *soft warfare* can be said to be compatible. Like soft power, soft warfare influences individual minds, public opinion, beliefs, emotions, etc. The various media hold the most important tools of soft warfare, and the administrators of soft warfare achieve their goals through media and various methods and techniques. The tactics are as follows:

- **Labeling**: In this tactic media change different terms to positive and negative expressions and attribute them to different institutions and people;
- **Censorship**: By removing some parts of the news and reporting others, this tactic creates ambiguities and incites rumors;
- **Character assassination**: Involves the exploitation of media system through exaggeration, defamation, dehumanization and misleading half-truths through which the reputation of a person will be tarnished;
- **Exaggeration**: A tactic by which a reality is shown to be true;
- **Terror and threat**: A tactic used to weaken the morale of an opponent and make him inactive;
- **Repetition**: Timed repetition, repeatedly reporting an issue so as to make it appear fresh and new, allows the message to seem more influential and penetrate the media space;
- **Spinning the message**: In this method, a message is "spun" by different kinds of tactics leading to the eventual indoctrination of its meaning;
- **Generalities**: Connecting thought or action with a certain concept so that the other party will accept it without specific reasons or evidence.
- **Rumor**: Rumor is the verbal transfer of a message to incite the beliefs of an audience and affect their morale (www.tabnak.ir/fa/news/68658)

Psychological operations

Generally, soft warfare – including any kind of psychological action or media propaganda – takes aim at a target society or group and its administrators become involved in a kind of *psychological operation* to achieve their goals. The concept of a *psychological operation* implies propaganda that is used by a country or group for affecting the behavior of other countries or groups based on political, economical, cultural or military logics. These logics may be divided into three categories depending on the positions of the two sides:

- **Two groups in one country**: The psychological operation is between two or more groups to achieve power, but because of its destructiveness the people of the country will be at a loss in the end;
- **Two non-belligerent countries**: The aim is to achieve economic and political privileges and both countries compete for their own national interests;
- **Two belligerent countries**: The desirability of changing the regime and / or making important changes in the target country is displayed in Table 1 (Ardestani, 2005: pg. 45).

Table 1: Goals of a psychological operation between two belligerent countries

The goals of psychological operations	Changing attitudes	Making citizens pessimistic about the government Showing that state authorities are inefficient Discrimination and lack of freedom in the target country The lack of development in the target country in comparison with other countries Suggesting that the aggressor country is a savior Exaggerating the power of the aggressor
	Changing behavior	Legal resistance against the government Illegal resistance against the government (riot) Not resisting against the aggressor country Helping the aggressor penetrate into the country

Generally, the goals of psychological operations are as follows:

- Being hopeless of removing opposing groups or the opponent country;
- Distrusting the opposite group or leaders of the country;
- Division and disagreement among target society (Soltanifar, 2006: pg. 11)

2. Information warfare

Related to soft warfare and psychological warfare (the psychological operation) is *information warfare*. Information warfare is the use and management of information in pursuit of achieving national or personal goals. It refers to actions aimed at accessing the enemy's information infrastructure by which information advantage is obtained. Information war includes all the measures used to obtain or employ information scope (information, information performances and information systems) and preventing the enemy from using them. The main purpose of information warfare is targeting the mind of enemy (military or civilians) by stopping, reducing, delaying, or manipulating information in terms of quality and quantity.

During the last two decades, the focus of information warfare has shifted due to the rapid progress of information technologies such as electronic networks. Two types of information warfare may be defined:

- Network warfare: This represents the war between people, communities and states, or aims at societies and non-military goals. Here the goal is to destroy thoughts and attitudes in a society and substitute others. Network warfare is accomplished in two ways: public opinion, and the opinion of elites. This type of war includes a wide range of propaganda diplomacy, psychological maneuvers and deception through local, media and penetrating into computer networks and databases. Disruption and manipulation of computer networks and information systems are included in this category.
- Cyber warfare: This is war with the aim of disrupting information and telecommunication systems, command and control systems, obtaining secrets, and espionage of enemy military forces in order to render them non-operational in battle or under normal conditions (Saduqi, 2001: pg. 128).

3. Media warfare

Media warfare is an important aspect of soft warfare and new international conflicts. Communication science experts believe that despite the fact that media warfare is mostly used during military wars, it also exists during other periods. Here, the most fundamental definition of media warfare is using media to weaken the target country by using the capabilities and capacity of media – including the press, news agencies, radio, TV, internet and propaganda principles – to defend national resources or invade the target country. Under these conditions media warfare is the only war that continues in peacetime conditions among different countries, unofficially, and each country should do its best to promote its political goals through media (Farahani, 2010: pg. 203)

Apparently, media warfare is observable in radio, TV, press interpreters, news agencies, journalists, news channels and internet sites. But the reality is that behind this journalistic façade there are funds allocated through formal budgets approved by a parliament, or allocated through information and security organizations and information intelligence services. It is possible that people who are bombarded by media warfare are relatively unaware of the "war" in their environment. Although aim of media warfare is to affect the performance of governments, it aims directly at a nation's people rather than targeting directly its government (Eidipour, 2008: pg. 127). Producers of news programs understand the mechanisms of media, and that they can be manipulated to bring war into people's homes and present restricted and censored perception of killing and suffering (Castells, 2006: pg. 526).

4. Media and national security

Today, there is a direct relationship between media and national security. National security issues can be investigated in relation to secure and unsecure factors, a distinction that helps to clarify the relation between internal security and communications, on one hand, and the effects of new communication media on an external dimension, on the other hand. New media in the information era challenge many previous assumptions and principles concerning national security. Developments in communications has made borders penetrable and created significant and problematic issues for governments, causing them to become increasingly sensitive and vulnerable.

Thus, there is an increasing need to determine new borders and recognize that today's security threats are different from traditional ones. The communications revolution has created new security dimensions at national and international levels. Among the national and transnational personal security threats that affect individuals are espionage activities, disclosing personal and governmental secrets, threats of defamation, violence and sex, subversive propaganda against opposing political systems, the spread of rumors and psychological warfare, propagation of lies and false news, insult and slander.

It should be noted that media can be the source of crises at different levels, but also help resolve these crises. The media can turn "Isn't" to "Is", and create a crisis or manage and eliminate it. In other words, media can generate legitimacy, participation, acceptability, reliability and stability, and it can also create crises around these factors for political system and societies. Media can exert influence on security issues, especially national security, in different ways … as in confrontations of legitimacy and acceptability in public opinion, questioning a regime's efficiency, destroying people's trust in their authorities and local media, disturbing communication systems between the government and its people. These are among of the common and new methods of subversion influence media can exert on security issues (Eidipour, 2008: pg. 21).

5. The Islamic Republic of Iran and media

Iranian culture, society and people experience the effects of new information and communication technologies and are being *digitalized* like all other communities, cultures and people. Although they are stepping forward slowly along this path, the direction and future implications of electronics and digitalization are the same as those experienced by other societies. On one hand, Iran belongs to the third world countries and, as Geoffrey Reeves believes (ermegi, culture and democracy), third world countries are suffering from the incorrect strategy of attempting to fully control media and the media's supervision process in order to achieve and maintain political stability. Such countries believe that stability will be realized through the lack of criticism and evaluation, and that citizens should follow the dominant political system in all affairs.

On the other hand, Iran is located in the Middle East where political, geographical and ideological issues have led to the creation of governments that do not accept that self-supervision by the media will insure political stability. Eftekhari (2002: pg. 89) states that Schwedler believes that Middle East governments do not allow media self-supervision media for security actions.

6. Iran's press

At the beginning of the Islamic revolution in Iran there were many issues concerning the range of freedom and performance that should be accorded the press. The Leader of Iran's revolution, Imam Khomeini, stated, "We respect such press that know what is the meaning of freedom of speech and freedom of press. It is said that people are free, does it mean that they can break a person's head?! Are they free to break the law?! Are they free to act against the rules?! Are they free to conspire against government?! These aren't freedom. Freedom is doing something in the realm of rules and reasoning" (Khomeini, sahifeye nor)

But other leaders of Iran's revolution had different views. Shahid Beheshti stated, "I think that in this situation, coercive actions for preventing some journals which are fighting against Islamic thought is not useful and, according to our experience, is neither good nor helpful for Islam". Finally, after controversial discussions involving different views on the performance and freedom of press, it was stated in Article 24 of the Constitution that, "…publications and the press have freedom of expression except when it is detrimental to the fundamental principles of Islam or rights of the public. Its details will be determined by law" (Atazadeh, 2002: pg. 26). In interpretation of these two issues – principles related to Islam and public rights – the following are examples of offenses as defined in the law:

- Insulting Islam;
- Publishing atheistic and anti-Islamic articles;
- Insulting the Supreme Leader of Iran and senior clergymen;

- Encouraging people and groups to commit actions against security, dignity and the interest of the Islamic Republic of Iran in the country or abroad;
- Revealing and publishing confidential documents and instructions;
- Publishing secret negotiations of the ISLAMIC CONSULTIVE ASSEMBLY and private courts of justice;

7. The Islamic Republic of Iran Broadcasting (IRIB)

The Islamic Republic of Iran Broadcasting Organization (the national media, as its name reveals) consists of two sections, Radio and TV. These two sections were established independently at first, with Radio being older than TV. Iran's radio began activity on April 23, 1940 and Iranian television on October 3, 1958. In 1971, the TV and Radio sections were combined into a single organization termed, "Iran's National Radio and TV". This organization had 100 local, international, provincial and urban radio and TV channels until 2006:

- 46 TV channels;
- 47 radio channels with 7 urban channels;
- The channels of Jam-e-Jam 1, 2, 3 are under the TV deputy;
- the channels of Al-Alam, Al-Kusar and Sahar are working abroad under subsidiary satellite channels;
- The radio channels include Iran, Farhang, Javan, Maaref, Quran, Salamat, Varzesh, Payam and Goftego, all of which are in the Radio Department;
- International channels include Sedaye Ashena in the Radio department; Arabic and African (in 4 languages of Arabic, Hausa, Swahili, and Hebrew); Middle Asia (13 languages such as Azeri, Assyrian, Armenian, Uzbek, Tajik, Talish, Turkmen, Turkish (Iran), Turkish, Kazakh, Cossack North Kurdish and Sorani, Georgian); Europe and America (8 languages of Albanian, German, Spanish, English, French, Italian, Russian and Bosnian); subcontinent and Eastern Asia (8 languages of Urdu, Bengali, Pashtu, Chinese, Indian, Japanese, Dari Persian, Malay), all under the direction of International Department.
- In total, there are 29 national TV channels, 19 provincial TV channels, and the international channels are broadcasting on the Internet;
- Of the 165 existing satellites in the world, 67 are in accessible in Iranian space and 8,597 TV channels and 2,707 Radio channels can be received through these satellites (Farahani, 2010: pg. 252).

The operation of this radio and TV organization and its general policy according Iranian law is as follows:

- Article 7: Islamic Republic of Iran Broadcasting belongs to all people of Iran and it should reflect the life of all kinds of people from different classes;
- Article 16: Iran broadcasting, with its permanent presence in the society, should reflect important social events truthfully and try to report the realities to people;
- Article 18: Iran broadcasting should present the latest news, and correct and important information of Iran and the world abroad, that are useful for most of the people as briefly and vividly as possible (Kazemi, 2002: pg. 133).

Despite these general policies regarding the regulation of Iran's broadcasting infrastructure, there are criticisms leveled by a number of experts and theorists. Some university theorists, for example, charge that "news censorship" occurs in the national media whereby some social realities are censored, and others are exaggerated, such that certain phenomena such as social distrust and the (distorted) attraction of a foreign audience occur, and especially in critical situations (Kazemi, 2002. Pg: 149).

8. Internet in Iran

The Internet created a new world by presenting services such as email, FTP file transfer, telnet, chat, the Gopher hierarchy network, the World Wide Web and news groups of Usenet (Ashena, 2002: pg. 217). The advent of the Internet in Iran occurred in 1991, but until the end of 1999 it was available only for governmental systems and only in 2000 did it became accessible to the private sectors. One of the first private companies to use email services in the private sector was Neda Rayaneh Company, an affiliate of the Tehran Municipality. This company started its activity in 1994 and attracted public attention (Anusheh, 2011: pg. 32).

Statistics indicate that in 2008 Iran had the highest number of Internet users in the Middle East (balatarin.com/permlink/2012). According to official reports from Iran's communication and information technology ministry, there are now more than 4.7 million internet users in the country and in comparison with 1.8 million users in 2002, an increase of 161%. Today the Internet is widely used by Iranian families and more than 7,000 Persian sites are active (Ziayiparvar, 2007: pg. 277).

9. Blogs

Blogs are one of the most important internet advantages for the Iranian people, such that Persian blogs in 2004-2005 were ranked first terms of the sheer number of blogs in the world; today, Persian blogs rank tenth in the world. More than 5 million Persian blogs with different services have been registered, of which about 450,000 are active (www.Sharifnews.com).

10. Social networks

Social networks are virtual gatherings in cyber space and the formation of these networks is made possible through email groups, blogs, chartrooms, forums and finding-friends sites such as Orkut. Internet-based social networks such as MySpace and Face book have attained considerable popularity among the American youth, and many Iranians are members of these social networks. The presence of Iranians on some sites such as "Orkut" has always been high: the number of Iranian members of this site is half that of Americans and twice that of Indians (www.Revayat.com.sysnews/cid/6777).

11. Iran and soft warfare

After the country's Islamic Revolution in January 1978, the Islamic Republic of Iran, according to its principle of *no East and no West in international order*, entered into confrontation with the great powers of the East and West, and especially the Western world and the United States. The leaders of Iran's revolution thought themselves to be supporters of weak peoples around the world, and the third world countries, and thus challenged the existing global order. This led to the development of challenging policies contrary to the West and the United States, and increased levels of conflict so that the position of the West and the United States were challenged, especially in the Middle East, and thus imposed various sanctions against Iran.

Accordingly, Iran's leaders today believe that the West and the United States are attempting to interfere in the domestic affairs of the country. These leaders believe that the imperialist world (America) is attempting to defeat the Islamic Republic of Iran through cultural invasion and soft warfare. The development and increasing presence of foreign Persian-language radios and Persian-language satellite TV stations are their evidence in this regard. There are currently more than 32 foreign Persian-language radio stations broadcasting for Iranians and other Persian speakers in the world; some estimate the listeners to number more than 150 million persons. Some of the major channels in this regard are:

- Radio BBC
- Voice of America (VOA)
- Radio Azadi (Radio Farda)
- Germany radio
- France radio
- Moscow radio
- Isreal radio
- China radio
- Japan radio
- Turkey radio
- Romanian radio

- Iran voice radio
- Radio Ashena
- Radio Pajvak
- Persian network radio

12. An example of radio warfare against Iran

Voice of Iran radio was heard for the first time in 1980 and this station claimed that it broadcasted from a free zone in Iran. In a report, this station claimed that Ayatollah Khomeini had obtained power by force from Bakhtiar, the legitimate President of Iran Sedaye Azad demanded that Imam Khomeini resign, otherwise he would be confronted with civil war.

13. Persian satellite TVs

There are now more than 15 TV stations, mostly in America, presenting programs for Persian speakers … including the residents of America, Iran, Tajikistan, Afghanistan and Europe. Some of these stations include Rangarang TV, Meli TV, Pars TV, Persian News Channel, Azadi TV, Tapesh TV, Iran TV channel, International Jam-e-Jam, Omid Iran TV, Persian Global Channel, and Voice of America TV (Ziayiparvar, 2007: pg. 269). Most reports from western journalists are related to political crises, gatherings, conflicts and confrontations between the government and the people, events such as human rights violation, the attempt to achieve nuclear weapons, massacre and breaching political and civil freedoms in Iran. These media try to create a prejudicial environment against Iran by broadcasting global news, or confidential or biased news, from foreign policy systems and intelligence services of their respective countries.

In this regard, the American congress assigned a budget of 40 million dollars to support civil and media activities in Iran. In a CNN interview by a correspondent in America with one of the administrators of Persian-language channels in Los Angeles, it was stated that a large part of this budget was attributed to Persian-language satellite TV stations in Los Angeles. These stations also are supported financially from by which are advocator of monarchy (reality.blogfa.com/post-321). In addition, according to an informed authority there are 97 satellites belonging to 15 countries in the world broadcasting programs for the Iranian people. Some of these satellites broadcast more than 200 TV channels in digital format. Some 600 TV channels are monitored by the Radio and TV organization, and a number of military, cultural and intelligence institutions. Some 15 channels are in the Persian language, and most of these present programs against the Islamic Republic of Iran and its social and cultural norms. The Western worlds' experts even attempted to put these channels in the UHF band, so that their reception would not require a satellite dish or receiver; as a result, any kind of television would be able to receive these satellite programs directly (Ziayiparvar, 2007: pg. 272).

14. American actions against Iran in internet space

After the end of the war of America and England against Iraq, America's foreign affairs ministry set up an internet site in Persian and stated its goal to be, "…bridging the gap between Iran and America" (todaywar.persiablog.ir/1391). This site disseminates the views of the foreign ministry of the USA and represents an attempt to conduct psychological warfare against Iran. It has a leading role in Internet warfare against Iran alongside Persian-language sites of the Voice of America and Radio Farda. The Standard Weekly Journal of Neo-conservatives, in its issue No. 18, July 2005 published a paper written by Jeffery Goodman entitled, "B plan for Iran". Goodman, by revealing the establishment of the Persian-language web site in the USA and its 3 million dollar budget, declared that this infringed on the rights of the democratic people of Iran, and implicitly declared that America's war with Iran had started (www.Revayat.ir.sys/cid/6777).

A few months after Iran applied its Internet filtering policy, it was stated that the American government concluded a contract with a company expert in breaking such filters as well as censorship of Internet content. The Anonymizer Company was directed by America to not allow censorship, and to prevent the Iranian people from having access to internet sites. According to a report on the Security Focus website, a free proxy is considered for the Iranian people in this contract, but this proxy is aimed at filtering porno sites as well (ziaiy.persianblog.com) A bill approved by the American Congress declared that an office termed, "Information Freedom in the World" would be established with a 50 billion dollar budget in order to assist peoples faced with internet censorship (Ziayi, 2007: pg. 314).

15. Internet in a suitcase

The New York Times newspaper writes, "… American government attempts to guide international action to create internet networks and 'alternative' mobile phones in repressive countries" (radio farad.com). The aim of this plan is to provide facilities for people and opponents of such regimes who do not have minimum communication freedom due to dictatorship in their countries. The New York Times noted that one of the methods that America has adopted is the creation of separate mobile phone networks in other countries, or investments to produce tools that can establish separate internet networks that can be snuck into target countries (www.Rediofarda.com/content/f3-usa-internet)

This plan is called i*nternet in suitcase* or *shadow internet* and it is one of the interfering actions and soft warfare tactics of America against certain Middle East countries, including Iran. The Iranian Foreign Ministry spokesman, Ramin Mehman Parast, considered this movement contrary to human rights and declared that the Iranian Cyber Army will respond to this attempt strongly (www.hamshahrionline.ir/news-138121.aspx)

16. Counter measurements of the Islamic Republic of Iran

16.1 The Iranian Cyber Army

One of the actions of the Islamic Republic of Iran in soft warfare and war in cyber space lies in forming a cyber army or passive defense. Mr. Gholamreza Jalali, the head of passive defense has said, "…this place is built to confront the threats of the Islamic Republic of Iran's enemies. …This base has been established by a collaboration between the Communication Organization, National Security Council and Ministry of Intelligence". He continued by saying that the, "…Army of passive defense will take some actions that are used without weapons and equipment for the stability of system and protecting human life". Jalali had previously stated that the Islamic Republic Cyber Army Base would be started soon and invited good-will hackers for collaboration (www.Digarban.com). Defense Tech is an American military and security institution that published an article entitled, "Iranian Cyber Army" which analyzed the Iranian Cyber Army. In this paper, the Iranian Cyber Army is introduced as subsidiary of the Islamic Revolution Guardians Army, and Iran is portrayed as one of five countries with the strongest cyber army with 2,400 persons involved and an estimated budget of 76 million dollars (www.Gerdab.ir).

16.2 National internet

Setting up a national internet is one of the concerns and goals of the Islamic Republic of Iran for soft warfare. In this regard, the Information and Communications Technology Research Institution has declared that its central activity concerning the national internet is reviewing and producing content, and providing multimedia services for the national network. This institution has developed plans for the security of information and sites; evaluation, analysis, security and upgrading of systems and communication networks; and in general, its primary goal is to set up a local internet network for Iranian users (www.Winbeta.Net).

The Communications and Information Technology Minister, Reza Taqipour, has stated that access to 20 Mb/Sec bandwidth for 10 million Iranian users is envisioned in the national internet / national information network, and according to estimations this bandwidth will be ready to meet the demands of users very soon. According to Taqipour, the most important achievement of the Internet National Network and National Data Center is security and protection against internet attacks. It is said that the design of the national internet network is completed, and the first phase of this project will be operational very soon (www.Mehrnews.com/fa/newsdetail.aspx).

16.3 National eMail

The Deputy of Iranian information Technology Company discussed setting up Iranian email at the address of www.iran.ir in the near future. Ali Asqar Ansari stated, "Now governmental managers used national email and they can go to the Iran electronics service www.iran.ir and sign up for email" (mail.iran.ir/mail). Mr. Ansari further stated that the most important characteristics of Iranian email is its high level of information security. Unlike non-Iranian email such as Yahoo, Gmail and Hotmail in which email information is transferred abroad, information in this email system circulates only in Iran (www.winbeta.net).

16.4 Filtering

Filtering is one of the practical measures and negative policies of the Islamic Republic of Iran in soft warfare and cyber space. The Minister of Post, Telegraph and Telephone of Iran, Dr. Moetamedi, has stated, “The principle of filtering is good and this issue is one of the problems of Iranian families but its instances are ambiguous. Instances of filtering should be defined well but we are not responsible for it and a responsible authority is needed”.A committee consisting of 3 people has been established by the Supreme Council of Cultural Revolution, including representatives from the Ministry of Islamic Guidance, Ministry of Intelligence, and the Radio and TV Organization for evaluating filtering issues. About 170 sites have been closed, of which 160 were related to anti-revolutionary, or anti-moral issues insulting holy concepts. Fewer than 10 political sites have been closed as well. Clearly, for such acts it is necessary that a bill be formulated by the judicial system concerning internet offences, which then should be approved by the government and introduced to parliament for legislation (Ziayipour, 2007: pg. 314).

16.5 Collection of satellite dishes

Article 10 which bans the use of satellite dishes since 1994 stipulates that the, “…Ministry of Culture and Islamic Guidance in collaboration with the Ministry of Post, Telegraph and Telephone and related organizations, are obliged to protect the cultural borders of Iran and families against vulgar and obscene programs through legal and international rules” (Legal group of TV and Radio, 2001, 175). Despite Iranian laws in which the importation, storage, installation, sale or use of a receiver and related satellite equipment is forbidden, the past years have witnessed a considerable increase in the use of satellite equipment and the number of satellite channel audiences has been soaring. Thus, sometimes the Iranian police obtains legal permission to enter residential buildings and confiscate satellite equipment and dishes, and the Iranian authorities believe that these measures are taken to protect the society against the West’s cultural invasion (www.mfatorehchi.blogfa.com/pst-241.aspx).

17. Conclusion

Given the emergence of new communication and information technologies, new concepts have been introduced to the political, military and communication sciences. Some concepts such as soft warfare, media warfare, information warfare, psychological operations and so on are employed by belligerent governments in order to achieve their national goals and interests. Undoubtedly, due to its technological capabilities and political and economical power the United States is one of the strongest proponents of soft warfare today.

The Islamic Republic of Iran, due to its international relations policies and confrontation with the West and the United States, is one of the countries involved in soft warfare. Thus, Iran has attempted to be more powerful by improving its technological capacity in communication and media. Measures such as the cyber army, launching satellites into space and setting up satellite channels such as Al-Alam, Sahar and Jam-e-Jam are some examples of this dedication. Some of the Islamic Republic of Iran’s measures are seen as actions restricting access in Iran to different foreign media being deployed in the field of soft and cultural warfare, and these negative measures bear heavier negative consequences than benefits. It seems in order to succeed in this field; the Islamic Republic of Iran is required to improve communication and media technologies to provide more services for at least some of its users and citizens. Some concepts such as *media knowledge*, *informing* and *scientific education* at the elementary levels are necessary, and negative policies such as restriction and filtering seem a poor choice. Finally, it is necessary that political and security views concerning media and communication systems be attenuated, and replaced by scientific and rational views … since political and security views may well be discredited by the Iranian people resulting in their adherence to foreign media.

References

Akhavan Kazemi, Bahram, 2002 “TV and Radio: political stability, national unity and security in media and political stability.” Tehran, Strategical studies research center.

Anushe, Ebrahim, 2011. “Living in cyber space: internet threats and opportunities.” Tehran, Aryaban.

Ashena, Hesamaldin, 2002. “Internet and political–social stability of Islamic Republic of Iran.” First Edition, Tehran, Strategical Studies Research Center.

Ardestani, Hassan, 2005. “Psychological operations and riot.” Journal of Psychological Operations, Second year, No. 8.

Eftekhari, Ali Asqar, 2002, “Political media stability in media and political stability.” Tehran, Strategical Studies Research Center.

Atarzadeh, Mojtaba, 2002. “The press and political stability in Islamic Republic of Iran.” In Media and Political Stability, Tehran, Strategical Studies Research Center.

Castells, Manuel, 2006. “Information age: The rise of the network society.” Translated by Aliqolian, Tehran, Tarh-e No.

Collins, John, 1991. “Big strategy: Principles and basics.” Translated by Kurush Baynder, Tehran, Ministry of Foreign Affairs.

Soltanifar, Mohammad, 2006. “The design of new scenario: Making news by principles of psychological warfare.” Two journals of media studies, first year, No. 3.

Saduqi, Moradali, 2001. “Information technology and national sovereignty.” Tehran, Ministry of Foreign Affairs.

Eidipur, Makan, 2008. “Media and national security.” Tehran, Islamic Azad University.

Farahani, Fatemeh, 2010, “Cultural development and national media,” Tehran, Islamic Azad University.

Legal studies group of TV and Radio, 2001. Legal challenges of broadcasting satellite programs, Tehran, TV and Radio publication.

Ziayipour, Hamid, 2007. “Soft warfare, media warfare.” Tehran, Abrar Moaser.

http.www.persianblog.ir

http.www.tabnak.ir/fa/news/69658

http.www.sharifnews.com

http.www.revayat.ir/sysnews/cid/6777

http.www.radiofarda.com/content/fa-usa-internet

http.www.hamshahri.ir/news-138121.aspx

http.www.digarban.com/node/4641

http.www.gerdab.ir/fa/news/503

http.www.winbeta.net

http.www.mfatorehchi.blogfa.com/post-241.aspx

http.www.mehrnews.com/fa/newsdetail.aspx

Recent Cyberwar Spectrum and its Analysis

Rabia Aslanoglu and Selma Tekir
Izmir Institute of Technology, Izmir, Turkey
rabiaaslanoglu@iyte.edu.tr
selmatekir@iyte.edu.tr

Abstract: War is an organized, armed, and often prolonged conflict that is carried on between states, nations or other parties. Every war instance includes some basic components like rising conditions, battlespace, weapons, strategy, tactics, and consequences. Recent developments in the information and communication technologies have brought about changes on the nature of war. As a consequence of this change, cyberwar became the new form of war. In this new form, the new battlespace is cyber space and the contemporary weapons are constantly being renovated viruses, worms, trojans, denial-of-service, botnets, and advanced persistent threat. In this work, we present recent cyberwar spectrum along with its analysis. The spectrum is composed of the Estonia Attack, Georgia Attack, Operation Aurora, and Stuxnet Worm cases. The methodology for analysis is to identify reasons, timeline, effects, responses, and evaluation of each individual case. Moreover, we try to enumerate the fundamental war components for each incident. The analysis results put evidences to the evolution of the weapons into some new forms such as advanced persistent threat. Another outcome of the analysis is that when approaching to the end, confidentiality and integrity attributes of information are being compromised in addition to the availability. Another important observation is that in the last two cases, the responsive actions were not possible due to the lack of the identities of the offending parties. Thus, attribution appears as a significant concern for the modern warfare. The current sophistication level of the cyber weapons poses critical threats to society. Particularly developed countries that have high dependence on information and communication technologies are potential targets since the safety of the critical infrastructures like; healthcare, oil and gas production, water supply, transportation and telecommunication count on the safety of the computer networks. Being aware of this fact, every nation should attach high priorities to cyber security in his agenda and thus behave proactively.

Keywords: cyberwar, Estonia attack, Georgia attack, operation aurora, stuxnet worm

1. Introduction

Among all centuries, a human being starting from his birth somehow got included in a group, community, society or nation. All these group titles, their functions and their legitimacy has changed throughout history; but the thing that remained the same is continual conflicts between different groupings, which conveyed the concept of war.

War is an inevitable fact of humanity. In every war, one recognizes some basic components like rising conditions, battlespace, weapons, strategy, tactics and consequences. Rising condition is growing reasons of conflict; battlespace signifies the realm of war whether it's air, sea, or land; weapon is an army's means of defence or offence; strategy can be stated as a high-level plan in order to achieve a defeat; and tactic is a way of using the appropriate weapons or resources to fulfill a strategy; and finally consequences are the victories or defeats, and effects on the opposing parties. Among these components, weapon is the transforming one by evolving from sword, arrow, spread to gun, rocket, nuclear weapons and medical weapons along the timeline.

In the 21st century, the war concept has experienced a paradigm shift in terms of battlespace and weapon components. Today the new battlespace is cyber space and the contemporary weapons are constantly being renovated viruses, worms, trojans, denial-of-service, botnets etc. The strategy is being adjusted through these weapons to damage core attributes of information security using propaganda, espionage or destruction of critical infrastructures. The addressed security attributes are confidentiality, integrity and availability which is known as CIA triad.

CIA is comprised of three criteria to evaluate information systems security (MU, 2008). The first criterion, confidentiality is preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information . A loss of confidentiality is the unauthorized disclosure of information. Integrity is guarding against improper information modification or destruction, and includes ensuring information non-repudiation and authenticity. A loss of integrity is the unauthorized modification or destruction of information. As for availability, it is ensuring timely and reliable access to and use of information. A loss of availability is the disruption of access to or use of information or an information system (Evan, 2004).

The more primitive cyber attacks were aimed at the availability attribute of information security and left observable symptoms. However, with the evolution of cyberwar weapons, the recent cyber incidents compromised confidentiality and integrity attributes in addition to the availability attribute and rendered themselves hard to recognize.
This study will cover Estonia Attack, Georgia Attack, Operation Aurora and Stuxnet Worm cases to put some evidences to the evolution of weapons of cyber war and their effects to security attributes.

2. Estonia cyberwar

Estonia is located in the Baltic Region of Northern Europe. It declared its formal independence on the 20th of August 1991 from Soviet Union. With the population of 1,34 million, it is the least populous member of Eurozone and referred as a developed country by United Nations (Wikipedia, 2011a).

Estonia attack is the first known public cyberwar case in terms of its effectiveness. It brought the country's IT infrastructure to a standstill though Estonia was one of the most developed countries in Europe with ubiquitous usage of information and communication technologies in all areas of life. All government organizations in Estonia have used X-Road (Figure 1) to interconnect with each other since 2001 and all citizens have an ID-card which allows them to connect with the goverment organizations and banks (Afrinic-11, 2009).

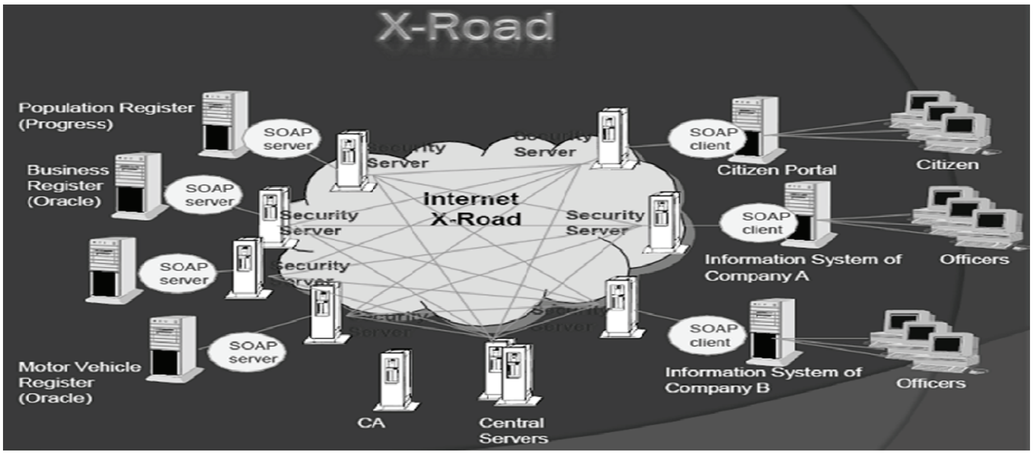


Figure 1: X-Road architecture (Afrinic-11, 2009)

Cyberattacks on Estonia began on April 27th 2007 since the Estonian Government moved a controversial Soviet-era World War II memorial from a square in the capital city of Tallinn to a more secluded location (Traynor, 2007). Protests erupted in Estonia and Russia, where Estonia's Moscow embassy was blockaded (Silverman, 2008).

A series of politically motivated cyber attacks targeting government portals, parliament portal, banks, ministries, newspapers and broadcasters of Estonia hit it and lasted three weeks. Typical attacks were phishing, email spam, web site defacing, syn/ICMP floods, botnets.

In the first wave of attacks, DDoS occurred to Estonia's ISP and governmental websites. Command of ICMP attacks posted to various boards, blogs and chats on Russian Internet. These commands converted into a batch file and uploaded to a web address (Afrinic-11, 2009).

In the second wave, Livejournal users have posted a list of email addresses of Estonia's parliament deputies. These posts were urging users to share the list of emails and caused multiple letters to be sent to the Estonia's deputies with "congratulations of the Victory Day". This action resulted millions of letters being sent and led to mail server mainframes' failure for 2 days.

In the third wave , Estonia's websites were attacked with various tools such as SQL injections (known vulnerabilities in Apache, PHP). Script kiddies were stoked into fervour by President Vladimir Putin's speech (Afrinic-11, 2009).

The attack heavily affected all network infrastructure; leaving damaged routers, changed routing tables, overloaded DNS servers, failed e-mail servers. Estonian Presidency and its Parliament, country's government ministries, political parties, two biggest banks, governmental ISP, telecom companies experienced interoperability problems.

After the attack, Estonia closed down the site under attack to foreign internet addresses and kept the sites only accessible to domestic users. 99% of bogus traffic coming from outside was cut and all ".ru" domain was blocked. Also, bots from DNS servers were identified and blockaded. Estonia's Computer Emergency Response Team (CERT) acted as a coordinating unit, concentrating its efforts on protecting the most vital resources. It persuaded ISPs around the world to blacklist attacking computers which overwhelmed Estonia's bandwith (Afrinic-11, 2009).

The methods used in this attack were not new. However, the country's small size and high reliance on information and communication technologies made the attack a significant threat. As Estonia had an established IT infrastructure, the incident was handled appropriately before the damage was grown. The weapons of this case are phishing, e-mail spam, web site defacing, Syn/ICMP floods, botnets, DDoS which damage availability. The strategy is to test IT infrastructure and make the digital services temporarily down in order to retaliate for the removal of the Soviet-era memorial and prove the Russian power. The consequence is the access failure to the existing information systems in the result of availability disruption (Table 1).

3. Georgia cyberwar

Georgia is an old Soviet Union member which is located at the crossroads of Western Asia and Eastern Europe with 4,7 million population (Wikipedia, 2011b).

Statistics about the Georgian ICT sector show that Georgia has 7 Internet users per 100 people (UNdata, 2006). Considering the geography of the region, Georgia has few options for Internet connectivity via land routes, namely Turkey, Armenia, Azerbaijan, and Russia. According to some sources, most of Georgia is, in terms of Internet infrastructure, dependent on Russia; more of Georgia's connections to the Internet pass through Russia than any other country, comprising nearly half of Georgia's thirteen links to the worldwide network (Today, 2008).

Conflict which caused the cyber attack against Georgia was started in August 2008 between Russia and Georgia over South Ossetia. On August 7th , Georgian forces launched a surprise attack against separatist forces in South Ossetia who was supported by Russia (Tikk, 2008). On August 8th , Russia responded to Georgia's act by military operations into Georgian territory, which the Georgian authorities viewed as Russia's military aggression against Georgia (MFAG, 2008).

By late August 7th, before the Russian invasion into Georgia commenced, cyber attacks were already being launched against a large number of Georgian governmental websites(Tikk, 2008), making it among the first cases in which an international political and military conflict was accompanied by a coordinated cyber offensive. On the August 8th, the President of Georgia, Mikheil Saakashvili, informed the international community of having begun mobilisation, and on August 9th, 2008, Georgia imposed a "state of war" (Saakashvili, 2008).

The methods of these attacks primarily included defacement of public websites containing Mikheil Saakashvili and the National Bank; and launch of DDoS attacks against government sites, important media sites, financial institutions (Tikk, 2008).

In this cyber incident, numerous targets and methods are similar to attacks used in Estonia. Several Russian blogs, forums, and websites spread a Microsoft Windows batch script that was designed to attack Georgian websites. Instructions of these downloadable scripts to ping flood Georgian government websites and lists of vulnerable Georgian sites were distributed on Russian websites and message boards. Emails of Georgian politicians were subjected to targeted attacks and spamming. Georgia has two main players on the Georgian Internet Access and services market; United Telecom and Caucasus Network. United Telecom of Georgia router was unavailable and incapable of providing service for several days. Caucasus Network was flooded with excessive queries. On August 9th, the National Bank of Georgia ordered all banks to stop offering electronic services.

After the attack, some of the damaged websites remained online and did not really make any changes to defend themselves. A few of them temporarily changed their IP addresses to loop back to the originating network in an attempt to thwart the attacks. A few others also changed hosts. The websites of the Ministry of Defence and the President were relocated to Tulip Systems, Inc., located in Atlanta, Georgia, USA, and the website of the Ministry of Foreign Affairs was moved to an Estonian server. The Office of the President of Poland provided their website for dissemination of information and helped to get Internet access for Georgia's government after breakdowns of Georgian local servers caused by cyber attacks. CERT Poland analyzed IP data and sent out abuse messages, while CERT France helped with collecting log files. Security specialists from CERT Estonia also visited Georgia in order to assist the local CERT by providing their know-how and experience. As it's apparent from the examples, international cooperation and assistance were offered, international awareness was raised, and media attention was drawn (Tikk, 2008).

Although the methods used in the attack were the same as Estonia case, the density of the damage was higher than it. Georgia couldn't manage the incident properly due to unsound IT infrastructure and high dependency on neighboring countries for internet connectivity. After all the most important disruption is that the timing of the cyber incidents coincided with the physical damages caused by the ongoing armed conflict and this situation resulted in discredit to the authority.

The weapons of this case are DoS, DDoS, web site defecement, TCP SYN floods, TSC RST flood phishing, e-mail spam which target availability attribute as was in Estonia cyber case but the intensity of the attacks was higher than Estonia. The strategy is to exploit unsound IT infrastructure in order to support the military conflict in cyberspace. All these heavy consequences occurred due to availability disruption (Table 1).

4. Operation aurora

On January 12th, 2010 Google publicly disclosed that they were under a highly sophisticated and targeted attack on their corporate infrastructure originating from China that resulted in the theft of intellectual property from Google. It claimed that the attackers were interested in accessing Gmail accounts of Chinese dissidents, as well. Google was not the only victim of this attack; at least twenty other large companies from a wide range of business including; the internet, finance, technology, media and chemical sectors have been similarly targeted (Drummond, 2010).

In Estonian and Georgian cases; cyber attacks occurred against users and after a while, they became aware. However, Aurora attack's landscape is against software; which exploited a browser vulnerability and occurred silently without user awareness. The attack leveraged a previously unknown vulnerability in Internet Explorer to compromise systems at Google, Adobe and more than 30 large companies. According to McAfee, primary goal of the attack was to gain access to and potentially modify source code repositories at these high tech, security and defense contractor companies (Kurtz, 2010). It completed its attack in six steps (McAfee, 2010):

- A targeted user received a link in email or instant message from a "trusted" source.
- The user clicked on the link which caused them to visit a website hosted in Taiwan that also contained a malicious JavaScript payload.
- The user's browser downloaded and executed the malicious JavaScript, which included a zero-day Internet Explorer (IE) exploit.
- The exploit downloaded a binary disguised as an image from Taiwan servers and executed the malicious payload.
- The payload set up a backdoor and connected to command and control servers in Taiwan.
- As a result, attackers had complete access to internal systems. They targeted sources of intellectual property, including software configuration management (SCM) systems accessible by the compromised system. The compromised system could also be leveraged to further penetrate the network.

Aurora employed an advanced persistent threat (APT) technique that proved extremely successful in targeting, exploiting, accessing, and exfiltrating highly valuable intellectual property from its victims (McAfee, 2010). APT is named by the United States Air Force analysts in 2006 in order to facilitate discussion of intrusion activities with their uncleared civilian counterparts (Daly, 2009).

Advanced means the adversary is conversant with computer intrusion tools and techniques and is capable of developing custom exploits (Daly, 2009).

In Aurora case, attackers gained initial access to the victim's network with a targeted spear phishing attack against the company. Several employees of the victim companies received an email that appeared to be from someone they trusted. However, the email contained a link to a Taiwanese website that hosted malicious JavaScript. The malware, in turn, exploited IE vulnerability. The exploit triggers when IE attempts to access memory that has been partially freed. In short, sophisticated attackers fulfilled the *advanced* criterion of APT (Binde, 2011).

Persistent means the adversary intends to accomplish a mission. They receive directives and work towards specific goals (Daly, 2009).

In Aurora case, after gaining a foothold in the victim companies, the attackers employed the exploited workstations to compromise other internal resources. The attackers then targeted SCM systems and exfiltrated source code to the attacker's command and control servers which was with innocuous-sounding domain names such as homelinux.org, ourhobby.com and servebeer.com (Lelli, 2010). In January 2010, Google was the first to publicly disclose loss of intellectual property. The attackers accomplished a specific mission meeting the *persistence* criterion of APT (Binde, 2011).

Threat means the adversary is organized, funded and motivated. Furthermore, there are objectives that may be political, economic (e.g., theft of intellectual property), technical or military (identification of weaknesses) (Daly, 2009).

The attacks traced back to two Chinese schools, Shanghai Jiaotong University and Lanxiang Vocational School. Jiatong hosts one of the top computer science programs in China. In 2010, it beat Stanford University in an international programming competition sponsored by IBM. Lanxiang is a large vocational school established with military support, training some computer scientists for the military. The school is operated by a company with close ties to Baidu, a strong Chinese competitor to Google. Sources within the schools denied organizational involvement in the attacks (Markoff, 2010). The adversaries, whatever their actual identities, demonstrated high motivation, were adequately funded, and were part of a structured organization and this meets the *threat* criterion (Binde, 2011).

APT is the weapon of this attack and the strategy is to exploit internet vulnerabilities in order to steal intellectual property and retaliate to Chinese human right activists by spying. After the incident, modification of source code in repositories, theft of trade secrets and unauthorized access to e-mails of Chinese human right activists prove the disruption of availability, confidentiality and integrity respectively (Table 1).

5. Stuxnet worm

Stuxnet increased attack sophistication level when it discovered in June, 2010. Months later Iran confirmed that centrifuges for uranium enrichment production at Natanz were affected and potentially damaged by it. Stuxnet was unique and did not follow traditional Web threat patterns and tactics (Clare, 2011). Also, it has apparently infected over 60,000 computers, more than half of them in Iran; other countries affected include India, Indonesia, China, Azerbaijan, South Korea, Malaysia, the United States, the United Kingdom, Australia, Finland and Germany. The virus continues to spread and infect computer systems via the Internet, although its power to do damage is now limited by the availability of effective antidotes, and a built-in expiration date of 24 June 2012 (Farwell, 2011).

Bruce Schneier's analysis on Forbes.com suggests air-gapped Windows systems were infected by the Stuxnet worm via USB, that four unpublished and highly valuable zero-day vulnerabilities were exploited, and that Stuxnet looks for a particular model of a Programmable Logic Controller (PLC) manufactured by Siemens, and its related controller software. Stuxnet installs its own driver in Windows using stolen certificates to legitimize itself, and checks back with two control servers for updates. It uses a peer-to-peer update scheme when it encounters itself, so the most current version is always utilized (Clare, 2011).

This was an unprecedented sophisticated attack that would have wide implications for future industrial systems because it has broken down common beliefs about control systems security on industrial infrastructures. Before the Stuxnet, there was an image that control systems were safe as far as new USB was used for data exchange and internet was not connected; and a virus could be monitored thanks to abnormal behaviours of related computer (Miyachi, 2011). Additionally, it occurred at an extremely critical time as industrial systems move towards the adoption of Internet based technologies and architectures (Karnouskos, 2011b). Its main target is industrial control systems with the goal of modifying the code running in PLCs in order to make them deviate from their expected behavior (Matrosov, 2010). This deviation would be small and only noticeable over a longer period of time. In parallel great effort was put by the Stuxnet creators in hiding those changes from the operators, even imitating “legitimate” data. To increase the success rate, a vast majority of security holes and tools was used such as rootkits, antivirus tricking, zero-day exploits, network discovery and peer-to-peer updates, process injection etc. Many of these are common on modern PCs but the sophistication of the attack was unprecedentedly well-planned and highly customized for specific industrial systems. It is believed that the development of such a highly sophisticated worm was a joint-effort with experts from different specializations and a huge investment in time and cost (Karnouskos, 2011a).

All of them make Stuxnet another Advanced Persistent Threat. The evidences are stated below:

Advanced: The original infection of the Windows computer may be done via simply plugging in a USB flash drive or from internal network if an infected machine exists. It uses stolen certificates in order to legitimize itself and then installs its own driver in Windows. When it encounters itself, it uses peer-to-peer update scheme.

Persistent: The target was solely Siemens SCADA systems targeting very specific industrial processes. Stuxnet infects project files of Siemens WinCC/PC S7 SCADA control software and intercepts the communication between the WinCC running in Windows and the attached PLC devices when the two are connected via a data cable (known as “man-in-the-middle” attack). It focused on identifying specific slave variable frequency drives attached to the Siemens S7-300 system. Furthermore it has been reported that it would only attack specific provider of those PLC systems. However in order to have a more specialized target, it monitors the frequency of the attached motors, and only attacks systems that spin between a specific range. Then it installs malware on the PLC that monitors the Profibus of the system and under certain conditions it periodically modifies that frequency, which results in that the connected motors change their rotational speed. Additionally it has installed the first known industrial rootkit which fakes industrial process control sensor signals, hence no alarms or shutdown is done due to abnormal behavior. This slowly deviating behavior in combination with the projection of “legitimate” data results in difficulty to assess what is malfunctioning and to pinpoint the faults before it is too late (Karnouskos, 2011a).

Threat: It is the first purpose-built worm designed to attack PLC, industrial control systems that help run critical infrastructure environments. As such, it can be hypothesized that Stuxnet was designed purely to attack PLCs and cause damage to the infrastructure they operate and, ultimately, to the people and organizations that depend on that infrastructure. Stuxnet is clearly an example of a stealthy worm developed by an adversary that spent a great deal of time and money on research and development. While the origins are still unknown, many experts feel that it was likely developed by a nation-state with nefarious intent driven by political rather than economic motivations (McAfee, 2011). Stuxnet’s design and architecture are not domain specific; it is a tool for APT. Hence with some modifications it could be tailored as a platform for attacking other systems e.g. in the automobile or power plants. Monitoring and control systems such as SCADA/DCS are responsible for managing critical infrastructures operate in these environments. Its highly sophisticated actions may prevent detection until it is too late (Karnouskos, 2011a). For Iran case, the strategy is to exploit control systems vulnerabilities in order to damage country reputation and avoid attribution. As a result of the attack Iran’s centrifuges were affected and potentially damaged and over 60,000 computers from other countries were infected. These happenings are indicators of availability, integrity and confidentiality disruption.

By the appearance of this worm, cyber security has become a high-priority item in agendas. Barack OBAMA’s the following statement supports this argument: “Cybersecurity is a matter of public safety and national security. We count on computer networks to deliver our oil and gas, our power and our water, public transportation and air traffic control.” (Obama, 2009). So high dependence on information and communication technology brings about new hidden weapons pointed at well-developed countries by exploiting their critical infrastructures. Stuxnet is the peak point of cyberweapons came.

6. Discussion

In Table 1, each war case is examined in terms of the associated war components. Considering the contents of the table and thus recognizing the tactical and weaponry differences, the examined war incidents can be divided into two groups: Estonia-Georgia and Aurora-Stuxnet respectively. In the former group, the main tactic is propaganda and the weapons are traditional cyber threats while in the latter, espionage is the common tactic and the weapons are referred as Advanced Persistent Threat.

Both Estonia and Georgia cases stand on political disputes. Although the weapons and consequences are partially similar to each other, Georgia experiences higher damage because of its unsound IT infrastructure. These cases are references of disruption in which masses of attack are traceable and familiar in landscape.

Commercial competition lies behind the Aurora case. It heats up the attack sophistication level by occurring silently without user awareness. Succeeding Aurora, Stuxnet is the peak point for cyberweapons thanks to its unprecedented attack combination and sophistication. Particularly this unique characteristic renders Stuxnet resistant to attribution. In both Aurora and Stuxnet cases, offensive parties were conversant with computer intrusion tools and techniques to accomplish a mission receiving directives, organized, funded, motivated for economic and political objectives exhibiting an Advanced Persistent Threat pattern.

Table 1: General picture of the cases

War	Estonia 2007	Georgia 2008	Aurora 2009	Stuxnet 2010
Rising Conditions	Political; elocation of a Soviet war memorial.	Political; occurrance of Georgian surprise attack against separatist forces in South Ossetia.	Global competition, political; Google's internet domination and Chinese human right activists' activities.	Political Keeping Iran's Uranium enrichment under control.
Battlespace	Cyberspace	Land & Cyberspace	Cyberspace	Cyberspace
Weapons	DoS and DDoS, defacement, e-mail and comment spam, Some targeted hacks using exploits/SQL injections.	DoS and DDoS, defacement, TCP SYN floods, TCP RST flood.	Advanced Persistent Threat.	Advanced Persistent Threat.
Goal	Retaliate for the removal of the Soviet-era memorial, prove the Russian power.	Support the military conflict in cyberspace.	Steal intellectual property, retaliate to Chinese human right activists.	Damage reputation, avoid attribution.
Strategy	Test IT infrastructure, make the digital services temporarily down.	Exploit unsound IT infrastructure.	Exploit internet vulnerabilities.	Exploit control systems vulnerabilities.
Tactics	Propaganda.	Propaganda.	Espionage.	Espionage.
Consequences	The disruption of access to related web sites and use of information of e-mail servers, bank system and telecom system occurred. (Availability disruption).	The disruption of use of information of e-mail servers, bank system and telecom system government sites, important media sites, financial institutions occurred. Higher intensity attacks compared to Estonia. (Availability disruption).	Modify source code in repositories, steal trade secrets and read e-mails of chinese human right activists (Availability, confidentiality and integrity disruption).	Iran's centrifuges were affected and potentially damaged, infected over 60,000 computers from other countries,too. (Availability, confidentiality and integrity disruption).

7. Conclusion

War is an inevitable fact of humanity. In the 21st century, the war concept has experienced a paradigm shift in terms of battlespace and weapon components. Today the new battlespace is cyber space and the contemporary weapons are constantly being renovated viruses, worms, trojans, denial-of service, botnets etc. Its current strategy is being adjusted through these weapons to damage core attributes of information security using propaganda, espionage or destruction of critical infrastructures.

This study aims to cover Estonia Attack, Georgia Attack, Operation Aurora and Stuxnet Worm cases to put some evidences to the evolution of weapons of cyber war and their effects to security attributes. The presented cyber attack cases are the important instances of the past five years. Most of the cyber weapons are not new in information technology but their unprecedented combination and sophistication have threatened well-developed countries in the matter of exploiting.

References

Afrinic-11 (2009) “Estonia Cyber Attacks 2007”, [online], http://meeting.afrinic.net/afrinic-11/slides/aaf/Estonia_cyber_attacks_2007_latest.pdf.

Binde B., McRee R. and O’Connor T. J. (2011) Assessing Outbound Traffic to Uncover Advenced Persistent Threat, SANS Technology Institute.

Clare, T. (2011) “2011 WebSecurity Report”, *White Paper*, Blue Coat.

Daly, M. K. (2009) “The Advanced Persistent Threat”, *23rd Large Installation System Administration Conference.*

Drummond, D. (2010) “A new approch to China”, [Online], *The Official Google Blog*, http://googleblog.blogspot.com/2010/01/new-approach-to-china.html.

Evan, D. L., Bomb, P. J. and Bement A. L. (2004) *Standards for Security Categorization of Federal Information and Information Systems*, Department of Commerce, Federal Information Processing Standards Publication: National Institute of Standards and Technology.

Farwell, J. P. and Rohozinski, R. (2011) “Stuxnet and the Future of Cyber War”, *Survival: Global Politics and Strategy,* 23-41.

Karnouskos, S. (2011a) “Stuxnet worm impact on industrial cyber-physical system security”, *IECON 2011-37th Annual Conference on IEEE Industrial Electronics Society*, 4490-4494.

Karnouskos, S. and Colombo, A. W. (2011b) “Architecting the next generation of service-based SCADA/DCS system of systems”, *IECON 2011 - 37th Annual Conference on IEEE Industrial Electronics Society*, 359-364.

Kurtz, G. (2010) Operation “Aurora” Hit Google, Others, [Online], *McAfee Blog* http://siblog.mcafee.com/cto/operation-%E2%80%9Caurora%E2%80%9D-hit-google-others/.

Lelli, A. (2010) “Trojan.hydraq technical details”*,* [Online], Symantec Corporation: Symantec Corporation, http://www.symantec.com/security_response/writeup.jsp?docid=2010-011114-1830-99&tabid=2.

Markoff, J. and Barboza D. (2010) “2 china schools said to be tied to online attacks”, [online] *The New York Times*, http://www.nytimes.com/2010/02/19/technology/19china.html?_r=1.

Matrosov, A., Rodionov, E., Harley, D. and Malcho J. (2010) “Stuxnet under the microscope”, *Technical Report,* ESET.

McAfee (2011) “Advanced Persistant Threat”, *Solution Brief.*

McAfee (2010) “Protecting Your Critical Assets, Lessons Learned from Operation Aurora”, *White Paper*, McAfee Foundstone Professional Services.

MFAG (2008) “Information for press”, Press and Information Department, [press released], 8 August 2008, http://www.mfa.gov.ge/index.php?lang_id=ENG&sec_id=461&info_id=7193&date=2008-0808&new_month=08&new_year=2008.

MU (2008) “Confidentiality, Integrity and Availability (CIA)”, [Online], University of Miami, http://it.med.miami.edu/x904.xml.

Miyachi, T., Narita, H., Yamada, H. and Furuta, H. (2011) “Myth and reality on control system security revealed by Stuxnet”, SICE Annual Conference, pp 1537-1540.

Obama, B. (2009) “Remarks by the Presidet on securing our notion's cyber infrastructure”, The White House, [press released], 29 May 2009, www.whitehouse.gov/the-press-office/remarks-president-securing-our-nations-cyber-infrastructure.

Saakashvili, M. (2008) “News”, Press Office of The President of Georgia, [press released], 8 Sept 2008, http://www.president.gov.ge/en/PressOffice/News?2273.

Silverman, J. (2008) “Cyber Attacks in Estonia”, [Online], HowStuffWorks, http://computer.howstuffworks.com/die-hard-hacker1.htm.

Tikk E., Kaska K., Rünnimeri K., Kert M., Taliharm A. and Vihul L. (2008) “Cyber Attacks Against Georgia:Legal Lessons Identified”, Nato, Cooperative Cyber Defence Centre of Excellence.

Today, S. (2008) “Georgia, Russia: The Cyberwarfare Angle” [Online], www.stratfor.com/analysis/georgia_russia_cyberwarfare_angle.

Traynor, I. (2007) “Russia accused of unleashing cyberwar to disable Estonia”, [online], The Guardian, http://www.guardian.co.uk/world/2007/may/17/topstories3.russia.

UNdata (2006) “Internet users per 100 population, 2006” [online], http://data.un.org/Data.aspx?d=MDG&f=seriesRowID:605.

Wikipedia (2011a) “Estonia” [Online], http://en.wikipedia.org/wiki/Estonia, [accessed on 02/01/12].

Wikipedia (2011b) “Georgia” [Online], http://en.wikipedia.org/wiki/Georgia, [accessed on 02/01/12].

Metrics Framework of Cyber Operations on Command and Control

Melanie Bernier[1], Sylvain Leblanc[2] and Ben Morton[2]
[1]Defence Research and Development Canada - Centre for Operational Research and Analysis, Ottawa, Canada
[2]Computer Security Laboratory, Department of Electrical and Computer Engineering, Royal Military College of Canada, Kingston, Canada
Melanie.Bernier@drdc-rddc.gc.ca
Sylvain.Leblanc@rmc.ca
Ben.Morton@rmc.ca

Abstract: The reliance of modern military forces on networks and information systems makes them susceptible to cyber attacks and highlights the importance of cyber operations. This increased awareness of cyber operations has led to a need for concept development and experimentation. Concept development and experimentation work must be assessed, which requires measurement and metrics. To date, little work has been done to measure the impact of cyber operations on military command and control. This paper will address this requirement by putting forward a framework for the measurement of the impact of cyber operations on the effectiveness of the command and control of military missions. There have been many research efforts to describe measurement in the following capabilities: Command, Control, Communications, Computers, Intelligence, Surveillance and Reconnaissance (C4ISR), Network Enabled Operations (Net Enabled Ops), and Command and Control (C2). While these related fields have strong links to cyber operations, none of the associated measurement efforts specifically address the particular measurement requirements of the cyber realm. We propose a metrics framework for cyber operations that is adapted from the measures development work of the US Department of Defence Director of Operational Test and Evaluation, which recommends conducting the assessment at the mission, task and system level. We pay particular attention to the mission and task levels, which describe what is being done, why it is being done, and how well it is being done. The framework elements are “Mission Objective”, “Desired Effects”, “Functions”, “Attributes”, and “Metrics”. This paper will describe how the framework measures the cyber effects described in *Simulation Approach for Military Cyber Operations* (also submitted to this conference). The major contribution of the paper will be the application of the attributes and metrics discussed in the related capabilities of C4ISR, Net Enabled Ops, and C2 to the measurement of cyber effects.

Keywords: measurement framework, metrics, attributes, cyber operations, cyber effects

1. Introduction

In the last decade, military forces around the world have been in a state of transformation, trying to keep abreast of the rapid rate of change in technology. Modern technologies, namely Information Technology Infrastructures (ITI), have provided the means for military decision makers to obtain information at ever-increasing speed and volume, which in turn has improved the richness and span of the information available. However this increased reliance on ITI systems has come at a cost, as it makes our military forces susceptible to cyber attacks and intelligence collection using cyber means. As the threat of cyber attacks increases and attacks become more sophisticated, the importance of cyber operations has been brought to the forefront for military force developers. Considering the use of cyber operations has led to a need for concept development and experimentation. In this paper we will discuss the requirement for measurements and metrics to assess the effects of adversarial cyber operations, particularly on command and control (C2). C2 is an important capability for military operations; adversary interruption of C2 would have a direct impact on mission effectiveness.

1.1 Command and control

According to the Canadian Department of National Defence terminology bank, command and control is defined as “The exercise of authority and direction by a commander over assigned, allocated and attached forces in the accomplishment of a mission”, which is aligned with the US definition in (US DoD 2004). A second definition for command and control proposed by Canadian defence scientists (Pigeau and McCann 2002), where the terms *command* and *control* are defined separately, is as follows: Command is “the creative expression of human will necessary to accomplish the mission”; and Control are “those structures and processes devised by command to enable it and to manage risk”. In general terms, C2 enables a commander to recognize what needs to be done in a given situation (i.e. make decisions) and to ensure that effective actions are taken. Therefore, this paper will only consider the activities of cyber operations that effect C2 functions.

According to the US Joint Command and Control Functional Concept (US DoD 2004), the basic C2 functions and process are depicted below in Figure 1(a) while Figure 1(b) illustrates the collaborative C2 process required to achieve shared understanding. The successful execution of these functions requires commanders to have the right information at the right time, in order to produce the desired effect. They also require subordinates to have a clear understanding of the overarching commander's intent. The move towards the collaborative C2 process implies a change from a *need-to-know* to a *need-to-share* information philosophy; this change is heavily reliant on a robust communications/computer network, which as previously mentioned, may be at risk of cyber attack.

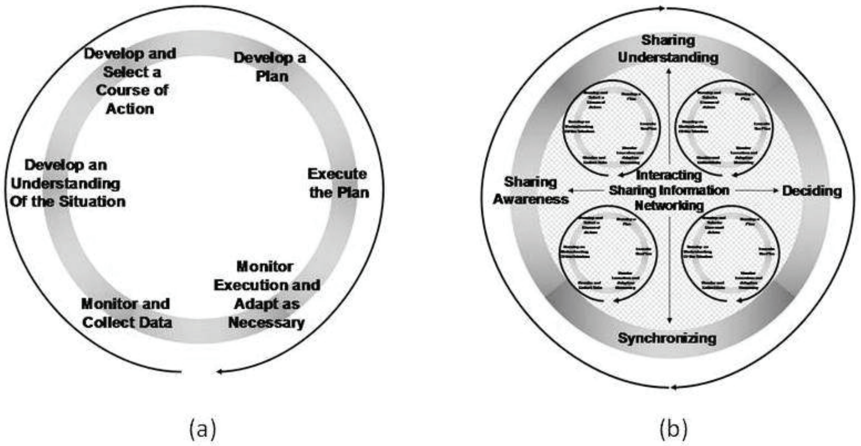

Develop and Select a Course of Action
Develop a Plan
Develop an Understanding Of the Situation
Execute the Plan
Monitor Execution and Adapt as Necessary
Monitor and Collect Data
(a)
Sharing Understanding
Sharing Awareness
Interacting Sharing Information Networking
Deciding
Synchronizing
(b)

Figure 1: (a) Basic C2 Process. (b) Collaborative C2 Process. (US DoD 2004) Note that the basic C2 process found in (a) is embedded in each of the four quadrants of (b).

The basic C2 capabilities are as follows:

- The ability to monitor and collect data on a situation
- The ability to develop an understanding of the situation
- The ability to develop and select a course of action
- The ability to develop a plan in order to execute course of action
- The ability to execute the plan including providing direction and leadership to subordinates
- The ability to monitor the execution of the plan and adapt as necessary

The collaborative C2 capabilities include

- The ability to network
- The ability to interact
- The ability to share information
- The ability to develop shared awareness
- The ability to develop shared understanding
- The ability to decide in a collaborative environment
- The ability to synchronize

1.2 Cyber effects on command and control

While military operations in the cyber environment are not new, advancements in ITI have blurred the lines between cyber operations and other traditional military activities. In this paper we consider cyber operations as a subset of activities that include elements of information operations, computer network operations, physical operations (i.e. land, air, maritime, space), psychological operations, electronic warfare, and signals Intelligence. Cyber operations can be either offensive or defensive which are defined as proposed in (Bernier and Treurniet 2010):

- Defensive cyber operations are actions taken in the cyber environment to protect one's own information and information flow and to maintain friendly decision makers' freedom of action in the cyber environment.

- Offensive cyber operations are actions taken in the cyber environment to deny the adversary's actual or potential use of, or access to, information or information systems and to affect their decision making process.

Therefore, we consider that cyber effects on C2 are any activities carried out through the cyber environment that affect the basic C2 functions or the collaborative C2 functions described in section 1.1. This paper will focus on adversarial offensive cyber operations which will provide insight on the effectiveness of friendly defensive cyber operations.

In their paper (Musman et al. 2010) propose a categorical description for cyber attack effects as shown in Table 1. In other related research by the authors, and also presented elsewhere at this conference (Morton et al. 2012), cyber attacks effects were related to a taxonomy of cyber attacks presented in (Chapman et al. 2011). By successfully applying the types of attacks from the taxonomy to these categories, the authors agree that the six proposed categories of effects are sufficient to gain an initial understanding of the effects of cyber operations on C2. Each effect can affect ITI resources themselves or the information that resides within them. Each effect can also have implications on the operational processes as well as on the operational data of the organization under attack (Musman et al. 2011).

Table 1: Cyber effect categories

Effect	Description	Implication
Interruption	An attacker causes an ITI asset to become unusable, unavailable, or lost for some specified period of time.	The ITI or information residing within it is unavailable for a specified period of time and the process will be unusable until the incident is recovered.
Modification	An attacker causes a modification of information, data, protocol, or software.	The information has been altered and as a result the processes that use this information may fail or produce incorrect results.
Degradation	An attacker causes degradation in the performance of an ITI asset.	The rate of information delivery is decreased resulting in the processes involved becoming slowed down.
Fabrication	An attacker causes information to be inserted into the system.	False information has been entered in the system and the process could include the insertion of false operational task that may interfere with legitimate operational tasks.
Interception	An attacker causes or takes advantage of information leaked from the system.	The information and/or the process, either via software or hardware, has been captured by the attacker.
Unauthorized Use	An attacker uses system resources for his own purpose.	This raises the potential for future effects on the information and the processes as well as unexpected outcomes on the processes.

2. Methodology

The goal of this section is to present a framework to identify possible metrics that can be employed to measure the impact of cyber operations on the effectiveness of the C2 of military missions. The framework is intended for use in experimentation to capture the data which will refine the emerging concepts of operation for cyber operations. In section 2.1, we will define the levels of metrics that we will be using and in section 2.2, we will describe the framework that will be used to develop the various metrics.

2.1 Metrics: Definition and characteristics

Measuring the impact of cyber effects on C2 and mission effectiveness requires a set of metrics that allows us to assess different attributes of C2 and their impact on mission outcomes. A large body of literature exist on the topics of metrics for C2 but the most notable are that of the NATO Code of Best Practice for C2 Assessment (NATO 2002) which presented a hierarchy for the metrics and the Joint Command and Control Functional Concept (US DoD 2004), which adapted the NATO hierarchy and proposed a list of potential metrics for C2. For experimentation purposes we will adopt the following levels of metrics:

- Measures of force effectiveness (MoFE), which measures the ability of a Commander to achieve the main goals in a given scenario, or the success of a mission.
- Measures of C2 effectiveness (MoCE), which measures the impact of C2 systems in the operational context.
- Measures of C2 Performance (MoCP), which measures the internal characteristics of C2 programs, initiatives, systems, etc.

The levels of metrics form a hierarchy where the lower-level metrics, MoCPs, feed into the Higher-level metrics, MoCE and MoFE, as depicted in Figure 2. The arrow in Figure 2 indicates that the value associated with each level of metrics becomes increasingly important to military commanders as we progress through the hierarchy.

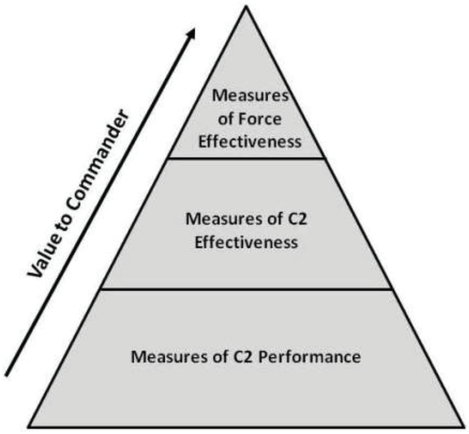

Value to Commander
Measures of Force Effectiveness
Measures of C2 Effectiveness
Measures of C2 Performance

Figure 2: Hierarchy of metrics

As described in (Garstka 2003) and (US DoD 2004) an attribute is a testable or measurable characteristic that describes an aspect of a system or capability. There are four categories of attributes proposed:

- *Objective* attributes: measure quality in reference to criteria that are independent of the situation.
- *Fitness-for-Use* attributes: measure quality in reference to criteria that are determined by the situation
- *Agility* attributes: measure the aspect of agility across the six dimensions of robustness, resilience, responsiveness, flexibility, innovation, and adaptation.
- *Concept Specific* attributes: measure unique aspect of some concept.

2.2 Metrics framework

Metrics need to have a purpose and must be linked to required information about the systems, processes or concepts being evaluated. One approach to help ensure the validity of metrics is the goal-question-metric (GQM) paradigm developed by Victor Basili (Basili 1992). This approach originated in software development, where it was developed for evaluating defects for a set of projects in the NASA Goddard Space Flight Center environment and involved a set of case study experiments. However, we believe that it is applicable to our purposes, as we are looking for a metrics framework for the purpose of experimentation. The GQM methodology is well described in (Perkins et al 2003) and includes the following steps:

- *Define goals*. Where goals should identify what needs to be accomplished relative to products, processes and resources; should be verifiable or measurable in some way; and should be defined in enough detail to be unambiguous.
- *Derive questions*. Where questions should only elicit information that indicate progress towards a specific goal; can be answered by providing specific information; and ask all the information needed to determine progress or completion of the goal.

- *Develop metrics*. Where metrics are the information needed to answer the derived question; each question may have multiple relevant metrics; and each metric requires two or more measurements for evaluation.

A second approach that was considered was (US DoD 2011), a standard operating procedure (SOP) for measures development as led by the Directorate, Operational Test and Evaluation. This SOP facilitates the development of mission and task measures to evaluate military capabilities used to achieve desired effects. The SOP provides a complete end-to-end process that decomposes missions and tasks into attributes and measures, and then links system attributes and measures to task performance and mission effectiveness. The measures framework they propose, as depicted in Figure 3, is a relationship diagram of capability key elements that identifies the basic questions of *who*, *what*, *why*, and *how*, and then connects measures to "*how capable* is the *who* and *how*?" and "*how well* is the *what* and *why?*".

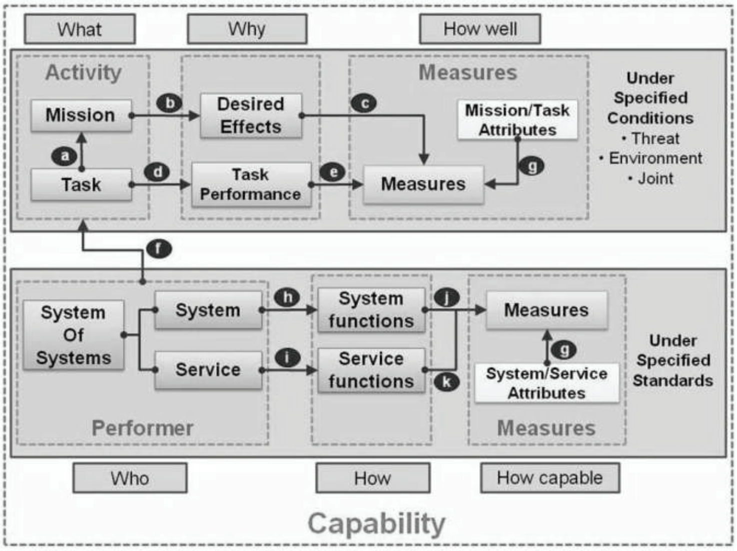


Figure 3: Measures framework relationship diagram, (US DoD 2011).

In this paper, we are more interested in the mission and task portion of the framework, as we are not generally concerned with the C2 systems themselves, but rather in the effects produced on the C2 processes and decision makers. For this reason we have adapted the US DoD framework to our needs and have related it the GQM paradigm and the hierarchy of metrics presented in section 2.1. The proposed metrics framework is depicted in Figure 4 where the mission objective and desired effect will define the goal, and the function and attributes will derive the questions which will in turn lead to the metrics. In the case of C2, the C2 functions and attributes will lead to a set of MoCP which will feed into the higher-level MoCE to assess the overall impact of C2 on the mission; finally, the ability of a commander to achieve the main goals will provide MoFE.

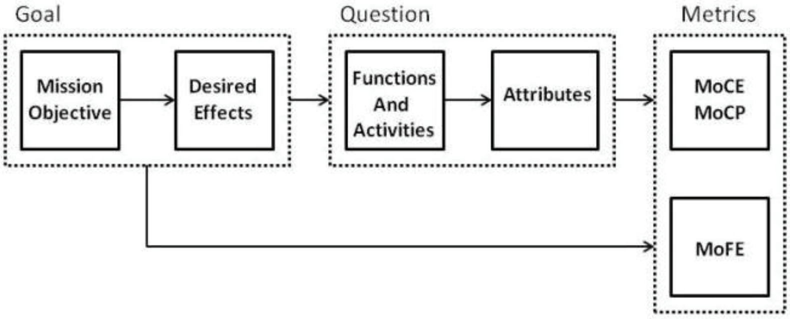


Figure 4: Proposed metrics framework

3. Metrics framework applied to cyber operations

This paper is not meant to provide a definitive list of metrics for cyber operations. Rather, this paper focuses on exploring possible metrics for quantifying the efficacy of a cyber operation once experimental data becomes available. In section 3.1, we will explore metrics that were previously developed for the field of C2 and other areas related to cyber operations. In section 3.2, we will provide an example of how to apply the framework in order to derive the required metrics for our purpose. Future work will involve testing the framework to derive metrics for assessing experimentation and exercises involving cyber operations concepts.

3.1 Metrics development in related fields

Although cyber operations have recently been receiving a great deal of attention, some of the activities now considered part of cyber operations have been around for some time. For example, information operations, influence operations, network enabled operations, even in Communications, Computers, Command and Control, Intelligence, Surveillance and Reconnaissance (C4ISR) processes all make use of the cyber environment. As activities in those fields are far from new, research in metrics development for these areas already exists. A survey of existing research was conducted. In this section we will assemble some of the research related to C2 which we think might be of use for cyber operations. Note that the intention is not to provide an exhaustive list of possible metrics but to provide examples which could assist in future metrics development work. For a complete list of metrics, the reader is referred to the respective sources.

3.1.1 Measuring the effects of network-centric warfare (Booz-Allen & Hamilton 1999)

The purpose of this paper is to identify and explore measures of effectiveness for the concept of network centric-warfare. In this paper, three categories of metrics are proposed: “reason metrics”, “physical metrics” or “belief metrics”. *Reason* metrics address the cognition domain that includes awareness, analysis and decision-making capabilities. *Physical* metrics measure the operational areas of move, strike and protect which occur within the dimensions of force, space and time. *Belief* metrics are more qualitative in nature than *Reason* and *Physical* metrics, and consider factors such as moral, experience, and the will to fight. This qualitative aspect makes *Belief* metrics the most difficult of the three categories to quantify. For our purpose of measuring cyber effect on C2, the most applicable is the *Reason* metrics as it considers the three operational area of “situational awareness”, “Command, Control, Communication and Computers (C4)”, and “information operations”. Table 2 provides a summary of the *Reason* metrics presented in (Booz-Allen & Hamilton 1999).

Table 2: Reason metrics (Booz-Allen & Hamilton 1999)

Attribute	Measure	Metric
Situational awareness effectiveness	Information integrity	Accuracy: The percentage of targets within a database that is current and classified correctly
		Completeness: The percentage of all targets who are current and classified correctly
		Consistency: The percentage of data in the database shared with other units over time
	Information precision	The average error in the objects location coordinates and the percentage of total objects identified/classified incorrectly
Situational awareness robustness	ISR coverage	The percentage of enemy area covered compared with the number of ISR platforms remaining
	ISR redundancy	The percentage of total area covered by each ISR type compared with the number of ISR platforms remaining
Situational awareness efficiency	Information timeliness	The length of time required to complete each phase of the information dissemination process
C4 effectiveness	Information accessibility	The time delay which friendly forces take to converge on an objective
	Information commonality and consistency	The percentage of units in the common situation map over time
	Lock-out	The number of enemy engagement opportunities compared with the course of actions available to a friendly commander
C4 robustness	Nodal redundancy	The number of nodes remaining functional after an adversary’s attack
	Link redundancy	The number of links remaining functional after an adversary’s attack
C4 efficiency	Information velocity	The operational response time compared with the network information delays

Attribute	Measure	Metric
	Network reliability	The percentage of time a network is operational and available to the warfighter
Information operations effectiveness	Synchronization of physical and mental effects	The percentage of an adversary's force affected as a function of time (duration of effect)

3.1.2 Network centric operations conceptual framework version 1 (Garstka 2003)

The objective of this document is to develop a set of metrics to assess the tenets of network centric operations (NCO). In order to achieve this, the authors developed a conceptual framework that identifies key concepts for NCO along with their relations and also identified attributes and metrics for each concept. A detailed discussion of the difference between NCO and cyber operations is out of the scope of this paper, but although NCO will use the cyber environment they are not synonymous with cyber operations. The NCO concepts include:

- Degree of networking
- Quality of organic information
- Degree of information "shareability"
- Degree of shared information
- Individual or shared awareness
- Individual or shared understanding
- Individual or collaborative decision making
- Quality of interactions

The majority of these concepts are measured using two sets of attributes, namely the *objective* attribute and the *fitness-of-use* attributes. Table 3 includes a sample of the attributes and metrics found in this document.

Table 3: Attributes and metrics from the network centric operations conceptual framework (Garstka 2003)

Attribute	Measure	Example of Metric
Objective attribute	Correctness	Extent to which information is consistent with ground truth
	Consistency	Extent to which shared information is consistent within and across Communities of Interest (CoI)
	Currency	Age of information
	Precision	Level of measurement detail of information item
	Extent	Proportion of information in common across force entities, within and across CoI
Fitness of use attribute	Completeness	Extent to which information relevant to ground truth is collected
	Accuracy	Appropriateness of precision of information for a particular use
	Relevance	Proportion of information collected that is related to task at hand
	Timeliness	Extent to which currency of information is suitable to its use
	Uncertainty	Confidence level (0%=uncertain, 100%=certain) of awareness
	Appropriateness	Extent to which decisions are consistent with existing shared understanding, command intent and shared team values
	Risk propensity	Extent of risk aversion in decision making

3.1.3 Joint command and control functional concept (US DoD 2004)

The purpose of this document is to provide the measurement framework for evaluating the C2 investment options needed to implement joint C2 in order to facilitate making C2 investment decisions. It also provides a basis for military experiments and exercises. While (Garstka 2003) discussed *objective* and *fitness-of-use* metrics, this document covers the *agility* and *joint C2 specific* attributes. *Agility* is said to be the overarching attribute of joint C2 as it permeates all aspects of the force. The characteristics of agility help shape the measures and metrics of the joint C2 concept specific attributes. Table 4 presents the *joint C2 specific* attributes and includes sample metrics found in the source document.

Table 4: Attributes and metrics from the joint C2 functional concept (US DoD 2004).

Attribute	Measure	Example of Metric
Superior decision-making	Appropriateness of the decision	Extent to which a decision is consistent with higher commander's intent, shared understanding and shared values
	Timeliness of decision	Extent to which currency of a decision is appropriate to the mission
	Currency	Time required to make the decision
	Extent	Percentage of C2 elements which accept the decision, participate in collaboration
	Flexibility	Number of feasible, suitable, and acceptable course of actions considered
	Innovativeness	Number of feasible courses of action developed
	Effectiveness	Extent to which commander's intent was achieved
Shared understanding	Extent	Proportion of C2 elements that share a given understanding
	Consistency of shared understanding	Proportion of key elements of shared understanding which are held in common
	Correctness	Percentage of key elements of shared understanding obtained that are consistent with ground truth
	Completeness	Percentage of key elements of shared understanding obtained
	Timeliness	Appropriateness of time required to achieve shared understanding in relation to mission need
Flexible synchronization	Adaptability	Time, effort and resources required to make changes
	Flexibility	Number and type of control mechanism available
	Synergy	Percentage of decisions that are conflicted, de-conflicted or synergistic
Simultaneous C2 processes	Currency	Time required to propagate critical information
	Synchronization	Percentage of sub-elements simultaneously involved in planning process, execution process
Dispersed command and control	Congruence with commander's intent	Percentage of actions which reflect the commander's intent
	Ability to execute the collaborative C2 functions across time and space	Percentage of dispersed C2 elements that were effective
Responsive and tailorable organization	Robustness	The ability to maintain effectiveness across the range of military operations
	Resilience	Time of effective performance without degradation
	Adaptability of the organizational structure	Number and type of C2 organizational structures available
	Responsiveness	Time required to change organizational structures
	Appropriateness	Match between organizational structure and task
Full spectrum integration	Understanding roles, goals, objectives and authority	Nature and the number of conflicts
		Percentage of compliance
	Accessibility of information	Number of times critical information is denied
	Extent of lexicon	Frequency of misunderstanding
	Congruence of command	Percentage of force elements in support of commander's intent
	Depth/breadth of interactions	Proportion of appropriate force elements who are able to participate across functions and echelons
	Coordination of C2 elements	Percentage of conflicted/de-conflicted synergistic C2 elements
Shared quality information	Relevant	Percentage of information relevant to task
	Accuracy	Confidence rating
	Usability across echelon	Interoperable (translatable or not translatable/ease of use)

Attribute	Measure	Example of Metric
	Timeliness	Extent to which currency of shared information is suitable in relation to mission need
	Completeness	Percentage of critical information shared
	Extent of sharing	Proportion of information in common across force elements, within and across CoI
	Consistency	Extent to which shared information is consistent within and across CoI
Robust networking	Quality of service	Resilience, modularity, reliability, secure, dependable, scalable
	Upgradable (forward compatibility)	Time and resources necessary to incorporate new technologies
	Maintainability	Time and resources necessary for routine use
	Reach (including mobility)	Proportion and distance of elements that are connected

3.2 Metrics framework for cyber operations

In order to demonstrate how the metrics framework proposed in section 2.2 can be applied to assess the impact of cyber effect on C2, we provide an example using a defensive cyber operation as depicted in Figure 5. The first step is to define the goal which includes mission objectives and desired effects. The mission objective is to protect one's own ITI against unauthorised activities and maintain freedom of action in the cyber environment for friendly decision makers. The desired effect, as an example, could be to reduce vulnerability to cyber attacks and minimise damage and recovery time from cyber attacks. The impact of cyber effects on C2 is related to the mission objective of maintaining freedom of action for decision makers. Measuring this impact (i.e. by assessing the effects on the C2 processes described in section 1.1) will provide information on how well the desired effects were achieved.

The next step is to derive a question that focuses on the relevant functions, that allows the selection of appropriate attributes and ultimately metrics for a given situation. Using the cyber effect of *degradation* as an example, where an attacker uses a denial of service attack against a server residing on the C2 network, the question would be: What C2 processes are affected by the degradation of the network? Going back to the C2 process defined in section 1.1, one can reason that all the collaborative C2 functions could be affected as well as the basic C2 function of providing direction and leadership to subordinates and the ability to monitor execution of the plan. We can then select the joint C2 specific attributes and metrics from Table 4, as appropriate, depending on the tasks performed. The exact attributes affected and associated metrics are highly dependent on the experimental scenario and will be investigated in future work.

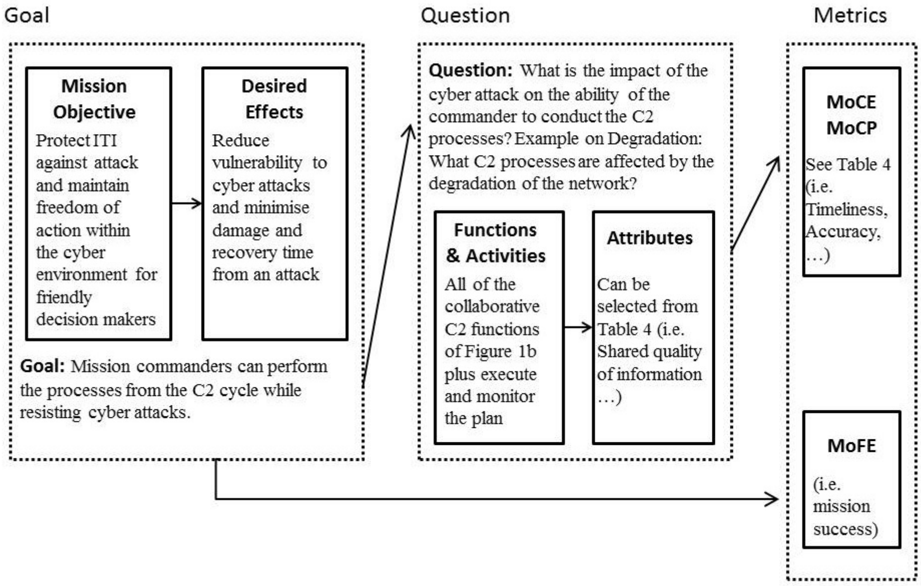

Goal
Question
Metrics
Mission Objective
Protect ITI against attack and maintain freedom of action within the cyber environment for friendly decision makers
Desired Effects
Reduce vulnerability to cyber attacks and minimise damage and recovery time from an attack
Goal: Mission commanders can perform the processes from the C2 cycle while resisting cyber attacks.
Question: What is the impact of the cyber attack on the ability of the commander to conduct the C2 processes? Example on Degradation: What C2 processes are affected by the degradation of the network?
Functions & Activities
All of the collaborative C2 functions of Figure 1b plus execute and monitor the plan
Attributes
Can be selected from Table 4 (i.e. Shared quality of information …)
MoCE
MoCP
See Table 4 (i.e. Timeliness, Accuracy, …)
MoFE
(i.e. mission success)

Figure 5: Application of the proposed framework to a defensive cyber operation

4. Conclusion and future work

Military forces of today are increasingly dependent on information technology and, as result, are more prone to cyber attack. This highlights the requirement for military forces to investigate the possibilities of integrating cyber capabilities into their operations. Concept development and experimentation provides the means to investigate these possibilities. An important aspect of experimentation is the assessment process. In this paper, we have proposed a framework for defining metrics that is adapted from the measures development work by the US DoD Directorate of Operational Test and Evaluation (US DoD 2011). The framework elements includes: "Mission Objective", "Desired Effects", "Functions", "Attributes", and "Metrics". We also described how the framework can be used to define measures to assess the impact of cyber effects on C2 and brought into play existing joint C2 attributes and metrics.

Future work in this area could involve implementing the framework to derive metrics for assessing experimentation and exercises involving cyber operations concepts. This effort would also be closely related to the work presented at this conference on "Simulation Approach for Military Cyber Operations" (Morton et al. 2012), which describes the effects of cyber attacks and argues for the inclusion of these effects in constructive simulations in order to educate senior military leaders. The integration of the two research avenues will provide further value in better defining cyber effects for C2 capabilities.

References

Basili, Victor. (1992) "Software Modeling and Measurement: The Goal Question Metric Paradigm", Computer Science Technical Report Series, CS-TR-2956 (UMIACS-TR-92-96), University of Maryland, College Park, MD, September.

Bernier, Melanie and Treurniet, Joanne. (2010) "Understanding Cyber Operations in a Canadian Strategic Context: More than C4ISR, More than CNO", In Czossek, C. & Podins, K. (eds.), Conference on Cyber Conflict Proceedings 2010, CCDCOE Publications, Tallinn, Estonia, pp 227-243.

Booz-Allen & Hamilton. (1999) "Measuring the Effects of Network-Centric Warfare", Company Report [online], http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA401399 (accessed 20/01/2012).

Chapman, Ian, Leblanc, Sylvain, and Partington, Andrew. (2011) "Taxonomy of Cyber Attacks and Simulation of their Effects", In Striker, A. (ed), Proceeding of the Military Modeling and Simulation Symposium 2011 (MMS 2011), Curran Associates Inc., Boston, Massachusetts, pp 73-80.

Garstka, John. (2003) "Network Centric Operations Conceptual Framework Version 1.0", Evidence Based Research, Inc.

Morton, Ben, Leblanc, Sylvain and Bernier, Melanie. (2012) "Simulation Approach for Military Cyber Operations", Proceedings for the 11th European Conference on Information Warfare (ECIW 2012), Laval, France.

Musman, Scott, Temin, Aaron, Tanner, Mike, Fox, Dick and Pridemore, Brian. (2010) "Evaluating the Impact of Cyber Attacks on Missions", In Armistead, E. and Cowan, E. (eds.), Proceedings of the 5th International Conference on Information Warfare and Security, Dayton, Ohio, pp 446-456.

Musman, Scott, Tanner, Mike, Temin, Aaron, Elsaesser, Evan and Loren, Lewis. (2011) "Computing the Impact of Cyber Attacks on Complex Missions", In Proceedings of the IEEE International Systems Conference (SysCon), Montreal, Quebec, pp 46-51.

NATO. (2002) "NATO Code of Best Practice for C2 Assessment", revised 2002, DoD CCRP Press, Washington, DC.

Perkins, Tim, Peterson, Ronald, and Smith, Larry. (2003) "Back to the Basics: Measurement and Metrics" CrossTalk: The journal of Defense Software Engineering [online], http://www.crosstalkonline.org/storage/issue-archives/2003/200312/200312-Perkins.pdf (accessed 20/01/2012).

Pigeau, Ross and McCann, Carol. (2002) "Re-Conceptualizing Command and Control", Canadian Military Journal, Spring, Vol 3, No. 1, pp 53-64.

US DoD. (2004) "Joint Command and Control Functional Concept", US Department of Defense [online], http://www.dtic.mil/futurejointwarfare/concepts/jroc_c2_jfc.doc (accessed on 16/01/2012).

US DoD. (2011) "Measures Development Standard Operating Procedure (SOP) Version 2" US Department of Defence, Joint Test and Evaluation Methodology – Transition [online], https://acc.dau.mil/adl/en-US/403465/file/56099/MeasuresDevelopmentSOPv2_2011-01-15.pdf (accessed on 16/01/2012).

Attribution in the Future Internet: The Second Summer of the Sisterhood

Matt Bishop[1], Mina Doroud[1], Carrie Gates[2] and Jeffrey Hunker[3]
[1]Dept. of Computer Science, University of California at Davis, USA
[2]CA Labs, New York, USA
[3]Jeffrey Hunker Associates, Pittsburgh, USA
bishop@cs.ucdavis.edu
sdoroud@ucdavis.edu
carrie.gates@ca.com
hunker@jeffreyhunker.com

Abstract: Attribution is the binding of data to an entity. An attribution framework is an infrastructure for managing attributes and their values. It consists of four components: a set of entities (actors) having an interest in attribution with respect to a transaction; a set of data to be attributed; the level of assurance with which values of attributes can be determined, and with which they can be associated with an entity; and a policy negotiation engine that actors use to negotiate an acceptable set of attributes and levels of assurance for their values in order to conduct a transaction (the "policy"). The actors include the sender and recipient, the sender's and recipient's organizations, ISPs, backbones, and political entities. This paper assumes that such a general attribution framework has been implemented. It examines the implications of such a framework upon the Internet, and upon transactions (specifically, the sending and receiving of packets) among actors. The embedding of attribution requirements in policies controlling communications between parties raises the question of who can communicate with whom. Specifically, how does the use and enforcement of policies based upon attributes affect users of the Internet? We examine this question in two contexts: that of the societal revolution known as "Arab Spring", and that of elections in the United States. We present requirements and the attributes that must be supplied to meet those requirements. We then examine some of the implications of supplying the attributes from the point of view of servers, clients, and intermediaries (such a ISPs and governments). We conclude with a discussion of when attribution is desirable, and when the inability to attribute actions is desirable.

Keywords: attribution, attribution framework, policy, enforcement, security

1. Introduction

Attribution is the binding of data to an entity. It usually arises in the context of identity. For example, much discussion has focused on attribution as a tool for identifying attackers (Burch and Cheswick, 2000; Pyun and Reeves, 2007). Other uses of attribution abound. In a distributed system, the location of nodes controlling a particular resource is an attribute of the resource; in an ordinary computer system, the role (job function) of each user is an attribute of that user. Indeed, the login name is an attribute of the entity that name represents.

The management of attributes depends on the nature of the attributes and the scope of their effect. On a Linux system, the attribute of "login name" with value "bishop" may be associated with the first author. On a file server, that same attribute-value pair may be associated with the user Michael Bishop. The scope of each pair is limited to the system on which the pair is defined, so the inconsistency is irrelevant; there is no conflict. But if the Linux system is trusted to authenticate users for the file server, then the inconsistency of attribute values creates a conflict, and the first author will have access to Michael Bishop's files.

The management of attributes requires an *attribution infrastructure* with four components. A set of *actors* specifies those entities with an interest in attribution for a particular transaction or set of transactions. Each actor has an associated policy describing which attributes it requires, and which attributes it will provide, to interact with other actors. An *attribution vector* lists the attributes for which values are desired, or a set of (attribute, value) pairs. Associated with each attribute value is a *level of (attribute) assurance* that describes how certain the reported value of the associated attribute is. Finally, the actors use a *policy negotiation system* to negotiate an acceptable set of attributes, values, and levels of assurance, or to conclude that what their policy allows is not acceptable to one or more parties. The details of this model are discussed in Bishop, Gates, and Hunker (2009).

As an example, consider the Linux system mentioned above. The attribution framework is implemented using a cryptographic hash function to bind the value "bishop" to the (external) entity using the computer. This binding occurs at login time, because the password that the entity types is hashed, and the hash compared to that associated with the login name “bishop”. If correct, the system performs the binding, and a kernel table maintains the binding. (More precisely, the binding is done using an integer that is associated with the login name; we omit the details for simplicity.) Here the actors are the external entity and the system. The attribute is “login name”. The attribute value after assignment is “bishop”. The level of assurance is that associated with the correct entity using the password. The policy associated with the entity is that the entity will supply the attribute-value pair “login name”-“bishop”, and the policy associated with the system is that it requires the provision of the value associated with the entity’s attribute “login name”, plus assurance that the value supplied is indeed associated with the entity. The entity supplying the password associated with the login name provides an acceptable level of assurance for this. (In some cases, the system may require two-factor authentication if the privilege level of the user is significantly high The policy negotiation system is constant; the system requires the correct attribute-value pair, or access is denied.

To illustrate the framework further, add to the above example the “trusted hosts” mechanism enabled by *rlogin* or *ssh* (Barrett ad Silverman, 2001). Recall that these two mechanisms have the system check the remote host from which the user is trying to log in. In this case, there are three actors: the external entity (Matt Bishop), the remote host (call it “nobhill”), and the Linux system being accessed. The attributes of the Linux system and of Matt Bishop are the same, but the system “nobhill” has associated with it the attribute “trusted by the Linux system” and the value “true”. The Linux system requires two attribute-value pairs, the first being “login name”-“bishop” and the second being “trusted by the Linux system”-“true”. The policy negotiation system now requires the second attribute-value pair be provided with some level of assurance (for *rlogin*, the assurance is obtained by looking the IP address up in a table of trusted hosts; for *ssh*, the assurance is based on cryptography). If the level is sufficient, the value of the attribute “login name” has sufficient assurance that the system accepts it without further checking. Again, the policy negotiation system is constant.

This paper assumes that a general attribution framework as described above has been implemented for the Internet. It, and the supporting infrastructure is built using existing Internet infrastructure, augmented as needed for the attribution framework and the integrity of both the framework and the information (attributes and values) that it contains. The attribution framework uses a distributed database similar to the Domain Name System. We also assume that the framework is correct; that is, when asked for an attribute-value pair for a particular entity, it supplies the information as required by the policy of that entity, and the information is conveyed correctly to the requester.

This paper examines the implications of this framework upon the use of the Internet, and upon transactions (specifically, the sending and receiving of packets) among actors. The embedding of attribution requirements in policies controlling communications between parties raises the question of who can communicate with whom. We are interested in situations where the attribution framework supports detection of violations of security policies. These may constitute attacks; they may result from mistakes; they may simply be due to carelessness or non-malicious factors.

We briefly review the framework structure. Next, we present two cases in which attribution is critical to the successful resolution of a specific problem. We then present some general thoughts on the role of attribution and how it affects both organizations and individuals.

2. Background on the framework

The attribution framework provides a basis for all the entities involved in a transaction to determine how to act (or to decline to act). The framework distinguishes between 4 classes of actors:

- An entity;
- The organization associated with that entity;
- The system(s) associated with that entity; and
- The government(s) associated with that entity.

Note this includes intermediate entities. So, if a message is sent from an entity in the United States to an entity in France via a hub in the Netherlands, the actors are the sender, all network providers in the U.S., the Netherlands, and France that the message transits, and the receiver; all the systems involved; all the organizations that run those systems and networks; and the governments of the U.S., the Netherlands, and France, including any political subdivisions (such as the state in the U.S. and the region in France). The governments apply their policies through laws; other entities apply them more directly, for example by policy-based routing or access control (as in the Linux example above).

Finally, the framework must support many types of attribution. The type sometimes is tied to the level of assurance. Consider entities interacting, each with a policy. They may require:

- *Perfect attribution*, in which the entities know each others' relevant attribute-value;
- *Perfect selective attribution*, in which one entity wants certain attribute-value pairs known to some entities but not to others;
- *Perfect non-attribution*, in which the entities do not want any other entity to know the attribute-value pairs;
- *Entity non-attribution*, in which an entity wants attribute-value pairs known but not that they are bound to the entity;
- *False attribution*, in which an entity will determine an attribute-value pair consistently—but the value will be wrong;
- *Randomized false attribution*, which is false attribution but the values determined are inconsistent;
- *Imprecise attribution*, in which the value of an attribute can eventually be determined to the level of assurance needed, but doing so takes so long the attribution is useless or determining the attribution costs more than the value of knowing the attribution; and
- *Unconcern*, in which the entities do not care about attribution.

In all cases, the level of assurance of the attribute-value pair must meet, or fail to meet, the entity's requirements.

3. Use cases

Each case begins with the goal—what is the problem? This drives the requirements, which in turn define the attributes and associated levels of assurance necessary to meet those requirements. Multiple actors are involved—-those who are dealing with the problem, and the intermediaries who pass the packets along, possibly augmenting them or taking other actions as needed. These may have differing, possibly mutually contradictory, requirements.

We do not specify implementation details. The programs and protocols used are incidental to this analysis. We note that they may require more attributes of the parties in order to succeed, but those attributes are products of the protocols and implementations and not of the requirements of a solution to the problem—and this section focuses on those solutions. Nor do we worry about the policy negotiation protocols; we simply note when one must exist, and any relevant properties.

3.1 Arab spring

The Internet enables communication across geographic and political boundaries. This provides a natural way for people in one part of the world to communicate with others. Indeed, the ubiquity of Internet service providers such as Google, and of social networking services such as Facebook, Twitter, and YouTube, mean that videos and text can "go viral" throughout the world.

Social activists have learned to exploit this availability very effectively. The "Arab Spring", a term for the uprisings in the Middle East, is a good example of this. A March, 2011 survey (Salem and Mourtada, 2011) showed that almost 9 out of 10 Egyptians and Tunisians used Facebook and other social networking sites to organize and spread awareness of the protests in those countries. As one activist put it (Kiss and Rosa-Garcia, 2011), they used Facebook to schedule the protest, Twitter to coordinate it, and YouTube to tell the world. The role that the social networks played was critical not only within the countries, but also in communicating events with the rest of the world.

During the Egyptian Arab Spring uprisings, the government attempted to block the peoples' access to the Internet. In response, Google and Twitter provided a "telephone-to-tweet" service. Google established 3 telephone numbers that people could call and record a message. Google then posted it to Twitter. Associated with each message was a hash tag indicating the geographic origin of the message, when that could be determined; otherwise, the message was posted without a hash tag.

Because the governments of these countries controlled the infrastructure, and attempted to crack down on the protests, the intermediaries are critical. The end actors (individuals and social media sites) have one set of requirements. The managers of the intermediate hosts and networks have another.

The requirements for the end actors were:

Anonymity: the tweets, and the recordings, must be anonymous.

Accuracy of origin: A hash tag indicating geographic origin, if present, must be accurate.

The requirements for the intermediate actors (governments or networks under the jurisdiction of the government) were:

Identity: the identity of the individual who sent the message.

Consider the transactions involved in the use of these social networks. First, the user must register with the social media services. Facebook requires that the user supply a real name and date of birth. The level of assurance is minimal; essentially, mere assertion provides sufficient assurance unless Facebook questions it. Facebook also checks users' IP addresses. If the user is using a very different IP address, Facebook will request additional assurance evidence (such as correctly identifying a friend's photo). Twitter and YouTube, on the other hand, simply require a name that the user will post under; unlike Facebook, the user need not identify herself at all, and indeed most users of those services use pseudonyms. So, the initial transaction requires the user to provide the attribute "name" and "date of birth", but the values of both typically have a low level of assurance. Similarly, the transaction requires that the server provide its identity to a level of assurance high enough to convince the user that it is the genuine social media site.

Associated with each message are the "from account" attribute indicating the account posting the message, and the "post to" attribute indicating where the message is to be posted. The user supplies these values; the server requires the second and may require the first. The assurance level for these attributes are both typically very low, because errors can be corrected, or the errors are deemed "harmless" (of course, this may not be how the poster feels).

The "telephone-to-tweet" service had different requirements. The only attribute the server requested was the country from which the message originated. In most cases, the telephone system would supply this information to an acceptable level of assurance. If the information were not available, no associated hash tag would be generated. In other words, the server would request the "country of origin" attribute, but if the client could not supply it, the policy negotiation mechanism rolled back to accepting the (unattributed) message.

The "telephone-to-tweet" service is an example of perfect non-attribution: none of Google, Twitter, nor any listener is to be able to identify the speaker from any metadata.

Now consider intermediaries. The intermediary ISP is an entity over which the government exercises *de facto* control. One set of requirements involve the ISP interacting with the server (actually, other entities that communicate with the server; in this case, those other entities have no requirements), and the other set involves the ISP interacting with the user client. The former having no special requirements, we examine the latter.

The goal of the ISP is to prevent communication with a set of servers that the government deems undesirable. The origin of the message is not relevant; where it is going, is. Thus, the requirement for this intermediate actor is:

Identity: the (Internet address of the) destination of the message

The attribute therefore is "destination address". The level of assurance required by the ISP is high, to ensure proper delivery of the message; as a side effect, the ISP can block the desired IP addresses.

The need for this level of attribution is instructive. One may circumvent this control by using a proxy. In this case, the messages transit the intermediate network with the destination being the proxy, so the intermediate network allows the messages through. The use of VPNs, encrypted communications channels, and mix routers such as Tor (Dingledine, 2011) encipher the information about the sender's identity and the ultimate destination. Thus, the intermediary cannot accurately attribute the origin of the message to an individual, or the destination identity to a site. Thus, it must block all such messages, allow all such messages through, or require use of another mechanism to identify the sender and destination.

3.2 Elections

Elections are the foundation of democratic and republican societies. Recently, many jurisdictions began exploring people voting over the Internet. We look at the use of the attribution framework in the context of elections within the United States. We focus on a key subset of the requirements:

- *Secrecy of the ballot*: a third party cannot associate a ballot with a voter.
- *Anonymity of the ballot*: the voter cannot prove that she cast a particular ballot to a third party.
- *Accuracy of the count*: the votes are counted correctly.
- *Authorization of the voter*: the voter submitting the ballot is authorized to do so.
- *Integrity of transmission*: the ballot is received and counted as cast.

A number of other requirements exist, but our exposition focuses on these.

United States residents live under at least 3 political jurisdictions (federal, state, and county) and in most cases many more (city, school board, and so forth). Each of the 50 states is responsible for holding its own elections; most delegate this responsibility to the counties, with the state having the ultimate authority to certify the results. For purposes of elections, the counties are divided into precincts, each of which matches one set of overlapping political jurisdictions. Each precinct conforms to a single set of candidates, so all voters in a precinct will vote using the same ballot. But two precincts may have different sets candidates, and thus two different *ballot types*. In Yolo County, California, for example, one election required over 100 different ballot types. In jurisdictions where more than 5% of the voters have a language other than English as their primary tongue, the ballots must also be printed in that language. For example, one election in San Francisco, California, required some ballot types to be printed in 7 languages. The location where the county counts the results is called *Election Central*.

This multiplicity of ballot types and languages means that three transactions are involved in elections. First, the voter registers to vote. Second, the voter receives the correct ballot from the ballot generator (most jurisdictions have strict formatting requirements, and require ballots to come from the county). Third, the voter marks his ballot and then transmits the marked ballot to Election Central, where the votes are counted.

Consider each transaction separately. The first transaction speaks to authorization of the voter. By the voter supplying attributes that uniquely identify her, the registration authorities can determine which ballot type she should receive, and set the "authorization to vote" attribute to a value that will enable her to obtain the correct ballot. (In practice, this is usually an address, because political jurisdictions in the United States are geographic.) The voter requires that the registration authority have the attribute "authorized to register voters" with value "true". The registration authority specifies the unique attributes it requires and the assurance evidence that the value of that attribute (in practice, the address)is correct (in practice, evidence that the prospective voter lives there). It in turn must possess the attribute "authorized to register voters" with value "true". The intermediate actors must allow this information to transit their systems and networks. Thus, the attributes that they require also must be present.

The second transaction occurs when the voter acquires the ballot. The voter first verifies that the ballot generation system is authorized to generate the ballots by checking the value of the attribute "authorized to generate ballots". The voter presents the value of the "authorization to vote" attribute, which the ballot server validates as having an acceptable level of assurance. That authorization is used to determine the ballot type that the voter requires. A ballot of that type, with the attribute "issued to authorized voter" and value a nonce is generated and sent to the user. Note that no attribute ties the ballot to the actual voter, which meets the requirement of secrecy of the ballot. But an obvious attack is for the voter to mark the ballot and copy it. To prevent this, attributes “access control” with value “originator-control”, “originator” with value “election-official”, “write access” with value “voter-issued-to”, and “read access” with value “election central, voter issued to” indicate that the ballot is not to be copied or marked by any entity other than the original voter.

The third transaction is the casting of the vote. The important artifact here is the ballot; it must come from the ballot server, and be voted by an authorized voter. The voter first contacts Election Central, and checks that the attribute “Election Central” is “true”. The voter then transmits her ballot to Election Central, which checks that the attribute “issued to authorized voter” has a nonce that is unused so far. It then processes the ballot and tallies the votes.

Thus, the attributes of interest here are:

- *Voter*. provides “identity”, “authorization to vote”; requires “authorized to register voters” from registration authority, “authorized to generate ballots” from the ballot generator, “Election Central” from Election Central
- *Registration authority*: provides “authorized to register voters”, “obtain ballot”; requires “identity”, “authorization to vote” from voter
- *Ballot generator*: provides “authorization to generate ballots” to voter; provides “issued to authorized voter”, “access control”, “originator”, “read access”, “write access” to ballot; requires “authorization to vote”
- *Election Central*: provides “Election Central”; requires “issued to authorized voter” from ballot
- *Ballot*: “issued to authorized voter”, “access control”, “originator”, “read access”, “write access”

The intermediate services and actors must simply pass messages on unchanged and without delay, regardless of the values of any attributes. (The lack of delay is required to ensure ballots arrive in time to be counted.)

4. Discussion

We have discussed in detail two different case studies where attribution, or non-attribution, would be desirable, and the effects of such attribution. What needs to be considered is the desirability of attribution by each of the different actors: end users, organizations, intermediary nodes and governments.

The governments have strong desires for attribution in the case studies presented. In the case of elections, the government as an entity desires a fair and lawful election. We note that this particular example assumes benevolence on the part of the government. In those cases where it does not want a fair election, that desire results from specific individuals within the government, not from “government” as a separate entity. Thus the government, in the case of elections, has a desire for proper attribution that matches the laws of that particular government (e.g., a citizen can only vote once; it must not be possible to match a vote to a voter, and so forth). Perhaps more importantly, the government desires that the electorate perceives the election to be fair, and the use of appropriate attribution technology can assist in creating such a perception. In the case of Arab Spring, the government also desired attribution, specifically of the individuals responsible for organizing the revolutionary activities. This is similar to the original push for attribution on the Internet, which was done by the government with regard to the ultimate identification of the actor(s) responsible for launching distributed denial-of-service attacks. In this case, the government is not interested in attribution to identify itself to others, but rather attribution to identify other actors. In general, the government as an entity is interested in the attribution of individuals performing particular actions (either against the government, or on behalf of the government). However, in some circumstances, a government might only need to know what (other) government to attribute an action to. For example, a government of a country under cyberattack might only care that the attack is sanctioned by a particular nation state, and not necessarily about knowing the individuals who performed the attack.

In the case of intermediary nodes, such as ISPs, attribution is likely not desired. In the same vein as ISPs not wanting to have the responsibility for recording traffic traversing their networks and maintaining that information for some amount of time as determined by government regulation, ISPs will also not likely wish to have the responsibility for providing attribution information. We note that such desire might change in the face of increased regulatory pressure or if sufficient profit is available; however, this seems unlikely given the current environment.

Organizations may wish to provide attribution to end users for whom they are providing services. For example, the use of an attribution framework will likely considerably reduce instances of successful phishing attacks. Thus, organizations such as financial institutions will likely be interested in providing such services. This could shift liability from financial institutions back to end users. Thus attribution might provide a benefit to those organizations. Similarly, organizations might desire attribution during negotiations with other organizations (such as when one organization desires to acquire another). However, we note that in this case the attribution is actually at the end user level, because the organizations will want to know the individual who is performing the actions in order to determine that this person has the appropriate authorization to perform the associated actions on behalf of their organization. On the opposite end of the spectrum, due to liability issues, organizations might not want to know the identity of the end user, and thus not desire that level of attribution. For example, during financial transactions, an organization (such as a store) might only want to know that the credit card information is valid, and not the identity of the user providing the credit card information. This ensures that liability remains with the end user and does not shift to the organization.

The case of the end user is much more complex. In an election, the end user may wish a vote attributed to him (to sell his vote, or to prove that he voted for a particular candidate or party). But it is against the government's interest (and may, in fact, be illegal) for such attribution to be possible. In the case of Arab Spring, the end users explicitly wanted to remain anonymous for fear of reprisals against themselves or their families. Thus the desires of an end user might be complex, and go against their organization's or government's desires. Often such conflicts are philosophical in nature. There isno "right" answer. In general, especially outside the political realm, an end user might want to provide attribution information to receive credit for some action (such as writing a book). End users may also desire anonymity in order to preserve privacy. This often elicits a claim that "if you have nothing to hide, you do not need privacy. In fact, many users do not want even benign activity tracked by any third parties (e.g., web browsing habits or search terms). Privacy is an extremely complex issue, and the reader is directed to the numerous papers written on the subject (for example, see Warren and Brandeis (1890) and Dwork (2008)).

5. Conclusion

The widespread deployment of an attribution framework, or set of frameworks, provides both benefit and risk. Its use simplifies some actions and relationships; it also makes more difficult the protection of privacy under some circumstances. The point of this paper is to argue that the implications of such a framework, even if perfectly implemented and supported, are unclear, and need to be considered as such frameworks develop and are deployed. Especially critical are the types of attribution (or non-attribution) the frameworks will support, and the level of assurance associated with the values of those attributes and the binding of the attributes to the entity. We need to consider the broad implications for social policy, and indeed for societies themselves.

References

Bishop, M., Gates, C. and Hunker, J. (2009) "The Sisterhood of the Traveling Packets", In *Proceedings of the New Security Paradigms Workshop,* PP 59–70.

Burch, H. and Cheswick, B. (2000) "Tracing Anonymous Packets to Their Approximate Source", In *LISA '00: Proceedings of the 14th USENIX Conference on System Administration*, Berkeley, CA, USA. USENIX Association, PP 319–328.

Dingledine, R. (2011) "Tor and Circumvention: Lessons Learned", *Advances in Cryptology—CRYPTO 2011,* Lecture Notes in Computer Science, 2011, Vol 6841/2011, Publisher: Springer Berlin / Heidelberg, PP 485-486.

Dwork, C. (2008) "Differential Privacy: A Survey of Results", In *Proceedings of the 5th International Conference on Theory and Applications of Models of Computation*, Lecture Notes in Computer Science, 2008, Vol 4978/2008, Publisher: Springer Berlin/Heidelberg PP 1–19.

Kiss, H., Rosa-Garcia, A. (2011). "Why do Facebook and Twitter facilitate revolutions more than TV and radio?" MPRA Paper No 33496.

Pyun, Y. and Reeves, D. (2007) "Strategic Deployment of Network Monitors for Attack Attribution" In *BROADNETS '07: Proceedings of the Fourth International Conference on Broadband Communications, Networks and Systems*, PP 525-534.

Salem, F. and Mourtada, R. (2011) "Arab Social Media Report: Civil Movements: The Impact of Facebook and Twitter", *Dubai School of Government*, Vol. 1, No. 2, May.

Warren, S. and Brandeis, L. (1890) "The Right to Privacy", *Harvard Law Review*, Vol. 4, No. 5 PP 193–220.

CloudComputing and Security

Abílio Cardoso[1] and Paulo Simões[2]
[1]Portucalense University, Portugal
[2]CISUC-DEI, University of Coimbra, Portugal
abilioc@upt.pt
psimoes@dei.uc.pt

Abstract: There is always a strong pressure on Information Technology (IT) to do more with fewer resources. Over the decades, this pressure to rationalize IT costs spurred a number of paradigms, technologies and buzzwords. Some of them failed to meet their promises, while others became successfully embed in IT practices and infrastructures, providing sizeable benefits. The paradigm of cloud computing is currently riding this wave, promising to be the next great revolution in IT. Cloud computing appears to have the right technological and market ingredients to become widely successful. However, there are some key areas where cloud computing is still underperforming – such as security. Availability, security, privacy and integrity of information are some of the biggest concerns in the process of designing, implementing and running IT services based on cloud computing, due to technological and legal matters. There is already an extensive set of recommendations for IT management and IT governance in general – such as the popular Information Technology Infrastructure Library (ITIL) guidelines and Control Objectives for Information and related Technology (COBIT) recommendations. However, the field of cloud computing remains poorly covered. ITIL and other general sources can be sometimes translated to the context of cloud computing, but there are many new challenges not addressed by those generic resources. Recognizing this state of affairs, a number of initiatives already started focusing on novel proposals specifically targeting cloud computing but, up to now, with no significant outcomes. In this paper, we discuss the security implications involved in the migration of IT services to the cloud-computing model, proposing a set of rules and guidelines to be followed in the process of migrating IT services to the cloud. This set of rules and guidelines largely builds on general ITIL recommendations, discussing how to extend/adapt them to the field of cloud computing and identifying which a number of novel areas not covered by current ITIL recommendations.

Keywords: cloud computing, security, ITIL

1. Introduction

The term “cloud computing” was coined in the fourth quarter of 2007, in the context of a joint project between IBM and Google [Vouk, 2008, Zhang et al., 2010b]. However, as also happens with many other emerging trends – and despite being a subject on which much has been written – there is no consensual definition of what cloud computing really means. As an example, in [Vaqueroet al., 2009] there is a list of at least 22 distinct definitions of cloud computing proposed during 2008.

One definition recognized by several authors [Grobauer et al., 2010, Khajeh-Hosseini et al., 2010, Shimba, 2010, Foster et al., 2008, Zhang et al., 2010a] is presented by the National Institute of Standards and Technology (NIST). NIST, adopting a broad scope, defines cloud computing as “a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.” [Mell and Grance, 2011].

Cloud computing is classified in four deployment models: public, private, hybrid and community. In the public model, the cloud resources such as applications, storage and computing, are owned by an organization external to the customer. Customers consume these resources over the Internet in general, either in a free or on a pay-by-use model. Furthermore, the customers share components and resources with others that are unknown to them in a multi-tenant environment. In addition, the customer data might be distributed by various regions around the globe.

In this model, the emphasis is on offering services targeting a wider client base while on the private cloud the focus is having more attention on customization and personalisation of cloud functionalities. Additionally, the service agreements are typically non-negotiable being the service terms entirely established by the cloud-computing service provider, despite there may be some negotiated service agreements. Thus and being the public model an overarching scenario in the cloud computing deployment models this paper the focus is on this model.

Traditionally, each of the aforementioned deployment models is divided into three layers (also known as service models), according to the services it provides to the users. These three layers include, on the first level, Infrastructure-as-a-Service (IaaS), where the user can afford, upon request, processor resources, storage and networking, among others. At this level, the user is required to have specialized technical knowledge and the provider delivers computing power/resources. On a second level, the Platform-as-a-Service (PaaS) layer allows users to implement their applications in the cloud, by using the programming languages and tools provided by the service provider. The third corresponds to Software-as-a-Service (SaaS), where the applications provided by the service provider run in the cloud infrastructure and are typically accessed using a Web browser.

With the large amount of resources it provides for developing and deploying applications and services, the cloud-computing paradigm is an attractive tool to upgrade, extend or replace many of the services hosted by the traditional data centre. Nonetheless, in order to completely fulfil its promises, cloud computing still needs to win the trust of involved stakeholders. A recent survey, which included more than 500 executives and IT managers from 17 countries, revealed that despite the benefits of cloud computing, there is more confidence in internal systems, due to safety threats and loss of control over valuable information. Another survey, from IDC [Gens, 2008], indicates that 74.6% of respondents point safety as the first challenge of cloud computing.

In the traditional data centre there is already an implicit need to trust hardware and software suppliers (as well as in outsourced and own staff), since each of the various components of the system (hardware, software, humans) may potentially compromise the security of information. Nevertheless, it is still possible to improve the protection of information by overlaying additional security schemes in order to obtain a more protected environment – even when there is less confidence in each supplier. A good example of this approach would be the installation of multiple firewalls and intrusion detection systems from different vendors (serially laid out).

However, in cloud computing the custody of information is handed over to a third party. Thus, there is a fundamental difference: while previously in traditional paradigms the information systems can be protected from one specific supplier, in cloud computing all these systems are typically managed by the service provider – resulting in the rather uncomfortable need of completely relying on the cloud service provider.

Whenever an organization analyses the possible migration of its IT services to the paradigm of cloud computing, availability, security, privacy and data integrity are on the top of considerations. These concerns relate with both technological and legal matters, bearing in mind that the service provider can be legally responsible for any security breach in a cloud-based service, but, nonetheless, the client is usually the most severely affected. Therefore, before moving any service to the cloud it is vital to properly understand and model the division of responsibilities, risks and potential impact between the client institution and the cloud service provider. Additionally, the customers must recognize that, despite shifting their IT infrastructure to the cloud they are still responsible for compliance, risk and security management. Otherwise, the expect benefits provided by cloud computing can be counterweighted by the involved risks.

Considering the general field of IT, the Information Technology Infrastructure Library (ITIL) [ITIL,], stands out as a widely recognized reference guideline for IT service management. Published by the Central Communications and Telecommunications Agency (CCTA) and, more recently, the Office of Government Commerce (OGC), ITIL provides a practical, no-nonsense framework for identifying, planning, delivering and supporting IT services to the business. Consisting of a set of good practices, described over five volumes known as Service Strategy, Service Design, Service Transition, Service Operation and Continual Service Improvement, ITIL is currently in Version 3 (known as ITILv3 and ITIL 2011 edition). Promoted by the English government for use in IT companies in 2007, and last updated in 2011, ITILv3 it has been rapidly adopted throughout Europe as the de facto standard for best practices in IT service delivery.

However, the field of cloud computing imposes novel challenges which are not properly covered by ITIL or similar initiatives, such as the Control Objectives for Information and related Technology (COBIT) recommendations. Even though such general sources can be partially translated to the context of cloud computing, there are several areas where the paradigm shift precludes straightforward application. Therefore, it is time to re-think how to handle the process of migrating IT services to the cloud, in order to develop a proper set of guidelines, recommendations and good practices. This paper contributes to such discussion, focusing in the specific area of security, analysing how to extend/adapt ITIL recommendations to the field of cloud computing.

The remaining sections of this paper are organized as follows. Section 2 compares the security challenges imposed by traditional IT services with the problems imposed by the concept of cloud computing. Section 3 discusses the migration of IT services to cloud computing, from the point of view of security. Section 4 discusses how the ITIL best practices could assist in the execution of such migration tasks. Finally, Section 5 concludes the paper.

2. The security challenges introduced by cloud computing

As already mentioned,security, privacy and integrity are some of the biggest concerns in the implementation and use of the cloud computing services [Armbrust et al., 2009]. However,d ata encryption, compliance with standards and service level agreements can be used to minimize security concerns.

From a technical point of view, the majority of security risks associated with cloud computing are already present in traditional data centres (or as argued by [Jansen, 2011] known problems cast in a new setting). Possibly, apart from very specific risks induced by server virtualization (which also exist, to some extent, in traditional data centres using server consolidation), most of the security risks are shared by both paradigms – for instance SQL injection, cross-site scripting, zero-day exploits of applications and operating systems, etc.

Virtualization does increase the impact of some of these risks, since successful attacks on the hosting machine (where the hypervisor is located) may potentially compromise every hosted virtual machine. However, such events can be reasonably avoided and/or controlled using appropriate protection mechanisms for the hosting machines. Faults on the virtualization platforms themselves are also an obvious risk, but up to now, there are very few examples of such faults (and even fewer of negative consequences of such faults).

In simple terms, availability means that an organization has its full set of computing resources accessible and usable at all times [Jansen, 2011]. Availability is also a major concern, even though there are no fundamental differences, from a strictly technical point of view, between traditional services and cloud-based services – except for the possible addition of more network links to the core of critical system components.

The majority of problems, therefore, are not inherently technical. They relate with the implicit need to trust external parties to maintain critical information and provide critical IT services. This need was already present in traditional IT services – whenever new equipment or new applications were deployed in the data centre, there was an implicit need to trust the associated providers. Nonetheless, there is a fundamental difference: in cloud computing it is much more difficult to manage the chain of trust, since there is no clear view – for the client institution – on the way the service is provided. The client only knows the service provider, and the whole web of subcontracted service components is usually opaque or, at most, not verifiable by third parties.

Under the cloud-computing paradigm, an organization relinquishes direct control over many aspects of security and, in doing so, confers an unprecedented level of trust onto the service provider [Jansen, 2011].

In order to maintain its systems protected, the customer needs to gather detailed information about the security-oriented requirements of its IT services, applications and data. This knowledge will be useful when migrating to cloud computing paradigms, since it allows comparing and evaluating traditional services with their cloud-based counterparts.

Before the customer can solve/mitigate the issues on security, he needs to perform a risk assessment to properly identify and evaluate the assets, threats and the possible countermeasures to implement.

Identify and evaluate assets

In traditional data centres, the customer assets encompass information, applications, hardware, network, installations and IT workers. However, the cloud-computing paradigm moves some of the responsibilities from the customer to the cloud provider. For instance, the cloud provider becomes responsible for hardware (in the case of IaaS) or applications and hardware (in the case of SaaS). Therefore, the customer should determine in advance, for the assets to move to the cloud, how valuable they are and what happens if, for instance, information becomes stolen or simply inaccessible.

Identify the threats

After proper identification and evaluation of the assets, it is important to recognise the threats that these assets may suffer. On cloud computing scenarios, security threats may arise from various sources, such as loss of availability, security flows on the cloud provider, other customers of the same cloud provider, attacks from external parties, etc. It is necessary to identify the threats that applications, data and virtual machines may suffer in the cloud.

Identify and apply countermeasures

After identifying the threats, the customer should apply the necessary measures to solve the problems encountered.

3. Security issues in the migration to the cloud

Cloud computing is the latest trend to partially or completely outsource IT operations to run a business from the public cloud that provides a flexible and highly scalable technology platform for an organization's business operations [Dhar, 2011]. The concerns faced by IT managers, when moving servers, applications and data to the cloud computing paradigm, can be grouped in two major sets: the so called *on-premise* and *off-premise*. The first includes all the issues that the IT managers must solve on their own, while the second encompasses the issues that IT managers must answer with the support of the cloud providers.

The *on-premise* set includes issues related with the complete identification of the IT customer solution. This group also encompasses, for instance, the complete and detailed list of all applications, the iteration among them, the identification of the data they use, the related security requirements, special hardware or software requirements, etc.

Data used by the applications also deserves the attention of the IT manager that needs to obtain a detailed definition of information that includes adopted security policies – namely in terms of availability, confidentiality, integrity, availability, access definition and redundancy.

Another important aspect is the infrastructure of network communications, namely external connections. This is a vital issue, since the access to the cloud computing services (and therefore the availability) depend on this communication line. The IT manager should identify the requirements and the costs of these connections, their security and the necessity of backup solutions.

The *off-premise* set encompasses the identification of the relevant providers that best answer to the necessities that were previously defined, as well as the identification of the services offered by each provider– in order to choose the more appropriate for the institution.

The identification of services includes those that are available in each of the service models (IaaS, PaaS and SaaS) and the characteristics of each service, namely the time needed to provision a service, to update, to expand and contract the resources and the service uptime. The organisation must validate what cloud model or models best address its concerns.

The cloud applications are grouped into three major groups, the applications migrated by customer, the applications made available by the cloud provider and the applications used by the provider to manage the cloud.

The applications migrated by the customer to the cloud are made available via IaaS or PaaS models. The customer should inquire the provider to obtain information regarding the degree of security and integrity provided to the applications. Additionally, information should be gathered regarding the backup copies in order to recognize its periodicity, location, validity and the possibility of make backup copies of data from the cloud among others. Moreover it is needed to identify what are the import and export facilities provided, specifically in terms of portability with the applications from other providers, what are the associated costs and what are the standards used to transfer the applications to and from another provider and from the customer institution to the cloud.

The applications made available by the service provider, typically via SaaS model, also need the attention of the team in charge of the move to cloud. The customer needs to recognize what customizations of applications are suitable to the institution needs, which the security measures are needed to assure the requirements, the details on how the information backup is processed, and by whom those applications are developed and tested.

The group of applications that the cloud provider uses to manage its services also need similar care from the costumer.

The data is another important issue in the cloud-computing paradigm requiring that the customer request information on where the information stored, since different locations may have different jurisdictions. Additionally, the customer needs to know who manage the storage, the cloud provider or a third party and what the costs are. The data security also needs special attention. Therefore it is need to know who have access to data, systems administrators, network managers or other employees, how often are the backup copies made and what types and levels of encryption that the provider can offer in order to ensure that data cannot be read notwithstanding was stolen.

As said, the security is on top of the top of concerns when moving services and data to cloud computing. Thus, the customer and the cloud provider must clearly identify what are their responsibilities and the responsibilities of other stakeholders in the process like staff responsible for the management of systems, including systems administrators, network managers and other employees. Additionally the customer needs to gather information concerning the provider employees such as training certifications and background.

To validate that the provider meets the applications security requirements previously identified, the customer need to request additional information form the provider regarding access management, namely the authentication mechanisms available and the level of integration that can be done with in-house authentication. The security issues should also encompass the cloud provider infrastructure that is, equipment, network, security standards and the provider security certifications. Knowing the security certifications of the various providers could assist the customer on the selection of the service provider.

The recording of the actions, the errors and the results (logs) by the service provider is an additional security issue that the customer should be concerned. Besides gathering the details of the information registered, the customer must also know how long the records are kept, identify who has access to them, in what way and as said by[Marston et al., 2011] confirm how long the information is stored to allow, if necessary, forensic analysis .

The Service Level Agreement defines a written agreement, negotiated between the customer and the service provider, which documents the agreed service levels and the respective costs. The SLA should include, in terms of security, service uptime, problem solution time, performance, response time, security measures, terms definition.[Cochran and Witman, 2011] emphasizes the importance of records and SLA when they state that the SLA should clearly identify the data that administrators have access and whether or not there are records of personal data. A more detailed discussion about the SLA can be found in [Kandukuri et al., 2009].

Another author [Patel et al., 2009] also highlight the SLA prominence in the cloud computing paradigm stating " As more costumers delegate their tasks to cloud computing service providers, the service level agreements between customers and service providers emerge as a key aspect".

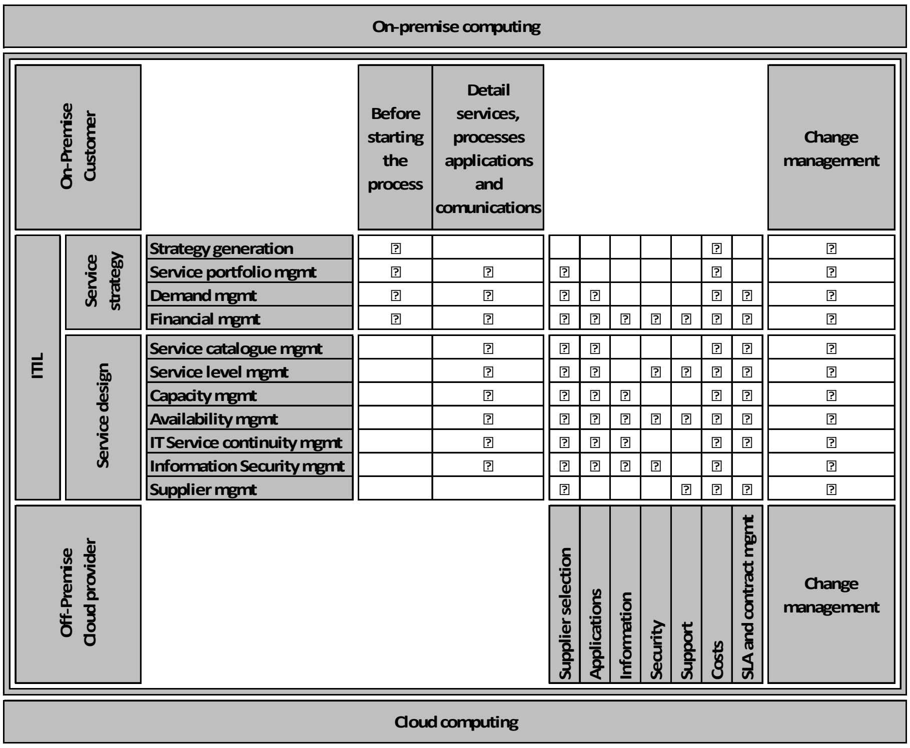
In short, the SLA forms an integral part of a client's first line of defence [Ramgovind et al., 2010].

4. ITIL and the security when moving to cloud

The migration to the cloud-computing paradigm is a complex task that needs the right tools to be more easily, with more control and in a more accurate way accomplished. Whereas in section 3 were presented the various activities and issues that must be engaged by the institution, with and without the cloud providers cooperation, in this section is presented, from the point of view of security, how ITIL could be the right tool to assist on the move to cloud computing.

The security issues presented are similar to the tasks and problems found in the management of the traditional IT services life cycle. Therefore and being the ITIL framework a set of good practices for the identification, planning, delivering and supporting of IT services, could be used to solve the security issues presented and support the institution in the development of those activities. Additionally ITIL facilitate communication between service providers and customers [Nehme et al., 2009], the two main actors in the migration to the cloud computing. Although it was published in July 2011 an update to the current version, 3 of ITIL (ITIL 2011 edition) did not seem necessary to recast the work as it still is not to our knowledge that has already been implemented in an institution.

Figure 1presents the relationship between the processes defined in the first two ITIL books and the activities and issues presented in order to obtain a broader perspective of the interaction between ITIL and the migration to the cloud computing. However, in the in the following paragraphs, are shown those that have a close relationship with security.



On-premise computing

On-Premise Customer

ITIL			Before starting the process	Detail services, processes applications and comunications	Supplier selection	Applications	Information	Security	Support	Costs	SLA and contract mgmt	Change management
ITIL	Service strategy	Strategy generation	⍰							⍰		⍰
		Service portfolio mgmt	⍰	⍰	⍰					⍰		⍰
		Demand mgmt	⍰	⍰	⍰	⍰				⍰	⍰	⍰
		Financial mgmt	⍰	⍰	⍰	⍰	⍰	⍰	⍰	⍰	⍰	⍰
	Service design	Service catalogue mgmt		⍰	⍰	⍰				⍰	⍰	⍰
		Service level mgmt		⍰	⍰	⍰		⍰	⍰	⍰	⍰	⍰
		Capacity mgmt		⍰	⍰	⍰	⍰			⍰	⍰	⍰
		Availability mgmt		⍰	⍰	⍰	⍰	⍰	⍰	⍰	⍰	⍰
		IT Service continuity mgmt		⍰	⍰	⍰	⍰			⍰	⍰	⍰
		Information Security mgmt		⍰	⍰	⍰	⍰	⍰		⍰		⍰
		Supplier mgmt			⍰				⍰	⍰	⍰	⍰

Off-Premise Cloud provider

Change management

Cloud computing

Figure 1: ITIL and the move to cloud computing

Service Catalogue Management

According to [Taylor, 2007] the objective of the service catalogue management is to manage the information contained within the service catalogue and to ensure that it is accurate and reflects the current details, status, interfaces and dependencies of all services that are being run or being prepared to run in the live environment. An accurate and consistent picture of the services made available by the institution is crucial to know what services could be moved to the cloud, what resources and security are needed by those services.

Service Level Management

The Service Level Management objectives embrace the monitoring and the improving of customer satisfaction with the quality of services delivered. The achievement of these objectives supportsthe cloud computing migration and maintenance. While monitoring the SLA the customer also certifies the quality of services provided by the cloud provider ensuring that the cloud services are provided in accordance with the contract. The security issues found in other ITIL process must be reflected in the SLA in order to have a written agreement with the provider in order to ensure that the customer's security needs are met.

Information Security Management

As stated in the previous sections the security is one of the major concerns with when adopting the cloud-computing paradigm. Therefore, the Information Security Management (ISM) is a key concern area being mandatory to ensure the integration between IT security and business security. The purpose of the ISM process is [Lloyd et al., 2007] to ensure that the security aspects with regard to services and all Service Management activities are appropriately managed and controlled in line with business needs and risks.

The Information security management could assist the migration to cloud computing in the same way it interacts with service level management, that is, support in determining of security requirements and responsibilities and their inclusion within Service Level Reports (SLRs) and SLAs.

The move to the cloud-computing paradigm entails a shift of some of security responsibilities from the customer to the service provider. However, the ultimate accountability, in terms of information security, always rests with the customer itself.

Availability Management

The goal of the Availability Management is to ensure that the level of service availability delivered in all services is matched to or exceeds the current and future agreed needs of the business, in a cost-effective manner [Lloyd et al., 2007]. Despite the ITIL Availability Management be a very important process in the migration to the cloud-computing paradigm, afterwards the services are migrated to the cloud it is difficult for the customer, to obtain the same information that it has in-house from service provider to maintain same level in the availability management. However, much of the responsibility of availability management is transferred to the cloud computing service provider.

5. Conclusion

Cloud computing encompasses various technologies, such as computing networks, virtualization, operating systems, used in the traditional data centres leading to that it may suffer some security problems associated with such technologies. While in the IT departments and datacentres the manager may install multiple firewalls and intrusion detection systems from different vendors (serially laid out) in order to protect the information in the cloud computing environment the customer must trust in the providers the security of his information and applications.

The migration to cloud computing has various security challenges. The use of the ITIL methodology either to obtain the necessary information either to point the appropriate methodology are an important assistance for the resolution of such issues.

The share of responsibilities between the customer and the cloud service provider governed by ITIL presented, as an approach to solving the security issues in the moving to cloud computing, has several advantages. The use of ITIL by both the customer and the provider facilitates their communication, making clear the interaction between them - since they speak the same language. Additionally this shared responsibility also comes from the point of view that each one has about security. The security is important, for the customer, to protect its data and applications and, in some cases, as a measure for the continuation of its business. For the service provider is a way to safeguard the assets entrusted to it and form the basis of the business. Moreover, security and trust between customer and supplier are crucial to the provider selection.

References

Armbrust et al., 2009 Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R. H., Konwinski, A., Lee, G., Patterson, D. A., Rabkin, A., Stoica, I., and Zaharia, M. (2009). Above the clouds: A berkeley view of cloud computing. Technical Report UCB/EECS-2009-28, EECS Department, University of California, Berkeley.

Cochran and Witman, 2011Cochran, M. and Witman, P. (2011). Governance and service level agreement issues in a cloud computing environment. *Journal of Information Technology Management*, 22(2):41.

Dhar, 2011 Dhar, S. (2011). From outsourcing to cloud computing: Evolution of it services. In *Technology Management Conference (ITMC), 2011 IEEE International*, pages 434 –438.

Foster et al., 2008 Foster, I., Zhao, Y., Raicu, I., and Lu, S. (2008). Cloud computing and grid computing 360-degree compared. In *2008 Grid Computing Environments Workshop*, pages 1–10. IEEE.

Gens, 2008 Gens, F. (2008). It cloud services user survey, pt.2: Top benefits & challenges. Available online at http://blogs.idc.com/ie/?p=210.Seen in: 2010.09.11.

Grobauer et al., 2010 Grobauer, B., Walloschek, T., and Stocker, E. (2010).Understanding cloud-computing vulnerabilities.*Security Privacy, IEEE*, PP(99):1–1.

ITIL, ITIL.Information Technology Infrastructure Library (ITIL).A>vailable online at http://www.itil-officialsite.com.Seen in: 2010.04.17.

Jansen, 2011Jansen, W. (2011). Cloud hooks: Security and privacy issues in cloud computing. In *System Sciences (HICSS), 2011 44th Hawaii International Conference on*, pages 1 –10.

Kandukuri et al., 2009 Kandukuri, B., Paturi, V., and Rakshit, A. (2009).Cloud security issues.In *Services Computing, 2009.SCC '09.IEEE International Conference on*, pages 517 –520.

Khajeh-Hosseini et al., 2010 Khajeh-Hosseini, A., Greenwood, D., Smith, J. W., and Sommerville, I. (2010). The Cloud Adoption Toolkit: Addressing the Challenges of Cloud Adoptionin Enterprise.*CoRR*, abs/1003.3866:1–10.

Lloyd et al., 2007 Lloyd, V., Rudd, C., and Taylor, S. (2007). *OCG Books ITIL - Service Design*. TSO (The Stationery Office).

Marston et al., 2011 Marston, S., Li, Z., Bandyopadhyay, S., and Ghalsasi, A. (2011). Cloud computing - the business perspective. In *System Sciences (HICSS), 2011 44th Hawaii International Conference on*, pages 1 –11.

Mell and Grance, 2011 Mell, P. and Grance, T. (2011).Cloud computing.Available online at http://csrc.nist.gov/-publications/nistpubs/800-145/SP800-145.pdf.Published date: September 2011.Seen in: 2011.10.22.

Nehme et al., 2009 Nehme, J., Persson, M., and Lahiji, S. (2009).*How can ITIL influence IT outsourcing*©.PhD thesis, JÖNKÖPING INTERNATIONAL BUSINESS SCHOOL.

Patel et al., 2009] Patel, P., Ranabahu, A., and Sheth, A. (2009). Service Level Agreement in Cloud Computing. *Cloud Workshops at OOPSLA09*, 1:1–10.

Ramgovind et al., 2010 Ramgovind, S., Eloff, M., and Smith, E. (2010). The management of security in cloud computing. In *Information Security for South Africa (ISSA), 2010*, pages 1 –7.

Shimba, 2010 Shimba, F. (2010).*Cloud Computing: Strategies for Cloud Computing Adoption*. PhD thesis, Dublin Institute of Technology,.

Taylor, 2007 Taylor, S., editor (2007).*ITIL – The Official Introduction to the ITIL Service Lifecycle*.TSO, London.

Vaqueroet al., 2009 Vaquero, L. M., Rodero-Merino, L., Caceres, J., andLindner, M. (2009). A break in the clouds: towards a cloud definition.*SIGCOMM Comput.Commun. Rev.*, 39(1):50–55.

Vouk, 2008 Vouk, M. (2008).Cloud computing: Issues, research and implementations.In *Information Technology Interfaces, 2008.ITI 2008.30th International Conference on*, pages 31–40.

Zhang et al., 2010a Zhang, Q., Cheng, L., and Boutaba, R. (2010a). Cloud computing: state-of-the-art and research challenges.*Journal of Internet Services and Applications*, 1:7–18. 10.1007/s13174-010-0007-6.

Zhang et al., 2010b Zhang, S., Zhang, S., Chen, X., and Huo, X. (2010b). Cloud computing research and development trend. *Future Networks, 2010. ICFN "10. Second International Conference on*, 0:93–97.

EU law and Internet Traffic Control Lost Between Privacy Rights and Freedom of Individual and Corporate Enterprise

Filipe Domingues Cerqueira Alves
Social Sciences Faculty – Portuguese Catholic University, Braga, Portugal
fcalves@braga.ucp.pt

Abstract: The European Union is facing a shift of legislative paradigm as far as cyberspace is concerned. Recent legislative movements in EU countries have sought to prosecute presumed illegal activities, mainly associated to file-sharing communities violating principles of intellectual property law. As the attempt to regulate and coordinate legislation on specific Internet abuses takes place, boundaries of privacy rights as they were previously understood are questioned. Yet, as France implements an independent authority with specific traffic monitoring powers and a generality of countries moves towards an additional taxation of physical devices considered as potentially promoting copyright violation activities, the ECJ, in its recent ruling in Case C-70/10 (Scarlet vs. SABAM), has precluded an injunction made against an Internet service provider which requires it to install a system for filtering all electronic communications passing via its services, in particular those involving the use of peer-to-peer software, complying with special and particularly strict requirements, with a view to blocking the transfer of files the sharing of which infringes copyright. Such ruling deeply contributed to the establishment of a milestone on this enduring process as it is now secure that a general traffic monitoring filter cannot by applied by an ISP and at its costs. Moreover, the *rationale* of the decision had clear implications on two major areas of Law. On one hand, privacy rights are clearly at stake since traffic monitoring cannot, *ab initio*¸ distinguish licit from illicit traffic and will provide ground for multiple privacy violations and abuses if not carefully regulated. On the other hand, ISPs and industry companies are concerned as the costs, expenses and burdens of such monitoring are bound to run on their side. Nevertheless, several questions concern the audiovisual industry in particular and the community in general. How can illegal Internet activities – not only file sharing – be monitored? Who can monitor them? What can be defined as abusive *vis-à-vis* user's privacy? How far can the monitoring obligations go so that they do not become an excessive restraint on freedom of individual and corporate enterprise? The quest for privacy rights' defenders has just only begun. This paper contributes for the answer of the previous questions while it attempts to approach a technical and legal crossed analysis of traffic monitoring alternatives, seeking to determinate whether the current legal establishment allows room for such strict regulation, as the audiovisual industry desires, or if intellectual property defense must be sought after by some means other than traffic monitoring.

Keywords: privacy rights; internet traffic control; fair balance; intellectual property rights; C-70/10

1. Introduction

We live times of shifting paradigms as far as our relation with technology is concerned. As such, while the general consumer tends to get acquainted and familiarised with the possibilities that technology provides, several society sectors struggle to face the speed of change.

In particular, the legal sector often finds itself lagging behind the exponential development of tools that the digital world offers consumers. Sometimes, legislative production cannot keep up the pace of real life developments. On other occasions, it prefers to wait and see as the situation unfolds and as courts trail way to new interpretations of law, connecting the dots between the law in books and the law in action.

The relation between intellectual property rights' (IPR) enforcement via Internet control and European Community (EC) law is currently in a similar standoff, as one can grasp from the *Commission Report on the Application Of Directive 2004/48/EC of the European Parliament and the Council of 29th April 2004 on the Enforcement of Intellectual Property Rights* (COM/2010/779) and from the *Public Hearing on Directive 2004/48/EC*, listing the wide array of challenges posed by the digital environment.

Several European Court of Justice (ECJ) decisions tend to shed some light on how can or cannot the intellectual property rights (IPR) be enforced by their holders. Meanwhile, member states' efforts on interpreting and implementing the several communitary directives that deal with the present matter seem to work as a factor of disharmonisation rather than harmonisation of law inside EU due to the amount of leeway each country is left as to the exact rules to be adopted by national legislation. In fact, efforts to prevent file-sharing communities from exchanging IPR protected material are challenging the boundaries of privacy rights as they were understood until so far as the attempt to regulate and coordinate legislation on specific Internet abuses takes place.

In its ruling from 24th November 2011, process C-70/10, *Scarlet Extended SA* vs. *Société Belge des auteurs, compositeurs et éditeurs SCRL (SABAM)*, the ECJ had the opportunity to clarify some points related to Internet traffic control.

2. The facts brought before the ECJ

A reference for a preliminary ruling under Article 267 of the Treaty of Functioning of the European Union (TFEU) was brought before the ECJ from the "Cour d'Appel de Bruxelles" (Belgium) concerning an Internet Service Provider (ISP) named Scarlet's refusal to install a system for filtering electronic communications which use peer to peer (p2p) file-sharing software infringing intellectual property rules.

SABAM is a management company that represents several branches of musical industry for purposes of copyright protection and authorisation of use of protected works by third parties. In 2004, SABAM concluded that Internet users took advantage of the services provided by Scarlet to download from the Internet pieces of SABAM's catalogue through p2p software, with no authorisation whatsoever and without paying any type of fees. Immediately, Scarlet cited SABAM for a proceeding before a court of first instance, claiming that, as an ISP, SABAM was the better positioned party to take measures destined to cease IPR violations by their clients. Firstly, SABAM sought a declaration of effective existence of IPR violations referring to its catalogue and, secondly, demanded *"Scarlet to bring such infringements to an end by blocking, or making it impossible for its customers to send or receive in any way, files containing a musical work using peer-to-peer software without the permission of the rightholders, on pain of a periodic penalty"*. Lastly, *"SABAM requested that Scarlet provide it with details of the measures that it would be applying in order to comply with the judgment to be given, on pain of a periodic penalty*[1] *"*.

Only when Scarlet brought an appeal against the first instance decision, favourable to SABAM, was the ECJ questioned whether the installation of a filtering system would be in breach of the provisions of European Union (EU) law on the protection of personal data and the privacy of communications, since such filtering implicates the processing of IP addresses which are considered personal data.

3. The attorney-general's (AG) opinion

The AG Cruz Villalón handed down his opinion: the ECJ was dealing with a matter of fundamental rights, the address of which would have to precede an analysis based on the proportionality of the measure at stake. Recalling fundamental principles stated on the Charter of Fundamental Rights of the European Union (CFREU) and on the European Convention on Human Rights (ECHR), the AG pointed out firstly that a filtering and blocking mechanism necessarily has to control the entire data packets either uploaded or downloaded from the users' station. Simultaneously, a systematic, universal and evolving method would be in consideration, forcibly comprising p2p and other type of communications that could involve any type of IPR violations, such as direct downloads from websites as the now infamous MegaUpload or RapidShare, just to name some.

Stressing that SABAM has not developed nor commented on the specific and technical *modus operandi* of the so called filter, the AG analysed the outline of a filter whose sole purpose would be to render impossible any form of upload or download by Scarlet's clients of material contained in SABAM's catalogue. Hence, SABAM's plea would constitute an obligation of result imposed on Scarlet, at its own expenses, and consequentially heavily burdening of its actions. Furthermore, the AG considered that adjudicating such an injuction would have an impact that would transcend the present case. In fact, a milestone could be set implying significant consequences to other ISPs and to other Internet users, for a permanent filter previously controlling every exchange of information that every general ISP client undertakes is deemed to interfere with general information exchange. It must be noted that the AG's opinion constantly and progressively alerted to the implications that the present ruling could have for the Internet business as we know it.

In legal terms, the AG believed that such a filter would be in serious contention with articles 7, 8 and 11 of the CFREU (the ones concerning respect for private and family life, protection of personal data, freedom of expression and information), being also here at stake property rights (article 17(2), in the

[1] See Case C-70/10: Judgment of the Court (Third Chamber) of 24 November 2011 (reference for a preliminary ruling from the Cour d'appel de Bruxelles (Belgium)) — Scarlet Extended SA v Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM), Official Journal C 025 , 28/01/2012 P. 0006 – 0007.

shape of IPR). A restriction to the former would be in progress, as considered by articles 8(2) and 10(2) of the ECHR.

Since an IP address is to be considered personal data and despite the propelled “restriction” being necessary in order to protect IPR, the AG ended up stating that SABAM’s plea could only be put in action if applied to every Scarlet’s end user, systematically and universally, permanent and perpetually, and with no guarantee of a fair proceeding aimed at personal data protection or communication confidentiality, denying opposition to the blocking of a determined file or appeal for such an action. Thus, he concluded for a negative answer to the question posed by the referring tribunal.

4. ECJ’s ruling

The ECJ quickly ruled out the possibility of an active monitoring of all electronic communications conducted on the network of the ISP as it would be opposed by article 15(1) of Directive 2000/31. Recalling the Case C□275/06 *Promusicae*, the ECJ appealed to a fair balance between fundamental rights by national authorities, concluding that no such balance could be achieved by a disposition that would infringe article 3(1) of Directive 2004/48 as it imposes an excessive burden on the ISP alongside with a potential violation of articles 8 and 11 of the CFREU, hence precluding an injunction made against an ISP which requires it to install a system for filtering all electronic communications passing via its services, in particular those involving the use of peer-to-peer software, complying with special and particularly strict requirements, with a view to blocking the transfer of files the sharing of which infringes copyright. It must be noted, still, that the ECJ refrained from addressing the compatibility of the fundamental rights at stake by referring only to the analysis of Belgium’s national law and it’s inapplicability to the case.

Such decision, while simple in its reasoning, laid path to the establishment of a clear borderline: general *ex ante* Internet traffic filters cannot be implemented on an ISP, at its own costs, since they are considered excessively costly and burdensome. Simultaneously, it endured the pending debate on the definition of limits to privacy rights confronting IPR. We will start by discussing the latter.

5. Comment

The fact that the ECJ recalled the *Promusicae* case is a clear symptom of how hesitant it is in making a decisive contribution to the debate that pits IPR holders against end users that download their protected works, with the ISPs closely watching in the corner.

In the *Promusicae* case, the ECJ brought the e-Privacy Directive 2002/58/EC into play, connecting it with the Data Protection Directive 95/46/EC, to conclude that “the protection of rights and freedom of the others” can indeed support restrictions on obligations of general confidentiality about private data imposed on the ISPs but, as the EU law currently stands, there is no such obligation on the member-states to enact or enforce the said restrictions. But *Promusicae’s* paragraph 68 (the one brought forward by the ECJ on the current matter) has a deeper meaning than at first glance. Actually, since the interpretation of the Directives by national law should be achieved through “a fair balance to be struck between the various fundamental rights protected by the Community legal order”, one cannot be oblivious that such interpretation has a large degree of appreciation by the member-states, both during the legislative procedure and in the judicatory appreciation of concrete cases by courts[2].

This outlook somehow conducts Europe to a “merely formal European constitutionalisation” (Kosta 2010), which has a devolutive effect back to national legislations as far as striking a balance between IPR and other fundamental rights is considered. Furthermore, it cannot be forgotten that ISPs’ fundamental rights such as property rights are also at stake, since a restrictive measure as the one proposed could pose an excessive weight on its property and freedom of establishment.

Recalling this principle back from *Promusicae*, the ECJ provided orientation on how cannot a balance between the concerned rights be struck – not at all by means of a general *ex ante* Internet traffic control filter. Yet, it is easily understandable why the ECJ refrained from proclaiming criteria that would serve as orientation on how can such a balance be struck, apart from the rather obvious and wide criteria that are implied in the ECHR and the CFREU.

[2] Refer to Jančić (2010: 430 – 461) for more information on the development of law interpretation in Europe and, in particular, in France.

During 2011, the Commission launched a series of initiatives[3] intended on ascertaining the panorama of IPR's enforcement, mainly based on the application of the 2004/48/EC Directive. The Comments on the Comission Report and the Public Hearing on the 2004/48/EC Directive made known a general perception of the Directive's inability to fulfil its goals and revealed a general concern on the role of intermediaries (namely, ISPs) on the matters of IPRs violations on digital environment, notably via file-sharing communities. It was also noted what could be described as a legal uncertainty to which by no means the ECJ contributes to soothe. But in the end, most of the stakeholders' opinions would converge: private users are concerned with their privacy, IPR management entities with their property and civil responsibility for IPRs violations via liability of the direct offenders: the 2004/48/EC Directive should be reformed.

However, inadvertently in the C-467/08 *Padawan* case, the ECJ might have set a new paradigm that can help the interpretation of EU law and its national application (Karapapa 2011). Stakeholders are well aware of the distinction that 2004/48/EC Directive draws between commercial scale violation of IPRs and private use of material protected by IPR. In fact, it is essential for the notion of IPR that there is, at some point, any kind of private usage of the protected works as the cultural and social development meant to be protected and stimulated by IPR can only occur if one effectively comes into touch with previous works. Hence, private usage and copying were considered up until now as a limitation to the reproduction right in form of a statutory license under article 5(2) of 2001/29/EC Directive. Implicitly, though, in *Padawan* paragraphs 52 to 56 the ECJ reasons that private copy must be treated as a user's right, which is in our opinion consentaneous with the appeal to a certain notion of justice that the AG Verica Trstenjak makes in her conclusions handled for the case. In fact, if one is bond to pay a levy based on a presumption that he will copy of protected works, one must consider himself entitled to do so, even if it is not his intention. If a presumption allows the Court to infer the practise of an unproved fact (the effective private copying) from an established reality (the purchase of merchandising that enables private copying or can be copied) by means of which the IPR managing entities are owed a compensation; if final consumers are not allowed to provide evidence that they did not copy any protected works in order to avoid the payment of such a levy; and if domestic private copying cannot be prevented for it would constitute an excessive privacy invasion; then, the user is legitimately entitled to privately copy the works in his possession.

While the AG Trstenjak clearly states (paragraph 78) that p2p file-sharing cannot be considered within the limits of "private copying" concept, it cannot be ignored that 1) users can make private copies of works; 2) sharing does not include, in the majority of cases, any kind of economical purpose. Both these considerations cumulated with the fact that generally monitoring an users' Internet connection is, as EU law stands, an illegal practice, seems the leave the end user not comprised within any legal provision that will stop him from privately copying protected works.

In fact, Internet connection is, as of today, a primary means of communication and socialisation, and it as even been considered by the French Constitutional Council as a fundamental right. Also, its connection with freedom of speech and expression is undeniable – here comprised the right to take part in the cultural life of the community (Akester 2010), as stated on article 10 ECHR, article 11 CFREU and article 27 of the Declaration of Human Rights. One of the main differences between democratic regimes and dictatorships is the control of Internet's flow of information. The latter tend to exert a rooted influence and impact on the free flow of information, with the undisputable consequences in terms of individual rights, quality of living for its citizens, establishment of pluralism, and generally, the level of development of a society. Free speech and freedom of information are what first comes to mind when one considers the formation of a democratic society that wholeheartedly respects human rights. Handing powers to control Internet connections to private parties or even to an administrative entity, with no judicial intervention, could also lead way to serious abuses related to censorship or wrongly intended use by information groups, hardly reviewable by courts.

The point is that the technical solutions that were envisaged were not at all compatible with the "fair balance" that needs to be struck between the different rights at stake. Even the expert that was invited to present its conclusions to the referring court in the SABAM process recognised that all but one of the solutions brought forward by the technical companies issued on the matter would block every p2p

[3] See the Comments on the Comission Report and the Public Hearing on the 2004/48/EC Directive and the Consultation on the Commission Report on the enforcement of intellectual property rights, available on http://ec.europa.eu/internal_market/iprenforcement/directives_en.htm.

communication regardless of its content[4]. Meanwhile, the only company that could provide a technical solution that would not pose a threat to the entire volume of communication that a user can process did not assure its functioning considered the amount of traffic an entire ISP generates.

Technically, filtering communications through the ISP is an excessive mean of controlling traffic, as it interferes with other fundamental rights related to freedom of speech and freedom of communication. Other traffic controlling forms seem rather ineffective: for instance, one could obligate the processing of certain traffic protocols via specific ports. That would leave unharmed other p2p software used not to disseminate file-sharing but to communicate, such as Skype, although it would be rather ineffectual when one considers the download via http or ftp protocol from file-sharing websites. These particular websites pose extremely acute problems as it is hardly feasible to control the content of every user's upload. Moreover, not all p2p software enabling file-sharing is used to illicit downloading. Through the use of such software, several artists can find free means of expressing their work, hence balancing the dominium of major labels and propelling innovation and creative industries, as well as the dissemination of uncensored information. The scope of such a measure would have to be considered too large as, obviously, the potential of p2p utilisation far surpasses its downsides on what concerns file-sharing.

In addition, the download of a file is an action located in the user's private sphere and does not fulfil the requirements of "communication to the public and making available rights", that are to be exclusively exercised by IPR holders. It can also be questioned whether the presumption of innocence and the fault principle are not being circumvented, since IPR management entities[5] state that without illicit downloads, users would buy their music on the shops. Still, that remains to be proved, for a subject may be willing to obtain a work for free but not at all willing to pay for it. Thus, virtually, no harm was ever to be made to the IPR management entity, since that user would never pay for the work he downloaded. Control of Internet traffic seems to be a hard task for the IPR holders.

Furthermore, there is a side of the industry that is constantly neglected on the discussion between privacy rights and IPR: the ISP.

The four EU freedoms comprise freedom of enterprise, as property rights are considered fundamental rights by the CFREU and the ECHR. The ISP's property rights would be seriously at stake if an injunction forcing the ISPs to monitor their traffic would be proclaimed. Despite such protection, serious breaches on fair competition between the ISPs could surge if such a proceeding would become standard. It is shouldn't come as shock that file-sharing is actually a cornerstone in the IT business. The more clients feel the urge to download and upload files, the more bandwidth they will crave for. More bandwidth and generalisation of connections benefit the ISPs as the increase in demand becomes notable. But the generalised demand of high-speed connections also provides the users the possibility of high speed/low price Internet connections. If the ISPs had to implement any type of filtering on their networks, the consequences would be unimaginable. Connection speeds could be lowered – and thus, promotion of communication and information harmed. Contractual breaches related with signed connection speeds and effectively offered connection speeds could arise. Prices would suffer an increase as well, due to the costs of such mechanisms, and the whole business system as we know it would have to go through considerable changes.

In addition, the financial burden of such mechanisms as intended by SABAM would run exclusively on the ISPs, which we tend to find an inadmissible interference in one's own business, hence excessively disturbing the freedom of enterprise. The costs and the logistics associated with such a mechanism would tend to take a large share of the ISP's resources, thus rendering connection progress unsustainable as faster communication speeds would imply more traffic, which would imply a larger volume of data to be analysed and so on.

If technical difficulties arise and if a judicial intervention is regarded as essential taking into account the fundamental human rights at stake when one deals with Internet connections, it seems to us that the path trailed by IPR management companies will not lead them to an appropriate protection of their interests. In fact, a viral sentence propelled by SOPA and ACTA oppositionists claims that *if the business cannot sustain its business model in the face of civil liberties, than it is the business model*

[4] See paragraph 21 of the sentence.

[5] For more on IPR management entities, see Dehin (2010: 220 – 237).

that must change, and not civil liberties. We are forced to concur. Indeed, a model where an audio CD can cost € 25,00 in a country where the minimum wage does not amount to € 500,00 cannot be considered balanced nor fair – we relate to the Portuguese reality. Major labels and IPR management entities must evolve and develop other business aside from their former core business: retail selling of their products.

In order to trigger both social and cultural development by conferring protection to the authors and knowledge to the consumers – "downstream creations often rely on the possibility of studying upstream creations" (Bonadio 2010) - we strongly believe that a non-legislative path must be trailed. It is also our view that ISPs must ally themselves with the IPR management entities, providing a fair system of checks and balances. For one, the ISPs could gather access to the IPR management entity's catalogue for a fair price and provide them to its customers via their own distribution channels, while pledging to help and effectively entail persecution of violation of those said catalogues. Quality of Service politics, such as traffic-shaping concerning certain protocols, could also pose a balanced solution as it would not imply a full monitoring of the connection, but only the allocation of the necessary amount of bandwidth to assure freedom of expression with no excessive restrictions on one's network performance. Thus, mass scale traffic of IPR protected material would be strongly refrained. Rights of freedom of expression and privacy would not be violated; the ISPs would not have their freedom of enterprise affected nor their business model would be harmed with excessive inconveniencies that would render their cost/profit relation non attractive.

6. Conclusion

An aura of legal uncertainty surrounds the protection of IPRs. As the ECJ remains shy on proclaiming broad principles that could provide the stakeholders a secure platform on which to operate, legislation and its application vary drastically from country to country (inside EU) as the ECJ places on the member states the task of interpretating EU law accordingly to the binding fundamental rights charts, striking a requested balance between rights at stake. Privacy rights seem to be deeply protected and rooted in our western conscience because of their strong connection with the development of pluralism and democratic societies. Additionally, technical measures seem inept to achieve IPR's managing entities' goals considering their severe interference with ISP property rights and freedom of enterprise. A joint action between the several stakeholders, mainly between the IPR management entities and ISPs, deploying a system of checks and balances that renders the protection of IPR more attractive to the ISP and its clients, indirectly, may be the correct path to follow. Conjugating such measures with private copy levies, could help diversifying the spectrum of legal offers that can be presented to end users and thus reduce the problem of IPR violation.

Acknowledgements

The author would like to publically express his gratitude to Prof. Vítor Sá (Portuguese Catholic University) for his valuable comments on the content of the essay.

References

Akester, Patricia. (2010) "The new challenges of striking the right balance between copyright protection and access to knowledge, information and culture", **European Intellectual Property Review**, vol. 32, no. 8, pp. 372 - 381.

Bonadio, Enrico. (2011) "File sharing, copyright and freedom of speech", **European Intellectual Property Review**, vol. 33, no. 10, pp. 619 – 631.

Dehin, Violaine. (2010) "The future of online legal music services in the European Union: a review of the Eu Commission's recent initiatives in cross-boarder copyright management", **European Intellectual Property Review**, vol. 32, no. 5, pp. 220 – 237.

Jančić, Davor. (2010) "The European Political Order and Internet Piracy: Accidental or Paradigmatic Constitution-Shaping?", **European Review of Constitutional Law,** no. 6, pp. 430 – 461.

Karapapa, Stavroula. (2011) "Padawan v SGAE: a right to private copy?", **European Intellectual Property Review**, vol. 33, no. 4, pp. 252 – 259.

Kosta, Vasiliki. (2010) "Internal Market Legislation and the Private Law of the Member States – The Impact of Fundamental Rights", **European Review of Constitutional Law**, no. 4, pp. 410 – 436.

Identification of Topics Targeted by Attackers

Manoj Cherukuri and Srinivas Mukkamala
Institute for Complex Additive Systems and Analysis (ICASA), Computational Analysis and Network Enterprise Solutuons (CAaNES), New Mexico Institute of Mining and Technology, Socorro, USA
manoj@cs.nmt.edu
srinivas@cs.nmt.edu

Abstract: The attackers often relied on using automated exploit kits to infect the legitimate websites with high traffic and inject malicious content into them. The compromised legitimate websites served the malicious content to its users who ended up getting infected. In this paper we present web crawling, inlink search, topic modeling and computational intelligent techniques to attribute the topics targeted by the attackers. We also identify the attack vector used by the attackers with respect to the topics, in targeting the internet users. A computational intelligent technique relying on Gibbs random sampling was used to extract the topics automatically from a set of webpages under study. The extracted topics are correlated with Google Trends to put forward some interesting properties that are helpful in detecting the malicious websites proactively. We identified the primary topics from the topic models generated by Gibbs random sampling across each month and analyzed the features of the targeted topics with respect to its importance then. We identified the words that were consistently targeted across the period of our study and analyzed the topics on which the attackers are always after.

Keywords: topic modeling, malicious websites, malware targeted topics, malware topic trends, malicious websites attribution, web security

1. Introduction

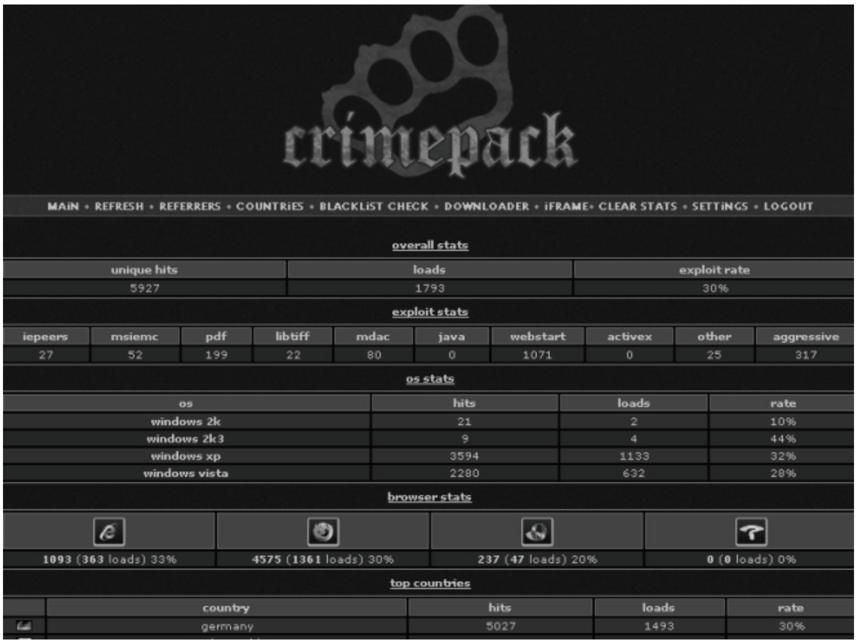
The websites hosting malware or malicious content are proliferated all over the web. Bleaken (2010) stated that 90% of the web based attacks happen through the compromised legitimate webpages. The brisk growth of the compromised webpages indicates the interest of the attackers on using the legitimate websites as a medium for the propagation of the malware and consequently exploiting the victims (compromised users).

The evolution of exploit kits like Crimepack, Neosploit, Phoenix Exploit, LuckySploit etc. had made the job of the attackers easier. These kits were developed with the intent of exploiting the vulnerabilities and injecting the malicious files into the victim's machine. For example, Paget (2010) stated that the Crimepack version 2.2.1 was sold at $400 in the wild. The attackers use the automated exploit kits to generate malicious scripts. These malicious scripts are hosted on compromised legitimate websites to exploit large number of vulnerable machines.

The Figure 1 is the screenshot of the Crimepack attack kit (Crimepack: Packed with Hard Lessons, 2010). It shows all the statistics required for an administrator to get an idea over the impact of the hosted attack. The "overall stats" section gives the total number of users that visited the webpages with the hosted malicious content and the number of cases where the exploit had been successfully loaded.

The "exploit stats" section gives the targeted exploits and the number of exploits that were launched successfully. The "OS stats" section gives the operating system of the machines that visited the infected websites and that have been exploited successfully. The "browser stats" section gives the number of visits from various browsers and that have been successfully exploited. The "top countries" section gives the total number of visitors from various countries and the number of visitors that are infected in the descending order.

Current mechanisms being used for leveraging the loss due to such attacks are blacklisting URLs (Uniform Resource Locators) and IP (Internet Protocol) addresses of the malicious websites. Adversaries employed domain fluxing techniques to overcome the blacklists. The ascension of the number of websites had made the traditional approach of crawling an infeasible solution for detecting such attacks in real time. In order to counter the growth of such malicious websites and hinder the losses due to such acts, the current mechanisms of crawling and blacklisting must be assisted with computational intelligent techniques.


crimepack
MAiN • REFRESH • REFERRERS • COUNTRiES • BLACKLiST CHECK • DOWNLOADER • iFRAME• CLEAR STATS • SETTiNGS • LOGOUT

overall stats

unique hits	loads	exploit rate
5927	1793	30%

exploit stats

iepeers	msiemc	pdf	libtiff	mdac	java	webstart	activex	other	aggressive
27	52	199	22	80	0	1071	0	25	317

os stats

os	hits	loads	rate
windows 2k	21	2	10%
windows 2k3	9	4	44%
windows xp	3594	1133	32%
windows vista	2280	632	28%

browser stats

1093 (363 loads) 33%	4575 (1361 loads) 30%	237 (47 loads) 20%	0 (0 loads) 0%

top countries

	country	hits	loads	rate
	germany	5027	1493	30%

Figure 1: Screenshot of Crimepack from its administrator perspective

In this paper we put forward an attack vector that helps the identification of the malicious websites proactively. We assessed the attack vector used by the cyber criminals with respect to the targeted topics. Previous works were done on the identification of the code features used by malicious scripts (Cova, Kruegel, and Vigna, 2010; Rieck, Krueger, and Dewald, 2010) which try to detect the maliciousness of the page on visiting but do not address the problem of accessing such webpages. Network hubs that acted as a major source for malicious hosting were detected (Kalafut, Shue and Gupta, 2010), but no work has been done with respect to the topics that are targeted by the attackers. In this paper we perform topic modeling on malicious websites to identify the attack vector with respect to the topics the attacker selects to compromise the legitimate websites. We believe this approach to be novel and prospective in protecting the users from web-based malware.

This paper is organized as follows: in section 2, we describe some background knowledge required for better understanding. In section 3, we discuss the method used for topics extraction. In section 4, we discuss the processes involved in our study. In section 5, we describe our dataset. In section 6, we discuss about the analysis of the dataset. In section 7, we conclude with the results.

2. Background

Topic modeling is based on the posterior distribution of the topic structures given the words of the document. Computing the posterior distribution for Bayesian approaches is a major limitation for usage as it involves the integration of high dimensional functions. Markov Chain Monte Carlo (MCMC) is one such type of methods addressing the limitation of Bayesian approaches simulating the direct draws from complex distributions. If the transition probability of a random variable among different states of the state space depends only on the random variable's current state then it is called a Markov process. Therefore Markov random variable state transition depends only on the current state and no information about the past is needed. This is represented by the following equation.

$$\Pr\left(X_{t+1} = s_j \middle| X_0 = s_k \dots\dots\dots X_t = s_i\right) = \Pr\left(X_{t+1} = s_j \middle| X_t = s_i\right)$$

A Markov chain refers to the sequence of random variables (X_0, …. ,X_n) generated by a Markov process. The transition of the state of a random variable from state i to state j in a single step is represented as P(i,j) and is given by the following equation.

$$P(i,j) = \Pr\left(X_{t+1} = s_j \middle| X_t = s_i\right)$$

The probability of transition from state i to state j in n steps is given by the following equation.

$$p_{j}^{(n)} = \Pr(X_{t+n} = s_j \mid X_t = s_i)$$

The Gibbs sampling is a specific form of MCMC, addressing the task of construction of the Markov chain for the required distribution. The main advantage of Gibbs sampling is to consider univariate conditional distribution i.e. all the random variables except for one are assigned fixed values. The univariate conditional distributions are easy to simulate rather than the complex joint distributions. Let us consider a bivariate random variable (x,y). The Gibbs sampling starts with assigning a random value for all of the random variables except for one. Let us assign a random value for y_0 in this case. The sampler proceeds by the following equations.

$$x_i \sim p(x \mid y = y_{i-1}) \qquad y_i \sim p(y \mid x = x_i)$$

3. Topic modeling

Topic modeling on a given corpus results in a set of topic models with each topic model consisting of the cluster of words that occur together frequently and describes the modeled topic. In this paper we use Gibbs Sampling as it provides a simple mechanism for discovering topics by considering the posterior distribution over the assignments of words to topics. The Gibbs sampling procedure considers each word token in the text collection and estimates the probability of assigning the current word token to each topic, conditioned on the topic assignments to all other word tokens. The primary advantage of using Gibbs sampling is that the assignment of the word token to a topic is dependent only on the current state of the remaining word token assignments. From this conditional distribution, a topic is sampled and stored as the new topic assignment for this word token. This conditional distribution is computed using the following equation.

$$P(Z_i = j \mid Z_{-i}, w_i, d_i, \cdot) \propto \frac{C_{w_i j}^{WT} + \beta}{\sum_{w=1}^{W} C_{wj}^{WT} + W\beta} \cdot \frac{C_{d_i j}^{DT} + \alpha}{\sum_{t=1}^{T} C_{d_i t}^{DT} + T\alpha}$$

where $Z_i = j$ is the probability of assigning a topic *j* to the word i. C^{WT}_{wj} and C^{DT}_{dj} are matrices of counts with dimensions W (Words) X T (Topics) and D (Documents) X T. The term C^{WT}_{wj} denotes the number of times the word *w* is assigned to topic j excluding the current instance *i*. The term C^{DT}_{dj} denotes the number of times a word token from the document *d* is assigned to topic *j* excluding the current instance i. Z-i refers to the topic assignments of all other words, wi and di refers to the current word and document instances respectively and "□" refers to all other known or observed information such as all other word and document indices w-i and d-i. This equation provides the probability that is not normalized and is normalized by dividing with the sum of the overall topics, T. All the results provided in this paper were obtained using the Gibbs Sampling for topic extraction which was proposed by Griffiths, and Steyvers (2004). A more detailed description of the model and the algorithm is provided by Steyvers, and Griffiths (2007).

4. Processes

Construction of our dataset is composed of four processes.

- The first process deals with the collection and classification of malicious webpages
- The second process deals with the construction of inlinks for the malicious webpages obtained in the previous process
- The third process deals with crawling and extracting the content of identified webpages in the first two steps
- The fourth process deals with topic modeling on the corpus extracted in the preceding step

4.1 Collection and classification of malicious webpages

The malicious webpages are collected from potential sources that are identified using the meta-search engines. These sources publish malicious webpages along with the metadata such as date, type of attack, executable name etc. A custom parser was written with respect to each such identified source to extract the malicious webpage and the date when it was reported. These sources are crawled to have an extensive collection of malicious webpages. Figure 2 is the pictorial representation of the process of collection and classification of malicious webpages.

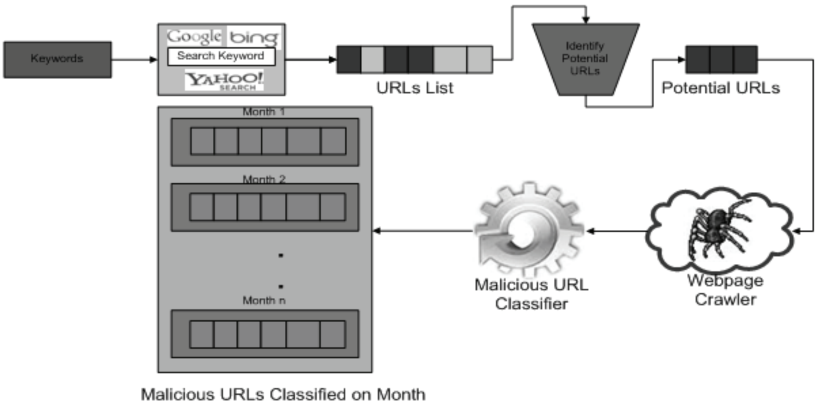


Figure 2: Process of collection and classification of malicious webpages

After collecting the malicious webpages, they are classified based on the month of a year in which they were reported. At the end of this process, we have a good collection of the malicious webpages distributed across different months of a year based on their reported date.

4.2 Construction of Inlink structure for malicious webpages

After the collection and classification of malicious webpages, the inlinks for these webpages are explored. Majority of the URLs (Uniform Resource Locator) listed on the websites publishing malicious webpages were identified to be injected into compromised legitimate websites. To make our collection exhaustive and more appropriate for performing topic modeling the inlinks for all the webpages collected from the potential sources are searched. Figure 3 is the pictorial representation of the process of the construction of inlink structure for the malicious webpages.

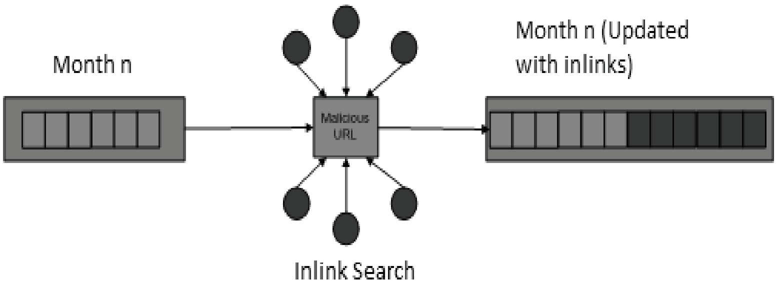


Figure 3: Process of construction of inlink structure for malicious webpages

We used Site Explorer Inbound Links API (Application Programming Interface) (2011) for searching the inlinks. The extracted inlinks are stored and are given the same date as that of the malicious webpage. At the end of this process, we have a good collection of malicious webpages and the inlinks associated with the malicious webpages distributed across different months of a year.

4.3 Crawling and content extraction

After the construction of the inlink structure for malicious webpages, all the unique URLs collected are extracted and are used as seeds for the process of crawling. A custom crawler was built for crawling and storing the response. A custom response analyzer was built to analyze the response headers and determine the status of the requested or crawled URL. The response analyzer is used to analyze the HTTP (Hypertext Transfer Protocol) status code in the response header. We analyzed the response header for status codes 404 and 403 and removed such URLs from our dataset to avoid further processing and the remaining URLs are passed on to the content extraction module for further processing.

The stored code for each of the URL under study is rendered in the browser to extract the content of the webpage. The code is executed using HTMLUnit (2010). HTMLUnit is an emulator for web browser which executes the HTML (HyperText Markup Language) and JavaScript but does not render the webpage visually. On executing the code in HTMLUnit, the content of the executed webpage is extracted and is stored.

Traditional parsing mechanisms fail to extract the content which is added dynamically through JavaScript and would not trigger some dynamic actions like redirection which fails to return the actual content rendered by the webpage. Usage of HTMLUnit helped in overcoming these limitations with a minimal effort. The pictorial representation of the process of crawling and extracting the content is shown in Figure 4. At the end of this process, we have the extracted content stored based on the month of a particular year.

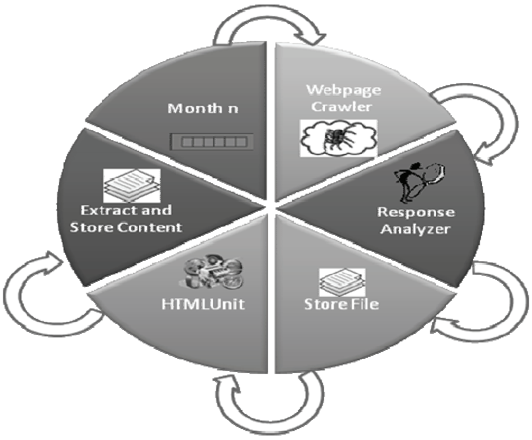


Figure 4: Process of crawling and content extraction

4.4 Perform topic modeling on the extracted content

After extracting the content from all the identified URLs, topic modeling is performed on the corpus. We use Machine Learning for LanguagE Toolkit (MALLET) for performing the topic modeling, implemented by McCallum and Kachites (2002), based on Gibbs sampling (described in Section 3). Figure 5 shows the pictorial representation of the process of topic modeling.

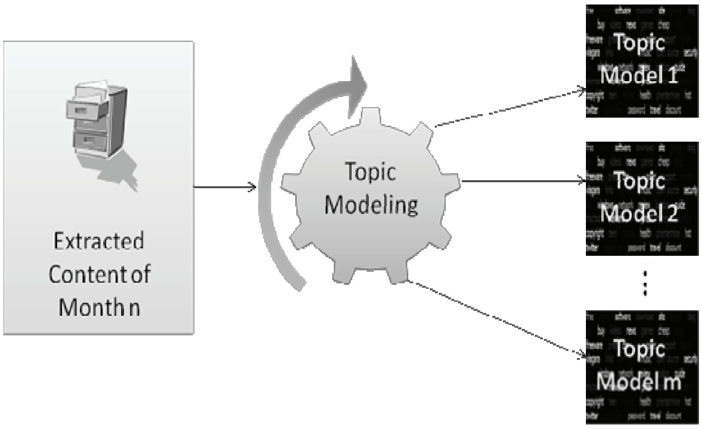


Figure 5: Process of performing topic modeling on the extracted content of a month of a year

5. Dataset

For our experiment, all the malicious webpages listed in the potential sources identified through meta-searches were extracted and classified based on the month of their reporting date. All the malicious webpages from the month of August to December of 2010 with both months inclusive were analyzed for our study. The inlinks for all the malicious webpages within our selection were extracted using the Yahoo Site Explorer Inbound Links API. All the identified URLs were crawled, emulated and stored the extracted content. We then performed topic modeling using the MALLET on the resulted set of documents. In order to have the topics neither coarse nor fine, we chose the number of topics to be 100 and the number of iterations to be 1000 for accuracy.

Table 1 gives an overview of the dataset used for this experiment. “Seed URLs” row represents the total number of malicious webpages identified from our data source across each of the five months considered for the study. “Total URLs” row represents the total number of webpages that were identified from the data source and on performing the inlink search. “Live URLs” row represent the total number of webpages that were live or active at the time of the experiment. The identified “Seed URLs” across various months indicate the growth of the number of malicious webpages with time.

Table 1: Overview of the dataset

	August	September	October	November	December
Seed URLs	272	280	399	717	1476
Total URLs	771	1204	2248	1722	9866
Live URLs	628	976	1962	1476	8829

6. Results

6.1 Similarity between Google trends and targeted topics

In this section, we analyzed the similarity between the Google Trends and the topics that were obtained on performing the topic modeling on the collected dataset. Google trends list the top 20 searches per each day. We manually extracted the Google trends for every day within the period considered for this study. We considered a similarity between the Google Trends and identified topics when more than 50% of the words in the Google Trends appeared in the identified words of atleast one topic. The similarity was measured by considering the top 50 words per topic and is shown in Table 3. The average correlation by considering the top 50 words across the period of five months was 61.4%. Thus about 61% of the malicious webpages can be detected proactively by monitoring the web based on the Google Trends. The probable reason for the attackers choosing such topics is to get huge number of victims within a short period of time (i.e. before they get detected).

Table 2: Describes the contribution of the top 50 words towards the correlation with Google Trends across each month

	Top 50	**%Top 50**
August	322	56.8
September	367	63.7
October	356	61.2
November	308	58.2
December	374	61

6.2 Distribution of the webpages across the topics

In this section, we analyzed the number of webpages covered by considering the top 10, 25 and 50 topics respectively across each month. This helps in getting an idea about the number of topics that are to be monitored for getting an effective detection rate. The number of webpages related to each topic was computed and then the top 10 or 25 or 50 topics were extracted based on the category under evaluation. After selecting the topics, all the unique webpages for the selected topics were considered and recorded the count of such number of webpages.

Table 3: Describes the contribution of the top 10, 25 and 50 topics towards the total number of webpages across each month

	Top 10	**%Top 10**	**Top 25**	**%Top 25**	**Top 50**	**%Top 50**
August	377	60	519	82.6	596	94.9
September	569	58.3	812	83.2	935	95.8
October	1137	58	1544	78.7	1806	92
November	833	56.4	1172	79.4	1397	94.6
December	5139	58.2	6956	78.8	8210	93

From Figure 6, it is evident that the percentage of the number of webpages related to the top ‘n’ topics with ‘n’ varying from 1 to 100 are almost the same. The growth in the contribution towards number of webpages is exponential initially and is saturated on crossing the midway along the X-axis. From Table 4, it is evident that the top 50 topics cover about an average of 93% of the total number of webpages. Therefore considering the top 50 topics that were modeled across each month would give an idea of the topics under attack and help in the detection of malicious webpages proactively.

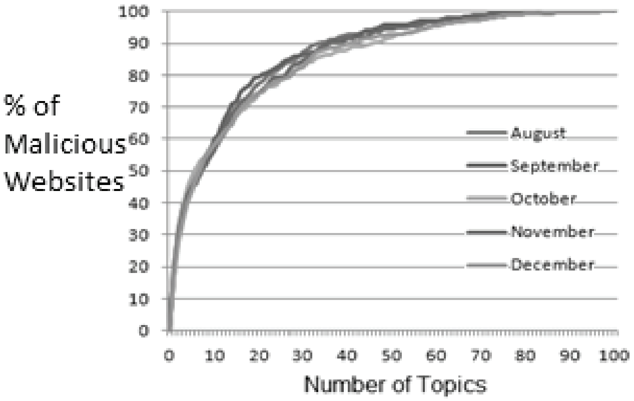


Figure 6: Graph showing the percentage of malicious webpages across each month from the top 'n' topics with 'n' varying from 1 to 100.

6.3 Top 25 Topics based on the number of webpages that were similar with Google trends

In this section, we depicted the topics that were targeted across each month of the period considered for the study. We considered a Google Trend to be similar, if we identified that more than 50% of the words of the Google Trend matched with the words across a topic of the respective month. The number of unique webpages associated with each Google trend was identified based on the modeled topics that were similar with the corresponding Google trend. The top 25 Google Trends based on the number of webpages related to it were identified for each month in the range of our study. These topics were identified to analyze the prominent topics that are targeted by the attackers.

Table 4: Top 25 topics, based on the number of malicious websites, which were similar to Google trends across each month within our study

August	September	October	November	December
dr laura n word	google instant	iq test for free	friday night lights j cole download	prisonplanet.com censored
free bilderberg films	google instant not working	dayna kempson video	black friday online deals	nfl network
florida department of education	watch nfl games online free	usa network ryder cup	girl talk all day download	lebron james south beach home
hgtv design star	connecticut home invasion	willow smith whip my hair music video	toys r us free shipping code	facebook cartoon profile pic news
paul jr designs	blu homes	ryder cup 2010 tv schedule	voting locations by zip code	watch the ball drop online
presidential address	facebook login problems	today show halloween costumes 2010	swiss legend men s commander collection watch	aj burnett divorce
kanye west power video	us open tennis live streaming free	government licensing internet	cam newton suspended	watch lunar eclipse online
love the way you lie music video	america s got talent top 4	airtel call home	macy s thanksgiving day parade 2010	school board shooting full video
bilderberg wide open	watch monday night football online	who owns my heart video	denair middle school	elizabeth edwards latest news
rashard mendenhall	delaware election results	erin andrews peephole video	top 50 most popular women on the web	convert utc time to local time
brandon spikes video	christopher plummer	tj lavin crash video	spiderman the musical	google colors
prelude to a kiss	sharktopus	world series game 1	the golden compass	google ebooks
fred savage dead	zabasearch	halloween music	carrie underwood	free shipping

August	September	October	November	December
		free	wedding	friday
the human centipede	apple tv	texas rangers world series tickets	2010 breeders cup results	metrodome collapse video
mine taylor swift video	facebook dns error	world series 2010	philadelphia marathon results 2010	morgan freeman dead
bill cosby dead	what is wrong with facebook today	rocky horror picture show	greg oden	is that so wrong julianne hough official video
james wilder jr	nfl player found dead	code maroon	censure definition	miley cyrus bong video
madison de la garza	facebook outage	columbus day parade nyc 2010	dengue fever	free coupons
de havilland dhc 3t	why is facebook down	ryder cup results	kat denning	google evil
lance cade	lone star tv show	ctu online login	keith olbermann suspended	rex ryan foot video
talladega nights	my generation tv show	oliver wendell holmes	what time does zenyatta race today	college football bowl games
drake miss me video	wbal tv	watch ufc 121 online	firebreather cartoon network	1999 time magazine person of the century
den brother	national coffee day 2010	christine o donnell debate	bupers online	restaurants open on christmas day
derek lee	steve martini book list	sweetest day 2010	matt hughes storm chaser death	is walmart open on christmas 2010
elin nordegren	blackberry playbook	daylight savings time 2010	cyber monday deals 2010	crystal harris playboy photos

The top 25 topics across each month of our study are listed in Table 5. The topics targeted by the attackers for the propagation of malware go by the trends or events with large concentration among the people. We will explain few topics targeted by the attackers in the period of our study. In August, the attackers targeted websites related to the controversy of “drlaura n word” which stirred the community then, and was related to racism where Dr.Laura called the ‘n’ word for 11 times over the radio. In September, the focus was on the Google Instant search which was launched in the month of September. In addition, the attackers also targeted the National Football League (NFL) and US Open viewers by hosting malicious content on webpages with the context of free links for watching these tournaments online. In the month of October the attackers targeted the Ryder Cup, a golf tournament which started in October and also the ‘Whip my Hair’ album from Willow Smith, an American child actress and singer. In November, the attackers targeted the black Friday deals and the free online links for viewing the Macy’s parade. In the month of December, the attackers targeted the topic of cartoon pictures for Facebook profiles which created waves in the end of November and the beginning of December. The topics that were targeted by the attackers give a clear understanding of the attack vector for attacking legitimate websites or hosting malicious content. The attackers are targeting the topics that attracts huge crowds of people like sports, public events etc. Thus crawling up for websites related to the events that draws the interest of large number of people and analyzing them would help in the detection of malicious websites proactively and efficiently(within less time).

6.4 Top 50 words targeted in the period of our study

In this section, we identified the words that were targeted by the attackers consistently over the period of our study and computed their importance by measuring their frequency across the whole period. We then identified the top 50 words and listed them in the Table 6. The picture gives us an idea about the topics that the attackers are always after. The words like free, nice, cheap, great, hot gave an idea that the attackers hosted malicious content on websites which attracts people interested in online shopping and those who look for deals over the internet. Couple of interesting terms identified in this study was ‘malware’ and ‘phentermine’. The occurrence of the term malware consistently across the period of our study showed the intent of the attackers about hosting malicious content on websites that are related to malware. Phentermine (2011) is an appetite suppressant to help in reducing weight. This term gave us an idea that attackers targeted topics specifically over a medical term rather than targeting a generalized term like diet, supplement etc. Therefore, we need to extract the broader context of the terms targeted by the attackers to get better results.

Table 5: Displays the top 50 words that were targeted consistently by the attackers across the 5 months considered for our study

1. free
2. online
3. software
4. download
5. site
6. google
7. blog
8. sex
9. buy
10. videos
11. news
12. games
13. cheap
14. shop
15. freeware
16. great
17. car
18. find
19. poker
20. support
21. forum
22. viagara
23. links
24. comment
25. music
26. open source
27. security
28. live
29. windows
30. network
31. review
32. location
33. guide
34. microsoft
35. malware
36. nice
37. price
38. money
39. insurance
40. copyright
41. teen
42. movies
43. health
44.phentermine
45. hot
46. twitter
47. rapidshare
48. password
49. travel
50. discount

7. Conclusions

In this paper we present a novel approach for performing topic modeling on the malicious websites to identify the attack vectors used by the cyber criminals with respect to the topics of the websites targeted by them.

We computed the similarity of the Google Trends with the topics targeted by the attackers within the period of our study and identified that about 61% of the topics that are targeted by the cybercriminals were similar with the Google Trends.

We identified the distribution of the webpages across the top 'n' topics with 'n' varying from 1 to 100 and identified that the number of webpages grew exponentially with 'n' up to 50 and saturated with 'n' above 50. Thus the top 50 topics proved to be effective in understanding the topics that were targeted by the cybercriminals.

We recognized the top 25 topics that correlated with the Google Trends and contributed the most number of webpages across each individual month within our study. These topics suggested that the cybercriminals targeted topics that are related to events or controversies with huge crowd attention.

We computed the top 50 words that were used consistently across all the months of our study and ordered them based on their frequency of occurrence. These words suggested the topics that were targeted by the cybercriminals irrespective of the period or month.

8. Future work

We would like to extend this study with time gap analysis, with respect to the occurrence of the trend and the malicious websites showing up related to the trends. In future, we also want to perform the similarity of topics targeted by the attackers with the past trends to get an idea about the topics targeted based on the past trends and the number of websites infected with respect to a trend across different geographic locations.

References

"Crimepack: Packed with Hard Lessons." (2010), Krebson Security [Online], Available: http://krebsonsecurity.com/2010/08/crimepack-packed-with-hard-lessons/#more-4340 [06 Jan 2012].

Bleaken, D. (2010) "Use of legitimate sites in malicious web attacks." Symantec [Online], Available:http://www.symantec.com/connect/blogs/use-legitimate-sites-malicious-web-attacks [23 Mar 2012].

Cova, M., Kruegel, C., and Vigna, G. (2010) "Detection and Analysis of Drive-by Download Attacks and Malicious JavaScript Code", *WWW 2010 - 19th International World Wide Web Conference*, North Carolina (USA).

Griffiths, T.L., and Steyvers, M. (2004) "Finding Scientific Topics", National Academy of Sciences, 101 (suppl. 1), 5228-5235.

"HtmlUnit." (2010), Gargoyle Software Inc. [Online], Available: http://sourceforge.net/projects/htmlunit/files/ [28 Dec 2010].

Kalafut, A.J., Shue, C.A. and Gupta, M. (2010) "Malicious Hubs: Detecting Abnormally Malicious Autonomous Systems", *IEEE Infocom Mini-Conference*, California (USA).

McCallum and Kachites, A. (2002) "MALLET: A Machine Learning for Language Toolkit", http://mallet.cs.umass.edu.

Paget, F. (2010) “An Overview of Exploit Packs.” McAfee [Online], Available:http://blogs.mcafee.com/mcafee-labs/an-overview-of-exploit-packs [23 Mar 2012].

“Phentermine.” (2011), Wikipedia [Online], Avaialble: http://en.wikipedia.org/wiki/Phentermine [14 Dec 2011].

Rieck, K., Krueger, T., and Dewald, A. (2010) “CUJO: Efficient Detection and Prevention of Drive-by-download Attacks”, Annual Computer Security Applications Conference (ACSAC), Texas (USA).

“Site Explorer Inbound Links API.” (2011), Yahoo Developer Network [Online]. Available: http://developer.yahoo.com/search/siteexplorer/V1/inlinkData.html [15 Jan 2011].

Steyvers, M., and Griffiths, T. (2007) “Probabilistic Topic Models” , T.Landauer, D. McNamara, S. Dennis, & W. Kintsch (Eds.), Latent semantic analysis: A road to meaning, Mahwah, NJ (USA).

Evaluation of Nation-State Level Botnet Mitigation Strategies Using DEMATEL

Christian Czosseck
Cooperative Cyber Defence Centre of Excellence, Tallinn, Estonia
christian.czosseck@ccdcoe.org

Abstract: Botnets have been recognised as a possible threat to national security, and over recent years national cyber security thinkers have started to draft national level strategies to reduce the threat posed. The steady increase in the number of infected machines and the damage caused by botnet-mounted attacks shows that the success so far has been limited. This research analyses nation-state and inter-state level botnet defence and mitigation strategies and ultimately evaluates their impact on the botnet threat by employing the *Decision-Making Trial and Evaluation Laboratory* (DEMATEL) method on empirical data collected via interviews from experts in the field. This paper develops and presents a system of nation-state level strategy groups and a simple model of effects they might have on the botnet threat. Based on this framework, the reciprocal influence of each element pair is identified, with the help of knowledgeable experts, and serves as the basis to conduct an analysis utilising the DEMATEL method. As a result we present a model of the influence that these strategy groups have on the botnet threat, identify strongly and weakly influential elements in this system and present a ranking based on these findings. This will lead to a recommendation as to which is the preferred strategy.

Keywords: botnets, DEMATEL, cyber defence, strategy evaluation

1. Introduction

Over recent years, cyber-crimes, and with them botnets in their sheer unlimited ways of usage, have risen from a primarily cyber-crime issue to nation-state security concerns. Besides“classical” spam and DDoS campaigns, executed for monetary gain, politically-motivated cyber-attacks, with botnets as their preferred means, are on a steady rise (Nazario, 2009). Users of botnets range from individuals and other organised groups up to nation-states(Ottis, 2010). Czosseck and Podins (2011) offer a generic taxonomy of users and usages of botnets based on recent history.

Under the lasting impression of the cyber-attacks against Estonia back in 2007, nation-states all over the world started to seriously recognise the cyber domain in a nation-state security context,and with it the newly emerging threats including botnet-mounted attacks. The increasing number of dedicated national cyber-security strategies reflects this.

These nations have their national interests challenged on multiple frontlines. Cyber-crime is now an organised, highly professionalised business offering well-developed exploits and malware to anyone willing to pay, ultimately hurting the economy on a large scale. As an example, the damage of cyber-crime to the UK economy is estimated to be £27bn per annum (Cabinet Office UK & Detica Ltd., 2011).Besides this, the very same technology and knowledge about vulnerabilities is sold via lawful channels, especially in the context of services dedicated to nation-state customers such as law enforcement, military or intelligence services(GTISC & GTRI, 2011).

The responses to this development are manifold and reach from technical solutions at one end to governmental actions on both national and international scales at the other.

A system of strategic options available to a nation-state actor is introduced in Section 2, followed by simplified effect model in Section 3 of this paper. As an empirical basis for further analysis, knowledgeable experts are interviewed and their answers analysed with DEMATEL in Section 4, motivating the discussion and conclusions of this paper.

2. Nation-state level botnet mitigation strategies

This research focuses on the nation-state level strategies understood as those instruments normally initiated, introduced or supported by governments, either because they have the unique authority to do so or they are in the position to facilitate it on a nation-state scale. Examples include, but are not limited to, state policies, changes to national legal frameworks or international collaboration e.g. in the framework of existing international organisations.

In the following research system of 10strategy groups of similar strategies is developed, motivated by existing practice and academic research. In this context, similar means different strategies which result in a similar effect, target similar stakeholders or use similar methods. They are encouraged by the findings of the ENSIA Botnet Study (Plohmann *et al*., 2011) and other studies, especially those by Dunn (2005) and Eeten *et al*. (2010), as well as the analysis of current national strategies, or they reflect examples of existing or academically discussed actions taken.

2.1 Promotion of dedicated and coordinated R&DPrograms

For cyber security to be developed effectively on a nation-state level, specialised and coordinated research becomes crucial. As Dunn (2005) argues, the fundamental issue and major challenge is the interdisciplinary nature of the research that needs to be conducted. While the existing research on IT-security is mainly technical by nature, this is not seen as enough to cover all aspects of the complex systems on hand, requiring a "*holistic and strategic threat and risk assessment at the physical, virtual and psychological level*". Anderson *et al*. (2008) argues the same.

This group of strategies reflects approaches such as the development or promotion of nation-state research agendas, the support of these with special grants or programmes, or the development of new/specialised curricula. In particular, by using instruments existing to govern the higher education system present in most nations, the availability of a specialised workforce can be positively affected.

2.2 Improvement of international law enforcement

Organised cyber-crime can be seen as the current root cause of the existing botnet threat. As the use of botnets for mostly criminal purposes is a highly lucrative business, a highly professionalised "underground economy" emerged. Botnet masters are highly motivated to make their investment resilient against take downs and are actively exploiting grey areas in international legal frameworks, missing cooperation between nations or bullet-proof hosting opportunities. Taking down botnets nowadays mostly implies an internationally-coordinated, timely effort between different law enforcement groups.

This group of strategies includes examples such as the Council of Europe Cybercrime Convention(Council of Europe, 2001),showing that international treaties are one way to mitigate obstacles in international law enforcement cooperation. Agreements between nation-states, legislation or regulations within supranational organisations such as the EU, or commitments of or recommendations to nations under the umbrella of international organisations are other possibilities to harmonise the legal frameworks. This group includes actions to address the "IP addresses are private data" issue in the EU, exploring/implementing exceptions to criminal offences for certain stakeholders in order to ease the legal risks of becoming active (e.g. the "Good Samaritan Law"),or exceptions from privacy concerns for IP exchange or reverse engineering of malware (breach of license issue), as encouraged in Plohmann *et al*. (2011).

2.3 End-user notification, support and good-behaviour incentives

As is commonly known and reflected in the findings of Eeten *et al*. (2010),infected end-users represent the largest part of the botnet population. There are many reasons for this, including the lack of general IT-security awareness, the use of stolen (and sometimes manipulated) software or general weaknesses to social engineering-based attacks. Often end-users are not aware of the infection or are not capable or willing to disinfect.

To support end-users, a proper notification is required in the first place. Additionally, (negative) incentives for self-cleaning, such as the introduction of walled gardens on an ISP level,are ways to increase pressure on end-users to encourage good practices. As these means are not available for free and also pose the concrete risk of alienating customers, ISPs are often reluctant to implement such means.

This group of strategies firstly covers activities aimed at establishing a system for notifying end-users about a present infection, and to help them in the process of clearing the infection from their systems. The Cyber Clean Centre in Japan(CCC, 2011)and the German Anti-Botnet-Advisory Centre(ECO, 2011) are examples of initiatives in which a joint private/public effort was made to assist end-users.

Secondly, governmental (e.g. legal) and successively ISP-based instruments to encourage good behaviour are included in this group. A government could support or enforce implementation of “walled gardens ”by introducing appropriate laws. Another method is the introduction of national acts penalising end-user misbehaviour, threatening them with e.g. disconnection from the Internet and suspension of Internet usage for a longer period of time. This has been possible in France since 2009(BBC, 2011),and is ultimately based on new EU legislation(European Comission, 2007).

2.4 ISP obligations and incentives to act

The empirical data presented and analysed in Eeten *et al.*(2010) highlights the central role ISPs are playing in the mitigation of botnets and their effects. They find that within the extended OECD, approximately 200 ISPs are covering about 80% of all Internet users, so governments are in a good position to tackle the problem by speaking to a relatively small group. However, they also identified that ISPs differ significantly with regards to the botnet activity within them, reflecting different security means applied by them.

This group of strategies reflects actions taken by nation-states to encourage ISPs to implement means and processes to pre-emptively mitigate botnet infections or their actions (beyond activities that directly target end-users). Service-based means, such as blocking port 25 as default for all retail customers, the implementation of network traffic monitoring and controlling, automatic botnet mitigation technology, as suggested by Asghari (2010), or the monitoring of traffic to well-known C&C servers are examples of additional actions which could be taken. There are various mechanisms for these, ranging from financial support or loans to active negotiations or regulations enforced by ISPs.

2.5 Awareness campaigns

Considering the increasing complexity of IT security threats and their mitigation solutions, it might be safe to assume that the general IT security awareness and the level of good behaviour is not on a adequate level to face the problem. Raising security awareness at all levels of society and explaining as well as encouraging the civic responsibility of everyone was identified as key by, for example, Plohmann *et al.*(2011). While related, this is different from end-user notification as it is a protective measure, helping to prevent an infection in the first place.

This group of strategies represents those measures taken to raise general public awareness on a broad and continuous basis, similar to campaigns for AIDS, smoking or drugs. This might include information portals dedicated to the user, as are present in many countries such as Germany(ECO, 2011) and Japan(CCC, 2011), but assumes that a substantial effort is made to reach citizens via multiple communication channels or media. Other ideas include obligatory classes in elementary and high-school or free/subsidised night courses.

2.6 Development of over-arching nation-state cyber security strategies

The mid-1990s shift from seeing information infrastructures primarily as a tool for getting a competitive advantage (especially in the business world) towards recognising national dependency on information infrastructures as a nation-state interest ultimately brought the protection of (critical) information infrastructures (CII) on the agenda of security policy, as elaborated by Dunn (2005).With the cyber-attacks against Estonia in 2007 and successive major cyber-related incidents and the emerging threat posed by hacktivism (Denning, 2001; Ottis, 2010), cyber security as an “extension” of CII protection emerged, forcing nations to re-evaluate their national security frameworks. Estonia was among the first to develop a dedicated national cyber security strategy, implemented in 2008, in response to the attacks suffered, making changes to their legal, organisational and strategic framework (Czosseck *et al.,* 2011).Many more countries followed suit, including Austria and Great Britain in 2009, Canada and Japan in 2010and France and Germany in 2011.

As such, this group of strategies reflects the process of developing and implementing a dedicated nation-state cyber security strategy or policy, and its subsequent reorganisation of nation-state responsibilities and authorities. This might include forming or further empowering specialised public bodies such as national CERTs, inter-ministry coordination centres like the German National Cyber Defence Centre, or centralisation of authority as in the Department of Homeland Security in USA.

2.7 Promotion and support of botnet hunting initiatives

A myriad of actors are currently investigating botnets, trying to monitor or infiltrate them, or to mitigate their effects on themselves or others. Botnets are an experimental subject of research and the basis for many business ideas.

Taking Microsoft's Digital Crime Unit, *a "worldwide team of lawyers, investigators, technical analysts and other specialists whose mission is to make the Internet safer and more secure through strong enforcement, global partnerships, policy and technology solutions"*(Microsoft, 2011) serves as an example of a private sector initiative to coordinate actions against botnets. While public bodies are involved (especially in law enforcement) they do not play a leading role.

This group of strategies is understood as those active efforts of a nation-state to encourage, facilitate and perhaps even financially support similar initiatives dedicated to providing intelligence on or taking down botnets. This could include establishing devoted public bodies or points of contact to become part of such a cooperation, (financially) supporting intelligence efforts necessary to gather take-down information, or maybe even issuing a "bounty".

2.8 Software developers' obligations or incentives

Developing software was always prone to (exploitable) bugs and flaws in the concept. While extensive research has been carried out on how to develop secure software, the reality shows that there is still a long way to go. There are many reasons for this, but economically-motivated time-pressure and a lack of security-aware programmers might be two of the more dominant issues involved. While different international IT security evaluation and accreditation frameworks, like Common Criteria (CCRA 2012) exist, they are not obligatory for any software to be sold, and even less so for software made available for free/open source.

This group of strategies includes efforts to increase the pressure on software developers to produce more secure (proven) code. There are a variety of instruments available, starting with the promotion of standards(such as Common Criteria), and successively the requirement for certified compliance with standards in public procurement. This might include the obligation to clearly indicate compliance to customers. Another method is the introduction of liability obligations to software developers, e.g. a mandate/incentive for software developers to release security patches to all users, including those using illegal copies, or the responsibility for disclosure and fast patching(Anderson *et al.*, 2008).

2.9 Obligation of cyber insurances

Another idea circulated among scholars for some time, and received a greater jolt of governmental encouragement in 2002; Richard Clarke, the former cyber advisor to the Bush administration, met with insurance companies in the US to lobby for the coverage of cyber-based risks by them (Risen 2010).While initial estimations for the development of this market were over-optimistic, it was estimated to cost around 0.5 billion USD in 2010 (Risen, 2010);a market emerged providing different types of coverage ranging from breaches of data, regulatory civil action, cyber extortion, virus liability and many more as presented by Wood (2011).

This group of strategies reflects the state-driven encouragement of cyber insurances and the possible introduction of the obligation to be insured. This obligation could especially aim for key industries that are identified by a nation-state as critical.

2.10 National or international partnership programmes and information exchange

It is generally accepted that botnets have become a global issue, and that the instruments for fighting them are mainly in the hands of private sector. Nonetheless, with state-sponsored espionage and already evolving military cyber capabilities, the role of the nation-states increases.

Private-public partnership programmes have been identified as key by many nation-states, as well as the need for collaboration between key stakeholders. They might provide a platform for consultation, cooperation and information exchange, a starting point to initiate and later facilitate joined initiatives or reduce tension from competitive market participants so that they jointly introduce measures seen as unpleasant for their customer base.

This group of strategies represents the active promotion of, participation in and contribution to national and international partnership programmes. Examples of this can be seen in the Australian Internet Security Initiative, established in 2006 (ACMA, 2005) or the Dutch anti-botnet MoU between ISPs, signed in 2007 (Evron, 2009). In an international context, the European Public-Private Partnership for Resilience Programme(European Commission, 2010)and the London Action Plan (L.A.P., 2005) serve as examples.

3. Shortfalls, threats or missing capabilities

In this section, a selection of key problems of the botnet threat is presented. They build upon the findings of the sources introduced in the last section. They are either existing shortfalls, missing capabilities or alternatively they are threats or existing problems. In both cases it would be "positive" if they are reduced by the strategies presented, and it would be "negative" if they are increased. They are the following:

- (Improving) **detection, monitoring and tracking of botnets** As Plohmann *et al*. (2011) identified, it is still assumed that many botnets are not detected at all and the existing methods to survey identified botnets are not sufficiently developed. As such,an improvement in this category will lead to better situation awareness, ultimately enabling more precise actions to be taken against botnets. In addition, the difficult problem of (technical) attribution might be reduced.
- (Reducing the) **existing botnet population** A strategy might have a direct influence on the existing population, leading to clean ups or at least the unavailability of these zombies for their bot master.
- (Reducing the risk of) **new infections and migration to new victim platforms** Some of the strategies might have a preventative influence, raising the bar for bot masters who want to launch new or further spread existing botnets.
- (Reducing profitability of the) **cyber-crime economy behind botnets** One major driving factor behind the current botnet issue is the fact that cyber-crime became highly profitable. A strategy might have a deterrent effect on people entering this "business", reduce the profit made or raise the arms race between good guys and bad guys to a level where the outcome is no longer worth the effort.
- (Reducing/deterring the) **botnet usageby APT or state sponsored espionage/CNO** Advanced Persistent Threat (APT) actors are increasingly reported as taking advantage of the existing botnet population, querying them for coincidentally infected zombies in or close to the target of their interest (GTISC & GTRI, 2011).Furthermore, there are an increasing number of cases where cyber-attacks are launched by actors who do not have a direct monetary interest but rather a state-driven political goal they wish to achieve. A strategy might have a deterrent effect on this type of botnet usage both now and in the future.
- (Reducing/deterring the)**botnetusage byhacktivism** Similar to state-sponsored activities, political goals are the driving factor behind hacktivism and have been introduced by e.g.Denning (2001). Similarly, strategies might have a deterrent effect on this group of people, who think and act slightly differently from"ordinary" criminals.
- (Inhibiting the)**development and proliferation of botnet technology** Botnet developers and the stakeholders trying to fight them have already entered a (mostly) technological arms race. Additionally, "botnets as a service" enables basically everyone willing to pay to get his hand on a botnet, dramatically increasing the access to them for everybody. The strategies presented in this article might have an influence on this proliferation.

4. DEMATEL analysis of the empirical data

Between 1972 and 1976, the Science and Human Affairs Program of the Battelle Memorial Institute of Geneva developed the *Decision-Making Trial and Evaluation Laboratory* (DEMATEL) method to research and solve clusters of complicated and intertwined problem groups called *problematiques.* Based on graph theory, problems can be planned and solved visually, dividing relevant factors into cause and effect groups to better understand causal relationships (Li & Tzeng, 2009). It has be successfully applied in different domains such as knowledge management (Wu, 2008), policy impact on SMEs (Shyu, 2008), financial investment strategies(Lee, 2009) or strategic cyber security (Geers, 2011). The method is constantly extended and combined with other methods such as the maximum mean de-entropy algorithm (Li & Tzeng, 2009), fuzzy logic approaches (Lin & Wu, 2008; Tzeng *et al.*, 2009) or causal loops (Jafari *et al.*, 2008). The four steps of the original DEMATEL method are: "(1)

calculate the average matrix, (2) calculate the normalized initial direct-influence matrix, (3) derive the total relation matrix, and (4) set a threshold value and obtain the impact-relations map", and are explained in detail in (Li &Tzeng, 2009).

4.1 Input matrix

Based on the 10 strategy groups and seven effects presented in the earlier section, a questionnaire was developed and sent to experts in the field. Every strategy group and effect was pair-wise compared, and the interviewees were asked to assess the influence one has on the other on a scale from 0 to 3 with the latter being the greatest influence. (The DEMATEL method allows for any positive number as input, but scales from 0 – 3 are very common.) In total 11 interviewees responded, covering seven countries and representing technical, strategic and legal viewpoints. Their individual answers were combined by calculating the average for every individual answer. This resulted in the initial input matrix presented as Table 1. In the following tables, strategies are abbreviated by S1 – S10 and effects by E1 – E7, matching the corresponding sub-section numbers.

Table 1: Average input matrix T, based on questionnaire

	S1 R&D Programs	S2 Intern. law enforcement	S3 End User	S4 ISP obligations &	S5 Awareness campaigns	S6 Cyber security	S7 Botnet Hunting	S8 SW Developers'	S9 Cyber Insurances	S10 Partnership	E1 Detection of Botnets	E2 existing population	E3 new infections	E4 Cyber Crime Economy	E5 APT/state-sponsored	E6 Hacktivism botnet	E7 Technology proliferation	**Total Influence**
S1	0,0	0,4	0,9	0,8	1,3	1,5	2,0	0,8	0,5	2,0	2,3	1,4	1,0	0,5	0,4	0,5	0,7	**17**
S2	0,5	0,0	0,7	1,9	1,0	1,5	1,3	0,5	0,7	2,0	1,2	1,3	0,8	1,9	0,5	1,5	0,4	**18**
S3	0,5	0,7	0,0	1,9	2,5	1,3	1,0	0,7	0,9	1,6	1,3	2,1	1,3	1,2	0,2	0,6	0,5	**18**
S4	1,0	1,3	2,5	0,0	2,4	1,6	1,4	0,7	1,4	2,0	2,7	2,3	1,8	1,5	0,6	1,0	1,2	**25**
S5	0,8	0,7	2,2	0,8	0,0	1,3	1,2	0,5	0,6	0,6	0,6	1,4	1,5	0,9	0,1	0,4	0,3	**14**
S6	2,4	2,1	1,0	1,3	1,9	0,0	1,3	0,9	0,6	2,3	1,3	1,0	0,6	0,8	0,7	0,9	0,5	**20**
S7	2,1	1,6	0,9	1,3	1,1	1,1	0,0	0,3	0,2	1,9	2,9	2,0	1,0	1,4	0,8	0,9	1,2	**21**
S8	1,5	0,2	0,8	0,6	0,9	0,4	0,3	0,0	1,1	0,8	0,4	1,3	1,7	0,6	0,5	0,4	0,8	**12**
S9	0,9	0,4	1,4	1,3	1,2	0,8	0,6	1,2	0,0	0,6	1,0	0,7	0,7	0,4	0,2	0,4	0,2	**12**
S10	1,6	1,9	1,6	1,2	1,2	1,5	1,5	0,4	0,2	0,0	2,0	1,3	1,1	0,6	0,5	0,4	0,5	**18**
E1	1,6	1,5	1,1	1,1	0,8	0,7	2,1	0,2	0,3	2,1	0,0	1,8	1,1	1,2	0,6	0,9	0,8	**18**
E2	1,0	1,2	0,9	0,8	1,3	0,7	1,7	0,4	0,3	1,6	2,1	0,0	1,8	1,9	0,8	1,3	1,0	**19**
E3	1,2	0,3	0,7	0,6	1,0	0,8	0,9	0,4	0,2	1,0	1,6	1,9	0,0	1,4	0,6	0,6	1,1	**14**
E4	0,6	1,3	0,3	0,4	0,4	0,9	1,0	0,2	0,2	1,3	1,4	2,2	1,5	0,0	0,1	0,1	1,3	**13**
E5	0,9	1,0	0,1	0,3	0,3	1,8	0,7	0,1	0,2	0,9	0,7	0,5	0,7	0,4	0,0	0,6	0,4	**10**
E6	0,5	1,2	0,4	0,6	0,8	0,8	0,6	0,1	0,1	0,7	0,5	0,6	0,3	0,3	0,4	0,0	0,3	**8**
E7	0,9	0,5	0,6	0,7	0,4	0,7	1,2	0,7	0,1	1,1	1,6	1,8	2,0	1,0	0,6	0,5	0,0	**14**
Level influenced	**18**	**16**	**16**	**16**	**19**	**17**	**19**	**8**	**8**	**23**	**24**	**24**	**19**	**16**	**8**	**11**	**11**	

4.2 Direct influence analysis

The sum of the individual influence levels in each row is presented in the column “Direct influence”. This expresses the total influence a strategy or botnet effect has, and allows for a first ranking of all strategies, as presented in Table 2.

Table 2: Total influence ranking of strategy groups

Strategy Group	Total Influence
S4 ISP obligations & incentives	25
S7 Botnet hunting	21
S6 Cyber security strategies	20
S2 Intern. law enforcement	18
S3 End-user	18
S10 Partnership programmes	18
S1 R&D Programmes	17
S5 Awareness campaigns	14
S8 SW Developers’ obligations	12
S9 Cyber Insurances	12

ISP obligations & incentives being at the top is as expected, as are strategies targeting cyber insurance or software developers being at the very bottom. But it is surprising that broadly launched awareness campaigns also rank far below the average, with a score of 17.

In a similar manner, the empirical data allows for a ranking of the level of total influence the botnet threats receive, as presented in Table 3.

Table 3: Total influence level on botnet threat

Effect on botnet threat	Level influenced
E1 Detection of botnets	24
E2Existing population	24
E3 New infections	19
E4 Cyber Crime Economy	16
E6 Hacktivism botnet usage	11
E7 Technology proliferation	11
E5 APT/state-sponsored usage	8

It becomes evident that addressing the more technical aspects of the botnet threat, meaning its detection and disinfections, is highly influential. On the other side, addressing the users of these botnets appears to be more limited.

4.3 Indirect influence

As Figure 1 illustrates, a strategy group can affect the botnet treat directly, but also indirectly. In the latter case it has a direct influence on another element, which in turn has a direct influence on the final element. This leads to the insight that a presumably strong direct influence of a given strategy group could be the result of multiple indirect influences. Over time, each of these indirect influences impacts every other element of the system, including itself.

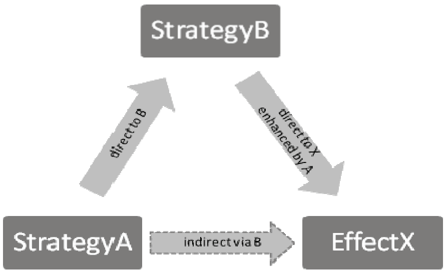


Figure 1: Indirect influence illustration

The DEMATEL method is capable of recognising this fact and can decompose direct and indirect influence by calculating the total relation matrix, Q = M x (**I** – M)$^{-1}$, where M is the normalised matrix of T (seeTable 1) and **I** is the identity matrix. Q is illustrated in Table 4.

Table 4: Total relation matrix, Q

	S1 R&D Programs	S2 Intern. law enforcement	S3 End User	S4 ISP obligations & incentives	S5 Awareness campaigns	S6 Cyber security strategies	S7 Botnet Hunting	S8 SW Developers' obligations	S9 Cyber Insurances	S10 Partnership programmes	E1 Detection of Botnets	E2 existing population	E3 new infections	E4 Cyber Crime Economy	E5 APT/state-sponsored usage	E6 Hacktivism botnet usage	E7 Technology proliferation	**Direct Influence**
S1	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**1,9**
S2	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**2,0**
S3	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**2,1**
S4	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**2,8**
S5	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**1,6**
S6	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**2,2**
S7	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**2,3**
S8	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**1,3**
S9	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**1,3**
S10	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**2,0**
E1	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**2,0**
E2	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**2,1**
E3	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**1,6**
E4	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**1,5**
E5	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**1,1**
E6	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**0,9**
E7	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	**1,6**
Indirect Influenc	**2,05**	**1,92**	**1,85**	**1,78**	**2,10**	**1,94**	**2,22**	**0,89**	**0,84**	**2,59**	**2,73**	**2,67**	**2,11**	**1,88**	**0,89**	**1,28**	**1,31**	

4.4 Findings based on the adjusted influence index

With the direct and indirect influence present, one can easily calculate the difference between the direct influences and the indirect influences that each single strategy group received. This "adjusted influence" expresses the remaining influence that a single strategy has on the system. Table 5 illustrates this.

Table 5: Remaining influence on the system per strategy group

Strategy Group	DirectInfluence	Indirect Influence	Adjusted Influence
S4 ISP obligations & incentives	2,85	1,78	1,07
S9 Cyber insurances	1,39	0,83	0,55
S8 SW developers' obligations	1,37	0,89	0,48
S6 Cyber security strategies	2,23	1,94	0,29
S3 End-user	2,11	1,85	0,25
S7 Botnet hunting	2,37	2,21	0,15
S2 Intern. law enforcement	2,02	1,91	0,11
S1 R&D programmes	1,98	2,05	-0,08
S5 Awareness campaigns	1,62	2,09	-0,48
S10 Partnership programmes	2,06	2,58	-0,53

The group "ISP obligations & incentives" again scores highest, confirming the initial observation and common assumption. What is more of a surprise is that the strategic groups *cyber insurances* and *software developers' obligations*, which were initially at the very end of the list, now rank second having about 50% less impact than the most influential group.

With about 25% of the impact of the most influential group, the groups *cyber security strategies* and *end-user obligations and good-behaviour* end in the third position, being ranking-wise the same, but in absolute terms are far lower in relation to the highest impact group.

The botnet hunting strategy group experiences another surprising rank change, now ending with only limited influence on the system, at around 10% of the best.

Looking at the end of the table, R&D programmes, awareness campaigns and partnership programmes do score negatively. This means that their initial assumed impact on the system is mainly a result of indirect influences by other elements of the system.

Table 6: Remaining influence on the botnet threat per effect

Effect on Botnet Threat	Direct Influence	Indirect Influence	Adjusted Influence
E1 Detection of botnets	2,07	2,73	-0,66
E2 existing population	2,10	2,67	-0,58
E3 New infections	1,63	2,11	-0,49
E4 Cyber Crime Economy	1,55	1,88	-0,34
E6 Hacktivism botnet usage	0,97	1,28	-0,32
E5 APT/state-sponsored usage	1,12	0,89	0,23
E7 Technology proliferation	1,65	1,31	0,34

Looking at the effects on the botnet threat also reveals some interesting findings. Table 6 illustrates the adjusted influences, calculated in the same manner. Most of the possible effects on the botnet threat do have a negative influence value, meaning that they are developing into less of a threat, so are moving in “positive” directions. The only exceptions are the usage of botnets for state actors and the proliferation of botnet technology. This means that the strategy groups discussed in this paper act as a driver for these two and we will see a steady rise in them.

5. Conclusions

Botnets have become major cyber weapons, threatening nation-states’ security and encouraging these nations to identify proper means to cope with them. This is a complex problem and, acknowledging that there are a multitude of factors to consider, this research has contributed in two ways. Firstly, it does so by establishing a framework of strategic options for nation-states to select from. Secondly, by applying the DEMATEL method on the data of the conducted survey, the system was analysed for the influence that each of the elements has. This allows for the ranking of the strategy groups, indicating the influence that each of them has on the botnet threat, as is presented in Table 5. The DEMATEL analysis revealed some interesting findings with regards to the influence order of the discussed strategic option.

The common opinion about the crucial and influential role that ISPs play in the fight against botnets has been confirmed. The greatest surprise is that *cyber insurances* and *software developers’ obligations* scoredso highly in the influence ranking, as they are commonly regarded as less feasible. The limitation of the conducted research is that it did not consider the difficulties one might face by implementing a certain strategy. On the other hand, this can turn into an advantage as implementation concerns are less likely to influence the findings, encouraging future research.

Also remarkable is the fact that the influence of the analysed groups drops in steps of roughly factor 2. The presented ranking by influence allows for easy selection of preferable strategies and can serve as an input for decision makers. With regards to the effect on the botnet threat, it is remarkable to see that state-driven usage of botnets and technology proliferation is not influenced in a mitigating way at all. While it might be assumed for the latter, it is surprising for the former.

Disclaimer

The opinions expressed here are those of the authors and should not be considered the official position of the NATO Cooperative Cyber Defence Centre of Excellence or NATO.

References

ACMA, 2005. *ACMA - Australian Internet Security Initiative.* ACMA. Available at: http://www.acma.gov.au/WEB/STANDARD/pc=PC_310317 [Accessed January 16, 2012].

Anderson, R. et al., 2008. *Security Economics and the Internal Market*, ENISA.

Asghari, H., 2010. *Botnet Mitigation and the Role of ISPs*. Delft University of Technology.

BBC, 2011. *BBC News - French downloaders face government grilling*. Available at: http://www.bbc.co.uk/news/technology-14294517 [Accessed January 11, 2012].

CCC, 2011. *Cyber Clean Center*. Available at: https://www.ccc.go.jp/en_index.html [Accessed January 11, 2012].

CCRA, 2012. *Common Criteria□: The Common Criteria Portal*. Available at: http://www.commoncriteriaportal.org/ [Accessed January 13, 2012].

Cabinet Office UK & Detica Ltd., 2011. *The Cost of Cyber Crime*.

Council of Europe, 2001. *Convention on Cybercrime*. Available at: http://conventions.coe.int/treaty/en/treaties/html/185.htm.

Czosseck, C. & Podins, K., 2011. An Usage-Centric Botnet Taxonomy. In *Proceedings of the 10th European Conference on Information Warfare and Security*. Tallinn: ECIW, pp. 65-72.

Czosseck, C., Ottis, R. & Talihärm, A.-M., 2011. Estonia after the 2007 Cyber Attacks. *International Journal of Cyber Warfare and Terrorism*, 1(1), pp.24-34.

Denning, D.E., 2001. Activism, hacktivism, and cyberterrorism: the internet as a tool for influencing foreign policy. In*Networks and netwars: The future of terror, crime, and militancy*, pp.239–288.

Dunn, M., 2005. A comparative analysis of cybersecurity initiatives worldwide. In *WSIS Thematic meeting on Cybersecurity, Geneva*.

ECO, 2011. *Anti-Botnet-Beratungszentrum*. Available at: https://www.botfrei.de/en/index.html [Accessed January 11, 2012].

Eeten, M.V. et al., 2010. The role of internet service providers in botnet mitigation: an empirical analysis based on spam data. *OECD Science, Technology and Industry Working Papers*, 2010/05.

European Comission, 2007. *COM (2007) 697*. Available at: http://ec.europa.eu/prelex/detail_dossier_real.cfm?CL=en&DosId=196418 [Accessed January 11, 2012].

European Commission, 2010. *European Public-Private Partnership for resilience –EP3R*. Available at: http://ec.europa.eu/information_society/policy/nis/strategy/activities/ciip/ep3r/index_en.htm [Accessed January 16, 2012].

Evron, G., 2009. *Dutch ISPs Sign Anti-Botnet Treaty - Dark Reading*. darkreading. Available at: http://www.darkreading.com/blog/227700601/dutch-isps-sign-anti-botnet-treaty.html [Accessed January 16, 2012].

GTISC & GTRI, 2011. *Emerging Cyber Threats Report 2012*,

Geers, K., 2011. *Strategic Cyber Security□: Evaluating Nation-State Cyber Attack Mitigation Strategies with DEMATEL* Faculty of Information Technology. TUT.

Jafari, M., Hesam, R. & Bourouni, A., 2008. An Interpretive Approach to Drawing Causal Loop Diagrams. In *Proceedings of the 26th International Conference of the System Dynamics Society: 20-24 July 2008; Athens Greece*.

L.A.P., 2005. *London Action Plan*. Available at: http://www.londonactionplan.com/ [Accessed January 16, 2012].

Lee, W.S.A.Y.H.C.-C.C., 2009. Financial Investment Strategy by DEMATEL and Analytic Network Process. *Network*.

Li, C.-W. & Tzeng, G.-H., 2009. Identification of a threshold value for the DEMATEL method using the maximum mean de-entropy algorithm to find critical services provided by a semiconductor intellectual property mall. *Expert Systems with Applications*, 36(6), pp.9891-9898.

Lin, C.J. & Wu, W.W., 2008. A causal analytical method for group decision-making under fuzzy environment. *Expert Systems with Applications*, 34(1), pp.205–213.

Microsoft, 2011. *Microsoft Digital Crimes Unit*. Available at: https://www.microsoft.com/presspass/presskits/DCU/ [Accessed January 13, 2012].

Nazario, J., 2009. Politically Motivated Denial of Service Attacks. In C. Czosseck & K. Geers, eds. *The Virtual Battlefield: Perspectives on Cyber Warfare*. 163-181: IOS Press, pp. 163-181.

Ottis, R., 2010. From Pitchforks to Laptops Volunteers in Cyber Conflicts. In C. Czosseck & K. Podins, eds. *Conference on Cyber Conflict Proceedings*. Tallinn: CCD COE Publications, pp. 97 - 108.

Plohmann, D., Gerhards-Padilla, E. & Leder, F., 2011. Botnets: Detection, Measurement, Disinfection & Defence. *Information Security*, p.153.

Risen, T., 2010. Can Insurers Protect The U.S. From Cyber-Attack? - Tom Risen - NationalJournal.com. *National Journal*. Available at: http://www.nationaljournal.com/njonline/no_20100208_9513.php [Accessed January 13, 2012].

Shyu, J.Z., 2008. Causal relationship analysis based on DEMATEL technique for innovative policies in SMEs. *PICMET 08 2008 Portland International Conference on Management of Engineering Technology*, (c), pp.373-379.

Tzeng, G.-H. et al., 2009. Fuzzy decision maps: a generalization of the DEMATEL methods. *Soft Computing*, 14(11), pp.1141-1150.

Wood, L., 2011. Got cyber insurance? *Network World*. Available at: http://www.networkworld.com/news/2011/102411-cyber-insurance-252145.html [Accessed January 13, 2012].

Wu, W., 2008. Choosing knowledge management strategies by using a combined ANP and DEMATEL approach. *Expert Systems with Applications*, 35(3), pp.828-835.

Telefonica – Potential Victim of American Espionage

Joey Dreijer[1], Matthew Roberts[2], Neera Jeymohan[2], Jeremy Julien[3], Tommy Karlsson[4], Raquel Cuesta[5], Nils Monning[6] and Elif Duru[1]
[1]Hogeschool van Amsterdam University of Applied Sciences, Amsterdam, Netherlands
[2]University of Salford, Salford, UK
[3]ESIEA, Laval, France
[4]Mid Sweden University, Sweden
[5]University of Alcala, Alcala de Henares, Spain
[6]University of Applied Sciences Bonn-Rhine-Sieg, Germany
julien@et.esiea-ouest.fr

Abstract: During the year of 2010, a set of reports were released that contained important information regarding European multinationals and local governments. These reports were published by Wikileaks and were described as 'cables'. These cables were sent from American embassies to the United States' government. Suspicions were raised that the United States might have been spying on European countries, companies and their respective partners. One of the companies being mentioned in the cable files is Telefonica. Telefonica is a major Spanish telecom, television and Internet service provider operating in Europe and Latin America. The United States could have political and/or economic reasons to spy the Southern American continent. The main question during our research is to find out what reasons the United States might have to spy on Latin American countries and if the rumors were true that they have indeed spied upon Telefonica. And if these rumors are true, what profits there to be found within the cables and other relevant sources? During our research, we found relevant data containing possible evidence related to American espionage. The data found within the cables contain information about America's interest in Telefonica and the pressure being put on their investments in Latin America. The cables contain quotes referring to America's interest in Venezuela and Cuba. The cables contain messages being sent to an American official, stating that the United States should be aware of Telefonicas vast growing market share and influences in Southern American countries. Other relevant cables contain quotes referring to 'informed sources'. Who are these sources and how did they obtain relevant information about Telefonicas expenses? Most of the other relevant cables refer to other investments being done in Latin America, especially in Brazil and Mexico. The United States have good knowledge about the investments being done in Latin America by Telefonica. Unfortunately, none of these cables contained actual proof that the US had directly spied on Telefonica. The cables only contain relevant information regarding America's interest in Telefonica. Even though these cables contain shady and maybe suspicious quotes; this cannot be identified as evidence regarding possible espionage.

Keywords: e-Discovery, Wikileaks, EDRM, espionage, Telefonica, forensics

1. Introduction

A non-profit organization by the name of "Wikileaks" is (in)famous for its full-disclosure of documents. Wikileaks is best known for its part during the Iraq war. Wikileaks published several classified documents related to U.S. strategies and allegations of possible U.S. abuse (BBC, 2010). During the year 2010, Wikileaks published a set of cables containing communication sent from US embassies all over the world including Europe to the United States government (WikiLeaks, 2010). These cables contained detailed information regarding investments in foreign countries. There are many European multinationals located in China, the United States, South America and other countries around the world. The cables contained information regarding the following ten European companies: Telefonica (Spain), Philips (Netherlands), Infineon (Germany), Cobham (UK), Thales (France), Alcatel (France), Siemens (Germany), Unilever (Netherlands), Datong (UK) and Nokia (Finland). In this paper only Telefonica will be discussed.

Suspicions were raised that the United States might have been spying on European countries, companies and their respective partners. In co-operation with several European universities, the European Commission of Trade and Industry wants to have this issue investigated to have a base for future actions. We as a student team tried through our investigative report to provide the necessary articles, search queries and jurisdictional aspects to provide the reader with a solid base-knowledge about the subject.

We as a student team will develop our own hypothesis based on our own research regarding possible espionage by the United States on Telefonica. The claim that espionage has been conducted needs to be provable and based on conclusive facts. It is not desirable to make an accusation or draw any sort of conclusion before doing a thorough research based on a method that is fair, provable and repeatable. Having a firm foundation to back up any claims regarding the use of immoral or illegal means in an attempt to gain an advantage over a rival is imperative. The main reason for the E-discovery research was initiated by a suspicion that espionage was instigated by the United States towards European based companies. This needs to be investigated. One of the main focuses of this paper is to prove, or disprove the suspicions raised by information revealed by Wikileaks and Cablegates that official American diplomatic personnel instigated and sent back confidential information to Washington. If it can be proved that this diplomatic intelligence has been going on for years this might have political and economic repercussions. But one must keep in mind that this intelligence that may have been gathered might have huge economic impact on European companies and trade.

Espionage is an illegal act according to international trade regulations and must be taken seriously. The fact that American embassies used their personnel as agents of covert intelligence or espionage is not acceptable according to agreements made about international trade.

Our main hypothesis, that is the basis of our investigation, is the allegation that the United States did indeed spy upon European companies for political and/or economical reasons. The company that is the main focal point for this investigation is Telefonica, a Spanish Telecommunication company with large investments in the growing economies in South America. Motivation for the United States to spy upon Telefonica could be both political and economic. South America has long been seen as the American “back yard” and America has for a long time been influencing South American politics, both directly and indirectly. The unstable political climate in South American countries and the growing economies makes these countries an easy and desirable target for American intelligence and business interests.

2. Basics of EDRM

Our research was planned according to the Electronic Discovery Reference Model (EDRM Project, 2005), which describes the stages being performed during an e-Discovery project. The official stages are divided into several subjects. The specific stages related to our research are explained below.

- **Information Management:** The information management track is currently left out of scope regarding our research. The information management track, if being included in e-Discovery, includes the policy of a company in which way data is being secured and stored.
- **Identification Track:** The identification track is the first stage to accomplish after an incident has happened and an investigation is being started. It includes the discovery of potential sources and to determine the research scope, breadth and depth. For our investigation, all of the cables that had been made publicly available by Wikileaks, were handed out on a DVD. All of the files located on the DVD also include the removed cables that no longer can be found on the Wikileaks website.
- **Preservation Track:** During the preservation process of the EDRM model all of the data has to be stored in a reliable way. The data should be protected against alteration or destruction. All of our data is located on a DVD and copied to our local laptops. The files have been verified by the use of MD5 hashes. Changing the content of a file will change the hash, and therefor proof that that file has been altered. All of the newly created data during our research have been MD5 hashed before use.
- **Collection Track:** The collection process describes the process of retrieving all of the (relevant) data. Please note; the collection track does not include the filtering of required data. A big chunk of files is gathered during this stage, but not yet filtered to a ‘high’ level. Some global keywords have been used to gather all of the data relevant to Telefonica (‘field of haystacks’), but have not yet been separated by the use of exact search queries.
- **Processing:** During the processing stage, the team is supposed to filter the volume of relevant data and converting for further review and analysis. The data gathered here should be relevant. This data can only contain specific and related information (limited specific ‘haystacks’).
- **Review:** During the review stage, the data gathered during the processing stage is ready to be reviewed. The data should be checked for relevance and privileges (per ‘haystack’).

- **Analysis:** The analysis stage of the EDRM model consists of evaluating the data for content and context. The appropriate team members should read through the relevant cables and find out if these cables could provide the evidence needed (i.e. 'find the neddle in the haystack"). The cables were reviewed by reading the actual file content and by the use of our own custom-built review application
- **Production:** The production stage includes the 'delivery' of evidence found in the relevant data and should provide a solid conclusion to the main question of the investigation.
- **Presentation:** The presentation stage includes the presentation of the final investigative report.

3. Identification track: America's interest

To slink down our research material, we had to filter our search queries. We started to wonder why Telefonica could be of interest to the United States. Telefonica is well known by many famous subsidiaries, such as O2, Vivo, Movistar and Terra Networks. Many of these subsidiaries are located outside of the European Union and are mostly based in Southern America (apart from Spain).

In our opinion, which is not yet verified during this first stage of our research, we think that the United States considers Latin America as an economical and partly political threat based on reasons we will discuss in *chapter 3-1* During this stage we found much less cables related to the European market. About 80% of our results included communication regarding Telefonica's communication with the Latin American countries. Hence our reasons to further investigate the American interest in Telefonica and searching for the relationship between the two in Southern American countries.

3.1 Why South America?

During our investigation, we found several articles and items regarding 'unique' American relations with some Southern American countries. In this chapter, we provide the reader with some of our findings based on US and South American history.

The Monroe Doctrine**:** (Monroe, 1823) The doctrine basically stated that any attempt to colonize or influence the South America and North America by any other western state world would be seen as an act of aggression towards the United States. The doctrine was founded in 1823, but claims have been made that the United States see, and indeed treat, the South America's as its 'backyard'.

The United States and Cuba: The United States and Cuba have more than a hundred year relationship that has stretched from direct control of the Island in the beginning of the 20th century to severing of all diplomatic ties and trade embargos (BBC, 2011). Tensions between America and Cuba have been bad the last decades shortly after the infamous missile crisis. Relationships can be described as frosty and hostile, and suspicions of spying and breaking humanitarian laws have been proclaimed from both sides.

Economical and Political relations: Countries such as Brazil and Argentina are big players on the import and export market towards the United States. Products originating from these countries are being sent all over the world. Both Brazil and Argentina are infamous for their possible corrupt political system. Because of the major increase of the common welfare in both countries, the United States may have economic reasons to spy on both Brazil and Argentina.

Venezuela is another country on our list. Venezuela is one of the major oil trading countries in Latin America. Venezuela plays an important role in the export of oil based products to Europe and America. Next to that, Venezuela does not have good political contacts with the US. Venezuela tried to prevent US intervention in Afghanistan and started to block US products from entering the country. Even though the political contacts have been improved throughout the years, the United States may have political and economical reasons to spy on Venezuela.

3.2 Obtaining all general keywords

As noted in chapter 3-1, there are many historical reasons to believe that the United States may have interest in spying on Telefonica because of their Latin American contacts. Before entering the next stage of the EDRM model, we had to sort our search priorities. Within Latin America, our primary research will be based on Cuba and Venezuela. Both countries have problematic political contacts with the United States. Brazil and Argentina are 'booming' markets and might be a place of interest for the US economic market. These two countries could be investigated during a later stage of our research.

We have to set up a global set of keywords that are being used to make our first ‘filtered’ content database. This stage of the EDRM model only includes the general keywords for a global understanding of the topic. It does not include any specific keywords that are necessary to find exact cable numbers related to American espionage. To build a set of global and general keywords, we have to get a basic understanding of Telefonica’s partners and subsidiaries. A diagram showing Telefonica’s general relationship can be found in the Annex of the final report.

The following list of keywords (Figure 1) seem important to us, based on the connections to Telefonica. All of the keywords are either general search options, specific names of company owners or employees that have been involved with Latin American investments, law and regulations and names of former companies that Telefonica invested in.

Random order: Telefonica, Movistar, military, Cuba, Castro, Valdes, embargo, Peter Erskine, Cesar Alierta, telecom, wireless, internet, television, World Trade Agreement political, economical, Spain, Nelson Ferrer, ECTECSA,Madrid, Brazil, AMDOCS, liberal act, SOX, Argentina, Ecuador, Mexico, Colombia, Carlos Lopez Bianco, Venezuela, Bellsouth, CVG Telecom, Chavez, Fernandez, , CTC Chile, CVG Telecom, Terra, Vivo, Suntech, Telcel, , Bellsouth, OTECEL, Moviles, CANTV, Ceragon, Santiago Fernandez, Lius Javrez Martinez, , Luis Abril, Ramiro Sanchez, Luis Miguel Gilperez, Ramiro Valdez, Jose Maria Alvarez Pellete, Gomez Gonzalez

General keywords: military, embargo, telecom, wireless, internet, television, political, economical, Spain, Madrid, Brazil, Argentina, Ecuador, Mexico, Colombia
Venezuela keywords: Venezuela, Bellsouth, CVG Telecom, Chavez
Cuba keywords: Cuba, Castro, Valdes, Fernandez, Nelson Ferrer, ECTECSA

Company keywords: Telefonica , CTC Chile, CVG Telecom, Terra, Vivo, Suntech, Telcel, Movistar, Bellsouth, OTECEL, Moviles, CANTV, Ceragon

Persons keywords: Peter Erskine, Cesar Alierta, Santiago Fernandez, Lius Javrez Martinez Carlos Lopez Bianco, Luis Abril, Ramiro Sanchez, Luis Miguel Gilperez, Ramiro Valdez, Jose Maria Alvarez Pellete, Gomez Gonzalez
Other keywords: World Trade Agreement, AMDOCS, liberal act, SOX

Figure 1: Keywords

4. Preservation track: Integrity and queries

Before continuing our project and starting to search through all of our cables, we had to make a copy of the files that were being used. All of the original files were already copied on a DVD by the University staff. Before copying these files from the DVD to our working stations, we had to make sure the data could not be altered and could be checked if any changes have been made. No live forensic methods have been applied to this research; only static non-volatile data has been used. The most common method used to verify alteration of files is called hashing.

During our research, several new databases matching our keyword list have been created and used for further investigation. These newly created databases have been hashed for alternation by the use of an MD5 hash. A hash is a unique identifier for a file’s content. If anything changes within the file, the hash should automatically change as well. It’s important to hash this file to provide enough evidence that the research material has not been altered during our investigation.

4.1 Key players and filter

To filter our first chunk of data, we have to choose several keywords to find as much data related to our primary search. To gather the first set of data, we used the PostgreSQL database with SQL search queries. A specific set of keywords was being used to filter the big chunk of cables (over 250.000) into a smaller more related database. All of the keywords shown below are related to the Latin American market. The European and Asian subsidiaries have been left out during our first (primary) filter. The key players being described in this chapter are mostly closely related to

Telefonica by means of partnership or purchases being done to obtain an already existing key-players on the Latin American continent.

Keyword	Reason
Telefonica	Main company name
Telcel	Former Bellsouth company, now belongs to Telefonica
Movistar	Telefonica Subsidiary
Vovi	Telefonica Subsidiary
Terra	Telefonica subsidiary
Imagino	Telefonica subsidiary

The exact query being used to gather information regarding these specific keywords is:

> *CREATE TABLE ClusterCompanies AS SELECT DISTINCT cableid FROM preprocessed p, to_tsquery(' Telefonica | Terra | Vivo | Telcel | Telefonica_CTC | CVG_Telecom | Movistar | Imagino ') q WHERE p.normalized @@ q;*

The above query searches through our entire cable database and creates a new table based on the above keywords. We're hoping to find most of the critical information based on these general keywords and to start finding more specific data during a later stage in our research.

5. Processing track: Reducing data stacks

During the processing stage of the EDRM model, we used the previously filtered data and divide it into several smaller chunks for easier review. The specific keywords gathered earlier in our research (see chapter 3) are now being used to separate relevant data.

We created several new databases, including two new 'tables' for statistics. Our first table represents how many times a keyword shows up in our newly created databases. The second table represents the relations between cables and specific keywords. This table shows every unique Cable ID of which the keyword is found in. For better search options and more reliable results, we used default PostgreSQL databases regular expressions and combine it with tokens that are being used for full-text search.

Based on the (categorized) keywords shown in chapter 3.2, several specific tables have been created, including the amount of appearances per keyword. By separating search options in a likewise way, evaluating and reviewing the cables can be done for each individual subject. This makes reviewing much easier during the next stage of the EDRM model, which will be discussed in chapter 5-1 and chapter 5-2.

Description	Appearances
General	120
Venezuela	34
Cuba	22
Companies	257
Other	25
Persons	851
Cables	1309

5.1 Reviewing track: Preparing data for analysis

As noted in chapter 5, we created specific tables for reviewing purposes during our research. During the review stage of the EDRM model we combined tables with all the relevant data. The idea is to obtain the exact Cable ID's and find any relevant connections between similar search results. The cable ID's can be used to find cable contents and match them by the use of a social network diagram. Every keyword found in a cable is compared to other relevant cables. It's easy to find connections between the matching cables by creating a diagram as shown below (Figure 2):

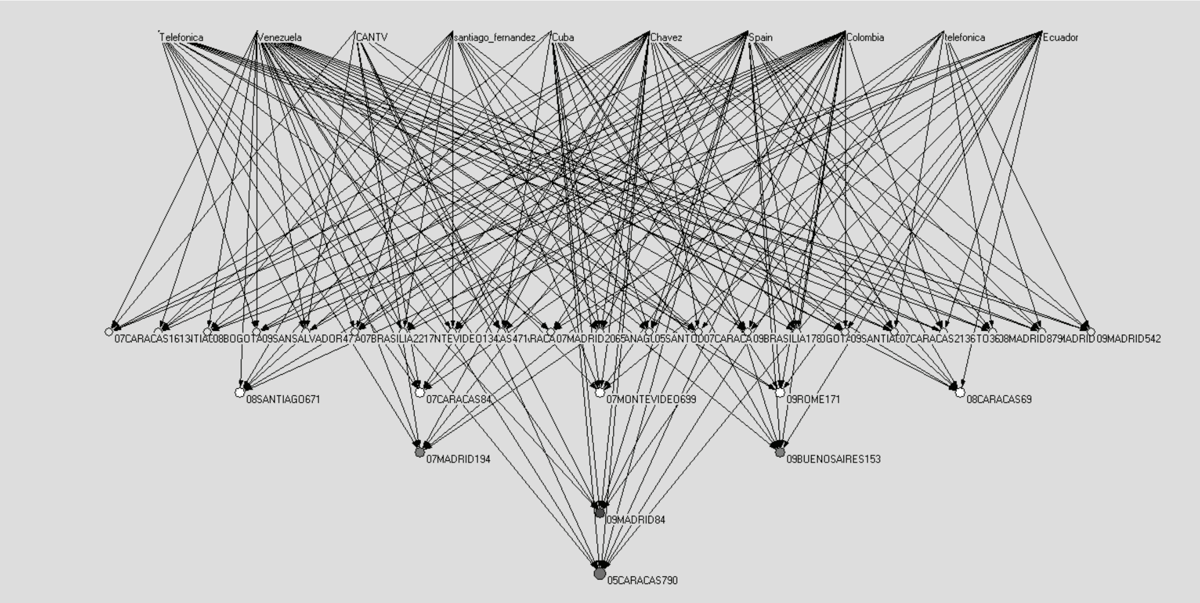


Figure 2 : Keywords and cables network

The above image is an example of how a possible keyword network could look like. The keywords are shown at the top of the image. The more keywords are found within a single cable, the possibility to find relevant information is higher. As shown in Figure 2, one single cable is found with all of the 9 keywords being used during the creation of this diagram. By matching these keywords with exact cable ID's, finding relevant cables for analysis is much easier.

Please note, cable importance is not evaluated by the use of our diagram. The diagram simply shows combinations of keywords found within cables. The importance of a cable is decided during a later stage of our research.

5.2 Social network analysis

Social networks play a crucial role in business operations of an organization. According to the theory of social networking an organization will usually create networks with other organization or partners of the same organization in different countries order to remain competitive in the market. Companies such as Telefonica are able to transfer knowledge and ideas for product innovation through these social networks.

The social network map as shown in the Figure 3 we made to visualize how Telefonica is able to expand their resources and gain necessary expertise resources for further development. For any organization during planning stage cultural difference continues to be an interesting factor affecting the improvement. So by creating networks among their own group of companies Telefonica is trying to influence governments and nongovernmental actors. According to our Social Network Analysis on Telefonica it's visible that the overall network is well connected and information sharing and decision-making is much easier for the top management. But at the same time the success of network depends mainly on the top players of the organization.

6. Analysis track: Reading and evaluating our findings

Although online services, such as those found on Cablesearch (European Center of Computer Assisted Reporting (Eccar), 2012), were employed in the initial stages, a bespoke tool was developed to better meet the requirements of the investigation. Due to team expertise, available resources and the requirement for a user friendly interface, a Windows application written in C# on Microsoft's Visual Studio .NET platform was decided upon.

The results of a search query are ranked according to keyword frequency and density. Including keyword density improves the relevancy by assessing how close to each other the key words are in the text. In addition to this, the ability to import and store pre-written SQL statements accompanied by a plain English description (to aid non technical users) has been incorporated. Users will be able to select a pre-set SQL query based upon the description and start looking for relevant cables.

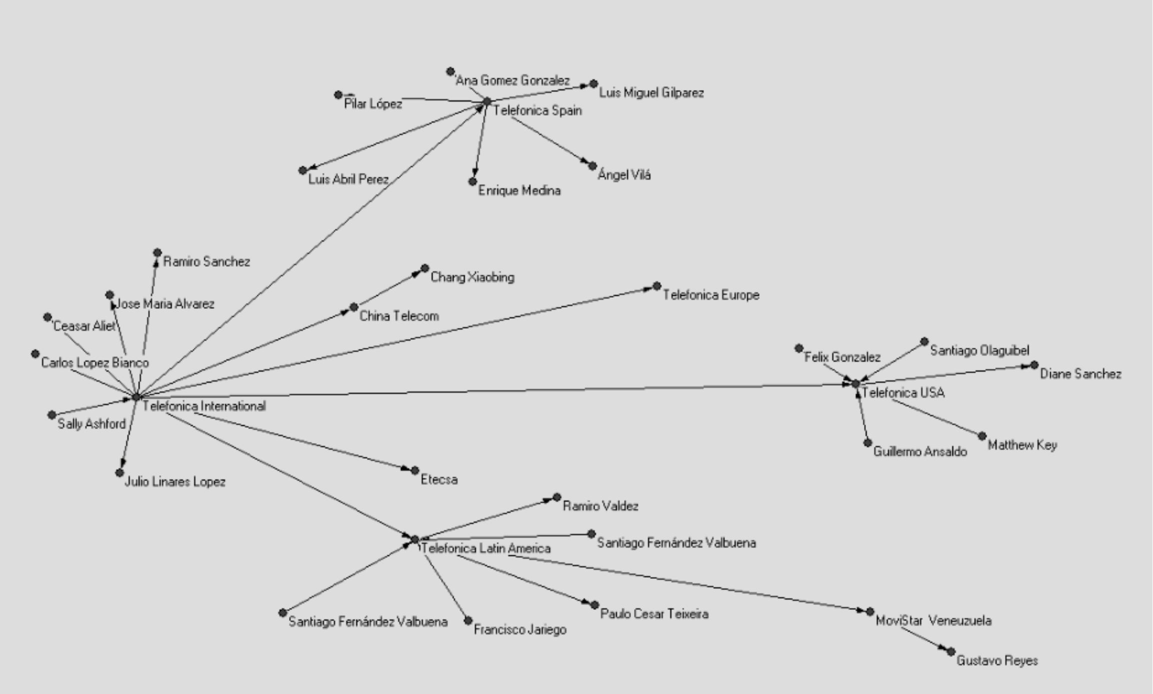


Figure 3: Social network map

Results of a search are displayed on a separate tab, listing each cable in a separate control and displaying the cable details and contents in separate fields. Required keywords are highlighted (yellow for required and a random colour for optional) to aid the location of relevant on information. This makes the reviewing process of the cables a lot easier. All of the relevant data is being highlighted for our law and information science students for further investigation.

6.1 Legal aspects

Industrial, economic or corporate espionage is a form of espionage conducted for commercial purposes. Economic espionage is conducted or orchestrated by governments and is performed on international scope, while industrial or corporate espionage is more often national and occurs between companies. There are suspicions that the US has been spying on Telefónica, but no solid evidence has been found yet. So we do not know if they use the potentially gathered information to make their own telecom businesses stronger and more valuable. If we assume that US has indeed spied upon Telefónica to gather important information to make their own economic position stronger, they would have violated a couple of (inter)national laws.

Firstly, the US is a member World Trade Organization. They have an international agreement called the “General Agreement on Trade and Services (GATS)” (World Trade Organization, 1995). The US and Spain have both signed this agreement. According to article 8 part 1 of the GATS, the US has acting inconsistent with the obligations and specific commitments. Espionage is against the principle of fair competition of the World Trade Organization, because the US can possibly abuse the information to gain a monopoly in Southern America. Besides part 1 of the article, part 2 describes that every country that signed this agreement, agrees that they publish specific and controversial operations. This includes possible espionage and other research related projects. By not publishing any related investigations the US have done over the last few years, they refuse to obey article 8 part 3 of the GATS.

Secondly, the US is a member of the International Convenant on Economic, Social and Political Rights. The first article describes: “All peoples have the right of self-determination. By virtue of that right they freely determine their political status and freely pursue their economic, social and cultural development” (Office of the High Commissioner for Human Rights (OHCHR), 1966). In conformity with article 5 part 1, nothing in the present Convenant may be interpreted as implying for any State, group or person any right to engage in any activity or to perform any act aimed at the destruction of any of the rights or freedoms recognized herein. If the US has spied the Telefónica they have limited the economic development in Spain en therefore violated article 5 part 1 of the International Covenant on Economic, Social and Political Rights.

Thirdly, the national civil law of Spain describes article 1.089 of the Civil Code. This article describes that any firm can ask for compensation due to the damage they have undergone by espionage. In conjunction with article 1.902 of the Civil Code, the spying company and/or country has to 'repair' the damage they have done to the opposing company (Ministerio de Justicia, Spain, 2009). Next to that, article 278.1 part 1 of the Código Penal is also relevant in this particular case. Article 278.1 part 1 describes that when there is any solid evidence regarding possible espionage, a company or country can be prosecuted. This penalty can be very rigid when a company used this information to make their own economic position stronger.

When enough evidence is gathered regarding American espionage, Telefónica has the right to go to the court of Spain. But besides going to the court there is another possible solution. Spain and US are members of the International Center for the Settlement of Investment Disputes (International Center for the Settlement of Investment Disputes, 2012). The primary purpose of this organization is to provide facilities for conciliation and arbitration of international investment disputes. Telefónica can use these facilities when there is an international dispute regarding to possible economic espionage by the United States.

7. Production track: Results

The cables below have been found during our research. These cables might emphasise the possibility that America spied on Telefonica. The content of the cable ID's can be found on the Wikileaks website. See Cablesearch (European Center of Computer Assisted Reporting (Eccar), 2012) for more information regarding the search functions made publicly.

Cable ID	Description
10BUENOSAIRES51	References to unknowns 'informed sources'
06MADRID205	Literal quotes, containing information regarding America putting pressure on Telefonica's investments in Venezuela. Contains America's list of Political interested countries.
09MADRID1146	The cable was about diplomatic connections between host country and Cuba.
09MADRID84	Cable containing detailed information about investments being done in Latin America.

Unfortunately, none of the above cables provide solid proof of American espionage. Though suspicions have been raised because of these specific cables.

Who are these 'informed' sources the US mentions in the cables? These sources tempt to know much information regarding Telefonica's future investments.

The cable relating to an Ambassadors visit to Madrid, clearly states that the US has major political interest in Cuba and Venezuela. This cable also contains important information regarding the American influence on Telefonica and the pressure being put on Telefonica's future investments in Latin America. Telefonica has been mentioned more than once in this cable.

So these cables provide solid explanations why the US has reasons to spy on Telefonica.

Many other cables, not mentioned here but available in the Annex of our final report, refer to the possible investments being done by Telefonica in Latin America. Because Telefonica owns a small part of the US telecom market, American espionage is even more likely to happen. There are several laws that allow the US government to investigate possible attempts of terrorism. Every company located in the US has to comply with these laws. Any data stored in the US, belongs to the US. This makes monitoring much easier without breaking any (inter)national laws.

Based upon both national and international laws (as mentioned in the Annex), our advice to the European Union is to enforce rules within the World Trade Organization based on the General Agreement on Trade and Services. Acts of possible espionage should be punished based on penalties or arbitration. A direct conflict with a country based on criminal law might be a bad idea because of the economical and political relationships with the US and the EU.

Secondly, to further investigate this case we advise the European Union to start a taskforce to gather more specific information about possible espionage on European companies. This taskforce should investigate similar events and use the report as a basis of further research.

References

BBC, 2010. *Wikileaks: Iraq war logs 'reveal truth about conflict'.* [Online] Available at: http://www.bbc.co.uk/news/world-middle-east-11612731

BBC, 2011. *Timeline: US-Cuba relations.* [Online] Available at: http://www.bbc.co.uk/news/world-latin-america-12159943

E-Discovery IP 2012 inc presentations, 2012. *E-Discovery IP 2012 Website.* [Online] Available at: http://oege.ie.hva.nl/~ediscovery/ip2012/

EDRM Project, 2005. *The Electronic Discovery Reference Model.* [Online] Available at: http://www.edrm.net/

European Center of Computer Assisted Reporting (Eccar), 2012. *Cable Search.* [Online] Available at: http://cablesearch.org/

Intelligence Report (inc. Annex) Telefonica *e-Discovery IP 2012 Website.* [Online] Available at: https://oege.ie.hva.nl/~ediscovery/ip2012/?page_id=579

International Center for the Settlement of Investment Disputes, 2012. *ICSID Convention.* [Online] Available at: http://icsid.worldbank.org/ICSID/StaticFiles/basicdoc/CRR_English-final.pdf

Ministerio de Justicia, Spain, 2009. *Spanish Civil Code.*

Monroe, J., 1823. *Monroe Doctrine.*

Office of the High Commissioner for Human Rights (OHCHR), 1966. *International Covenant on Economic, Social and Cultural Rights.* [Online] Available at: http://www2.ohchr.org/english/law/cescr.htm

PostgreSQL Global Development Group, 2012. *PostgreSQL.* [Online] Available at: http://www.postgresql.org/

Telefónica, 2012 *Telefónica.* [Online] Available at: http://www.telefonica.com/

WikiLeaks, 2010. *Secret US Embassy Cables.* [Online] Available at: http://wikileaks.org/cablegate.html

WikiLeaks, 2011. *The Spy Files Wikileaks.* [Online] Available at: http://spyfiles.org/

WikiLeaks, 2011. *The Spyfiles.* [Online] Available at: http://wikileaks.org/the-spyfiles.html

WikiLeaks, 2012. *WikiLeaks.* [Online] Available at: http://wikileaks.org/

World Trade Organization, 1995. *General Agreement on Trade in Services (GATS).* [Online] Available at: http://www.wto.org/english/docs_e/legal_e/26-gats_01_e.htm

From Perception Management to Communication Strategy

Arto Hirvelä
National Defence University, Helsinki, Finland
arto.hirvela@mil.fi

Abstract: Strategic communication is developing into an identified function of a successful information-age military operation. However, it is a concept which is still often misunderstood in the military. Leaders from strategic to tactical level must operate in an information environment to deliver the same message to the intended audiences. To address this challenge through unified action, a whole-of-government approach and concept known as strategic communication has emerged. Strategic communication is a concept which unites efforts of governmental organisations to influence intended key audiences in support of national interests. The concept tries to answer challenges posed by changes in the information environment; the increased flow of information; the increased number of networks and reach of media; the increased value assigned to information, and the greater impact of e-media. Governments have influenced key audiences in support of national interests throughout history. This influencing has had different names like propaganda, psychological warfare or operations and perception management. The question is, do we keep inventing the same things again and again or is there really a major difference? Have previous terms, such as perception management, gained negative status and need to be replaced as a result? Do we need a new term for describing how we affect the minds of others? According to Clausewitz, war is an act of policy. A strategic communication concept tries to get tactical level operators to work towards strategic level ends. Previously all different levels; strategic, operational and tactical, has had their own objectives which were not necessarily related. Tactical level actions have even worked against political objectives. So, have we got back to Clausewitz's theory of warfare as a continuation of politics? This theoretical paper clarifies the strategic communication concept and its relation to other similar terms and connection to the international politics.

Keywords: perception management, strategic communication, communication strategy, Clausewitz, international politics

1. Introduction

The power of information as a tool in international politics has been acknowledged in national security strategy papers at least from the 1980s. Information has been characterised as a key strategic instrument for shaping fundamental political and ideological trends around the globe. (Kuehl, 2000: 277) So it is no surprise that the other political tool, the military, has embraced the importance of influence and has developed means, such as information operations, to wage war in the information environment. The military has again found Clausewitz maxim that "even when the enemy is no longer able to prosecute the war and his land is conquered, still the war, that is, the hostile feeling and action of hostile agencies, cannot be considered as at an end as long as the will of the enemy is not subdued also" (Clausewitz, 2008: 37).

Nation states use information as an instrument to achieve political ends. Government agencies and the military are among the organisations that employ information as a means to influence. The military has its own objectives that it strives to achieve using also information actions. Defeating the opponent on the battle field is merely the first phase of a military operation. Ultimately success, the achievement of political objectives, depends on successfully influencing the population to accept the new state of affairs. (Trent and Doty, 2009) Unfortunately sometimes these military information actions contradict with political objectives. Strategic communication aims to prevent these contradictions and shape the information environment at all levels towards the political objectives. To military commanders, the most important factor when employing information actions is the commander's intent regarding the political objective of a given operation (Darley, 2009: 111).

2. Power of states

In contemporary world politics, Realism views the state as the most important actor on the world stage as it answers to no higher political authority. States have supreme power over their territory and populace and no other actor is above them to govern the global system. (Kegley and Blanton, 2011: 32) Realism sees world politics as a struggle for power. (Kegley and Blanton, 2011: 33) This struggle does not mean that there is kinetic war going on, but that there may be a war of perceptions that aims to change the behaviour of others to achieve both internal and international political objectives.

In Clausewitz's (1989) well-known theory of warfare, war is an act of force to compel the enemy to act according to our will. It is not merely an act of policy but a true political instrument and a continuation of politics. On the other hand, according to the diversionary theory of war, political leaders initiate conflict abroad to increase national cohesion and approval of leadership. So, domestic crisis becomes internationalised because leaders who experience internal opposition are inclined to provoke an international conflict in the hope that their citizens will become less rebellious if their attention is diverted to the threat of external aggression. (Kegley, 2008: 428) Conflict does not necessarily need to be physical; it may only be of ideas and perceptions.

Despite reason for conflict, according to Clausewitz, the most important, far-reaching and comprehensive decision that political and military leaders have to make is to establish what kind of war they are embarking on. (Clausewitz, 1989: 89) Is the conflict conducted physically with hard power or is it going to be a battle of perceptions with soft power? The Cold War was not a very kinetic war, even though arsenal was a big part of it, or at least the perception of arsenal. During the Cold War, when strategic warfare basically meant nuclear warfare, at least in the United States and the Soviet Union, battle was more a question of perceptions. It was a race about who was perceived to be the strongest. Like Clausewitz (2008) mentions in the book On War “a major victory may be achieved through psychological effects”. During the Cold War, each side saw itself as virtuous and peace-loving, whereas the other was seen as untrustworthy, aggressive and ruled by a corrupt government. (Kegley and Blanton, 2011: 13) Because of that, end of the Cold War could not be explained by the concept of Realism, as the main concern of Realism is how vulnerable, self-interested states survive in an environment where they are uncertain about the intentions and capabilities of others. The explanation for the end of the Cold War came from Constructivism, the main concern of which is how ideas and identities develop, change, and shape world politics. (Kegley and Blanton, 2011: 47)

Changing the behaviour of other states requires an understanding of the target audience. In the beginning this understanding requires a state-level analysis. Analysis consists of the authoritative decision-making units that govern a state’s foreign policy processes and the international attributes of that state, e.g. the type of government, level of economics and military power, and the number of nationality groups which both shape and constrain leaders’ foreign policy choices. The analysis is a process whereby the state makes its decisions regarding war and peace and its capabilities for carrying out those decisions. (Kegley and Blanton, 2011: 19) How a target audience is perceived and what means there are available determine whether politics will be supplemented with hard power or soft power. Strategic communication is needed regardless.

3. Definitions of influence

Influencing different audiences in support of national interests has had different definitions throughout history. This influencing has been called propaganda, psychological warfare or operations, and perception management. Strategic communication is different from previous means of influence.

Strategic communication is defined as “focused ‘national’ government efforts to understand and engage key audiences in order to create, strengthen or preserve conditions favourable for the advancement of ‘national’ interests, policies, and objectives through the use of coordinated programs, plans, themes, messages, and products synchronized with the actions of all instruments of national power (US Joint Forces Command, 2010) or as NATO policy defines it “the coordinated and appropriate use of NATO communications activities and capabilities – Public Diplomacy, Public Affairs (PA), Military Public Affairs, Information Operations (Info Ops) and Psychological Operations (PsyOps), as appropriate – in support of Alliance policies, operations and activities, and in order to advance NATO’s aims”. (NATO Bi-SC, 2010) There is a considerable difference between the definitions, as the first one is more action-oriented and the second one concentrates more on capabilities. Basically strategic communication is about the need of states to control the influencing messages from strategic to tactical level. This is not a new issue, as there has been and still are other similar definitions about affecting target audiences’ perceptions and, ultimately, behaviour.

Formerly, the term perception management was used to describe the means which were used to target the human dimension in politics and conflict. The definition of perception management is “Actions to convey and/or deny selected information and indicators to foreign audiences to influence their emotions, motives, and objective reasoning as well as to intelligence systems and leaders at all levels to influence official estimates, ultimately resulting in foreign behaviours and official actions favourable to the originator’s objectives. In various ways, perception management combines truth projection, operations security, cover and deception, and psychological operations.” (Joint Publication 1-02, 2001) The definition does not include a demand for coordination, but that need has been identified. In order to be effective, the means of perception management need to be applied synergistically and with political, economic and military elements of power (Dearth, 2000: 153). Compared to the idea of strategic communication, perception management is more aggressive and concentrates on what to do, whereas strategic communication also addresses what not to do. Perception management also targets just foreign audiences, but strategic communication may target any audience — friendly, neutral or hostile.

Psychological operations have been closely tied to the strategic terms of influence. Definitions for psychological operations are “planned psychological activities designed to influence attitudes and behaviour affecting the achievement of political and military objectives” (NATO APP-6, 2009) or “planned operations to convey selected information and indicators to foreign audiences to influence their emotions, motives, objective reasoning, and ultimately the behaviour of foreign governments, organizations, groups, and individuals. The purpose of psychological operations is to induce or reinforce foreign attitudes and behaviour favourable to the originator’s objectives” (Joint Publication 1-02, 2001). Especially the U.S. definition has common characteristics with the definitions of perception management and strategic communication, which is understandable as psychological operations are one of the means for both. There has been debate about whether strategic communication is an entity with an organisation, as with psychological operations, or a way of thinking. Strategic communication is more of a process and about coordination, so the military will not have organisations similar to psychological operations to conduct it. Strategic communication is not a stand-alone capability either. The integration of strategic communication in to all aspects of policy development, the comprehensive planning process and execution of a plan, supported by the evaluation effectiveness, is critical to success.

Propaganda is a term that is avoided in matters of influence. The term has gained a very negative reputation and has been replaced with other terms that refer to a similar idea but under a different name, as for instance public diplomacy. The definition for propaganda is “any information, ideas, doctrines, or special appeals disseminated to influence the opinion, emotions, attitudes, or behaviour of any specified group in order to benefit the sponsor either directly or indirectly” (NATO APP-6, 2009). An adapted definition of public diplomacy is “the totality of measures and means to inform, communicate and cooperate with a broad range of target audiences world-wide, with the aim to raise the level of awareness and understanding about the sponsor, promoting its policies and activities, thereby fostering support for the sponsor and developing trust and confidence in it” (NATO Bi-SC, 2010). Unlike propaganda, public diplomacy concentrates on building a positive image or brand of the sponsor, and employs cooperation to develop trust. So the correlation between propaganda and public diplomacy is a bit clumsy, but serves to illustrate how states have struggled to exert influence using the same tools but with different names throughout history. Public diplomacy is considered, as are psychological operations, to be one element of strategic communication.

In crisis prevention and management, the term comprehensive approach is used to describe a concept for security policy development. The principal determinants of security policy development are not only military, but social, economic, ecological and cultural, all of which can be best influenced through multinational cooperation. Comprehensive approach is not a term for influencing like those mentioned before, but it has many similarities to strategic communication. The Comprehensive Approach “incorporates government and non-government actors, using all appropriate national and international levers of power, to shape a regional environment and create stability. It seeks to broaden the context of pre-crisis, crisis and post-crisis management by comprehensively engaging all relevant (civil and military) ministries / departments, agencies and organizations in an interagency framework”. (MNE5, 2008: 8) Both concepts seek similar coordination and cooperation through all levels and actors. While the comprehensive approach concentrates on multinational crisis prevention and management, including non-governmental organisations, strategic communication seeks to advance national or organizational interests. Both concepts try to connect actors at all levels both horizontally and vertically as depicted in Figure 1.

The Comprehensive Approach is relevant, as is strategic communication, to the strategic, regional and local levels, and is applicable from pre-crisis situations to post-conflict reconstruction and through the transition of responsibility to local authorities. (MNE5, 2008)

With the previous terms that describe influence it did not matter if the perception of the world was realism or constructivism-oriented. With strategic communication it does matter whether we consider it as a comprehensive attack aiming for identity transformation.

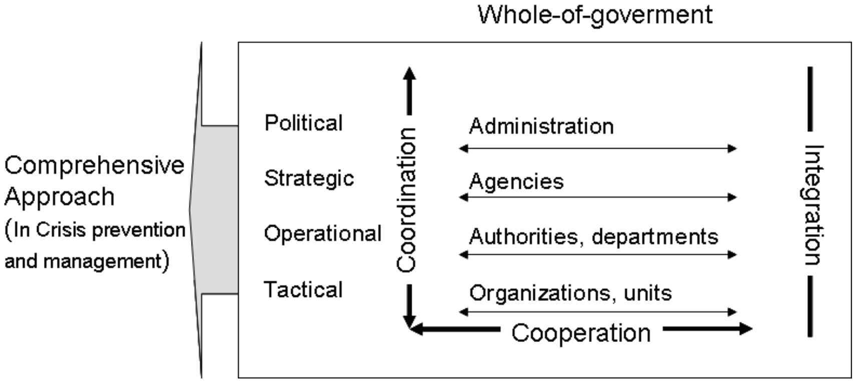


Figure 1: Connection between actors at different levels

4. Strategic communication

In the operational environment, tactical level actions have sometimes worked against political objectives. Even individuals, so called “strategic corporals”, have caused politically negative effects. The amount of these incidents has not been great, but usually negative effects have been. And that has not been the only problem: leaders at different levels have not been told what the political or strategic objectives are, or how political leaders want their nation to be perceived by locals in an operation area, by its own citizens or by the global community. When different agencies also have diverse objectives which are not known to others, there is a danger of releasing conflicting messages. The concept of strategic communication tries to get actors and agencies at all levels to work towards strategic level ends by synchronising messages. It tries to connect actors at all levels both horizontally and vertically.

The information environment is a complex entity. Because the information environment and all the actors in it are not controlled by the military in an operational area, it is difficult to create themes, messages and actions at all levels that both achieve desired effects and avoid undesired effects.(US Joint Forces Command, 2010: I-4). Even with the complete implementation of strategic communication there are still elements that do not support strategic themes. To achieve political goals and address strategic communication issues commanders are urged to develop a communication strategy. (US Joint Forces Command, 2010: I-5) Communication strategy also helps when dealing with issues that are not directly under the commander’s decision making power. The concepts of information environment and strategic communication are depicted in Figure 2.

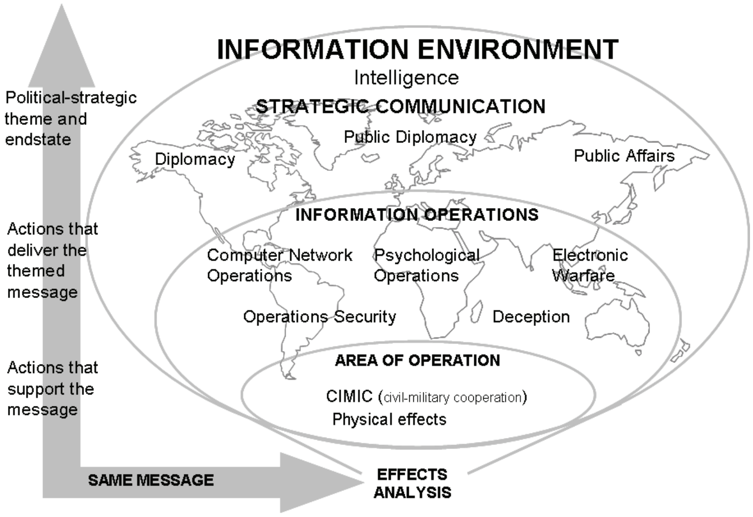


Figure 2: Information environment and strategic communication

The effects on the information environment consist of all direct and indirect results as well as the outcomes and consequences that follow from a specific action. Direct effects usually occur immediately and are easily recognisable, whereas an indirect effect may occur at a much later date and will not be so obvious. Information operations are military actions that deliver strategically constructed messages and uphold narratives. Information operations should be planned based on its effects (EBO) and not as it used to be done in conventional warfare, where the target was chosen first and then the means of affecting it.

Strategic communication puts great emphasis on understanding target audiences. This is because meaning attributed to the message or action is determined by the receiver and not by the source. Without an understanding of the target audience, a message may have contradictory effects to that intended by the source. This is important if strategic communication is considered from the viewpoint of Realism and Clausewitz's theory of warfare, i.e. if strategic communication is a way of coordinating the attack against another nation's identity. Just as with a kinetic attack, the characteristics of the target must be known for the attack to be effective and without negative side effects. But, if strategic communication is considered from the viewpoint of Constructivism, which is concerned with how ideas and identities shape world politics, strategic communication is more a coordinated effort to show how exemplary a certain nation is, and why everybody should admire this model nation and follow its lead. The analysis of the target audience is not so important then, as the aim is to present the most desirable qualities of the nation. The presented image of the nation must be based on the truth, or there will be negative effects if it is exposed as "propaganda".

Effect-based planning is essential in information operations as finding all the possible means of attaining a particular result and avoiding negative effects. This is especially important when operating in networks where it is not always possible to ascertain if the target is owned by an adversary or some other third party. Actions have meaning. Everything that happens in the area of operation sends a message to different audiences, and that message probably holds different information to different receivers. Strategic communication emphasises the need to both control actions and to consider from the strategic point of view what should not be done. Messages, and actions, may be tailored at a tactical level but they must be strategically alike (Trent and Doty, 2009).

5. Conclusion

Strategic communication is a concept. From the viewpoint of Realism and Clausewitz's theory of warfare, strategic communication is a way of coordinating an attack against another nation's identity. And as Realism sees world politics as a struggle for power, strategic communication is just a toolbox with a new name but the same tools, especially since there are terms that describe attempts to influence the behaviour of foreign audiences. Also "the revelation" that all governmental agencies and the military must work towards the same goal is not a new one. But strategic communication does not need to be just another version of perception management. Especially since strategic communication also targets friendly audiences, it cannot be described just as an attack.

From the perspective of comprehensive approach, when all the actors, military and non-military, work towards same ends in crisis prevention and management, and at the same time strategically communicate best national qualities and practices to create better living for all, Better control of national elements and prevention of negative effects is a good thing. If we want to show a good example, we have to live according to it. Problems with strategic communication are usually not communication problems, but policy, coordination and execution problems. Every time a strategic corporal decides to kick down a door of a religious building or leaflets are spread like trash from the air all over the village that has just been cleaned by villagers because they were inspired by other PSYOPS team to do so, we lose our credibility and liability. Messages are not some kind of magic bullets that may be fired and then we just wait for the effect despite what we actually do. Like Nietzsche wrote, if the outcome of action is bad, it is all too easy to loose the correct perspective on what has been done. (Nietzsche, 2009: 19) Good intentions with wrong actions result in negative effects. The road to hell is paved with good intentions.

From the perspective of Constructivism, strategic communication needs to build a truthful and tempting model of society and communicate that to the target audiences. With increased electronic social cooperation and interactivity, the ability to communicate with most of the world's population has become easier. With social media, a type of mass individualism has become possible. (Virilio, 2010: 89) In cyberspace it is possible to challenge the old norms and leadership. For strategic communication to be effective it is not enough merely to make information available. Communities must work hard, so that people would be able to understand each other. People are, after all, the principal elements of the information environment.

References

Clausewitz (1989) *On War*, edited Michael Howard & Peter Paret, New Jersey: Princeton University Press.

Clausewitz (2008) *On War*, Translated by J.J. Graham, Digireads.com Publishing. p 97, Available http://books.google.fi [18 Nov 2011]

Darley, W. M. (2009) Clausewitz's Theory of War and Information Operations, *Ideas as Weapons - influence and Perception in Modern Warfare*. David, G. J. and McKeldin III, T. R., Washington D.C., Potomac Books.

Dearth, D. H. (2000) Shaping the 'Information Space': Perception Management in Peace, Crisis and War, *Cyberwar 3.0: The Human Factor in Information Operations and Future Conflict*. Dearth, D. H. and Campen, Alan (Eds), Virginia, AFCEA International Press.

Joint Publication 1-02 (2001) *Department of Defense Dictionary of Military and Associated Terms*.

Kegley, C. W. Jr. (2008) *World Politics, Trend and Transformation*. Boston, Wadsworth.

Kegley, C. W. Jr. and Blanton, S. L. (2011) *World Politics, Trend and Transformation*. Boston, Wadsworth.

Kuehl, D. (2000) The Information Component of Power and the National Security Strategy, *Cyberwar 3.0: The Human Factor in Information Operations and Future Conflict*. Dearth, D. H. and Campen, A. (Eds). Virginia, AFCEA International Press.

MNE5 (2008) *The Information Factor within a Comprehensive Approach to Multinational Crisis Management*, Second Draft v2.0, Bonn, MNIOE White Paper.

NATO APP-6 (2009) *NATO Glossary of terms and definitions*.

NATO Bi-SC (2010) *Information Operations Reference Book*. Version 1.

Nietzsche, F. (2009) *Ecce Homo - How to become what you are*. A new translation by Duncan Large. New York, Oxford University Press.

Trent, S. and Doty, J. L. (2009) Marketing: An Overlooked Aspect of Information Operations, *Ideas as Weapons - influence and Perception in Modern Warfare*. David, G. J. and McKeldin III, T. R. Washington D.C., Potomac Books.

US Joint Forces Command (2010) *Commander's Handbook for Strategic Communication and Communication Strategy*. Version 3.0. Suffolk, Joint Warfighting Center.

Virilio, P. (2010) *University of disaster*. Translated by Julie Rose. Cambridge, Polity Press.

What Does the Concept of "Ambidexterity" Mean in the Current Military Planning Process?

Aki-Mauri Huhtinen
Finnish National Defence University, Helsinki, Finland
aki.huhtinen@mil.fi

Abstract: How do organizations survive in the face of change? This is a key question for Western military organizations after the Iraq War and its consequences. All human crises are manmade because of we are human beings. The spreading of individual risk also increases systemic risk. The root cause of the problem is what has been termed "rational irrationality" – behavior that, on the individual level, is perfectly reasonable but that, when aggregated in a complex system, produces calamity (Alpaslan & Mitroff 2010, xvii). From the perspective of organizational adaptation and learning, March (1991) argues that a significant number of competencies needs to be learnt and unlearnt during each and every process of change. According to Birkinshaw and Gibson (2011, 2004), in many sports, ambidexterity is a competitive advantage. Footballers are encouraged to use both left and right foot; left-handed batsmen have a slight advantage against right handed bowlers; the southpaw boxer presents a rarely encountered challenge to a boxer with an orthodox stance; some ambidextrous tennis players even use both hands, separately, to play strokes during a rally. And while some individuals are naturally two-handed or two-footed, many work hard to gain an advantage by practising until they master ambidexterity. The challenge for public security and safety organizations is that with terrorism and changes brought on by cyber-security they are faced with their greatest challenge since the end of World War Two. Not only are the structures and operating procedures undergoing change but also attitudes and values are pressed on by a changing society. Rational black and white thinking no longer functions when immigrants, various ethnic backgrounds, social media and the operating mechanisms and values of market economy force their way into the training grounds of military bases and battlefields. This article examines the usefulness of the concept of ambidexterity as part of the Comprehensive Approach planning and decision-making process adopted by Western military organizations.

Keywords: ambidexterity, planning process, comprehensive approach (CA)

1. Introduction

In the great wars of the 20th Century ideologies fought one another. This materialized in the immense struggles between the nation states. The level of material and human destruction was unprecedented. Industrial warfare caused an unprecedented level of material and human destruction. The Cold War and nuclear weapons paralyzed the application of military plans in the organization. Tactical level activities re-emerged in the 1990s when the information revolution penetrated the battlespace in the shape of the military-industrial complex. Actually, in the technological revolution of warfare (Military Revolution Affairs, RMA), the question is whether it is still a fact that the great powers (USA, Russia, China, Israel, Turkey) need to maintain a Weberian hierarchical and organized government and a power policy machinery using new technology, or to accept the post-modern idea of the fragmentation and pulverization of government and power as part of the global information age. (see Hill 2010)

The contemporary challenge is that the armed forces operate in global networks and globalization does not support the idea of self-regulation in the rule formation of a social network. The internet used to be a network and a set of values shared by researches and the armed forces. When it turned into a commercial network and a citizen highway, it crossed the traditional boundaries of self-regulation and commonality. When the Internet expanded from a network and set of values of researchers and armed forces into a citizen highway and commercial network it crossed the boundaries of self-regulation and communality. Then again, a novel form of self-regulation can be seen in the volunteer moderators of different social media conversation forums. How does this idea of a collective need for internal control fit into the activities of a military organization?

In the classic way of thinking, the military command must arrange and coordinate everything the army, navy or air force needs to operate (food supply, sanitary service, system of military justice, and so on.). Also, the command enables the military organizations to carry out their core mission, which is to inflict the maximum amount of death and destruction on the enemy within the shortest possible period of time and at a minimum loss for itself. (van Greveld 2003, 6)

According to van Greveld (2003, 263), the term "system", one of the contemporary buzzwords is appropriate in the technological environment in which we live, but the fact is that all systems that are not purely mechanical, they consist, at least partially, of people, and that those people are by no means entirely individual, but not purely products of the system either. From Plato to NATO, the history of command in war consists essentially of an endless quest for certainty – certainty about the state and intentions of the enemy forces. (ibid., 264) Certainty itself is best understood as the product of two factors, the amount of information available for decision making and the nature of the task to be performed. (ibid., 265)

The focus of this article is the so-called Comprehensive Approach (CA) model that western security organizations, primarily NATO, have started to use when leaning towards the fact that the world's power machinery has become fragmented and pulverized. This model is based on systems thinking and tries to shed some light on the so called "ambidexterity" phenomenon in military operations.

2. The concept of "ambidexterity"

According to O'Reilly and Tushman (2007, 10), central to the adaptive process are the notions of a firm's ability to exploit existing assets and positions in a profit producing way and to simultaneously explore new technologies and markets; to configure and reconfigure organizational resources to capture existing as well as new opportunities. This capacity has been referred to either as exploration and exploitation or ambidexterity.

The term "organizational ambidexterity" was coined by Duncan (1976) as a tool for managing trade-offs between the conflicting demands of exploration and exploitation. Since the seminal work of March (1991) on the dual nature of organizational learning there has been more discussion on how to be efficient and innovative simultaneously. This has been reflected in building a firm's abilities on how to both exploit existing resources and to explore potential opportunities in an ambidextrous manner.

The two key elements for strategic thinking are creative and critical thinking. Operating effectively requires leaders to learn quickly and also to adapt quickly when necessary. Creativity requires developing new ideas and concepts that are effective in resolving the situations at hand. Military professionalism requires that leaders are creative and critical in their thinking. In this regard, the biggest obstacle is the military organization's hierarchical nature and cultural norms. Reflective scepticism is a forbidden method. Dialogue and discussion do no to belong to the military culture and are not part of its prevailing discourse. (Hill 2010; Kuronen 2011). Doublethink or "ambidexterity" also means the ability to believe in two opposing ideas at the same time. It describes the phenomenon of saying one thing while meaning another (Coker 2010, 35). It is the state of being equally adept in the use of both left and right appendages (such as the hands). It is one of the most famous varieties of cross-dominance. People that are naturally ambidextrous are rare, with only one out of hundred people being naturally ambidextrous. A person's ambidexterity is often a qualitative factor in determining the skills.

According to O'Reilly and Tushman (2007, 11), exploitation is about efficiency, increasing productivity, control, certainty, and variance reduction. Exploration is about searching, discovery, autonomy, innovation and embracing variation. Ambidexterity is about doing both. From a strategic perspective, achieving long-term success requires that firms possess not only the operational capabilities and competencies to compete in existing markets, but also the ability to recombine and reconfigure assets and organizational structures to adapt to emerging markets and technologies.

NATO and the western armed forces have long been seeking a solution for controlling the powers of crises such as the ones in Afghanistan and Iraq. Pure military power seems to be ruled out since kinetic weapons cannot be used to create local trust and democracy. The elimination of individual dictators and their governments has not led to a democratization process among the citizens. The challenge for military operations has been fitting them into the political process.

According to O'Reilly and Tushman (2004), ambidextrous organizations establish project teams that are structurally independent units, each having its own processes, structures, and cultures. However, the teams are still integrated into the existing management hierarchy. One of the most important lessons is that ambidextrous organizations need ambidextrous senior teams and managers—executives who have the ability to understand and be sensitive to the needs of very different kinds of businesses.

According to Raisch et al. (2009), differentiation refers to the separation of exploitative and explorative activities into distinct organizational units, whereas integration refers to the mechanisms that enable organizations to address exploitative and explorative activities within the same organizational unit. The tension relates to the question of whether ambidexterity manifests itself at the individual or organizational level. Ambidexterity research usually describes organizational mechanisms that enable ambidexterity, such as formal structures or lateral coordination mechanisms. The third tension relates to static versus dynamic perspectives on ambidexterity. Finally, the fourth tension relates to internal versus external perspectives on ambidexterity. To sum up, managing a paradox requires a creative way that captures both extremes rather than a simple either/or trade off. However, it is still unclear how the tensions between differentiation and integration should be managed.

3. The concept of "resilience" within the concept of ambidexterity

Organizational ambidexterity signifies a firm's ability to manage tensions. Especially in a security organization the ability to manage different kinds of crisis is crucial.

Resilience engineering describes such individual, an organization, or system that has, in relation to safety and security, the ability to foresee, observe, react and learn. In traditional risk evaluation certain dangers are anticipated and expected, whereas resilience anticipates changes which include an increased likelihood of an accident. Resilience engineering questions and compliments the old thinking and operating models in security. It looks for ways to expand the organization's ability to create processes that are sturdy but flexible, even when a crisis is threatening. Traditional risk evaluation methods cannot do that. (Costella et al. 2009)

According to March (2009), exploration includes things captured by terms such as search, variation, risk taking, experimentation, play, flexibility, discovery, innovation. Exploitation includes such things as refinement, choice, production, efficiency, selection, implementation, execution. As a result, maintaining an appropriate balance between exploration and exploitation is a primary factor in system survival and prosperity. In studies of organizational learning, the problem of balancing exploration and exploitation is reflected in distinctions made between the refinement of an existing technology and the invention of a new one. Effective selection among forms, routines, or practices is essential to survival, but so is the generation of new alternative practices, particularly in a changing environment. Typically, the military organizations' tasks in western states have changed from passive territorial defence of one's own territory into the active coalition and shared crises management away from one's territory. But like March says (2009), what is good in the long run is not always good in the short run. What is good at a particular historical moment is not always good at another time. What is good for an organization is not always good for the larger social system of which it is a part of. Firstly, the lessons learned from Iraq and Afghanistan are that the local people are the centre of gravity, not the individuals. Secondly, the Taliban is not the only enemy of the local people. The people are also threatened by inadequate governance, corruption and abuse of power.

Comprehensive Approach (CA) planning requires continuous self-control from the actors as well as the regulation and management of the various information networks and strains. Therefore the bureaucracy of a military organization is no longer based on positional authorities but on self-regulating mechanisms that are utilized on the economic markets and instilled in individual soldiers. Those military leaders who are too set in their ways to adapt to new operational procedures will be replaced either by technology or by civilian experts whose status will never rival that of the officers. Also we need new kinds of concepts or metaphors to explain this self-control of military actors. One such concept is "resilience"**.** It is an outcome of successful adaptation to adversity **(**Reich et al. 2010, 4). The characteristics of the person and organization may identify the resilient process, but only if they lead to healthier outcomes following stressful circumstances, like battlespaces for example. Resilience is a dynamic process whereby individuals exhibit positive behavioural adaptation when they encounter significant adversity, trauma, tragedy, threats, or even significant sources of stress.

"In other words, none of the problems that are parts of a mess can be taken apart and analyzed independently of all the other problems that constitute the mess. A mess is a complex system of problems. For instance, the "military problem", or better yet the "military mess", doesn't exist apart from other messes such as crime, health care, poverty, real estate values, and so on. The concept of mess is essentially synonymous with the concept of a "system". In fact, there is no such thing as a single crisis that is not embedded in a system of other crises. All crises are messes, but all messes are not crises" (Alpaslan & Mitroff 2010, xiv).

In the CA environment we have to plan and prepare for the simultaneous occurrence of multiple crises or catastrophes. One also needs to consider how the effects of multiple crises can reinforce one another so that the overall result is worse than if merely one of the crises occurred. The fact that this process is complex, uncertain, and never perfect is not an excuse for doing nothing (Alpaslan & Mitroff 2010, xvi).

A typology has been proposed by Simsek et.al. (2009) that divides ambidexterity on a 2x2 matrix into four subcategories along the temporal and structural axes:

- Harmonic ambidexterity - whereby exploration and exploitation are pursued simultaneously within a single organizational unit (ibid. pp. 869; 881-2);
- Cyclical ambidexterity - in which organizations engage in long periods of exploitation interspersed with sporadic exploration (pp. 882-4);
- Partitional ambidexterity - based on dual independent structures where exploratory and exploitive parts of organization are separated (pp. 884-6);
- Reciprocal ambidexterity - using pooled interdependence of organizational units that are exploratory and exploitive inputs and outputs (pp. 886-7).

Cyclical ambidexterity poses challenges to designing and studying appropriate strategy processes for organizations. The question of how strategizing in-between the exploration and exploitation extreme takes place might be a rather straightforward one were it not related to several elements of strategizing that are completely at the opposite ends of the scale.

Table 1: Typology of the practices studied, based on Crossan et.al (1999) model

Unit of analysis	Types of practices	Example practices observed
Individual	Intuiting	Experiences described, Images described, Metaphors used.
Individual-group	Interpreting	Language used, Cognitive maps, Conversation/ dialogue, Meetings.
Group-organization	Integrating	Shared understandings, Interactive systems.
Organization	Interpreting	Diagnostic systems, Rules and procedures.

According to Birkinshaw and Gibson (2011; 2004), organizational ambidexterity is the difficult act of balancing between two diametrically opposed organizational qualities – adaptability and alignment. Adaptability is about focusing on the future. It is the ability to respond to change, to be nimble, to progress. Alignment is about maximising the present, leveraging existing ideas, exploiting markets. The organization that successfully reconciles both is rewarded with a significant competitive advantage. On the one hand, organizations must be adaptable. They must be able to quickly seize new opportunities and rapidly adjust to new situations. They must avoid complacency. An adaptable company is nimble, innovative and proactive. On the other hand, as well as adapting to new circumstances organizations need to make the most of an existing situation. This is where the quality of alignment is important. Alignment is about exploiting proprietary assets, rolling out existing business models quickly and stripping costs out of existing operations.

According to Hollnagel (2008), a resilient organization has the ability to anticipate disruptions, pressure and their consequences. What this means is that an organization has to look further ahead than just at the present or the near future. Anticipation gives the organization time to prepare for different kinds of situations. The organization has an ability to observe its own activities and environment flexibly and to adapt its activities. Adaptability and flexibility mean that the basis of shabits. Observation enables surviving what are or what might be critical issues in the near future. The organization has the ability to react to various disturbances and to regular and irregular threats. The actual situation often differs from the expected or imagined one and therefore, ready-made solutions do not exist for all possible situations. In such situation the organization has to be able to adapt its actions so that they meet the prevailing circumstances. However, this has to be done so that the actions are in line with the organisation's own requirements and the available resources. The organization has the ability to learn from the experience; just gathering information on accidents and close calls is not enough. Time and resources are needed to learn from the experience.

4. The concept of comprehensive approach – ambidexterity?

In contemporary warfare and especially in the essence of military organizations, the question is whether military leadership is justifiable as one of the society's functions or should it be seen as a factor which influences and changes the society. The aim of good government is to use invisible power so that the masses under control feel "less controlled" (New Public Management). Getting citizens involved and empowering them, as well as increasing their interaction with public administration is a trend of the information age. This point of departure needs new models for planning, such as the Comprehensive Approach (CA). In CA, control is based on interaction between actors and interactivity in systemic networks. CA strives to support the actors' self-governing that is, a governing model in which each individual can experience their own legitimacy when participating in the common process. Everyone can have his or her voice heard without having to follow an unfamiliar process. The legitimacy of the process is not reducible to one concept of rationalization. In CA, the common good does not overshadow the good of the individual. (see Rintakoski and Autti 2008)

Due to the increasing complexity of the world peace, security and the development of different kinds of crises are more interconnected than ever. In this situation, there has to be closer cooperation and coordination among international organizations. The necessary requirement is that the different organisations play their respective, complementary and interconnected roles in crisis prevention and management. The essential foundation for a Comprehensive Approach is the willingness of individuals, units, departments and organizations to collaborate with one anothers (Johnson 2010; Smith 2002). It is worth nothing to the spirit of the military-industrial complex that civilian organizations have their own plans and agendas and they might not have the resources or they may not be willing to help in the military planning process unless there is something in it for them. There has to be an incentive for the civilian organizations to participate.

According to Hollnagel (2008), one of the perquisites for resilience is being on guard constantly.It prevents the actor from becoming too self-satisfied. Resilience requires a realistic view of one's own skills. It requires knowledge about what has happened, what is happening and what will happen as well as knowledge about what must be done. According to the resilience engineering point of view security and safety are what the socio-technical system creates – and not so much what kind of system it is or what it has. Security and safety are not properties of the system in the sense that the security that has been achieved would be a permanent state. CA strives for a situation in which the shared goal determines the means that can be used in the attempts to reach that goal. CA pyrkii juuri tilanteeseen, jossa yhdessä sovittu toiminnan päämäärä sanelee keinot vaikuttaa päämääriin. Only after an agreement has been reached on the goal and on the means, it is possible to identify the actors and construct the organisation which will be used in the action. The activity is described by the organization's capability at that time.

According to March (2009), two distinctive features of the social context are considered. The first is the mutual learning of an organization and the individuals in it. Organizations store knowledge in their procedures, norms, rules, and forms. They accumulate such knowledge over time, learning from their members. At the same time, individuals in an organization are socialized to organizational beliefs. Such mutual learning has implications for understanding and managing the trade-off between exploration and exploitation in organizations. The second feature of organizational learning considered here is the context of competition for primacy. Organizations often, especially military ones, compete with each other under conditions in which relative position matters.

CA is a one "ambidextrous" attempt to use politico-military power in order to gain control over complicated and chaotic, foreign culture-based areas of operation like Afghanistan and Iraq. CA is an attempt to meet this challenge. It strives to direct the powers of different actors towards a common goal and the achievement of a controlled effect a controlled effect. CA includes order and control in a defined area as well as powers functioning in time and systemic relationships between them to achieve the desired effect. At the politico-strategic level, steering focuses on international events and at the operational level, it is a question of steering the organization's administration (see Foucault 2003). In the Comprehensive Approach process, power is not a characteristic of the process, but rather the effect it produces. The objective of CA's use of power is to limit powers. According to the principles of CA, the process requires parallel planning and implementation, because measuring the effects provides grounds for the continuous reorganization of the activity.

The CA's core questions are: are operations progressing according to plan and are they having the anticipated outcomes? If not, what needs to be adjusted? Traditionally, these questions have been addressed by reports through the chain of command and these reports have usually been of a qualitative nature and based on subordinate commanders' skill and experience. In NATO and especially in the US, the culture has been that of quantitative measures. With evidence-based quantitative methods there is an abundance of data but there have been great difficulties in linking the data to the high level goals. The old saying that everything that counts cannot be counted and everything that is counted does not count still holds true. The question is to find a balance between experience-based (leadership) and evidence-based (management) methodology. (Johnson 2010; Smith 2002).

According to Birkinshaw and Gibson (2011; 2004), the combination of adaptability and alignment is a difficult one to maintain. The difficulty is to find the right balance The consequences of not getting it right, however, are costly. Some companies pay too much attention to alignment at the expense of adaptability. The result is a good short-term performance at the expense of the long-term success. With a short-sighted approach the changing business environment will eventually catch you up.

5. Conclusion

The mutual learning between exploration and exploitation in organizations still continues when we use CA. There is not enough practical experience in military organizations one of the use of the Comprehensive Approach, but its idea supports western military organizations' tasks in the changing environment. We can find the idea of ambidexterity and the resilience process in CA.

According to O'Reilly and Tushman (2007, 13), the key factors needed to succeed in exploitation demand a short-term time perspective, efficiency, discipline, incremental improvement and continuous innovation. The alignment of competencies, systems, structure and culture to execute this strategy is completely different from the alignment needed for exploration. The key success factors emphasize a longer term perspective, more autonomy, flexibility and risk taking and less formal systems and control.

In conclusion, in the 20th Century there was an attempt to turn the battle itself into an industrial type of operation. However, this was impossible because of the lack of a suitable communication system. The essence of war, to quote Clausewitz once again, is a confrontation with the enemy's independent will (van Greveld 2003, 187). The current aim of command is based on General Montgomery's idea of a "Phantom" system of liaison officers who used cars and aircraft to visit every part of the battlespace of operations and report back directly to the HQ. Today we use computers and social media (ibid., 191). CA is based on the idea of systems of system. All information flows and will be shared within different kinds of complex and adaptive systems. Yet the question remains: how do we harmonise the humane in the CA process?.

A proper command system should be able to set goals for itself, and then strive to attain those goals despite the clear realization that things will go wrong, but it should also have the confidence that when they do go wrong, the system will be able to overcome the obstacles (van Greveld 2003, 194-195).

According to Birkinshaw and Gibson (2011; 2004), the mastering of ambidexterity requires some topsy-turvy thinking. The problem is that most organizations approach the ambidexterity challenge from a wrong direction. Classical organizations, like military ones, believe that this so-called structural separation is necessary because the two sets of activities are so different that they cannot coexist. Separation leads to isolation. This is dangerous in the 24/7 information age's networks and social media pressures. Therefore, organizations have experimented with alternative structures that combine elements of both adaptability and alignment. This is still a top down approach.

These structures still rely on the people in charge of directing the work, and deciding on how best to divide up the time of their employees between another set of activities.

According to Birkinshaw and Gibson (2011; 2004), ambidextrous individuals in organization take the initiative and are aware of opportunities beyond their own jobs; ambidextrous individuals co-operate, and seek opportunities to combine their efforts with those of others; ambidextrous individuals are brokers who build internal networks; ambidextrous individuals are natural multitaskers. To conclude, we have to ask how ready the classical homogenous military organisation will be for this kind of definition of military personnel.

By useing CA western military organizations try to achieve a new military organizational culture. CA process encourages the military leaders to become performance managers, a combination of stretch and discipline; and social support, a combination of support and trust. Performance management is concerned with stimulating people to deliver high-quality results and making them accountable. Social support is about providing people with the security, support, and latitude that they need to perform consistently with their highest potential. A burn-out context, like military ones, puts so much emphasis on performance management that the social support systems are neglected, or never put in place. It is necessary to combine the concept of resilience together with the concept of ambidexterity. (Johnson 2010; Smith 2002).

As an “ambidexterity” concept, Comprehensive Approach (CA) illustrates the professional community, the self-control of which tolerates opening up its own activities to include external actors. Actually, CA means that security is built together with the opponent. The motto "know your enemy" has changed into: “communicate with your opponent”. Because the operational environment of the struggle is cyberspace, the form of combat is communication. People are the center of gravity. We combat among people. We cannot commute to a physical fight; instead, we must communicate. The bullets are changing into words, concepts and pictures. The target is not only some individuals, but entire networks. Publicity makes communication continuous and open. The border between attack and defense disappears along with the boundary between what is public and private. A good example of this is Wikileaks or the defense of an open source code.

References

Alpaslan, Can M.; Mitroff, Ian I. 2011. Swans, Swine, and Swindlers. Coping with the Growing Threat of Mega-Crises and Mega-Messes. Stanford University Press.

Birkinshaw, Julian; Gibson, Cristina (2011) “The Ambidextrous Organisation”. The Advanced Institute of Management Research (AIM), UK.

(2004) “The Antecedents, Consequences, And Mediating Role Of Organizational Ambidexterity”, Academy of Management Journal, (47), 209-226.

Costella, M .F., Saurin, T. A., Guimarães, L. B. (2009) ”A method for assessing health and safety management systems from the resilience engineering perspective “. Safety Science, Volume 47, Issue 8, October 2009, pp. 1056-1067.

Crossan, M. M., Lane, H. W., & White, R. E. (1999). “An organizational learning framework: from intuition to institution”. Academy of Management Review, 24 (3), 522-537.

Duncan, R. B. (1976). “The ambidextrous organization: Designing dual structures for innovation”. In L. P. R. H. Kilmann, The management of organization design: Strategies and implementation. (pp. 167-188). New York: North Holland.

Foucault, Michel. 2003. “Society Must Be Defended”. Lectures at the Collège de France 1975-76. Penguin Books.

Hill, Charles. (2010). Grand Strategies. Literature, Statecraft, and World Order. London: Yale University Press.

Hollnagel, E., (2008). “Safety Management–Looking back or looking forward”. In Resilience Engineering Perspectives: Remaining sensitive to the possibility of failure. Ashgate Studies in Resilience Engineering. Ashgate, pp. 63-77.

Johnson, Thomas F. (2010) ”The Comprehensive Approach and the Term of “EBAO”. The Three Swords Magazine 17/2010, pp. 10-13. Available at http://www.jwc.nato.int/files/17_10_Magazine.pdf 11.1.2012.

Kuronen, Tuomas. 2011. Ritual in constructing strategic leadership mythologies. Aalto University publication series. Doctoral Dissertations 123/2011. School of Science. Department of Industrial Engineering and Management. Institute of Strategy. Espoo.

March, J. G. (1991). “Exploration and Exploitation in Organizational Learning”. Organization Science , 2, No 1, 71-87.

O’Reilly III, Charles A. & Tushman, Michael L. (2007) ” Ambidexterity as a Dynamic Capability: Resolving the Innovator’s Dilemma” (http://www.hbs.edu/research/pdf/07-088WP.pdf, 29.12.2011).

2004) “The Ambidextrous Organization”. Harward Business Review.

Raisch, S., & Birkinshaw, J., Probst, G., Tushman, L. M. (2009) ”Organizational Ambidexterity: Balancing Exploitation and Exploration for Sustained Performance” Organization Science, Vol. 20, No. 4, July–August 2009, pp. 685–695.

Raisch, S., & Birkinshaw, J. (2008). “Organizational Ambidexterity: Antecedents, Outcomes, and Moderators”. Journal of Management, 34 (3), 375-409.

Reich, John W.; Zautra, Alex J.; Hall John Stuart (eds.). 2010. Handbook of Adult Resilience. New York: The Guilford Press.

Rintakoski, Kristiina and Autti, Mikko (2008 eds.). Comprehensive Approach. Trends, Challenges and Possibilities for Cooperation in Crisis Prevention and Management. Seminar Publication. Crisis Management Initative. Helsinki: Edita Prima.

Simsek, Z., Heavey, C., Veiga, J. F., & Souder, D. (2009). "A Typology for Aligning Organizational Ambidexterity's Conceptualizations, Antecedents, and Outcomes". Journal of Management Studies, 46 (5), 864-894.

Smith, Edward 2002. Effects Based Operations. Applying Network Centric Warfare in Peace, Crisis, and War. Information Age Transformation Series. CCPR Publication Series.

van Greveld, Martin. 2003. Command in War. London: Harvard University Press.

The Susceptibility of the South African Media to be Used as a Tool for Information Warfare

Anna-Marie Jansen van Vuuren[1], Joey Jansen van Vuuren [2] and Suna Venter [3]
[1]University of the Witwatersrand, Johannesburg South Africa
[2]Council for Scientific and Industrial Research, Pretoria South Africa
[3] Private
anna-marie.jansenvanvuuren@wits.ac.za
jvvuuren@csir.co.za
suna.venter@gmail.com

Abstract: Many theorists refer to the "digital revolution" when they refer to social media and new media technologies. Internet use can also aid journalists in the mainstream media to improve their traditional reporting in terms of speed and feedback. However media practitioners should also recognize the negative consequences and ethical implications of these new media platforms, as the sources distributing information on social media sites such as Twitter may not be truthful and accurate. Journalists should be aware that these sites can be used by enemies of the state to distribute false information. The purpose of this article will be to investigate whether the South African media is at risk of being used as a tool for information warfare. The paper adopts an argumentative analytical approach on case studies with the intention to sensitize journalists to the possibility that different forces may try and exploit their weaknesses in order to influence social opinion with potentially destabilizing effects. In conclusion the paper ends with an overview of some challenges the mainstream broadcast media have to overcome to prevent being used as weapons by the enemies of the state.

Keywords: media influence; agenda-setting; Twitter; radio; social media; behaviour influencing

1. Introduction

Journalists are increasingly drawing on the internet and social media as news sources. The key advantage of the social media such as Twitter is that it gives journalists the opportunity to publish 'breaking news' around the clock. News organizations have adopted social media sites such as Twitter because its speed and brevity makes it ideal for breaking scoops. However, these attributes raise the question of whether this could be called journalism, as true reporting is not just merely stating facts, but rather verifying the information, interpreting and contextualising stories (Manoim 2006). Social media sites may *contain* journalism, but journalists and the audience should be aware that propaganda can be part of the content and therefore could be used as a weapon for information warfare. Information warfare for the purpose of this article is defined as the use of information or information technology during a time of crisis or conflict to achieve or promote specific objectives over a specific adversary or adversaries {Princeton Univerity}. The information on Facebook and Twitter often originates from partisan sources that have vested interests and ulterior motives (Maher 2006). This raises questions about the reliability of information gained from social media sources. The verification of the information distributed through these channels remains difficult and therefore false information can be disseminated. Information Warriors can potentially use these platforms to influence the population negatively against a government as seen in the recent Arab Spring {Comninos, 2011}..

2. The role of the social media in a network society

Media theorists have argued for a long time that the mass media transmit the ideas that constitute the basis upon which a society is formed and developed (Hassan 2004). However, this has been a highly contested theory since the development of the "network society", a term coined by Castells and Van Dyke to express the reality of a 21st century information society that is characterised by its flexible, informal and less hierarchical networks (McQuail 2000). McLuhan states with his "the medium is the message" theory that technology shapes news (McLuhan 1994). In turn, whilst many users regard technology as 'neutral', Hassan argues that there is an ideological bias embedded in every tool (Hassan 2004). According to him the audience perceives the world through the instruments and technologies they use. Theorists such as Chandler criticise this theory for its narrow, reductionist viewpoint (Chandler 2007). Whatever the case may be, the social media revolution has changed the role of media practitioners and presents new challenges for journalists (Hirst, 2011}.

The Internet and social media are valuable resources for journalists. In the field of distribution, new technology has made the production process more cost effective (Hoskins, McFadyen, & Finn 2004).

A key advantage of the social media is that it acts as an early news alert system to give journalists the opportunity to publish breaking news around the clock.

The Mobility 2011 research project by World Wide Worx have shown that the mobile phone habits of South African phone users have evolved dramatically in 2011 since smart phones, mobile applications and the mobile Internet entered the mainstream. At the time of writing almost 83% of South African households have mobile phones with 39% of urban netizens and 27% of rural netizens browsing the Internet on their phones (Jansen van Vuuren, Grobler, & Zaaiman 2012). Thus the South African audience uses mobile platforms such as Twitter and the internet to access news on their mobile devices.

According to Rossouw, the South African media initially saw the expansion of the web as a huge threat, but since then "they have started to ride the digital wave and seize the opportunities created by it" (Rossouw 2006). Since then most newsrooms have been re-engineered into content providers that offer their news collection to other media platforms (Underwood 1995). Therefore media companies have transformed themselves to be cross media "profit centers" (Fink 1996).

Some of the key role players in the South African media industry differ about the role of the social media. The head of news research at the South African Broadcast Corporation (SABC), Izak Minnaar, suggests that journalists should use the Internet more for assistance in their reporting (Minnaar 2007). According to Maher some of the strongest forms of citizen journalism are happening on sites operated by the digital arms of traditional media (Maher 2007a). Maher argues that since the commercialised media conglomerates mostly run the traditional media, citizen journalism upholds the true ideals of the fourth estate. In the past readers had to write letters to the editor to express their opinions to reporters, whilst social media now gives the audience a direct communication line to share opinions, contacts and content with each other online (Acceleration-Media 2007). Thus in this globalised era, it is possible for the *consumers* of the content to interact with the *producers* of the content (Moerdyk 2007). As a result "modern" consumers play an important role in the network society (Fink 1996).

As new media technologies have developed, it has changed the techniques in which reporters tell stories (Powell 1993). Blogging, podcasting and social media such as Facebook and Twitter are some of the prominent new technologies that are practically implemented by reporters in their storytelling. In 2007 the Afrikaans national PBS radio station, *Radio Sonder Grense* (RSG), launched a website that offers diverse podcasts additional to the current "live streaming audio" services (Steyn 2007). These podcasts varies from short newscasts to long-form programmes that consumers can download. The webmaster, Herman Steyn states that the website offers interactivity opportunities for their audience. The station's current affairs shows, *Monitor &Spekrum*, also have their own Facebook page on which journalists report on news as it happens. Listeners can also upload their own photographs and videos of newsworthy events on this site (MonitorenSpektrum 2012).

The Online video portal, YouTube, features clips from international news networks such as CNN. Even our local South African investigative journalism programme, Carte Blanche, can be downloaded from this site. Many of these clips have been uploaded on the site without permission of the authors or producers (News24 2007). Thus, the quality of the journalism or "citizen reporting" on social media websites is the subject of many debates within newsrooms and media platforms.

In explaining *Mail & Guardian (M&G) Online*'s multimedia strategy, Maher (2007) states that they want to contribute to democracy instead of entrenching the elites. Therefore social media has been called the "watchdog of the watchdog" because it monitors the mainstream media (Holian 2007). However, the question that arises then is, who will act as a watchdog and monitor the quality of information distributed via social media sites?

According to Maher (2007) ethical standards was a prominent consideration when *M&G Online* created their blog, *Thoughtleader*. In the end they created a hybrid between a group blog and a more traditional editorial site where they would be able to monitor the content that is published on the site. *M&G Online*'s aim was to have an open media platform while still maintaining their reputation as a quality news source (Maher 2007b).

Maher argues that the debate between citizen journalism and the mainstream media is irrelevant, because they both aspire to the same ideals. However, it must be taken in account that blogging and tweeting often runs counter to journalistic practice. Some information on these social media sites originates from partisan sources that have vested interests and ulterior motives (Maher 2006). Social media sites such as Twitter and Facebook may *contain* journalism, but journalists and the audience should be aware that it could be used as a weapon for information warfare. Therefore journalists that use social media platforms, as sources, must still apply all the “checks and balances” associated with the standard practices of journalism (Maher 2006). Still, many theorists argue that this cannot be called journalism, as true reporting is not just merely stating facts, but rather interpreting and contextualising stories (Manoim 2006).

Though Editors and newsroom managers are grateful for the positive aspects of technology and digitisation, they are also concerned about the social media’s negative consequences (Fink 1996). News distributed by social networking sites is generally written in directory style without context or interpretation (Pavlik 1998b), This is especially relevant in the case of Twitter. It has become a recent trend in newsrooms for journalists to make use of Twitter, a micro-blogging site that restricts posts to 140 characters or less, as a news-dissemination channel and a reporting and source-building tool (Smith 2011). As Farhi states: “Whether they are reporting about it, finding sources on it or urging viewers, listeners and readers to follow them on it, journalists can’t seem to get enough of the social networking service” (Fahri 2009). Carafano argues that the use of social networking tools to facilitate discussion and the exchange of information on an international scale is a well-established phenomenon (Carafano 2009). However, Farhi (2009) criticises Twitter as being an “overhyped technology” and “cultural technofad”. According to a new-media consultant, Craig Stoltz, news organizations have adopted Twitter because it’s speed and brevity makes it ideal for breaking scoops to Twitter-savvy readers: “Twitter works best in situations where the story is changing so fast that the mainstream media can’t assemble all the facts at once” (Farhi 2009). Another advantage of Twitter is that it attracts the type of audience the media wants: “those who are interested in, and engaged with, the news” (Farhi 2009).

The South African News Service “Eye Witness News” (EWN) that broadcast on the commercial radio stations owned by Primedia, was one of the first news organizations in the world to use Twitter as a news outlet (Katopodis 2012). The editor-in-chief, Katy Katopodis, explains that EWN has various accounts on Twitter, including ‘EWN Updates’ (an automated feed that ‘tweets’ the stations’ news headlines”), “EWN Reporter” (reporters give insights on the stories that they follow whilst they are covering it) and “EWN” Sport. EWN is regarded as an agenda-setter, as many other South African news outlets tend to report on a story after it was made public on one of the EWN radio stations. As EWN use both the social media and the broadcast media as platforms, they tend to report on matters faster than many other news outlets such as newspapers (Katopodis 2012).

Thus, whilst social networking tools such as Facebook and Twitter are increasingly becoming “agenda-setters” for what the media reports on, the reliability of the information distributed through these channels remains difficult to verify. Therefore they foresee the need for new editorial skills to suit this new production environment (Pavlik 1998a).

3. Reliability of social media such as Twitter

In his article on how social networking shaped Iran’s Election Protests, Carafano (2009) explains that when the Iranian government cracked down on the traditional media the world turned to social-networking tools in support of: street journalism, the mobilization of the Iranian diaspora and organizing the activists. He argues that the lessons of the Iran crisis illustrate the challenges of operating in a Web 2.0-enabled world. According to Carafano, the key challenge of employing social networks such as Twitter is information assurance: “… ensuring the right information gets to the right person at the right time, while making sure that the information provided is credible, understandable, and actionable” (Carafano 2009).

Silverman (2010) raises awareness on credible reporting with a case study in which an erroneous “tweet” was sent to almost 2 million followers (Silverman 2010). In April 2010 MSNBC.com tweeted on their Twitter account (@Breaking News) that there was an *indication* that the Icelandic volcano, Hekla, has begun erupting. The news spread at an immense speed. Although the journalist, Alex Johnson, used the word “indicate”, to show that the report was unconfirmed, his followers on Twitter treated the report as a fact. Johnson later followed it up with a tweet to emphasise that the eruption has not been confirmed, but at that stage the message had been sent out to millions of people (Silverman 2010). In the end when it turned out that the volcano did not erupt Johnson had to issue a correction on Twitter (Silverman 2010).

According to Johnson it is a problem with breaking news on Twitter sites to find a balance between speed and sourcing (Silverman 2010). He explains that at certain times news is sent out without proper attribution of "who" or "what" the source of it is (Silverman 2010). Therefore questions arise about the reliability and credibility of news that are spread through social media and raises the following concern: In this new era where any person has the potential to be a reporter by spreading news through the use of new media technologies such as Facebook and Twitter – Is there potential for "incorrect" or "biased" news to be spread by the state and enemies of the state? (Smith 2011).

4. Social media as a tool for information warfare

Since 9/11, terrorism and international conflict, especially the wars in Iraq and Afghanistan have been the most popular subject in North American news (Ecoo 2009). Ecoo argues that at the most critical state of American history the mass media has failed in its capacity as a watchdog of the people as it only reported one side of the War story. The South African authors De Beer and Merril (2004) agree that the media has an important role to play in war reporting and to be sensitive to unverified information especially in a time where the media are drawn to sensationalism, violence, wars and political controversies.

Carafano argues that Web 2.0 technologies have an important role to play in a range of activities related to national security, from public diplomacy to communication with citizens during catastrophic disasters: "Government must become practiced in effectively employing these technologies, battling malicious actors online, and ensuring the resiliency of the global open network of free debate made possible by social networking tools" (Carafano 2009). This raises the question of what the role and responsibility of the media is.

The desire of news outlets in the current competitive environment to "break the news" first can result in an external force placing a strategic message on the South African news agenda. Jantunen argues, "the legitimacy of warfare is one of the key themes in news reporting in the 21st century" (Jantunen 2011). She then explains that the language used in news coverage of war reporting can paint a different picture of own and enemy action: "The enemy typically behaves in an immoral and cowardly way, as the enemy is demonized and the 'self' is glorified," (Jantunen 2011). Therefore, through simple lexicon, it is argued that a journalist can describe its own country as leading a "liberation operation" whilst the enemy has instigated "a brutal attack". In spite of this, external forces that are tired of being represented as the "villain" in their enemy's war narrative are now using social media to take matters into their own hands.

One of these external forces is the Islamist militant group in Somalia, Al-Shabaab that launched their own Twitter account in December 2011 to give their point of view of the military conflict. Al-Shabaab's Twitter feed, @HSMPress, has a self-description that states: "Harakat Al-Shabaab Al Mujahideen is an Islamic movement that governs South & Cen. Somalia & part of the global struggle towards the revival of Islamic Khilaafa" (Smith 2011). At the time this article was written, they already had 11 257 followers, demonstrating that "in the 21st century, no radical insurgency or martyrdom operation is complete without a social media platform run from California's Silicon valley, even if Somalia is one of the world's poorest and most anarchic countries" (Smith 2011). Smith describes Al-Shabaab, who has links to al-Qaeda, as fighting the weak, UN-backed Somali government with sophisticated media operations. This includes sending out press releases with photos in well-written English. Al-Shabaab has already allegedly exaggerated the numbers of wounded civilians and its military victories on its twitter feed, but as Smith states: "While their fighters wage war with bombs and bullets, Al-Shabaab is locked in an online propaganda war with Kenya, using weapons of 140 characters each" (Smith 2011). The spokespersons for both Al-Shabaab and the Kenyan army frequently trade insults on the social network and according to a Kenyan human rights activist, Hassan Omar Hassan, "the flurry of tweets obscures the paucity of information about actual operations by the Kenyan military since it entered Somalia in October," (Smith 2011). Hassan states that Kenyans need more accuracy in war reporting to be able to make an honest judgment about the war (Smith 2011). This scenario demonstrates another case study in which journalists have to check the facts they receive on Twitter as the information can be obscured to suit the interests of the party who sends it out. The media should also contextualize the information they report on in such a way that the audience can form their own opinions of what is *really* happening out there.

5. Twitter as a tool for information warfare in the South African context

Some classic examples of the misuse of Twitter (with potential dangerous consequences) in the South African scenario involve the health of the country's 93-year old former president, Nelson Mandela. In an incident in December 2011, a well-known radio host, Bob Mabena, announced erroneously during a live radio broadcast that Mandela had been hospitalized. The source of this information, Mabena later said, was a broadcast on a privately owned South African news channel, eNews - which turned out to have been a rebroadcast of a news story done earlier in the year when Mandela was indeed in hospital. Ten minutes after making the announcement, and still not aware of his error, Mabena tweeted the following under his Twitter account @bob959: "Breaking news - Nelson Mandela is in hospital" (Masungwini 2011). Mabena has more than 2000 followers on Twitter, among them influential figures in the South African media community, such as City Press assistant editor Adriaan Basson and international relations spokesperson Clayson Monyela. Within minutes other Twitter users including some media organisations picked up the rumour. Less than an hour after the initial tweet, the Johannesburg newspaper *The Times* tweeted on its Twitter account, @TimesLIVE: "Former president Nelson Mandela is in hospital, according to reports. Watch this space for more information" (BBC-News 2011). Monyela, who often tweets in his capacity as government spokesperson, also picked up the tweet. On this occasion he wrote: "Pray for Tata", referring to Mandela (Katopodis 2012). The editor-in-chief of popular private radio news service Eyewitness News, Katy Katopodis argues that it was really Monyela's tweet that created a problematic situation in this case - because he "gave credibility to an unsubstantiated rumour" (Katopodis 2012). Katopodis says that Eyewitness News (which has more than 26 000 followers on its Twitter account) took a conscious decision in the aforementioned case not to retweet the rumour before seeking confirmation from the presidency. She says she tries to teach her news staff to always ask themselves the question: "Am I being used, and by whom?" when dealing with breaking news and so-called tip-offs. Katopodis admits however that the popular EWN has made mistakes of its own in the past. In 2008 the news service erroneously reported the death of former Zambian president Levy Mwanawasa, after a reporter received a telephone call claiming that he had died. They tried to follow it up by phoning the Zambian High Commission, whose response was that they would get back to EWN. Therefore, when EWN received a telephone call later on confirming the report, and was prepared to go on record with the information, they assumed it was a response to their initial call to the High Commission (Katopodis 2012). "Later on it surfaced that this person did not work for the Zambian High Commission. It was someone with an agenda, most probably with an incentive to cause harm," (Katopodis 2012). This incident, however, helped EWN to get a system of checks and balances in place for dealing with so-called 'breaking news'.

6. Consequences of false rumours on Twitter

The Twitter frenzy around the Mandela rumours, however, proves that other media institutions in South Africa are still susceptible to false information feeds via Twitter. The December 2011 incident was not the first surrounding Mandela. In January of the same year a tweet from @lebolukewarm, a private user, reading: "RIP Madiba", sparked a similar Twitter frenzy (Zarella 2011). Social media expert Dan Zarella compiled statistics showing that within 90 minutes of the first tweet the rumour was being retweeted 100 times per minute (see Figure 1).

In September of the same year a rumour again surfaced on Twitter that Mandela had died. In that instance, news outlets were quick to seek confirmation, and soon after the rumour surfaced the assistant editor of the weekly newspaper City Press, Adriaan Basson, tweeted: "Almost certain Mandela death is twitter hoax. Zelda [la Grange] says she's heard nothing like that and [a] family member told @SaPolitics1 [political journalist Cedric Mboyisa] Tata is fine." La Grange, who spent years as Mandela's personal assistant, later tweeted: "You want a rumour to stop? Stop speaking about it. Thank you!!" (Pasley-Banks 2011).

This touches on another aspect of the twitter rumour: perpetuation through denial. In all three case studies the #mandelahashtag continued to trend on Twitter long after official denials were issued - and many users were slow to catch on to the fact that the initial rumour was untrue (Zarella 2011).

The ruling African National Congress, in its statement railing against those behind the latest Mandela hoax, perhaps summed up the problem the best: it described the false news of Mandela's death as akin to creating "an atmosphere of panic and anxiety in the country", and continued to say that: "Those behind this hoax are certainly people without any interest in the political and economic stability of South Africa, which we very much owe to the immense contribution by comrade Nelson Mandela – the country's first democratically-elected President" (Mail&GuardianOnline 2011). This could easily be seen as an information warfare attack.

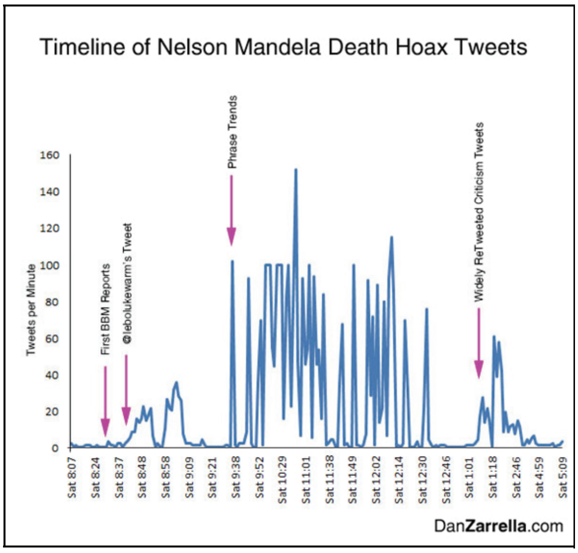


Figure 1: Timeline of Nelson Mandela death hoax Tweets (Zarella **2011)**

7. Recommendations for journalists on using Twitter in a responsible manner

Silverman uses the erroneous Hekla volcano tweet as a case study to advise journalists to use "backchannels" on Twitter (Silverman 2010). According to him these back channels will allow journalists to share details of their reporting process and interact with their readers, especially in cases were errors were made in reporting news on Twitter. "The use of the separate editors account suggests a model for thinking about how to correct an errant tweet and deal with similar challenges on Twitter and elsewhere", (Silverman 2010). Silverman explains that the "backchannel" (editors blog or special Twitter account) fulfils the role the ombudsman does in traditional newspaper – in explaining to the audience why certain choices were made and responding to their questions. A Facebook page could also be used to provide instant context and explanation. However, he does warn that even by creating a backchannel it does not guarantee that followers of the news outlet will use it (Silverman 2010).

According to the editor-in-chief of South Africa's Eyewitness News (EWN), Katy Katopodis, the service does not usually get its "cue's" or "tip-offs" from the social media. If a journalist does see an interesting news item on Twitter, it is part of the editorial policy that they have to find three alternative sources confirming the report (Katopodis 2012). Katopodis agrees that using Twitter as a news source carries the risk of being used as a tool for information warfare, because any person can create an account on the social platform and use it for their own agenda. She advises: "If journalists don't want to fall into the trap of publishing unverified information, they have to return to the old values of journalism that includes checking your sources" (Katopodis 2012). Katopodis admits that in the current fast paced news environment some steps in the process are frequently left behind, but that the journalist and editor should use information gained from social networks sensibly, and also be responsible about the manner in which they "tweet" about the news (Katopodis 2012).

Attribution is another problem with Twitter. According to Katopodis any person can create an account on Twitter and pretend that they are someone else. Thus a journalist should check with a news source if the account belongs to him/her and if a specific post can be attributed to the news source (Katopodis 2012).

If a journalist does make a mistake in its reporting on Twitter, the journalist should repeatedly correct it. Silverman explains that the journalist whose tweet on the Hekla volcano situation was blown out of proportion offered multiple corrections on Twitter. "Since Twitter messages flow by in a constant stream, it is important to repeat your corrections," (Silverman 2010). He advises that one should repeatedly send out corrections whilst the mistaken information is being retweeted. "When something is retweeted, it takes on more authority among people and search engines – so your job in issuing a Twitter correction is to get it retweeted as much as possible," (Silverman 2010).

If a journalist has not confirmed the source or validity of a fact, he/she should begin a tweet with words such as "UNCONFIRMED", "DEVELOPING" or "EARLY REPORT" (Silverman 2010). However, after the incident with the death of the former Zambian President that was discussed in an earlier part of the article, Katopodis advises her journalists against posting any information on Twitter that has not been confirmed. EWN also has a journalist that constantly monitors the tweets that gets posted under EWN's name (Katopodis 2012).

Silverman also argues that Twitter and other new media platforms have a responsibility of creating features on their sites that enable corrections or popularizing standards for indicating unconfirmed or corrected reports. He suggests that the new Annotations function of Twitter could perhaps be used for this purpose (Silverman 2010).

Katopodis confirms that although her news service always wants and needs to be the first ones to report on a breaking news story, it would never be at the expense of the truth: "We would much rather err at the point of caution, and take it slower, to ensure that no mistakes are made," (Katopodis 2012).

8. Conclusion

Many theorists refer to the "digital revolution" when they refer to social media and new media technologies. Internet use can also aid journalists in the mainstream media to improve their traditional reporting in terms of speed and feedback. Consequently the "new" or social media has also been called "the watchdog of the watchdog" as it contributes to the democracy of journalism. However media practitioners should also recognize the negative consequences and ethical implications of these new media platforms, as the sources distributing information on social media sites such as Twitter may not be truthful and accurate. Journalists should be aware that Twitter and Facebook could be used by enemies of the state as tools of information warfare in distributing false information.

References

Acceleration-Media. (2007). New Marketing opportunity unfolds as social media takes off in SA [Electronic Version]. *Bizcommunity.com*. Retrieved August 16 from www.bizcommunity.com.

BBC-News. (2011). Nelson Mandela hospital tweet prompts newspaper apology. Retrieved 20 January 2012, 2012, [online]http://www.bbc.co.uk/news/technology-16363417

Carafano, J. J. (2009). All a Twitter: How Social Networking Shaped Iran's Election Protests. *Douglas and Sarah Allison Center for Foreign Policy Studies* Retrieved 2012/02/02, 2012, www.heritage.org/Research/Technology?bg2300.cfm

Chandler, D. (2007). Technological or media determinism, *Dr Daniel Chandler's homepage*.

Comninos, A. (2011). User-generated content and social networking in the Arab spring and beyond. *policy briefs on the mobile internet from a human rights perspective,* May 2011. Retrieved 20 January 2012, from http://www.apc.org/en/node/12432/

De Beer, A. S., & Merrill, J. C. (Eds.). (2004). *Global Journalism: Topical Issues and Media Systems* (Fourth ed.). Boston, MA: Pearson Education Inc.

Ecoo, C. F. (2009). *The pen and the sword: Press, war, and terror in the 21st century*. Thousand Oaks, USA: SAGE.

Farhi, P. (2009). The Twitter Explosion. *American Journalism Review*(April/May 2009).

Fink, C. (1996). *Strategic Newspaper Management*. Needham Heights, MA: Allyn & Bacon.

Hassan, R. (2004). *Media, Politics and the Network Society*. Berkshire, England: Open University Press - McGraw-Hill Education.

Hirst, M. (2011). *News 2.0 Can journalism survive the Internet?* NSW, Australia: ALLEN & UNWIN.

Holian, D. B. A. P. (2007). Indiana Dialogue Video conference - The US Media and Political coverage. In A.-M. J. v. Vuuren (Ed.).

Hoskins, C., McFadyen, S., & Finn, A. (2004). *Media Economics – Applying Economics to New and traditional Media.* Thousand Oaks, California: Sage Publications, Inc.

Jansen van Vuuren, J. C., Grobler , M., & Zaaiman, J. J. (2012, March 22). *The influence of cyber security levels of South African citizens on National Security.* Paper presented at the ICIW2012, Seattle, /USA.

Jantunen, S. (2011). *Strategic Communication and Revolution in Military Affairs: Describing Actions and Effects*.

Katopodis, K. (2012). Interview about social media usage at Eyewitness News (EWN).

Maher, V. (2006). The War is Over - A Frank Discussion about ways the mainstream media can work with blogs [Electronic Version]. *Vincent Maher* from http://www.vincentmaher.com.

Maher, V. (2007a). The role of new media in traditional journalism. In A.-M. J. v. Vuuren (Ed.).

Maher, V. (2007b). Where is social media going [Electronic Version]. *Bizcommunity.com*. Retrieved May 3 from http://www.bizcommunity.com.

Mail&GuardianOnline. (2011). ANC condemns Mandela death reports. Retrieved 20 January 2012, 2012, [online]http://mg.co.za/article/2011-01-16-anc-condemns-mandela-death-reports

Manoim, I. (2006). Tomorrow's news. In *Changing the Fourth Estate: Essays on South African Journalism* (pp. 247). Cape Town: HSRC Press.

Masungwini, N. K., M. . (2011). Bob Mabena's embarrasing newsbreak about Madiba. *Sunday World* Retrieved 20 January 2012, 2012, [online] http://www.sundayworld.co.za/news/2011/12/31/bob-mabena-s-embarrassing-newsbreak-about-madiba

McLuhan, M. (1994). *Understanding media - The extensions of man*. Massachusettes: MIT PRESS.

McQuail, D. (2000). *McQuail's Mass Communication Theory*. Thousand Oaks, California: Sage.

Minnaar, I. (2007). Computer-aided journalism. In A.-M. J. v. Vuuren (Ed.).

Moerdyk, C. (2007). Newspapers are dead, long live newspaper publishers [Electronic Version]. *Bizcommunity.com*. Retrieved August 28 from www.bizcommunity.com.

MonitorenSpektrum. (2012). Monitor en Spektrum - Facebook Group. [online] http://www.facebook.com/groups/309042425804305/346698288705385/?notif_t=group_activity

News24. (2007). YouTube tries out new system [Electronic Version]. *Sci-Tech News*. Retrieved August 23 from http://www.news24.com/News24/Technology/News.html.

Pasley-Banks, C. (2011). Twitter abuzz over Mandela rumours. Retrieved 20 January 2012, 2012, [online]http://www.news24.com/SouthAfrica/News/Twitter-abuzz-over-Mandela-rumours-20110913

Pavlik, J. V. (1998a). *New Media Technology*. Needham Heights, MA: Allyn & Bacon.

Pavlik, J. V. (1998b). *New Media Technology - Cultural and commercial perspectives. 2nd Edition.* Needham Heights, MA: Allyn and Bacon.

Powell, A. C. I. (1993). Getting the picture. In *Demystifying Media Technology*. Mountain View, CA: Mayfield Publishing.

Princeton Univerity. (2012). Information Warfare. *WordNet A lexical database for English* Retrieved 15 Feb 2012, [online], http://wordnetweb.princeton.edu/perl/webwn?s=information%20warfare

Rossouw, A. (2006). Journalism and the Internet. In *Changing the Fourth Estate: Essays on South African Journalism* (pp. 247). Cape Town: HSRC Press.

Silverman, C. (2010). Eruption, Interrupted: What's the best way to correct an errant tweet? *Columbia Journalism Review*.

Smith, D. (2011). Al-Shabaab and Kenya's army at war - on Twitter. Retrieved 14 December 2011, [online]http://mg.co.za/article/2011-12-14-alshabaab-and-kenyas-army-at-twar

Steyn, H. (2007). RSG website. In J. v. Vuuren (Ed.).

Underwood, D. (1995). *When MBA's rule the newsroom*. New York: Columbia University Press.

Zarella, D. (2011). Anatomy of a Twitter death hoax. Retrieved 20 January 2012, 2012, [online] http://danzarrella.com/anatomy-of-a-twitter-death-hoax-rip-nelson-mandela.html#

Governance of CyberSecurity in South Africa

Joey Jansen van Vuuren, Jackie Phahlamohlaka and Louise Leenen
Defence Peace Safety and Security: CSIR, Pretoria, South Africa
jjvvuuren@csir.co.za
jphahlamohlaka@csir.co.za
lleenen@csir.co.za

Abstract: It is each government's responsibility to provide oversight on national security, which includes human security for its citizens. Recent declarations from the UK and USA governments about setting up new cybersecurity organisations and the appointment of cyber czars reflect a global recognition that the Internet is part of the national critical infrastructure that needs to be safeguarded and protected. South Africa still needs a national cybersecurity governance structure in order to effectively control and protect its cyber infrastructure. Structures need to be in place to set the *security controls* and policies and also to govern their implementation. It is important to have a holistic approach to cybersecurity, with partnerships between business, government and civil society put in place to achieve this goal. The aim of this paper is to propose an approach that South Africa could follow in implementing its proposed cybersecurity policy. This paper investigates different government organisational structures created for the control of national cybersecurity in selected countries of the world. The main contribution is a proposed structure that could be suitable for South Africa, taking into account the challenges of legislation and control of cybersecurity in Africa, and in particular, South Africa.

Keywords: cybersecurity, national security, governance, policy implementation. cybersecurity awareness toolkit

1. Introduction

Around the world cybersecurity challenges give rise to serious national security alarms. There is an international drive by various governments to either develop and implement, or review existing cybersecurity policies. From the United States of America's (USA) point of view, the policies include strategies and standards regarding the security of and operations in cyberspace, and encompasses the full range of threat reduction, vulnerability reduction, deterrence, international engagement, incident response, resiliency, and recovery policies and activities, including computer network operations, information assurance, law enforcement, diplomacy, military, and intelligence missions as they relate to the security and stability of the global information and communications infrastructure. The USA has created a Cyber Command (CYBERCOM) under the Strategic Command led by the head of the National Security Agency (NSA), who reports directly to the President. The main reason stated was that the current capabilities to operate in cyberspace have outpaced the development of policy, law and precedent to guide and control these operations.

Developing nations such as South Africa focus more on the increase of connectivity and neglect the risks that accompany the connectivity. The over reliance on cyberspace compelled the USA to start all its cybersecurity initiatives. Developing nations will have no option, but to join in the race for cybersecurity policy development and implementation. They need to satisfy themselves, as well as instill the confidence across their nations, that the networks that support their national security and economic wellbeing are safe and resilient. Statistics also has shown that despite a low Internet penetration rate, South Africa ranks third in the world after the USA and United Kingdom (UK) on the number of attacks in a country (Amit, 2011).

In its cybersecurity policy (SA Goverment Gazette, 2010), South Africa has acknowledged that it does not have a coordinated approach in dealing with cyber security. Whilst various structures have been established to deal with cybersecurity issues, they are inadequate to deal with the issues *holistically.* There are some interventions to deal with cybercrime, but to have an efficient cyber security strategy there is a need for a partnership between business, government and civil society. South Africa's efforts to ensure a secured cyberspace could be severely compromised without this holistic approach.

As part of the cybersecurity strategy and implementation, we propose a cybersecurity governance structure and an implementation model based on the Cyber Security Awareness Toolkit (CyberSAT) (Phahlamohlaka et al, 2011) that is underpinned by key National Security imperatives as well as by international approaches. Our proposal draws on several analyses derived from international trends and comparing them with key elements of South Africa's cybersecurity policy.

Section 2 contains an overview of the evolution of cybersecurity structures and policies in a number of countries. In Section 3 we draw on these international approaches to craft a proposal for cybersecurity structures and the implementation of a cybersecurity policy for South Africa. The paper is concluded in Section 4.

2. International approaches

2.1 Estonian approach

Estonia is seen as the world's first victim of cyber war, although web traffic was already jammed during the Kosovo war 10 years ago. When Estonia came under cyber attack in 2007, the country realised the necessity of a cyber defence policy. Multiple botnets were used to conduct Distributed Denial of Service (DDoS) attacks against critical national infrastructure, media, telecommunications and the main banks. Websites were also defaced and a significant portion of the economy and government ground to a halt. Although it was suspected that the culprits were Russian nationals, the Russian government did not want to assist in the search for these cyber attackers (Boyd, 2010). These attacks resulted in NATO creating the NATO Cyber Defence Research Centre in Tallinn, a county in Estonia, in 2008, where research and operations take place to counter future activity of this sort. In addition, Estonia adapted its governmental structures due to the realization of the importance of cybersecurity. A National Cyber Security Council was formed as part of its National Security (Tiimaa-Klaar 2010).

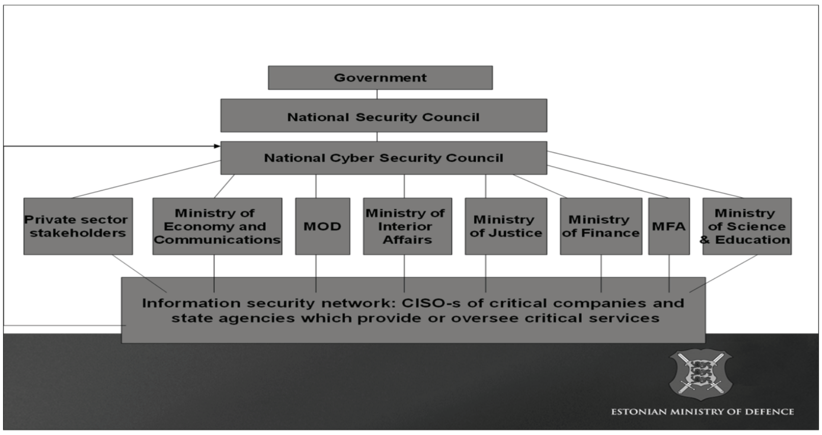

Government
National Security Council
National Cyber Security Council
Private sector stakeholders
Ministry of Economy and Communications
MOD
Ministry of Interior Affairs
Ministry of Justice
Ministry of Finance
MFA
Ministry of Science & Education
Information security network: CISO-s of critical companies and state agencies which provide or oversee critical services
ESTONIAN MINISTRY OF DEFENCE

Figure 1: Estonian cyber security structure (Tiimaa-Klaar)

In 2009, the NATO Computer Incident Response Capability was founded in Mons, Belgium, with intrusion detection and prevention capabilities for NATO networks.

2.2 USA approach

The USA took note of cyber war scenarios and threats they could face from countries with advanced cyber warfare capabilities and thus established the Office of the Assistant Secretary for *Cyber Security and Communications (CS&C)* as part of *Homeland Security* in 2006. This organization is discussed later in this section. In reaction to the cyber attacks worldwide and, in particular, an attack on South Korea (United States. Executive Office of the Presiden,t 2009), the USA embarked on a program to emphasise these cyber issues. President Obama announced that he will make cybersecurity the top priority for the 21st century. He reiterated this vision when he said that cyber-infrastructure is a strategic asset and stressed the need for the appointment of a national cyber adviser to report directly to the president during a summit on national security at Purdue University. He further stated that the USA needs to coordinate efforts across the federal government, to implement a truly national cybersecurity policy and tighten standards to secure information, from all the networks, federal government and personal networks of civilians (Jansen van Vuuren et al, 2010).

As a result, the USA created the CYBERCOM led by the head of the NSA in October 2009. The cyber units associated with each branch of the military fall under the control of the head of CYBERCOM and the NSA. The CYBERCOM will support the Director of the Defence Information Systems Agency (DISA), which in turn has input into a Joint Operations Centre that will be the core of operations under the command of a Deputy Cyber Commander.

Outside the military, the National Cyber Security Division (NCSD) within the USA Department of Homeland Security (DHS) bears responsibility for overall cybersecurity in the USA. It oversees the US-CERT (Computer Emergency Readiness Team) and coordinates activities between public and commercial security groups as part of their mandate. In addition, the DHS operates the CS&C which is concerned with protecting critical information infrastructure. There also exists a National Cyber Security Centre that is responsible for the central coordination of the many organisations within the USA government that deal with cybersecurity. It is still unclear how these cybersecurity offices will work with the Department of Defense (DOD) CYBERCOM.

During the hearing for the appointment of the first head of CYBERCOM, Senator Carl Levin posted three scenarios from the USA side on the responsibilities of cyber defence in the USA. The scenarios as well as responses to them, can be summarised as follow (Stienon 2010):

- If the legal framework under which the USA military operates is used during a traditional operation against an adversary, the commander will execute an order approved by the President and the Joint Chiefs that would presumably grant the theatre commander full leeway to defend USA military networks and to counter cyber attacks that emanate for the attacking country.
- In the case where cyber attacks emanate from a neutral third country, additional authority would have to be granted.
- In a case of a major attack during peace time against computers that manage critical infrastructure, routing the attack through computers owned by USA citizens and routers inside the USA, it will most probably be the responsibility of the Department of Home Affairs and the Federal Bureau of Investigation, but there is no clear guidance in this regard.

From the discussion of the above scenarios, it is clear that this new CYBERCOM needs some research to determine the assignment of responsibility for setting up policies on how the USA should deal with cyber attacks. The creation of USA Cyber Command resulted in other countries following suit as discussed in the following subsections.

The USA cyber Organisation (Figure 2) makes provision for the separation of control of private networks and that of the security sector and is mostly controlled by the Department of Homeland Security.

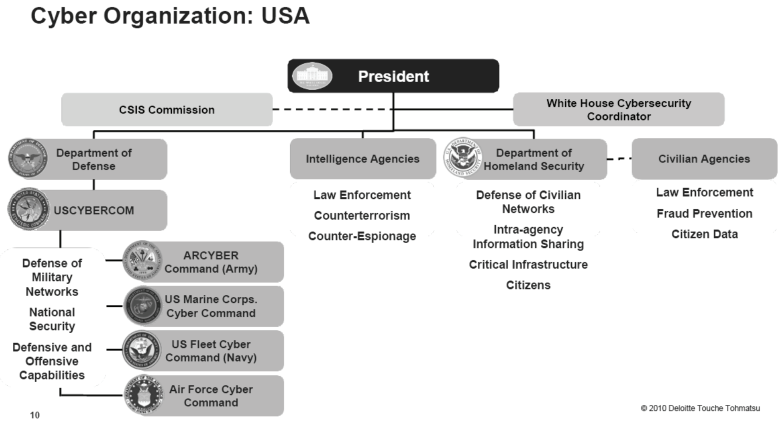


Figure 2: USA cyber organisation (Deloitte & Touch 2010)

The CS&C office consists of three divisions:

- *National Cyber Security Division*: Works collaboratively with public, private, and international entities to secure cyberspace and USA cyber assets.
- *Office of Emergency Communications*: Integrates and coordinates government-wide efforts addressing interoperable emergency communications.
- *National Communications System*: Works with the public and private sectors to ensure continuity and restoration of communications for the Nation in times of domestic emergencies.

The *National Cyber Security and Communications Integration Center (NCCIC)* is a center responsible for the production of a common operating picture for cyber and communications across the state, and local government, intelligence and law enforcement communities and the private sector. The NCCIC is operated within the DHS's CS&C as part of the National Protection & Programs Directorate. Operational elements include the US-CERT, the Industrial Control Systems Cyber Emergency Response Team, (ICS-CERT), National Coordinating Center for Telecommunications (NCC) and DHS Office of Intelligence & Analysis. The NCCIC integrates information from all partners including the Department of Defense, Department of Justice, Federal Bureau of Investigation, Secret Service, and the NSSA, private sector and non-governmental partners. During a cyber or communications incident, the NCCIC serves as the national response center.

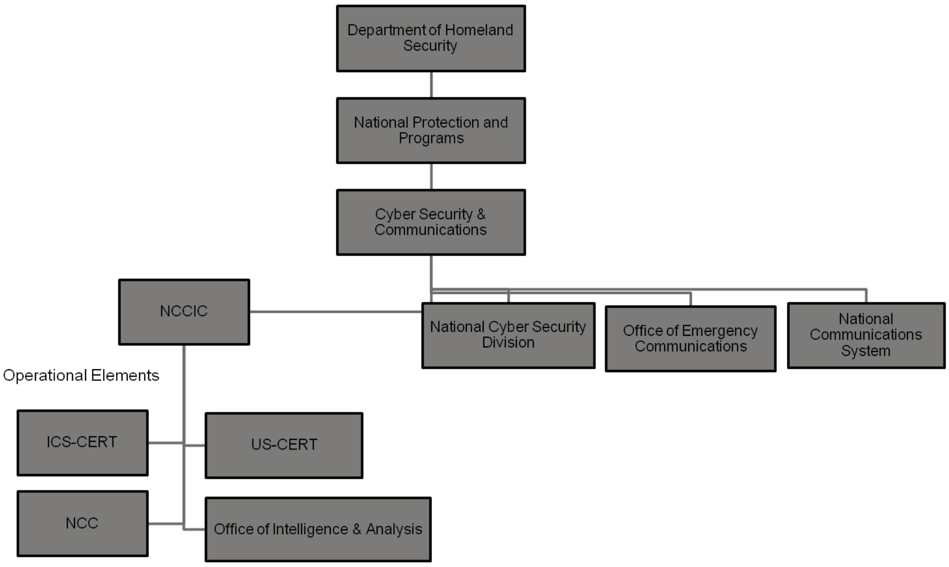

Department of Homeland Security
National Protection and Programs
Cyber Security & Communications
NCCIC
National Cyber Security Division
Office of Emergency Communications
National Communications System
Operational Elements
ICS-CERT
US-CERT
NCC
Office of Intelligence & Analysis

Figure 3: USA homeland security structure

2.3 South Korean approach

South Korea, a country with advanced IT developments, experienced a DDoS attack in July 2009 and experts indicated that it was politically motivated and revealed weaknesses in the national Internet security. A total of 26 domestic and foreign sites were attacked, included the Korean presidential office, government and defence sites. Thousands of infected personnel computers were turned into zombies spreading malicious codes with connection requests to websites, which in turn, paralysed the websites creating this DDoS attack. In addition, malicious code were spread that overwrote the infected PCs' hard drives which could have resulted in massive loss of data and information (Jansen van Vuuren et al. 2010).

North Korea was blamed for a wave of attacks against USA and South Korean websites, but since botnets were used in the attack the true orchestrator of the attack remains unclear. Trojan-based attacks targeted at South Korean government agencies dating back to 2004 were blamed on Chinese hackers rumoured to have the support or perhaps even the involvement of the Peoples' Liberation Army. More recently, North Korean hackers were suspected of stealing a secret USA-South Korean war plan from South Korean systems. Some reports suggested that the hack was done by the use of an insecure memory stick. This cyber attack resulted in the Ministry of Defence in South Korea launching a cyber warfare command centre (mimicking the USA defensive steps), designed to fight against possible hacking attacks blamed on North Korea and China (Zorz 2010). The Centre, which along with a cyber police force, is charged with protecting government organisations and economical subjects from hacker attacks. The centre consists of 200 technical staff members, who are tasked to identify and counter the threat of Chinese hackers and others responsible for the reported 95,000 hacking attacks the country's military networks face every day. It is interesting to note that North Korea already started 20 years ago with the training of cybersecurity experts. It is believed that North Korea has more than 1000 skilled cyber hackers (Zorz, 2010); (Leyden, 2010).

The latest attack in March 2011 targeted 40 institutions in South Korea including banks and financial regulators, as well as military facilities and facilities controlled by the USA forces in South Korea, and the presidential office. The on-line trading system was temporarily shut down under the force of the attack but a spokesperson of the South Korean president indicated that no damage was done. The attacks were done by 11000 zombie computers, very similar to the 2009 attacks (Duncan, 2011; Evron, 2008). As mentioned above, South Korea established their Cyber Warfare command in December 2009. The South Korean cyber organization is shown in Figure 4 (Deloitte, 2010).

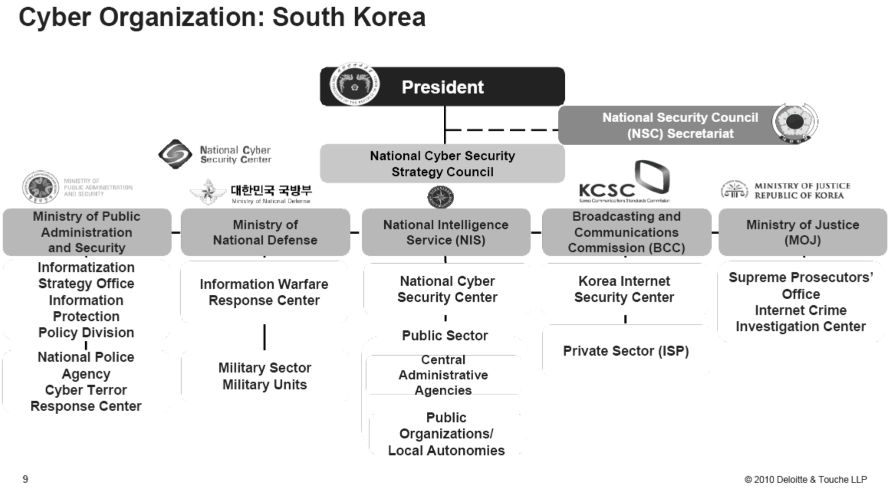

Cyber Organization: South Korea
President
National Security Council (NSC) Secretariat
National Cyber Security Strategy Council
National Cyber Security Center
대한민국 국방부
KCSC
MINISTRY OF JUSTICE REPUBLIC OF KOREA
Ministry of Public Administration and Security
Ministry of National Defense
National Intelligence Service (NIS)
Broadcasting and Communications Commission (BCC)
Ministry of Justice (MOJ)
Informatization Strategy Office Information Protection Policy Division
National Police Agency Cyber Terror Response Center
Information Warfare Response Center
Military Sector Military Units
National Cyber Security Center
Public Sector
Central Administrative Agencies
Public Organizations/ Local Autonomies
Korea Internet Security Center
Private Sector (ISP)
Supreme Prosecutors' Office Internet Crime Investigation Center
9
© 2010 Deloitte & Touche LLP

Figure 4: Korea cyber security structure (Deloitte & Touch 2010)

It is important to note the similarity between the structures of Estonia and the USA with the separation between government, defence and the private sector.

2.4 UK approach

The UK‘s head of MI5 gave a written warning in 2007 to 300 UK companies that they were likely targets of hacking attempts by the Chinese Government. He confirmed that UK Government systems had also been attacked. This was the first time that such an event had been publicly acknowledged in the UK. Other nations as Germany and Belgium also indicated that they had experienced similar attacks. The UK‘s defence minister stressed the need to build robust cyber defences in November 2010 after a Romanian hacker cracked the Royal Navy's Website.

The Government Communication Headquarters (GCHQ) is a British Intelligence Agency responsible for providing signals intelligence as well as providing advice and assistance to UK Government departments and the Armed forces on the security of their communications and information technology systems. It operates under the Joint Intelligence committee. The CESG, originally the Communications-Electronics Security Group, is a branch of the GCHQ that provides the cyber security assistance to armed forces and government departments. They are also responsible for cryptography and to secure critical parts of the UK national infrastructure. In addition, the CESG is the UK national technical authority for information assurance: it primarily advices government and armed forces staff tasked with handling and processing official information, as well as agencies and firms carrying out work for the government.

The increase in expense at a time of economic cutbacks, was justified by stating that future battles will be fought not just on the ground, but in cyberspace. The role of cyber-tactics in offensive actions against enemy states, not just defensive concerns, was also acknowledged (Allan, 2010).

With the publication of the UK Cyber Security Strategy in June 2009 it became clear that the UK's growing dependence on cyberspace, results in the security of cyberspace becoming even more critical to the health of the nation and the protection of national critical infrastructure. Currently, all the approaches to cyber attacks are reactive. The current onslaught of attacks is always one step ahead of the "defender". As a result, Great Britain decided to establish a dedicated team of computer experts that will monitor, analyse and counter hostile computer-based assaults in an attempt to defend the country against cyber attacks (Phahlamohlaka et al, 2011). Lord West, the Security Minister, admitted that the UK already has its own online attack capability. "It would be silly to say that we don't have any capability to do offensive work from Cheltenham and I don't think I should say any more than that." The Cyber Security Operations Centre (CSOC) was set up in conjunction with the Office of Cyber Security, the government computer security agency with its primarily co-coordinating role in the defence of critical IT systems, such as those at utilities or financial institutions. The centre will also have an offensive role to conduct cyber attacks on those posing a threat to the security of the critical infrastructure (Espiner, 2010). Whitehall officials said that the UK and USA will be co-coordinating as there are a close relationship between GCHQ and its USA equivalent.

The UK government also initiated a cyber security hub that will enable the exchange of Cyber security threats by the public and private sectors (Nguyen, 2011).

2.5 Republic of China's approach

In the 1990s, China realised that it needed to develop an alternative way of fighting wars in order to even the odds of defeating a likely opponent with their outdated technologies. The government's relied heavily on cyber warfare to attack modern targets. China was also the first country to start with the formation of cyber□warfare units. In 2000, a series of high-technology combat exercises by the People's Liberation Army (PLA) was suspended when a computer hacker attacked the military's network (Stokes et al, 2011). It is not clear if they are responsible for both private and public networks. Since 2003, China has worked on developing the capability and acquired new technology, reducing the time to design and build new systems.

China’s General Staff Department (GSD) Third Department and its counterparts in the Air Force, Navy and Second Artillery, oversee the vast infrastructure for monitoring and collection of information inside China. GSD Third Department is specifically responsible for network surveillance and intelligence. It controls several operational bureaus responsible for technical reconnaissance. (Stokes et al, 2011). The focus of the GSD Third Department’s signals intelligence, historical lack of an offensive role, and its large staff of trained linguists and technicians, make it well suited for oversight of the computer network defence (CND) and computer network exploitation (CNE) missions in the PLA (Krekel, 2009). The Third Department is comparable to the USA National Security Agency (Stokes et al, 2011).

The Fourth Fepartment most probably has the computer network attack responsibilities. Both the Third and Fourth Departments are said to jointly manage a network attack and defence training system (Stokes et al. 2011). The Fourth Department has set up the Blue Army that will be responsible for offensive cyber attacks as well as defensive actions. As early as 2010, China identified a need to establish a National Cyber Command similar to the USA CYBERCOM due to the need for the prevention of cyber attacks by the Ministry of Home Affairs and the Ministry of Defence to prevent cyber attacks (Guardian, 2010).

As a result of the USA CYBERCOM, the PLA has initiated a dedicated department in December 2010, the Information Security Base, to tackle cyber war threats and protect information security (Hsiao, 2010). Its goal is to gather information and to safeguard confidential military information. However, an officer in the General Staff headquarters told the Global Times: "It is a 'defensive' base for information security, not an offensive headquarters for cyber war.” (Guardian, 2010).

In addition, China has a National Computer Network Emergency Response Technical Coordination Center (NCNERTCC) in Beijing. This team released a report claiming that more than 4,600 Chinese government websites had their content modified by hackers in 2010, an increase of 68 percent over the previous year.

3. Cybersecurity governance strategies for RSA

In this section we consider requirements for establishing cybersecurity governance structures and give a proposal for such a structure in Section 3.1.

A cyber security implementation plan must be implemented on a national level to improve national security levels regarding ICT risks and misuses. To effectively implement a national strategy for a cyber security policy you need an effective approach and culture (Ghernouti-Helie 2010). This include:

- Political will and national leadership to ensure that the plan receives governmental support. It therefore must be supported by
- the justice system and policing with a legal frameworks that supports police to combat cyber crime on national and international level;

- cybersecurity capacity that include organisational structures, human capacity as well as the use of technical and procedural cybersecurity solutions; and
- cybersecurity culture and awareness training of citizens

The USA policy review team suggested that at a minimum, the following elements must be considered (Phahlamohlaka et al, 2011).

- *Governance* structures for policy development and coordination of operational activities related to the cyber mission across the executive branch. This element will typically include the review of overlapping missions and responsibilities that are the result of authority being vested with various departments and agencies.
- *Architecture* that will include the performance, cost, and security characteristics of existing information and communications systems and infrastructures as well as strategic planning for the optimal system characteristics needed in the future. This element will typically include standards, identity management, authentication and attribution, software assurance, research and development, procurement, and supply chain risk management.
- *Norms of Behaviour* will include those elements of law, regulation, and international treaties and undertakings, as well as consensus-based measures, such as best practices, that collectively circumscribe and define standards of conduct in cyberspace.
- *Capacity Building that will include* the overall scale of resources, activities, and capabilities required to become a more cyber-competent nation. This element will typically include resource requirements, research and development, public education and awareness, and international partnerships, and all other activities that allow the government to interface with its citizenry and workforce to build the digital information and communications infrastructure of the future*.*

Structures at national level should exist to sustain the effective cybersecurity solution for all. These structures include adequate organisational structures which should take local cultures, particular economic contexts, country size, ICT infrastructure development and users in consideration. National as well as international needs must also be considered.

Ghermouti-Helle (2011) also argues that the building of capacity should be based upon the understanding of the role of cybersecurity's actors (including their motivation, their correlation, their tools, mode of action, and the generic relevant security functions of any security actions. These considerations will be the underlying principles to be applied for organisational structures to be effective and to determine the kind of tools, knowledge, and procedures necessary to contribute to solving cybersecurity problems. Efficient partnerships between public and private sectors linked to cybersecurity organizational structures, dedicated to support operational proactive and reactive activities linked to cybersecurity management at a national level in turn, should exist.

Based on the South Africa's constitution, the key national security imperatives must be aligned with the governing principle, principle 98 of the South African Constitution, which states very clearly that "National Security must reflect the resolve of South Africans as individuals and as a nation, to live as equals, to live in peace and harmony, to be free from fear and want, and to seek a better life" (Constitution of the Republic of South Africa, 1996). The human security aspect is therefore central to South Africa's perspective on national security. The modern definition of national security defines national security in terms of the respective elements of the power base of a state (Jablonsky 1997). Jablonsky identifies two such elements, the natural determinants and the social determinants. The natural determinants (geography, resources, and population) are concerned with the number of people in a nation and with their physical environment. Social determinants (economic, political, military, psychological, and informational) on the other hand concern the ways in which the people of a nation organize themselves and the manner in which they alter their environment.

The strategy discussed by Ghermouti-Helle as well as the USA cybersecurity policy strategy argue for a holistic approach in the implementation of the Cybersecurity Policy. Phahlamohlaka et al (2011) also argue that the philosophical position; the fundamental premise on which cybersecurity policies are developed is an absolute necessity. This is because cyberspace is a socially constructed, man-made space and therefore a cross-cutting social dimension of national power. At the core of any cybersecurity awareness initiative must therefore be the realisation that no full proof technological protection is possible in a socially constructed space. We argue that the holistic approach to cybersecurity policy that South Africa is looking for is likely to be enhanced by this philosophical position and understanding (Phahlamohlaka et al. 2011). As a cross-cutting social determinant of national power, a cybersecurity awareness programme developed with national security in mind could be confined to the economic, political, military, psychological and informational dimensions. It is these dimensions that constitute their proposed Cyber Security Awareness Toolkit for national security (CyberSAT).

3.1 Cybersecurity governance

In this section we present a proposed cybersecurity governance structure for South Africa based on similar structure in other countries.

Estonia established the Cooperative Cyber Defence Centre of Excellence (CCD COE), a NATO-approved think-tank, whose mission is essentially to formulate new strategies for understanding and preventing on-line attacks (Stienon, 2010). In addition, they developed and implemented their cyber security strategy. Estonia's cyber security strategy seeks primarily to reduce the inherent vulnerabilities of cyberspace in the nation as a whole. The strategy is governed by a structure with a National Cyber Security Council reporting to government. All ministries that are responsible for different aspects of cyber security report to this council. They also differentiate between private sector, government and the military.

The USA created the CSIS commission at the highest level with the White House Coordinator representing the president. The USA cyber organisation makes provision for the separation of control of private networks by the DHS and that of the security sector. The security sector is managed by the CYBERCOM under the Strategic Command led by the head of the NSA.

In South Korea, the Ministry of Defence launched a cyber warfare command centre. Along a cyber police force, the centre is charged with protecting government organisations and economical subjects from hacker attacks. Their structure has at the highest level the National Security Strategy Council and a distinction is also made between security networks and private networks.

The UK established the GESQ and CSOC with the motivation that future battles will be fought not just on the ground, but in cyberspace. As a key part of their Cyber Security Strategy, the UK government also initiated a cyber security hub that will enable the exchange of cybersecurity threats by the public and private sectors.

The Chinese approach is mostly done from a military perspective with the establishment of the Information Security Base which that may serve as the PLA's cyber command. The NCNERTCC in Beijing is responsible for monitoring government websites. It is uncertain who is responsible for the private sector.

It is clear that nations and governments are responding to the cybersecurity challenges by setting up institutional coordination, control and response mechanisms. Linked to the institutional arrangements are also research, development and innovation plans. These national structures responsible for cybersecurity must also lead the capability building processes that will ensure collaboration on international level to achieve the goals identified by global cyber security policies. As seen from the literature, it is important that the cybersecurity be controlled on a very high level, as in the case in the USA, Estonia and Korea and other countries.

The proposed RSA structure (Figure 5) provides for a national body (National Cyber Security Council) reporting to the president as done in the USA, Estonia, UK and Korea. There is also a separation of the civilian and the security networks. The difference between our proposed structure and those of the USA and Estonia is that the government networks will also be controlled by the security services. The official structure for the control of cyber security is still debated in South Africa. Pressure is applied for control by State Security and thus the National Intelligence Agency. As seen in the literature with the establishment of the Cyber Command in the USA, the private sector questioned the fact that the military will play such an important role in the process. The same concerns on privacy of data might be in South Africa if State Security controls cybersecurity, and therefore also civilian networks. The concerns raised in the USA where whether the NSA will overshadow the civilian cyber defence efforts and on what assistance for civilian cyber defence there will be. Some concerns were laid to rest with the assurance that the Department of Homeland Security will be responsible for federal civilian networks including the dot-gov, and that CYBERCOM will only assist the Department of Homeland Security in the case of Cyber hostilities as a response to an executive order (Burghardt, 2012).

The model proposed by Phahlamohlaka et al (2011) for the implementation of the proposed Cybersecurity Strategy in South Africa is the CyberSAT that makes provision for policy decisions and the determinants of national power. Although the toolkit is based on policy elements from the South African environment, the determinants of national power are generic, and thus the toolkit could be adopted for cybersecurity awareness raising by other countries when national security considerations are pertinent.

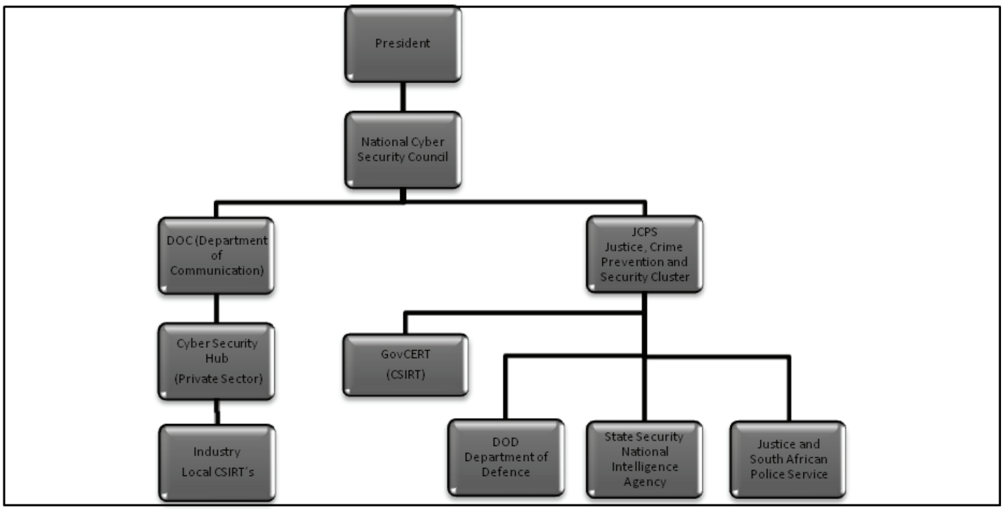

President
National Cyber Security Council
DOC (Department of Communication)
JCPS Justice, Crime Prevention and Security Cluster
Cyber Security Hub (Private Sector)
GovCERT (CSIRT)
Industry Local CSIRT's
DOD Department of Defence
State Security National Intelligence Agency
Justice and South African Police Service

Figure 5: Proposed cyber security structure South Africa

4. Conclusion and future work

This paper gives an overview and analysis of cybersecurity organisational structures in the USA, UK, Estonia, South Korea, and China. Based on the result, we proposed a cybersecurity organisational structure for South Africa. In addition, a methodology for the implementation of the cybersecurity strategy and policy in South Africa is also considered. An organisational structure for effective governance was proposed as well as the Cyber Security Awareness Toolkit for national security (CyberSAT) as an operational guideline that could be used in the implementation of South Africa's proposed cybersecurity policy. In order to implement a cybersecurity strategy South Africa needs a formal approach to describe the cybersecurity environment.

We are in the process of developing an ontology for the cybersecurity strategic environment which we will use to support the implementation process. An ontology is a technology that allows one to encode a shared understanding and representation of a domain.

References

Allan, D. (2010) *Defence Minister to stress need for cyber-defence*. Retrieved 15 February, 2011, [online] http://www.techwatch.co.uk/2010/11/09/defence-minister-to-stress-need-for-cyber-defence/

Amit, I. I. (2011) *Information Security Intelligence Report: A recap of 2010 and prediction for 2011*. Retrieved 5 February, 2011, [online] www.Security-Art.com

Boyd, C. (2010) *Why Estonia Is the Poster Child for Cyber-Security*. Retrieved 5 February, 2011, [online] http://news.discovery.com/tech/why-estonia-is-the-poster-child-for-cyber-security.html

Burghardt, T. (2012) *The Launching of USA Cyber Command (CYBERCOM)*, Offensive Operations in Cyberspace. Retrieved 24 February 2012, [online] http://www.globalresearch.ca/index.php?context=va&aid=14186

Constitution of the Republic of South Africa. (1996) *Chapter 11, Governing Principle 198*. Retrieved. from Deloitte & Touch. (2010) *National Cybersecurity Strategies*. Paper presented at the GOVCERT.NL symposium

Duncan, G. (2011) *New cyberattacks hit South Kore.a* Retrieved 5 March, 2011, [online] http://www.digitaltrends.com/computing/new-cyberattacks-hit-south-korea/

Espiner, T. (2010) *UK's cyberdefence centre gets later start date*. Retrieved 21 February, 2011, [online] http://www.zdnet.co.uk/news/security-threats/2010/03/10/uks-cyberdefence-centre-gets-later-start-date-40082405/

Evron, G. (2008) *Estonian Cyber Security Strategy Document: Translated and Public*. Retrieved 15 February, 2011, [online] http://www.circleid.com/posts/estonian_cyber_security_strategy/

Ghernouti-Helie, S. (2010). *A national strategy for an effective cybersecurity approach and culture*. Paper presented at the 2010 International Conference on Availability, Reliability and Security.

Guardian. (2010) Chinese army to target cyber war threat (Publication.: http://www.guardian.co.uk/world/2010/jul/22/chinese-army-cyber-war-department)

Hsiao, R. (2010) China's Cyber Command? *China Brief, Volume: 10 Issue: 15*.

Jablonsky, D. (1997) National power. *Parameters, 27*, 34-54.

Jansen van Vuuren, J., Phahlamohlaka, J., & Brazzoli, M. (2010) The Impact of the Increase in Broadband Access on National Security and the Average citizen. *Journal of Information Warfare, 5*, 171-181.

Krekel, B. (2009) *Capability of the People's Republic of China to Conduct Cyber Warfare and Computer Network Exploitation*: DTIC Document.

Leyden, J. (2010) *South Korea sets up cyberwarfare unit to repel NORK hackers.* Retrieved 4 March 2011, [online] http://www.theregister.co.uk/2010/01/12/korea_cyberwarfare_unit/

Nguyen, A. (2011) Government launches cyber security hub pilot [online]

Phahlamohlaka, L. J., Jansen van Vuuren, J. C., & Coetzee, A. J. (2011). *Cyber security awareness toolkit for national security: an approach to South Africa's cyber security policy implementation*. Proceedings of the First IFIP TC9/TC11 Southern African Cyber Security Awareness Workshop (SACSAW), Gaborone, Botswana.

SA Goverment Gazette. (2010) *South African National Cyber Security Policy*.

Stienon, R. (2010) *Seven Cyber Scenarios that should keep you up at night.* [online] http://threatchaos.com/

Stokes, M. A., Lin, J., & Russell Hsido, L. C. (2011) *The Chinese People's Liberation Army Signals Intelligence and Cyber Reconnaissance Infrastructure: Project 2049*.

Tiirmaa-Klaar, H. (2010). *International Cooperation in Cyber Security: Actors, Levels and Challenges*. Proceedings of Cyber Security 2010, Brussels.

United States. Executive Office of the President. (2009) *Cyberspace Policy Review, Assuring a Trusted and Resillient Information*. Retrieved 12 February 2011 [online] http://www.whitehouse.gov/assets/documents/Cyberspace_Policy_Review_final.pdf

Zorz, Z. (2010) South Korea preparing for cyber war. Retrieved 5 February 2011, 2010, [online] http://www.net-security.org/secworld.php?id=8722

Security and Safety Education in the Czech Republic and eSEC-Portal User Requirements

Roman Jasek, Radek Vala and David Malanik
Tomas Bata University in Zlín, Zlín, Czech Republic
jasek@fai.utb.cz
vala@fai.utb.cz
dmalanik@fai.utb.cz

Abstract: Education in the field of security and safety is at different level in each country and it is divided into many distinct areas. Teaching of the certain fields of security is strategic due to countries' geographical location, local business or cultural and social aspects. The European Union countries are currently missing some kind of online public database, which would cover the security and safety field and bring together students, researchers and experts interested in the subject. The main aim of the preparation of portal "eSEC - Competency Based e-portal of Security and Safety Engineering" (eSEC-portal), is to establish a web system, which would serve as a new tool in the process of learning, for both students and professionals, in fields of security studies, safety studies and crisis management. One of the challenges, while establishing the web portal, was to analyse the conditions of teaching security and safety in the selected EU countries. That is why the first part of this paper focuses on current educational situation in the Czech Republic representing an EU member, however, it is above all intended for international audience such as students and teachers outside the Czech Republic. The second part introduces a qualitative SWOT analysis of education in security and safety and connection with services, which will be offered by the prepared eSEC-portal. The third part predicts and quantitatively analyses the profiles of potential future eSEC-portal users together with their possible requirements. Users are split into several groups such as students, pedagogues, scientists and experts. The analysis discusses the user requirements for content, interactivity, user-friendly extremity and graphical surroundings of eSEC-portal´s modules, all in the connection to their already existing ICT skills. Input data have been collected through on-line questionnaires and the sample has been represented by 144 respondents. The results of this user requirements analysis, enriched with similar data gathered in another five EU countries have been fundamental for functional design and final content of the eSEC-portal. The certain limitation of the study lies in the research sample, which consists mainly of respondents who are familiar with the Information security area. Conclusion discusses the other challenge for eSEC-portal developers and maintainers which is in keeping the portal alive.

Keywords: security safety education eLearning web portal eSEC

1. Introduction

The field of security and safety is highly discussed in area of education recently. While in the past the content of printed material was relevant for several years, nowadays the information changes much faster. Security and safety studies include number of different study subjects and whole topic is broadly interdisciplinary. Moreover, education in this field is at different level in each country of European Union and students, teachers, researchers and experts are currently missing some on-line public database, which offers relevant, current and complex information about each area of security and safety. Fortunately, the expansion of the Internet to public domain enables new forms of education. The new approach presented with the on-line information database should eliminate the two main problems of current information state. Currently, the biggest disadvantage in connection with education process is length of the period of publishing new information and findings. Ristvej et al (2010) states that 'the better possibility is the interval of three months (scientific and professional journals), but the worse possibility is twenty-four months (several international conferences). Second biggest disadvantage is limited number of possible readers and limited feedback from them, which is very important for self reflex of the author.'

Newly prepared on-line portal "eSEC – Competency Based e-portal of Security and Safety Engineering" (eSEC-portal) should address people from this field of interests and offer a new eLearning tool which will solve the disadvantages listed above.

One of the challenges, while establishing the web portal, was to analyse the conditions of teaching security and safety in the selected EU countries. First part of this article summarizes one part of the document Study analysis, which was created as a particular research of educational systems of different EU member countries. It is focused on current educational situation in the Czech Republic (authors' homeland) representing an EU member, however, it is above all intended for international audience such as students and teachers outside the Czech Republic.

The second part introduces a qualitative SWOT analysis of education in security and safety in mentioned country, which proves possible interests in prepared eSEC-portal. The third part quantitatively analyses the profiles and possible requirements of potential future eSEC-portal users to predict general portal requirements.

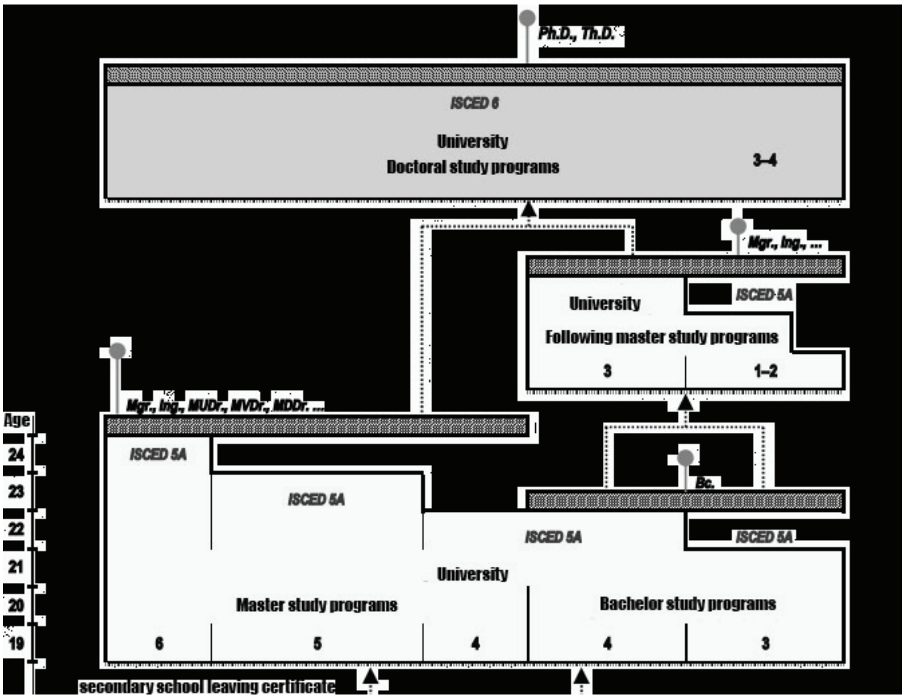


Figure 1: Schema of higher education in the Czech Republic (nicm.cz, 2011), with ISCED level description

2. Higher educational system in the Czech Republic

For the purpose of establishing eSEC-portal general educational system in the Czech Republic as an EU member country was analyzed. These findings are part of the document Analysis study (eSEC Consorcium, 2010), which is published on the temporary webpage of prepared eSEC-portal (www.esecportal.eu). Information below confirms that the Czech Republic should be a fully-fledged EU partner country in the area of education and is able to offer study opportunities and mobility programs for bachelor, master or Ph.D. students. This research is one of the important parts of eSEC-portal project, especially for setting up the study programs and mobility on-line database. This database has ambitions to become the unique EU information source, for students, teachers or researchers, interested in international collaboration or study.

2.1 University/higher educational system

According to legislation of the Czech Republic, higher education institutions are of two types – **University** and **Non-university**.

University type of higher education institutions provides all levels of study programs (Bachelor's, Master's and Doctoral) and carry out associated scholarly, research, development, artistic or other creative activities.

Non-university type of higher education institutions usually offer only Bachelor's degree programs, but never Doctoral programs.

Study programs consist of three grades: **Bachelor**, **Master** and **Doctoral** study program and each of these programs may be attended in a day form study (full-time), a distance form or a combination of both.

The minimum requirement for admission to a higher education institution is secondary education completed with a school-leaving examination. The requirement for admission to a Master's program following on from a Bachelor's program is the successful completion of the Bachelor's program. The requirement for admission to a Doctoral program is the successful completion of a Master's program. Schema of higher education with recommended time period of study is shown in the figure 1.

2.1.1 Bachelor's study programs

These types of programs are focused on professional training and provide a basis for studies in Master's study programs. The programs last 3-4 years and graduates acquire the qualification. Bachelor's degree studies end with a final state examination and part of which is usually the defense of a thesis.

The graduates receive these academic degree: *Bachelor* (Bc.) or *Bachelor in the field of art* (BcA.).

2.1.2 Master's study programs

These types of programs aims to provide theoretical knowledge based on the latest scientific findings, research and development, at mastering their applications and to develop creative skills. These study programs follow on from Bachelor's degrees. Their standard length is 1-3 years. If the character of the study programs so requires, accreditation can be granted to a Master's degree programs (4-6 years long), which does not follow on from a Bachelor's one. Therefore the standard study program is between three and seven years. Master's degree studies end with a state examination part of which is the defense of a thesis.

The academic degrees awarded are: Magister (Mgr.), Magister in the field of art (MgA.), Engineer (Ing.), Engineer Architect (Ing. arch.). Students of medicine, veterinary medicine and hygiene are the exception. They finish their studies with a state examination – Rigorous Exam – and they are awarded the degree: Doctor of medicine (MUDr.), Dentist (MDDr.), or Doctor of veterinary medicine (MVDr.).

2.1.3 Doctoral study programs

These types of programs can follow the completion of a Master's program. It is aimed at scientific work, independent creative activity in the area of research and development or at independent theoretical and creative skills in art. It is offered solely in universities and lasts 3-4 years. Doctoral studies finish with a state doctoral examination and defense of a thesis. The degree for all fields of study is: *Doctor* (Ph.D.) or *Doctor Theology* (Th.D.).

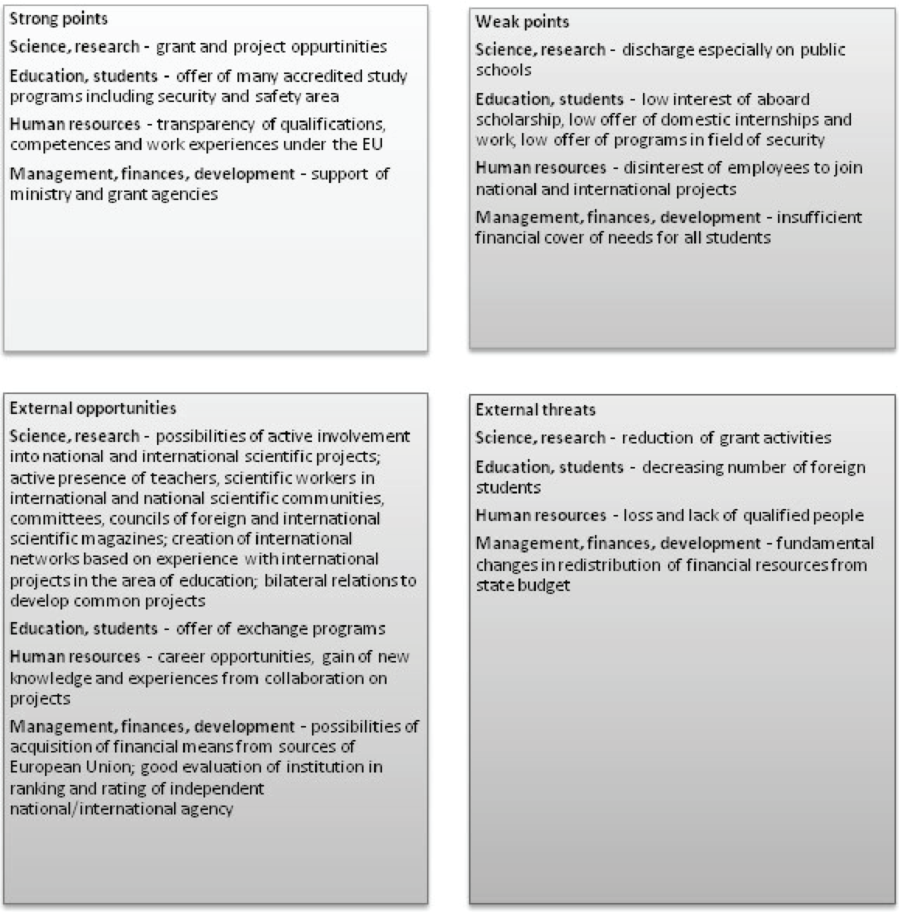
Having been awarded the academic degree, „Magistr" students can sit a state examination Rigorous Exam which includes the defense of a thesis – Rigorous Work. Those who have passed the exam and successfully defended their thesis are awarded the degree of doctor; the abbreviation differing according to the field of study: Doctor of Laws (JUDr.), Doctor of philosophy (PhDr.), Doctor of natural sciences (RNDr.), Doctor of pharmacy (PharmDr.), Doctor of theology (ThDr.), Doctor of catholic theology (the degree is – licentiate) – (ThLic.).

2.1.4 Main security and safety study program areas in Czech Republic

It is possible to study in these areas of security in Czech Republic: Private security, Fire protection, Economic security, Environmental security, Health safety, Industry security, Transport infrastructure security, Energy infrastructure security, Information security and International security.

2.1.5 Czech Republic as a contributor in security and safety field

Especially Tomas Bata University in Zlín (TBU), Faculty of Applied Informatics as a partner of eSEC-portal is the main portal content contributor from Information security and Health and Occupation Safety. Researchers from TBU have created a part of *Basic Security and Safety Glossary of Terms* (GT) from areas mentioned above, which should be according to the user requirements (Chapter 4.1) part of on-line eSEC-portal. GT is a kind of basic keyword-description dictionary from different areas of security and safety, where the definitions were taken from the relevant sources (ISO/EIC etc.) and the purpose is to bring one reliable source with definitions from different areas of security and safety and also in 7 different languages.



Strong points

Science, research - grant and project oppurtinities

Education, students - offer of many accredited study programs including security and safety area

Human resources - transparency of qualifications, competences and work experiences under the EU

Management, finances, development - support of ministry and grant agencies

Weak points

Science, research - discharge especially on public schools

Education, students - low interest of aboard scholarship, low offer of domestic internships and work, low offer of programs in field of security

Human resources - disinterest of employees to join national and international projects

Management, finances, development - insufficient financial cover of needs for all students

External opportunities

Science, research - possibilities of active involvement into national and international scientific projects; active presence of teachers, scientific workers in international and national scientific communities, committees, councils of foreign and international scientific magazines; creation of international networks based on experience with international projects in the area of education; bilateral relations to develop common projects

Education, students - offer of exchange programs

Human resources - career opportunities, gain of new knowledge and experiences from collaboration on projects

Management, finances, development - possibilities of acquisition of financial means from sources of European Union; good evaluation of institution in ranking and rating of independent national/international agency

External threats

Science, research - reduction of grant activities

Education, students - decreasing number of foreign students

Human resources - loss and lack of qualified people

Management, finances, development - fundamental changes in redistribution of financial resources from state budget

Figure 2: SWOT analysis of education in the area of security and safety in the Czech Republic

3. SWOT analysis of education in the area of security and safety in the Czech Republic

Before starting the eSEC-portal project each project partner member performed the SWOT analysis of security and safety education in their homeland. As inputs of the SWOT analysis both data from student questionnaires and researchers opinions was considered. Results of the SWOT analysis are shown in figure 2.

We can see that one of the noticeable external opportunity in security and safety education in the Czech Republic is the possibility of collaboration with institutions from other countries. For example for scientist and researchers could be very interesting to be involved into national or international scientific project or to present their work more often on international level. However, the real situation shows, that students or researchers are not very active currently in international projects, due to lack of information about these opportunities. This is one of the potential benefits of prepared eSEC-portal website, to bring this information about educational institutions in other countries or events from the security and safety area.

4. Analysis of eSEC-portal user requirements

To obtain user requirements for designing and developing an on-line portal within project eSEC 320 questionnaires were sent to respondents in the Czech Republic from which 144 (45%) returned completed. Time period of survey realization was May-June 2010.

The questionnaires were designed to obtain especially these objectives:

- Justification of need to create eSEC portal from view of target groups.
- Requirements of target groups on individual modules of eSEC portal in terms of contents, form and significance.

4.1 Evaluation of questionnaires and general user requirements

Questionnaires contained 13 questions. The text below is divided into three sections. The first section describes general information about respondents the second shows respondents' opinions connected with education or eLearning the third discusses which portal functionality or modules respondents prefer and the fourth section discusses which objectives eSEC-portal should fulfill according to the questionnaires.

4.1.1 General information about respondents

In terms of respondents' age the most frequent is the group up to the age of 30 which makes 86% of all respondents, the second most common group was people between 31 and 40 years (only 6%). Proportion of users by their gender showed that there were 81% of male and 19% of female respondents. From the perspective of practical experience in the field, 63% of respondents have none in the area of security and safety. This result is caused by the fact, that there were 76% of student, but it can be stated that some of them already have practical experience in this field. Other significant number is a proportion of 30% respondents who have experience in the range of 1 to 10 years. In terms of profession the second largest number (in addition to students) is 11% of teachers and then 10% of experts from practice.

4.1.2 Respondents' opinions connected with education or eLearning

From evaluation of contingency "What are, in your view, deficiencies by yourself used electronic information sources, considering respondent´s profession", can be stated that currently available electronic sources do not provide complex information, that respondent requires to his/her activity/study. Moreover there is insufficient updating of published information, resources often do not provide interpretative dictionary of stated technical terms and require membership in a particular organization/institution or university. 49% of respondents prefer attendance/distant eventually combined form of education, 31% of respondents prefer self – study, 16% eLearning and classical textbooks' texts and 4% prefer eLearning with limited distant approach.

4.1.3 Respondents' portal function/modules preferences

From preparing modules of eSEC-portal, potential future users, respondents of the survey, would most welcome these modules or functionality: professional journal focused on practice; preparing events – trainings, seminars, exhibitions, conferences; published technical works/textbooks/surveys; thesis and qualification works; supply and demand on labor market; list of educational institutions and terminological directory.

4.1.4 Evaluation of questionnaire regard to set aims

From the first goal point of view (justification of need to create eSEC-portal in terms of target groups), as many as 87% of respondents have not met during their study/profession with a software/ tool/service, which would help them choose study stay/fellowship/subjects based on by themselves required parameters. As 59% of survey respondents stated, at most activities employment/ study requires further form of lifelong education, which is, however, voluntary and they have an interest in it. After adding 6% of respondents who stated that education is compulsory, eventually it is required by law, this group makes together 65% of respondents. Other supporting argument for creating eSEC-portal is, that as many as 80% of respondents have no particular idea at which institution they would obtain required education, if they decide for fellowship or foreign stay.

Form the second goal perspective (Requirements of target groups on individual modules of eSEC portal in terms of contents, form and significance.), potential future eSEC-portal users would most welcome from preparing eSEC portal modules these ones: professional journal focused on practice, preparing events – trainings, seminars, exhibitions, conferences, workshops, published technical works/ textbooks/surveys, thesis and qualification works from division/field, supply and demand of labor market and list of educational institutions from the field. According to individual target groups, students would in addition welcome terminological dictionary, chat and blogs. Teachers and researchers would also prefer scientific peer reviewed journal.

5. Conclusion

Within the project co-funded by the European Commission 'eSEC – Competency Based e-portal of Security and Safety' a research of current state of higher education in the Czech Republic (and other project partner countries) was carried out. eSEC Consortium had published the complete results on the temporary webpage www.esecportal.eu, in Study Analysis document. In addition, questionnaire survey for relevant data collection was performed to confirm that the respondents are highly interested in some eLearning information database, such prepared eSEC-portal is.

External opportunities from the SWOT analysis (Chapter 3) of security and safety education in the Czech Republic shows that there is high demand for international collaboration. Scientist and researchers are interested into possibility of active involvement into international scientific projects, communities, committees, etc., but the real situation shows that their activity is limited by the information and communication lack. From the management, finances and development point of view the international collaboration brings to the involved institutions a possibility of acquisition of financial means from sources of European Union, and moreover, these involved institutions are able to get better ranking from independent national or international agencies.

Opportunities listed above also confirm that prepared eSEC-portal is able to create a great tool for networking and cooperation between international universities, companies or students and there are a lot of potential users in the Czech Republic.

Analysis of future eSEC-portal user requirements was carried out using the questionnaire survey (Chapter 4) with two main objectives: Firstly, to justify the need to create eSEC portal from view of target groups, and secondly, to obtain requirements of target groups on individual modules of eSEC portal in terms of contents, form and significance. In addition, the user module and functionality preferences were obtained and the respondents most wanted these modules: professional journal focused on practice; preparing events – trainings, seminars, exhibitions, conferences; published technical works/textbooks/surveys; thesis and qualification works; supply and demand on labor market; list of educational institutions and grant agencies and terminological directory. These results show the usefulness of prepared eSEC-portal, and the possible strength of this unique tool, which is in the worldwide quick and easy accessibility of information from discussed field. Possible restriction of the international European information portal should be language and translation. The localization of prepared portal should be in English. It means, every visitor or possible user, should at least be able to understand English. However, Welcome word, Terms and Conditions will be translated into several languages of project partner members. Moreover other important information should be added to portal in native language of every user, but this could lead to creation of country-specific areas in the portal. One of the future challenges of partner countries involved in creation and developing of eSEC-portal is, to keep this portal alive. It is necessary to promote this new eLearning tool and to attract new contributors. Not only universities (students and researchers) but also companies from filed (experts), should be interested, thanks to the possibility to introduce themselves or to find possible workers with appropriate competencies.

Acknowledgements

This paper is part of project Competency Based e-portal of Security and Safety Engineering – eSEC – number: 502092-LLP-1-2009-1-SK-ERASMUS-EMHE, funded by the European Commission with support from the Education, Audiovisual and Culture Executive Agency (EACEA) and reflects only the views of the authors

References

eSEC Consorcium (2010). *Study Analysis,* Available: http://fsi.uniza.sk/kkm/esec/study_analysis.pdf [10 Jan 2012]

eSEC Consorcium (2011). *European Basic Security and Safety Glossary of Terms,* ISBN: 978-80-554-0328-1

Chráska, M. (2007) *Metody pedagogického výzkumu: základy kvantitativního výzkumu*, Praha: Grada Publishing a.s.

Národní informační centrum pro mládež (nicm.cz), (2011) Schéma vzdělávacího systému České republiky, [Online], Available: http://www.nicm.cz/schema-vzdelavaciho-systemu-ceske-republiky [6 Jan 2012]

Pahl, N. and Richter, A. (2009) *SWOT Analysis - Idea, Methodology And A Practical Approach*, Norderstedt: GRIN Verlag.

Ristvej et al. (2010) 'eSEC - Competency Based e-portal of Security and Safety', *Proceedings of the 7th International ISCRAM Conference – Seattle*, USA

Explaining Politico-Strategic Cyber Security: The Feasibility of Applying Arms Race Theory

Eli Jellenc
Verisign (iDefense Security Intelligence), London, UK
ejellenc@idefense.com

Abstract: This paper applies existing theories of arms races to explain key problems of cyber security among nation-states. The motivating empirical problem (on which current theoretic approaches exhibit no grasp) is the rapid and pervasive increase in cyber security preparations and malicious activity in a politico-strategic context. Moreover, policy debates concerning cyber security are fragmented, often incoherent, and lack consensus on how even to judge the effectiveness of policies relative to problems. Primarily to blame for such confusion is the absence of a common, tested conceptual framework. To date, neither scholars, nor policymakers, nor industry professionals have succeeded in deriving robust theoretic approaches to unify various islands of useful empirical research on the geopolitics of cyber security. However, such theoretic models and approaches are not only available, but show promise for adaptation to specific cyber security problems. Cyber security is a new and complex class of issues, but it is, in the end, technologically mediated social behavior, and as such, it can be fruitfully studied as such. The increasing significance of cyber security to the international system makes such inquiry necessary to help inform the behavior of states, firms, and other stakeholders. The research effort summarized here hypothesizes that a cyber arms race is indeed underway, and all available findings provide no significant disconfirming challenge. In fact, the approaches to cyber security by the world's major powers (and many minor ones) indeed exhibit all the features of a novel, multilateral arms race of hitherto unseen complexity: a global cyber arms race. Moreover it is improbable that anything short of revolutionary legal or diplomatic initiatives will prevent severe, near-term increases in cyber conflict activity. More generally, this research shows how existing theoretical work in political science, sociology, and communications theory can lend new rigor to the study of cyber security as a geopolitical issue.

Keywords: cyber warfare, cyber conflict, cyber espionage, cyber security governance, arms races, social science theory

1. Introduction

Cyber threats can no longer be accurately described as an "emergent" international security problem; rather, they already exert significant, occasionally primary, influence upon today's most relevant international security developments, and with ever-increasing frequency and consequence. The current politico-strategic cyber threat environment has become so pervasive, complex, and dynamic as to constitute a salient destabilizing force both within many nations and across the international system: a truly global security problem regarding which no state has yet derived any solution. This condition emerged extremely rapidly, with precursors dating back to early1990s, but, in earnest, only since 2007.

The single most relevant, yet hitherto overlooked, theoretic framework to help explain cyber security dynamics among nations is the research tradition on arms races. A few mainstream journalists (Riley and Vance 2011; Glenny 2011), academics (e.g., Goldsmith 2010), and some public-facing cyber security researchers (Goldstein 2010, Hypponen 2012) have used the terms "arms race "to describe geopolitical cyber security developments, but always informally and uncritically; none refer to the topic's vast scholarly literature. Recently, McAfee Corp. surveyed policymakers and business leaders, of whom 57% (of n=250) believe that states face a cyber arms race (SDA 2012). While telling as a barometer of opinion, surveys alone cannot answer whether formal arms race dynamics actually inform cyber security statecraft. Thus, three core research questions present themselves:

- Are states in a cyber arms race?
- If so, what are its characteristics and dynamics?
- What do these entail for states' strategies and future trajectories?

1.1 Methods

Because this essay's basic purpose is to show the operation of formal arms race dynamics insets of empirical phenomena (viz., states' cyber espionage and conflict behavior), the primary method used is process tracing of strategic-interaction sequences, which matches the concrete facts of states' behavior to the general theoretic elements of arms race models. Process tracing through event catalogues (Tilly 2004) is also an obvious methodological choice because arms races are phenomena that operate via mechanisms of reciprocally interactive causation, the presence of which supports the arms race hypothesis, while an absence would falsify it.

Beyond the core argument of this essay, it is clear that the authors' empirical findings derive from additional investigative and analytical methods (only summarized herein due to word-count requirements). Structured case study comparisons and simple statistical correlation helped to establish the rapid intensification of cyber conflict capabilities development. Of course, formal game theory underpins the entire edifice of arms race theory. These methods were employed in the larger analysis of which this essay merely represents a summary of findings; the author will provide them upon request.

2. What is an arms race?

2.1 Definition and theoretic elements of arms races

Arms races are pervasive, complex phenomena in geopolitics that, in their essence, reflect the influence upon states' military postures of the "security dilemma": the International Relations thesis specifying that states' efforts to increase their own security often make others less secure, which provokes reactive counterbalancing, and so on. Arms races have proven amenable to precise analysis using formal frameworks, methods, and models now comprising a venerable tradition of theoretic and practical inquiry. Despite great variety in perspective and method, an essential theoretic core (Rathjens 1969), first made explicit in Lewis Richardson's *Arms and Insecurity* (1960), lends coherence to the diverse literature. Arms races minimally consist of these elements (Gray 1971):

- Two or more states, each perceiving potential or actual antagonism
- Each state considers force posturing with direct reference to other states' postures
- Competition occurs quantitatively (men, weapons) and/or qualitatively (men, weapons, organization, doctrine, deployment)
- Increases and improvements are rapid (relative non-hostile interaction patterns)

Singer (1970) later outlined relevant sequences that characterize arms races; his empirical findings showed that states follow the prescribed sequence in response to or anticipation ofother states' corresponding steps along the same sequence:

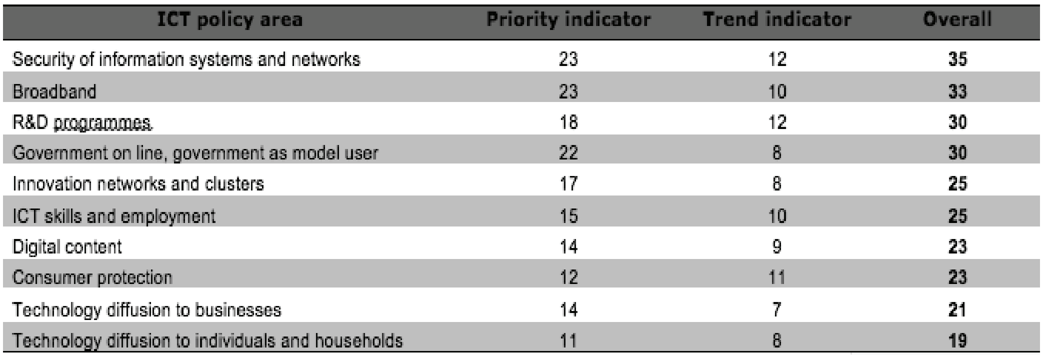
- Weapons Acquisition (i.e. capabilities development):
- *Government leaders' expression of necessity (entails assessment of adversary's developments)*
- *R&D and testing*
- *Budget appropriation*
- *Contracting*
- Weapons Deployment:
- *Policymakers' strategizing*
- *Emplacement and maintenance*
- Weapons Use:
- *Planning*
- *Doctrinal formulation*
- *Field maneuvers and "live fire" testing*
- *Small-scale employment in actual conflict*

Such are the steps constituting the "racing" in an arms race, and whom ever is "ahead" at the onset of conflict will, *ceteris paribus*, "win". Thus, to demonstrate the current occurrence of a cyber arms race, the next section shows that the empirical events in states' cyber security developments since 2005 correspond to the theoretic sequences of interaction above.

2.2 Evidence that arms race theory applies to current geopolitical cyber security

Cyber security now exerts a transformative influence on the strategic dynamics of the international system and on some nation-states' internal security, both reflecting and amplifying nation-states' increased prioritization of cyber operations. The following catalogue of the most important developments in geopolitical cyber security over the last decade shows a trajectory of accelerating hostility and armament that states have undertaken with explicit reference to one another and to the threat environment holistically. The evidence shows a clear arms race dynamic: early experimentation leads to ominous successes → victims, attackers, and emulating states react by increasing interest → new capabilities spur intensified operations→ further incidents compel greater concern → the cycle iteratively escalates.

As of2005, only 5 states (the US, France, China, Russia, South Korea, and Israel) included politico-strategic cyber security among their strategic concerns in any official documentation, and then only marginally. By 2008, only the US had developed a comprehensive government cyber strategy, the CNCI, which was deeply classified (NSC 2009). As of 2012, however, cyber security stands among the top-5 national security priorities of every strategically relevant and technically developed state, over 25 in all. Moreover, it has also become the single most important *technology* policy issue, as shown in the following OECD findings:



ICT policy area	Priority indicator	Trend indicator	Overall
Security of information systems and networks	23	12	**35**
Broadband	23	10	**33**
R&D programmes	18	12	**30**
Government on line, government as model user	22	8	**30**
Innovation networks and clusters	17	8	**25**
ICT skills and employment	15	10	**25**
Digital content	14	9	**23**
Consumer protection	12	11	**23**
Technology diffusion to businesses	14	7	**21**
Technology diffusion to individuals and households	11	8	**19**

Figure 1: OECD rankings of technology policy priorities (2011)

The reasons for this shift are straightforward. Early cyber espionage campaigns, such as "Titan Rain" and "Moonlight Maze" (late 1990s), showed expert communities in and outside of government, especially in the US, China, and Russia, that cyber threats were becoming geopolitically significant; these countries' early experiences help explain why they became first-movers in developing capabilities, doctrine, and organizations to conduct cyber operations. States that closely follow strategic developments in the US (i.e. France, the UK, and Australia)or in China (i.e. Taiwan, South Korea, and North Korea) became the next cohort to develop serious cyber capabilities, following Singer's standard arms race sequence.

Nationalistic DDoS attacks in April, 2007, by Russian entities against Estonian government and corporate ICT systems galvanized international attention upon politically motivated cyber threats .In 2008, increased alarm followed the cyber attacks against Tbilisi, coinciding precisely with the advent of the Russo-Georgian War. Since these "game-changing" attacks, dozens of government studies worldwide and independent policy analysts in the US (Lewis 2012), Europe (Dunn-Cavelty 2012), and elsewhere have documented a sharp and global increase of the objective frequency of salient, consequential cyber attacks.

Since 2007, no less than 15 states have established cyber command headquarters or dedicated military units focused on cyber defense and offense (Dunn-Cavelty 2012; Demchak 2011; Nakashima 2012; ENISA 2011). In 2007, governments' spending on cyber security stood at less than $10 billion globally; by 2012, it exceeded $50 billion with further increases certain (Dunn-Cavelty 2012). The US alone spends over $5 billion annually. The UK designated $1 billion over the 4 years from 2010 to 2014, and in 2011, Iran committed a similar amount. Indeed, every strategically relevant country reports growing cyber security budgets, the most oft-counted metric in formal analyses of arms races.

To respect this publication's brevity standards, and because the evidence base of such phenomena is vast, only a few illustrative examples are necessary here. Numerous scholarly and industry resources provide ample documentation of the trends listed above across over 20 nation-states, thus rendering unnecessary a comprehensive recapitulation of evidence (McAfee 2012; EIU 2012; Jellenc in Verisign-iDefense 2011; Lewis 2011; Kilmburg 2011; Carr 2011; Deibert, et al. 2009, Deibert, et al. 2010). Moreover, dozens of official sources from multiple governments cite the rapid, destabilizing ascent of cyber security as a primary justification for intensifying their own efforts and budgetary appropriations (for brevity's sake, citing only several extensive sources suffices here to evidence the argument: Giles 2009; US DoD 2011; US State Department 2011; Lynn 2010; US Office of the President 2010; China Daily 2012; ANSSI 2011; German Interior Ministry 2011; UK Cabinet Office 2011; South Korean National Security Committee 2011; Indian MoIT 2011).

As of early 2012, geopolitical cyber security competition constitutes a primary issue complicating the strategic dynamics of almost all enduring dyadic rivalries between powerful states, including:

- US and Russia (NCIX 2011; Giles 2010, 2012)
- US and China (Mulvenon 2009; Thomas 2010; Krekel 2010; Stokes, et al 2011)
- US and Iran (Reuters 2012)
- US and North Korea (Reuters 2012)
- South Korea and North Korea (Laurence 2011)
- South Korea and China (Lewis, 2012)
- Russia and China (Interviewee 2010)
- Japan and China (Muncaster 2011)
- India and China (Villeneuve, IWM 2010; *Times of India* in Lewis 2012)
- India and Pakistan (AL Jazeera 2010)
- Russia and Georgia (Arquilla 2012)
- Iran and Israel (Katz 2012)
- Syria and Israel (Follathand Stark 2009)

Driven by precisely the developments above, states have actively formulated cyber security policies and strategies in explicit reference to one another's activities and intentions. The following states adopted national cyber security strategies prior to 2011 (note: these refer to the creation of cyber defense strategies exhibiting a national security, not counter-crime, focus):

- the US (2008 and 2010)
- Russia (Giles 2010)
- The UK (UK Cabinet Office 2009)
- Australia (2008)
- Canada (Public Safety Canada 2009)
- Estonia (Estonian MoD 2008)
- Singapore (2009)
- Switzerland (2010)

Conspicuously absent from this list is the PRC. Extensive, credible analysis has found that the PRC fields a mature cyber security strategy under development since the mid-1990s, though never officially articulated(Thomas 2009; Mulvenon 2009; Stokes, et al, 2011). In any case, the PLA's justification for building a robust information warfare capability is rooted in its apprehensiveness regarding US forces' performance in the First Gulf War and pursuit of RMA-oriented developments (Mulvenon 2009). Subsequent modifications to these policies or strategies flowed from policy debates containing specific references to the capabilities and intentions of other leading states (NCIX 2011, Russian MoD, 2012). Moreover, the following nation-states adopted or updated their strategies or working drafts in 2011 or early 2012, in every case referring to other states' capabilities as a reason for escalation:

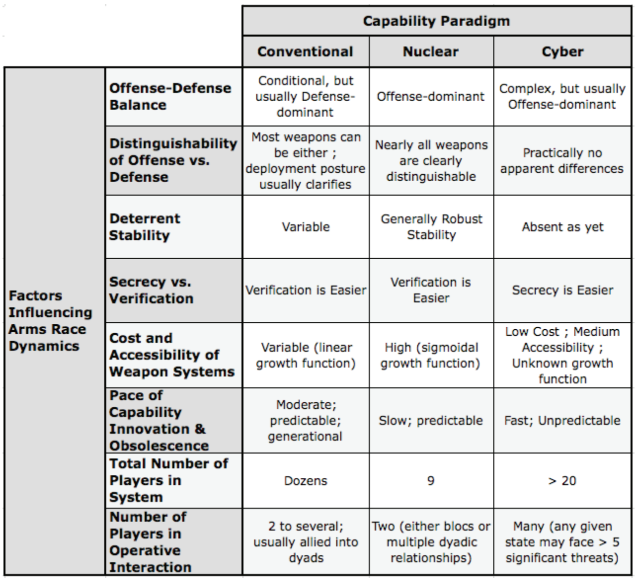
- Argentina (Ortiz 2011)
- Austria (Mader 2011)
- Czech Republic (Czech MoI 2011)
- Denmark [draft] (Klimburg 2011)
- France (ANSSI 2011)
- Finland [draft] (O'Dwyer 2011)

- Germany (German MoI 2011)
- Israel (Muhareb 2011)
- India (Indian MIT 2011)
- Japan [draft] (NISC 2011)
- Latvia (Klimburg, Timaa-Klarr 2011)
- Lithuania (ENISA 2011)
- Netherlands (Dutch MoSJ 2011)
- New Zealand (NCSC 2011)
- Norway [draft] (O'Dwyer 2010)
- Poland [draft] (Klimburg, Timaa-Klarr 2011)
- Russia (Russian MoD 2012)
- South Africa (Guy 2011)
- South Korea (Valdez 2011)
- The UK [update] (Say 2011)
- US State Department (US DoS 2011)
- US DoD (US DoD 2011)

Beyond official political institutions, geopolitical cyber security implicates private sector actors as intensely as governments. A sizeable literature now catalogues the private sector's rush to provide the cyber capabilities that governments increasingly demand (Riley 2011; Grow 2011; Gross 2011; Wikileaks 2011), another metric widely applied in past arms race studies. The evidence above demonstrates that dozens of states' behaviors reflect an arms race dynamic, fulfilling the core 4 theoretic elements and Singer's phases. Furthermore, extensive researches found no significant evidence to falsify the hypothesis that arms race processes characterize geopolitical cyber security today.

3. Varieties of arms races: today's cyber arms race is which?

Beyond the conceptual bedrock, the literature on arms races (for broad surveys, see Schelling 1960 & 1966; Glaser 1995, 2000, 2004; Downs 1991; Evangelista 1988; Fischer 1984) identifies other factors that amplify or mitigate the intensity, stability, and conflict-proneness of arms races. The most relevant factors are listed below with examples from three past armament paradigms :



		Capability Paradigm		
		Conventional	Nuclear	Cyber
Factors Influencing Arms Race Dynamics	Offense-Defense Balance	Conditional, but usually Defense-dominant	Offense-dominant	Complex, but usually Offense-dominant
	Distinguishability of Offense vs. Defense	Most weapons can be either ; deployment posture usually clarifies	Nearly all weapons are clearly distinguishable	Practically no apparent differences
	Deterrent Stability	Variable	Generally Robust Stability	Absent as yet
	Secrecy vs. Verification	Verification is Easier	Verification is Easier	Secrecy is Easier
	Cost and Accessibility of Weapon Systems	Variable (linear growth function)	High (sigmoidal growth function)	Low Cost ; Medium Accessibility ; Unknown growth function
	Pace of Capability Innovation & Obsolescence	Moderate; predictable; generational	Slow; predictable	Fast; Unpredictable
	Total Number of Players in System	Dozens	9	> 20
	Number of Players in Operative Interaction	2 to several; usually allied into dyads	Two (either blocs or multiple dyadic relationships)	Many (any given state may face > 5 significant threats)

Figure 2: Ancillary characteristics arms races across 3 primary capabilities paradigms

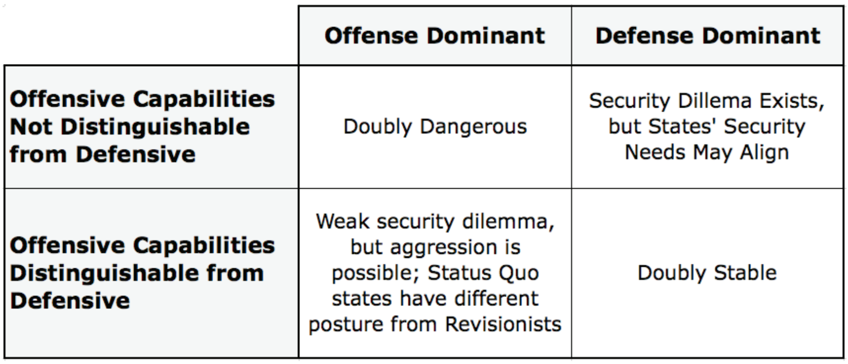
These summary findings emerge from a more detailed study (available upon request) than word-count restrictions for this essay permit to be presented. However, to summarize, the current cyber arms race exhibits, for every relevant factor in the table above, characteristics encouraging greater intensity, instability, and proneness to conflict. The following sections outline the most consequential of these factors.

3.1 Characteristics of offense-defense balance

Two aspects of the Offense-Defense balance interact to yield important consequences for the character and endstate of arms races:

- Offense-Defense Dominance: when offensive capabilities dominate, the security dilemma intensifies, and armament proceeds more urgently, thus incentivizing pre-emptive attacks which increase the probability of conflict.
- Distinguish ability of Offensive from Defensive Forces: offensive and defensive capabilities that are indistinguishable tend to increase uncertainty of force-effectiveness. Secrecy and uncertainty, even if intended as defenses or deterrents, tend to increase mistrust and insecurity, making conflict more likely

The interactive effects of these factors are examined thoroughly in Jervis's classic article, *Cooperation Under the Security Dilemma* (Jervis 1978).



	Offense Dominant	**Defense Dominant**
Offensive Capabilities Not Distinguishable from Defensive	Doubly Dangerous	Security Dillema Exists, but States' Security Needs May Align
Offensive Capabilities Distinguishable from Defensive	Weak security dilemma, but aggression is possible; Status Quo states have different posture from Revisionists	Doubly Stable

Figure 3: Interaction effects of offense-defense capability dominance and distinguish ability (adapted from Jervis 1978)

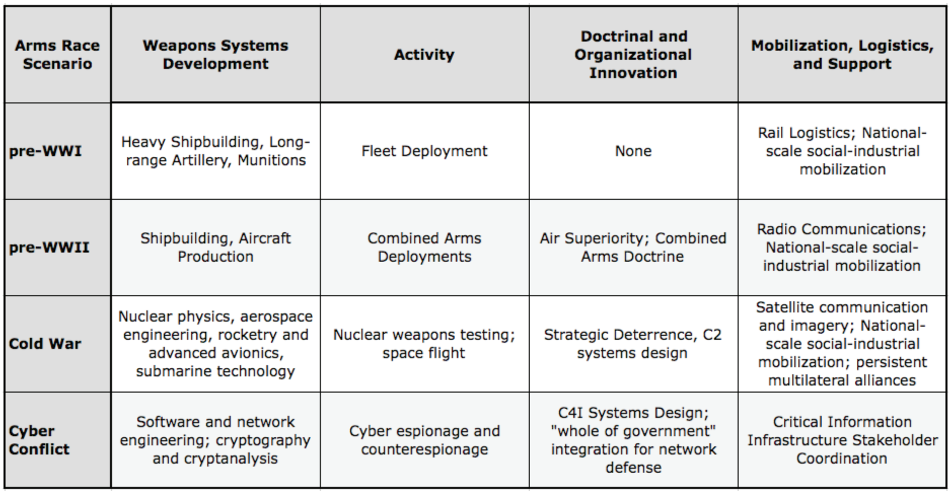
Unfortunately, the dominance of offensive cyber operations is one of the few points of consensus across the expert literature on geopolitical cyber security. Almost every well-informed analysis of the offense-defense balance at the macro-scale concludes that offense does, indeed, carry advantages (see, *inter alia*, Geers 2011; Kramer, et al. 2009; Carr 2011; Klimburg 2011; Libicki 2009). One US National Research Council expert reflects the conventional wisdom of most policymakers, stating, "since you don't know how to do good defense, you can't prevent offensive dominance. And you can't do good deterrence because effective retaliation is hard, so if you want to take advantage of cyberspace, you will do offensive operations for non-defensive purposes" (Lin 2010).

Distinguishing offensive cyber attack capabilities from defensive ones is practically impossible, not so much because the two types capabilities *appear* to be similar, but because they *are*, in fact, composed of almost exactly the same elements. The equipment, the software, core skills of personnel, and many tactics are all essentially identical or profoundly similar. Thus, for a state with any defensive capability whatsoever, credibly proving the absence of offensive capabilities founders on the ease of hiding, disguising, or rapidly converting into offensive systems.

3.2 Technical aspects of cyber security privilege secrecy and uncertainty

The following two sections examine the importance of secrecy and espionage-oriented activity as properties emerging from complex interactions among the technical and socio-organizational systems in the current cyber arms races. The following graphic lists the empirical features of the present cyber arms race compared to previous ones:

The intrinsic importance of secrecy of cyber capabilities subtly influences most aspects of a cyber arms race. Specifically, once one is aware of a particular tactic, it becomes much easier to defend against it. This is not the case regarding conventional or nuclear armaments; of course, awareness usually helps, but merely knowing an adversary's air force strength does not necessarily translate into an effective defense. With cyber capabilities, however, awareness of an opponent's capability constitutes a significant proportion of overall defensive competence; this is why that which we usually classify as “cyber espionage ”is, in fact, the crucial element of“ cyber conflict ”capabilities.



Arms Race Scenario	Weapons Systems Development	Activity	Doctrinal and Organizational Innovation	Mobilization, Logistics, and Support
pre-WWI	Heavy Shipbuilding, Long-range Artillery, Munitions	Fleet Deployment	None	Rail Logistics; National-scale social-industrial mobilization
pre-WWII	Shipbuilding, Aircraft Production	Combined Arms Deployments	Air Superiority; Combined Arms Doctrine	Radio Communications; National-scale social-industrial mobilization
Cold War	Nuclear physics, aerospace engineering, rocketry and advanced avionics, submarine technology	Nuclear weapons testing; space flight	Strategic Deterrence, C2 systems design	Satellite communication and imagery; National-scale social-industrial mobilization; persistent multilateral alliances
Cyber Conflict	Software and network engineering; cryptography and cryptanalysis	Cyber espionage and counterespionage	C4I Systems Design; "whole of government" integration for network defense	Critical Information Infrastructure Stakeholder Coordination

Figure 4: Comparison of systemic arms races over the past 120 years (sources: Murray and Knox 2001; Kramer, et al, 2009)

3.3 Evidence: Cyber espionage is secretive, offensive, and counts for most of what we call “cyber conflict”

Capable nation-states and their proxies now, systematically and towards strategic ends, conduct cyber espionage against international economic institutions and private companies to pursue strategic macroeconomic advantages in addition to purely politico-strategic goals (NCIX 2011). Since 2005, “Advanced Persistent Threats” (APTs), a euphemism for state-conducted cyber espionage, purportedly from China, now constitute an officially acknowledged threat against (at the very leaser least) the US (NCIX 2011; USCC 2010, 2011; Blair 2011; McAfee 2011), the UK (Leppard 2007), France (Grow citing Wikileaks 2011),Germany (Grow citing Wikileaks 2011), Japan (Muncaster 2011), Taiwan (Lai 2011), and India (Deibert, et al, 2009, 2010), among others. Likewise, Chinese and even US officials (Nakashima 2009; Lewis 2010) acknowledge that the US also conducts information operations against the PRC and others.

The complexity of strategic competition in cyberspace emerges from the fact that states conduct cyber operations against multiple other states while simultaneously resisting infiltration by multiple others, some of whom are not overt enemies. In 2009, a leaked UK Army Intelligence Report warned that at least two NATO allies were among more than 20 nations who had hacked restricted UK government systems (Rayment 2009). One month later, *Der Speigel* published findings that the *Bundesnachrichtendienst* (BND, Germany’s preeminent intelligence agency) had infiltrated upwards of 30 countries, some of them nominal allies (von Holger 2009).The history of arms races contains few cases of armed threats from both allies and adversaries at once; as such, models of today’s cyber arms race must be adapted to consider allies’ cyber espionage as a threatening element.

Both offensive and defensive cyber conflict capabilities arise fundamentally from prior success in cyber espionage praxis (Campen 2000; Cullather 2006; Williams, et al, 2010), as the following cases suggest:

- Russia’s disruption of Georgian government systems in 2008 (Arquilla 2012).
- Israel’s cyber attacks against air defense systems to support airstrikes against Syria’s nuclear facility in 2008 (Eshel 2010)
- The Stuxnet worm’s sabotage of Iranian nuclear centrifuges

Stuxnet, in particular, exemplifies a hybrid threshold between cyber espionage and covert cyber conflict; F-Secure’s Mikko Hypponen accurately notes, "…we now have proof that states are investing serious resources into the development of next-generation viruses" (Glenny 2011).In short, cyber espionage is the “running” of the cyber arms race, just as nuclear physics research and rocketry constituted the “running” of the Cold War nuclear arms race.

It is no surprise, then, that intelligence agencies occupy privileged positions of authority and functional competence regarding cyber security in nearly every state, and in direct proportion to each state's overall maturity in strategic cyber security. Every state, partially excepting Japan, listed above in Section 2.2 as having developed national cyber strategies also shows evidence that their intelligence communities act as centers of gravity regarding cyber strategy formation and capabilities development.

3.4 The crucial complication of a high number of players

The most unique feature of today's cyber arms race, however, is that the number of relevant participants reaches into the dozens, making this the most populous arms race on record (Dipert 2010),in contrast to the rival dyads or allied dyadic blocs that characterized previous arms races. No significant historical arms races have consisted of more than a handful of states, so nearly all formal arms race models posit only low-n player populations.

The few exceptions are those in which game theorists have used computer simulations to model the abstract efficacy of various strategies in mixed-motive games resembling arms rages, altering the number of players from 2 to thousands. All such large-n simulations (Axelrod 1984, 1990; Ball 2002) suggest the following strategic principles:

- Complex interactions amongst large-n populations' strategic choices lead to unpredictable results.
- Large-n arms race dynamics tend towards the absence of stable static equilibria, instead favoring dynamic, multiple equilibria (Ball 2002).
- Players must thus alter strategies at critical times to remain viable.
- Large-n arms races undergo occasional periods of ineluctable instability.

Groundbreaking work by mathematical complexity theorists has further illustrated the disruptive and unpredictable potential of calculated attacks against large, scale-free networks, such as the Internet. Their analyses suggest that large-n arms races, if they ever devolve into even minor conflict, could produce widespread (if temporary) and severe consequences (Albert, Barabasi, et al., 2000; Cohen 2001) affecting vast numbers of players.

For the reasons above, a large-n population of players suggests that the current cyber arms race is deeply intractable and prone to complex shifts in instability. States' actions and reactions thus simultaneously influence one another and the entire population of shifting threats, rather than against only one or a handful of others; this causes perceived threats and consequent incentives for armament to grow geometrically rather than linearly.

4. Conclusions and predictions

Having established that arms race theories are indeed applicable to cyber security dynamics, it becomes possible to explore how a more nuanced grasp of such problems can inform further inquiry and policy planning.

Arms races can end in three ways: conflict, mutual amity via trust-building, or exhaustion/outstripping of some players. From the analysis above, conflict is a likely outcome for many interacting dyads and multi-actor sets in the current cyber arms race. Now that most major states are committing serious resources and will to pursue strong, versatile cyber capabilities, every currency unit thus spent by one nation constitutes a threat to multiple adversaries and even some allies simultaneously. Concordantly, with little to curb escalatory dynamics, arms race theory suggests that information operations among states will accelerate in frequency and severity. Thus, a key prediction is that the future of cyber conflict will grow more complex and dangerous.

Additional study is necessary to determine additional implications of arms race models for the study of cyber conflict. Beyond this, further inquiry into communication theory's intersection with socio-political organization and strategic interaction can help in refining frameworks to assess states' effectiveness in "doing" cyber security, thus informing not only the mechanics of the cyber arms race, but also what may be done about it. Four obvious starting points are:

- Actor Network Theory (especially that of Bruno Latour), to model the core principles of cyber-enabled states and societies

- "Complex Systems Theory of Social Communication" (e.g.Leydesdorff, 2000), to explain how individual organizations function, structure themselves, and relate to their broader social and macro-global context
- "Securitization" theory (Waever, 1995; Taurek 2010; Erickson 2001), to model how the various stakeholders within states prioritize cyber security and form policy responses.
- "Communication Power" frameworks (Castells, 2009), to model incentives for conflict and cooperation across vast digital networks; also posits an incipient theory of "power" in networked organizations that can help explain why some cyber strategies perform better than others.

The work of theorizing cyber security is only beginning, but fortunately, there are ample resources ready-to-hand from all across the social sciences. Without reference to these established theoretic traditions, cyber security analysts risk wasting precious time and unnecessarily increasing the error rate of their assessments and predictions.

References

(ANSSI) AgenceNationale de la Sécurité des Systems d'Information (National Information Systems Security Agency). (2011) Défenseetsecurité des systems d'information – Stratégie de la France (Defense and Security of Information Systems – Strategy of France), February 15, 2011. ANSSI, Paris, http://www.ssi.gouv.fr/IMG/pdf/2011-02-15_Defense_et_securite_des_systemes_d_information_strategie_de_la_France.pdf

(NCIX) Office of the National Counterintelligence Executive. (2011) Report to Congress on Foreign Economic Collection and Industrial Espionage, October, Government Publishing Office, Washington DC, http://www.ncix.gov/publications/reports/fecie_all/Foreign_Economic_Collection_2011.pdf

(SDA) Security and Defense Agenda Staff, (2012) Cyber Security: The Vexed Question of Global Rules, Report and Survey Conducted for McAfee, Inc., January 30, Brussels, http://www.securitydefenceagenda.org/Portals/14/Documents/Publications/SDA_Cyber_report_FINAL.pdf

AL Jazeera. (2010) "India and Pakistan in cyber war", AlJazeera.com, December 4, http://www.aljazeera.com/news/asia/2010/12/20101241373583977.html

Alexander Klimburg. (2011) "Mobilising Cyber Power," Survival, vol. 53, no. 1, International Institute of Strategic Studies, London, UK

Arquilla, John. (2012) "Cyber war is already upon us", Foreign Policy, February 27, http://www.foreignpolicy.com/articles/2012/02/27/cyberwar_is_already_upon_us?page=full

Axelrod, Robert. (1984) The Evolution of Cooperation, TKTK

Ball, Philip. (2002) Critical Mass: How One Thing Leads to Another, FGS Press, London

Barabasi, A.-L., R. Albert, H. Jeong. (2000) "The Internet's Achilles' heel: Error and attack tolerance in complex networks", Nature, vol. 406, pp. 378–382, July

Blair, Dennis C. (2010) "Annual Threat Assessment of the US Intelligence Community for the Senate Select Committee on Intelligence",Office of the Director of National Intelligence, Washington, DC

Campen, Alan D., and Douglas H. Dearth. (2000) Cyberwar3.0 : human factors in information operations and future conflict, AFCEA International Press: Fairfax, VA

Carr, Jeffrey. (2011) Inside Cyber Warfare, 2nd Edition, O'Reilly: California

Castells, Manual. (2009) Communication Power, Oxford University Press: London

Cimbala, Stephen J. (2011) "Nuclear Crisis Management and 'Cyberwar': Phishing for Trouble?",Strategic Studies Quarterly (Spring 2011):117-31

Cohen, R., K. Reez, D. Ben-Avraham, and S. Havlin, (2001) "Breakdown of the Internet under intentional attack, " Phys. Rev. Lett. vol. 86 , no. 16, pp. 3682–3685,

Cornish, P, R Hughes, and D Livingstone (2009) "Cyberspace and the National Security of the United Kingdom", London, UK: Chatham House.

Crosston, Matthew D. (2011) "World Gone Cyber MAD: How 'Mutually Assured Debilitation' Is the Best Hope for Cyber Deterrence." Strategic Studies Quarterly (Spring 2011):100-16.

Cullather, Nick. (2006) "Bombing at the Speed of Thought: Intelligence in the Coming Age of Cyberwar," Intelligence and National Security , vol. 18, no. 4

Deibert, Ron and Rafal Rohozhinski, et al. (2009) Access Controlled. MIT Press: Cambridge, MA

Demchak, Chris C. (2011) Wars of Disruption and Resilience: Cybered Conflict, Power, and National Security. Athens, Georgia, USA: University of Georgia Press.

Dipert, Randall. (2010) "The Ethics of Cyberwarfare", Journal of Military Ethics, Vol 9, No 4,

DoD Staff. (2011) "Department of Defense Strategy for Operating in Cyberspace", June, Department of Defense, Washington, D.C.

Downs, George. (1991) "Arms Races and War," in Tetlock, Philip E. , Jo L. Husbands, Robert Jervis, Paul C. Stern, and Charles Tilly, eds., Behavior, Society, and Nuclear War, Vol. 2 Oxford University Press, New York, NY, pp. 82–84

Dunn-Cavelty, Miriam. (2012) "The militarization of cyber security as a source of global tension", in Andreas Wegner, ed., Strategic Trends 2012, ETH Zurich CSS, Zurich. http://www.sta.ethz.ch/Strategic-Trends-2012/The-militarisation-of-cyber-security-as-a-source-of-global-tension

Economist Staff. 2010. "The future of the internet: A virtual counter-revolution." The Economist, September 2.

ENISA. (2011) Lithuania Country Report, ENISA Country Report Archive, http://www.enisa.europa.eu/activities/stakeholder-relations/files/country-reports

Eshel, David. (2010) "Cyber--Attack Deploys in Israeli Forces", Aviation Week, September 15,

Espiner, Tom. 2009. "UK launches dedicated cybersecurity agency - one of whose functions will be to develop a cyberattack capability." ZDNet UK online, July 25.

Evangelista, Matthew A. (1988) Innovation and the Arms Race: How the United States and the Soviet Union Develop New Military Technologies, Princeton University Press, Camden, NJ

Fischer, Deitrich. (1984) "Weapons Technology and the Intensity of Arms Races", Conflict Management and Peace Science, vol. 8, no. 1, Fall, pp. 46-69

Follath, Erich and Von Holger, Stark. (2009) "How Israel Destroyed Syria's Al Kabir Nuclear Reactor", Der Spiegel. Nov. 2, http://www.spiegel.de/international/world/0,1518,658663-2,00.html.

Geers, Kenneth. (2011) "Sun Tzu and Cyber War", Cooperative Cyber Defense Center of Excellence Occasional Article, February, Tallinn, http://www.ccdcoe.org/articles/2011/Geers_SunTzuandCyberWar.pdf

German Federal Ministry of the Interior. (2011) Cyber Security Strategy for Germany, March 11, http://www.bmi.bund.de/SharedDocs/Downloads/DE/Themen/OED_Verwaltung/Informationsgesellschaft/cyb er_eng.pdf;jsessionid=365A25B8FF75170FF9566570016DDEA9.1_cid165?__blob=publicationFile

Giles, Keir. (2011) "Information Troops: Russia's Cyber Command", Presentation at the 3rd International Conference on Cyber Conflict, June 9, NATO Cooperative Cyber Defense Centre of Excellence, Tallinn, Estonia, http://conflictstudies.academia.edu/KeirGiles/Papers/718775/_Information_Troops_-_a_Russian_Cyber_Command

Glaser, Charles L. (2000) "The Causes and Consequences of Arms Races," in Nelson W. Polsby, ed., Annual Review of Political Science, Vol. 3, pp. 251–276.

Glaser, Charles. (2004) "When are Arms Races Dangerous?",International Security, vol. 28, no. 4, p. 44-84

Glenny, Mihsa. (2011) "The Cyber Arms Race Has Begun", The Nation, October 11, http://www.thenation.com/article/163923/cyber-arms-race-has-begun

Goldsmith, Jack. (2010) "Can We Stop the Global Cyber Arms Race", New York Times. February 1. retrieved from http://web.mit.edu/ecir/pdf/goldsmith-race.pdf

Goldstein, Guy-Philippe. (2010) How cyberattacks threaten real-world peace, TedxParis Conference Presentation. February. http://www.ted.com/talks/guy_philippe_goldstein_how_cyberattacks_threaten_real_world_peace.html

Gross, Joseph M. (2011) "A Declaration of Cyber War", Vanity Fair, April, http://www.vanityfair.com/culture/features/2011/04/stuxnet-201104

Grow, Brian and Marc Hosenball, (2011) "In cyberspy vs. cyberspy China has the edge", Reuters Special Report, April 14, http://www.reuters.com/article/2011/04/14/us-china-usa-cyberespionage-idUSTRE73D24220110414

Guy, David. (2011) "Cyber security policy will go before cabinet for approval this year", DefenceWeb, accessed October 10, 2011, http://www.defenceweb.co.za/index.php?option=com_content&view=article&id=13783:cyber-security-policy-will-go-before-cabinet-for-approval-this-year&catid=48:Information%20&%20Communication%20Technologies&Itemid=109

Horowitz, Michael C. (2011) "Information Age Weaponry and the Future of Security in East Asia", Global Asia, vol. 6, no. 2, Summer, http://www.globalasia.org/V6N2_Summer_2011/Michael_Horowitz.html?PHPSESSID=d3b725fb17906592b9b49f2aac21ac08

Jervis, Robert. (1978) "Cooperation Under the Security Dilemma", World Politics, iss. 2

Katz, Yaakov. (2012) "IDF building elite hacker teams amid cyber threat", Jerusalem Post, January 27, http://www.jpost.com/Defense/Article.aspx?id=253487

Klimburg, Alexander and HeliTirmaa-Klaar. (2011) "Cybersecurity and Cyberpower: Concepts, Conditions and Capabilities for Cooperation and Action within the EU", June, Study for the European Parliament. http://www.evi.ee/lib/cyber.pdf

Knox, M. and Murray, W., eds (2001). The Dynamics of Military Revolution, 1300–2005. Cambridge: Cambridge University Press.

Kramer, Franklin D., Stuart H. Starr, and Larry K. Wentz (eds.) Cyberpower and National Security, NDU Press: Washington DC

Krekel, Brian. (2009) Capability of the People's Republic of China to Conduct Cyber Warfare and Computer Network Exploitation, Northrop Grumman Corp., McLean VA

Lai, Anthony. (2011) "Balancing the Pwn Trade Deficit: APTs in Asia", Presentation at DEFCON Conference, July, https://media.defcon.org/dc-19/presentations/Lai-Wu-Chiu-PK/DEFCON-19-Lai-Wu-Chiu-PK-APT-Secrets-2.pdf

Laurence, Jeremy. (2011) "North Korea hacker threat grows as cyber unit grows", Reuters, June 1, http://www.reuters.com/article/2011/06/01/us-korea-north-hackers-idUSTRE7501U420110601

Leppard, David. (2010) "China bugs and burgles Britain," Times Online, January 31, http://www.timesonline.co.uk/tol/news/uk/crime/article7009749.ece

Lewis, James A. (2011) "Cybersecurity: Assessing the Immediate Threat to the United States", Testimony Before the US House Oversight and Government Reform Committee. May 25, Washington DC, http://csis.org/files/ts110525_lewis.pdf

Lewis, James A. (2012) “Significant Cyber Incidents Since 2006”, Center for International and Strategic Studies, Washington DC, February 1, http://csis.org/files/publication/120307_Significant_Cyber_Incidents_Since_2006.pdf

Libicki, MC. 2009. Cyberdeterrence and Cyberwar. Washington DC: Rand Corporation.

Lynn, William J. (2010) "Defending A New Domain: The Pentagon's Cyberstrategy." Foreign Affairs, September/October, http://www.foreignaffairs.com/articles/66552/william-j-lynn-iii/defending-a-new-domain

Mader, Georg. (2011) “Austria Unveils New Security Doctrine Amid Neutrality Concerns”, Jane’s Defence Weekly, 8, March

McAfee. (2011) Operation ShadyRat. McAfee Whitepaper, September 7, https://kc.mcafee.com/corporate/index?page=content&id=KB72605

Indian Ministry of Information and Technology. (2011) National Cyber Security Strategy of India, January, http://www.mit.gov.in/content/cyber-security-strategy . Accessed September 29, 2011.

Ministry of the Interior of the Czech Republic. (2011) Cyber Security Strategy for the Czech Republic for the 2011 – 2015 Period, August, Czech Ministry of the Interior, Prague.

Muhareb, Mahmoud. (2011) “Israel and Cyber Warfare”, Doha Institute Book Review, September 29, http://english.dohainstitute.org/Home/Details/5ea4b31b-155d-4a9f-8f4d-a5b428135cd5/c82f6a5e-6ba7-40c0-ba42-819b34167108

Mulvenon, James. (2009) "PLA Computer Network Operations: Scenarios, Doctrine, Organizations, and Capability," in Kamphausen, Roy, David Lai, and Andrew Scobell (eds.), Beyond the Strait: PLA Missions Other Than Taiwan, Strategic Studies Institute, U.S. Army War College, Carlisle, PA

Muncaster, Phil. (2011) “Japanese parliament hit by cyber attack from China”, V3 News, October 25, http://www.v3.co.uk/v3-uk/news/2119840/japanese-parliament-hit-cyber-attack-china

Murray, Williamson and Macgregor Knox. (2001) The Dynamics of Military Revolutions 1350 -2050. Cambridge University Press, London

Nakashima, Ellen (2012) "U.S. Accelerating Cyberweapon Research." The Washington Post, March 13, http://www.washingtonpost.com/world/national-security/us-accelerating-cyberweapon-research/2012/03/13/gIQAMRGVLS_story.html

Nakashima, Ellen. “China Proves to be Aggressive Foe in Cyberspace.” Nov. 11, 2009. The Washington Post. http://www.washingtonpost.com/wp-dyn/content/article/2009/11/10/AR2009111017588.html.

Netherlands Ministry of Security and Justice. (2011) Dutch Cyber Security Strategy 2011, ENISA Security Strategy Archive, February, http://www.enisa.europa.eu/media/news-items/dutch-cyber-security-strategy-2011/view

New Zealand National Cyber Security Center. (2011) National Cyber Security Center of New Zealand, June, http://www.ncsc.govt.nz/

Northrop Grumman. (2010) Capability of the People's Republic of China to Conduct Cyber Warfare and Computer Network Exploitation, report prepared for The US-China Economic and Security Review Commission (USCC), Northrop Grumman Corporation, McLean, VA

(NSC) National Security Council. (2009) The Comprehensive National Cybersecurity Initiative, White House Cyber Security Website, http://www.whitehouse.gov/cybersecurity/comprehensive-national-cybersecurity-initiative , accessed February 1, 2012

O’Dwyer, Gerard. (2010) “Norway Drafts Cyber Defense Initiative,” DefenseNews. January 26,

O’Dwyer, Gerard. (2011) “Finland to Develop Cyber Defense ‘Counterpunch’,” DefenseNews. October 20, http://www.defensenews.com/story.php?i=8003134&&s=TOP

Ortiz, Javier Ulises. (2008) “Argentina: The Challenge of Information Operations”, IOSphere, http://www.au.af.mil/info-ops/iosphere/08special/iosphere_special08_oritz.pdf

Public Safety Canada. (2009) Canadian National Cyber Security Strategy, June

Rathjens, George W. (1969) "The Dynamics of the Arms Race", Scientific American, April

Rayment, Sean. (2009) “Britain Under Attack from 20 Foreign Spy Agencies Including France and Germany”, The Daily Telegraph, Feb. 8, http://www.telegraph.co.uk/news/newstopics/politics/defence/4548753/Britain-under-attack-from-20-foreign-spy-agencies-including-France-and-Germany.html

Reuters. “Pentagon Sees N Korea Cyber Threat, 2012 Provocations”, (2012) Reuters.uk, March 28, http://uk.reuters.com/article/2012/03/28/uk-usa-korea-north-idUKBRE82R1B420120328

Richardson, Lewis Frye. (1960) Arms and Insecurity, Quadrangle Press: Chicago

Sample, Susan G. (1997) “Arms Races and Dispute Escalation: Resolving the Debate,” Journal of Peace Research, Vol. 34, No. 1, February, pp. 7–22;

Say, Mark. (2011) “Cabinet Office backs trusted computing”, The Guardian, October 21, http://www.guardian.co.uk/government-computing-network/2011/oct/21/cyber-security-strategy-trusted-computing

Schelling, Thomas (1960). The Strategy of Conflict, Harvard University Press, Cambridge, MA.

Schelling, Thomas. (1966) Arms and Influence, Yale University Press, New Haven, CT.

Schweller, Randall. (1996) “Neorealism’s Security Bias: What Security Dilemma?”,Security Studies, Spring, Vol. 5, no. 3, p. 117

SecDec Group. (2011) Collusion and Collision: Searching for Guidance in Chinese Cyberspace, September, SecDev Group, Toronto, http://www.scribd.com/doc/65531793/Collusion-Collision

Singer, J. David. (1970) “The Outcome of Arms Races”, Proceedings of IPRA 3rd General Conference, IPRA: Oslo

Stokes, Mark A. Jenny Lin, and L.C. Russell Hsiao. (2011) The Chinese People's Liberation Army Signals Intelligence and Cyber Reconnaissance Infrastructure, November, Project 2049 Institute, Arlington, VA, http://project2049.net/documents/pla_third_department_sigint_cyber_stokes_lin_hsiao.pdf

Swiss Federal Council. (2010) Bericht des Bundesrates an die Bundesversammlungunber die Sicherheitspolitik der Schweiz, June 23, p. 32

Taurek, Rita. “Critical Approaches to Security: Telling the story of securitization theory,” Proceedings of the Central and Eastern European Studies Association Convention, University of Tartu, Estonia, June 25-26, 2006. http://www.ceeisaconf.ut.ee/109100

UK Cabinet Office, Cyber Security Strategy of the United Kingdom, June, 2009, at http://www.cabinetoffice.gov.uk/media/216620/css0906.pdf

US Department of Defense. (2011) Strategy for Operating in Cyberspace, July 14, http://www.defense.gov/news/20110714cyber.pdf

US Department of State. (2011) International Strategy for Cyberspace: Prosperity, Security, and Openness in a Networked World, May, The Office of the President of the US, Washington DC, http://www.whitehouse.gov/sites/default/files/rss_viewer/international_strategy_for_cyberspace.pdf

Valdez, Adrienne. (2011) “South Korea Outlines Cyber Security Strategy”, FutureGov Asia-Pacific, August 13, http://www.futuregov.asia/articles/2011/aug/13/south-korea-outlines-cyber-security-strategy/

Verisign analyst interviews with Japanese MoD and National Information Security Council leaders (identities protected). Tokyo, Japan, September 26, 2011.

Villeneuve, Nart, Information Warfare Monitor and Shadowserver Foundation. (2010) Shadows in the Cloud: An Investigation into Cyber Espionage 2.0, Information Warfare Monitor, Toronto, April 6, http://www.forbes.com/sites/firewall/2010/04/06/shadows-in-the-cloud/

Von Holger, Stark. (2009) “BND Infiltrated Thousands of Foreign Computers”, Der Spiegel, March 7, http://www.spiegel.de/netzwelt/web/0,1518,611954,00.html

Wæver, Ole (1995) “Securitization and Desecuritization”. In Lipschutz, R.D. On

Security. New York: Columbia University Press

Wamala, Frederick. (2011) ITU National Cybersecurity Strategy Guide. September, May, United Nations, Geneva, http://www.itu.int/ITU-D/cyb/cybersecurity/docs/ITUNationalCybersecurityStrategyGuide.pdf

Williams, Phil, Timothy Shimeal, and Casey Dunlevy. (2010) "Intelligence Analysis for Internet Security," Contemporary Security Policy vol. 23, no. 2

Towards an Automated Security Awareness System in a Virtualized Environment

William Aubrey Labuschagne[1] and Mariki Eloff[2]
[1]Defence, Peace, Safety and Security, Council for Scientific and Industrial Research, Pretoria, South Africa
[2]School of Computing, University of South Africa, Pretoria, South Africa
wlabuschagne@csir.co.za
Eloffmm@unisa.ac.za

Abstract: A majority of African Internet users do not have access to the Internet. The lack of infrastructure in rural areas affects Internet usage. Since costs are high and the bandwidth low, these factors encourage users to access the Internet using shared resources. This is an efficient solution to access the Internet. However users might not be aware of the security threats that exist on using shared resources. Many companies provide security solutions to automatically protect resources on the network and security awareness training to users. This ensures that users are aware of the security threats and provide methods to mitigate them. These measures are useful in a corporate environment where funds exist to enable these security solutions. Public platforms, for example Internet Cafes and schools, allows multiple users to access the Internet using shared resources. This implies that multiple people will use the same computer to perform required tasks. Numerous security threats exist within the Internet sphere that could affect users utilizing shared resources these include but are not limited to viruses, keyloggers and phishing attacks. This shared environment could provide a platform that promotes the spread of virus infections. Users using these platforms should be made aware of these threats and monitor the effectiveness of the security awareness campaign. This paper proposes a system used to address these issues from a single platform. The Shared Public Security Awareness (SPSA) system is an automated virtualized system used to determine the current security awareness levels of users on a shared platform accessing the Internet. The system uses virtual machines to provide users with access to the Internet, assess the security awareness levels of the users, determines if any web browser components were infected by web based malware during browsing sessions, provides users with access to security related material affecting the users and provide reports on online behaviour. This paper evaluates the proposed SPSA system as a mechanism to conduct a security awareness campaign in a shared resource environment while providing a capability to analyze the online behaviour of users that affects the security of this environment.

Keyword: internet cafes, security awareness, security training, virtualized environments, cyber literacy, internet

1. Introduction

The Internet provides a vast range of information resources and services which form part of everyday life. Usages include but are not limited to searching for information, conducting business, paying bills and the purchase of goods. Moreover the development of human capital has been identified as an important economical performance indicator in rural areas (Agarwal, Rahman & Errington 2009).This can be attained with access to knowledge available on the Internet. However the high adoption and use of the Internet by citizens introduced an opportunity for cyber criminals to utilize this platform to coordinate cyber attacks with the intention to cause damage. Kim identified a comprehensive list which includes loss of money, defamation, invasion of privacy, physical harm, loss of time and psychological damage (Kim et al. 2011). Most companies provide security measures against these attacks for their employees. In most instances employees require to attend security awareness training programs to equip them with strategies on how to mitigate these cyber attacks when encountered. Furthermore the network infrastructure is secured with expensive security solutions. Therefore, people working at companies are best equipped against cyber attacks. Users in rural areas are in a disadvantages position. In most instanced these users do not have ownership of resources to access the Internet. The cost of access to the Internet and equipment inhibits ownership of resources like computers. The need to access the Internet was addressed with entrepreneurial initiatives which provide access to the Internet with the use of shared resources. This implies multiple users using the same computer to access the Internet. An example of this implementation are schools and Internet cafes. However, users sharing the same resources could assist in the spread of malware infections. In the event of discovering a malware infection at these establishments, the services provided need to be suspended which has an effect on revenue for the owners. Another issue which could be encountered at these establishments is security literacy.

Most of these users are not aware of the cyber threats that are devised and deployed by criminals. Security awareness programs are used to educate the users and provide them with measures to identify and mitigate the threats encountered. Grobler studied the cyber awareness initiatives in South Africa (Grobler et al. 2011). She reported on initiatives by the Council for Scientific and Industrial Research (CSIR), the University of Pretoria (UP), the University of Fort Hare (UFH) and Nelson Mandela Metropolitan University (NMMU). The CSIR collaborated with the University of Venda to raise security awareness in the rural areas by developing content which addresses cyber security topics and training community members which then in turn will train the community. The UP project, PumaScope, equips students with the required security knowledge to educate scholars at identified schools. UFH tested the proficiency levels of the user in a particular area. NMMU addresses educating users through the use of games and eLearning platforms to provide access to security awareness content for a wider audience. A need has been identified to provide an automated platform which incorporates the core ideas of the mentioned initiatives a platform that could be used to determine the proficiency levels of the users and provide access to resources to improve security awareness in rural areas.

This paper looks at the design of an automated tool, Shared Public Security Awareness (SPSA) system, which promotes security awareness in rural areas where the community uses shared computer resources to access the Internet. These resources can be located at schools or Internet café where access to the Internet is provided through the use of shared computers. Establishments would be used throughout the paper that references the communal area where the shared computer resources are located. The deployment of the SPSA system addresses three primary functions: The first function is to provide the capability to conduct a security awareness program which consist of assessing the literacy of the users and deliver the security awareness topics to the users. The second function analyzes the online behaviour of users and the collection of malware which would assist in developing strategies which addresses the security threats encountered at these establishments. The third function provides a turnkey solution which automates the functionality of the SPSA system with limited intervention from personal to administrate the system.

The rest of the paper is organized as follows: section 2 summarizes research related to the component identification of the SPSA system. The main contribution: the design of the SPSA system is outlined in section 3. Conclusions and future work are discussed in section 4.

2. Related research and underlying concepts

The SPSA virtualized and collection system requirements are discussed in this section. These establishments provide resources which enables user's access to the Internet through the use of web browsers. Cyber attackers have adopted attacking strategies which include automated exploitation of computer systems without the intervention of the user. The resources used by these establishments must be protected against possible attacks originating from the Internet. Also a mechanism is required to identify the threat and evaluate the actions performed by users which initiated these attacks. The system should exhibit the following capabilities:

- A robust and automated architecture which ensures availability and configurability of the system. This is achieved with the implementation of virtualization and customization of existing systems (See Section 2.1).
- The identification of threats originated from users visiting malicious web sites, accomplished with the collection and analysis of data generated during browsing sessions (Section 2.2).

2.1 Virtualization, automation and customization

The SPSA system underlying architecture consists of virtual machines. Bell defines a virtual machine as software that functions as a computer without physically being a computer (Bell, Lintumaa 2011). The use of virtual machines provides numerous of advantages.

The implementation of virtualized environments is cost effective. England proposed a model for deploying virtual machines as a securing mechanism for the enterprise desktop (England, Manferdelli 2006). Some organizations require users to conduct classified work. In these organizations the users will be provided with two physical computers: one to conduct normal duties and the other for classified duties. This is not cost effective. The use of virtual machines would allow both functionalities to be conducted within a virtualized environment and provide the required security measures.

Virtual machines can be controlled programmatically with the use of scripting language which automates the process of operations which include start-up and shutdown. Light proposes the use of scripts to control virtual machines within an automated sandbox (Ligh et al. 2010). He also described the malware analysis cycle with the use of virtualization which is supported by Harlan (Harlan 2005). The cycle described by Light is adapted for the SPSA system. A baseline virtual machine is created. A copy is made of the baseline virtual machine and then loaded daily for usage at these establishments. This will ensure that uninfected virtual machines are deployed for use every day. It also provides the opportunity to examine the virtual machines for possible infections; this is achieved by storing the virtual machine used during the day.

The added benefit of virtual machines is the efficiency of restoring to a state which users can use to access the Internet after malware infections. An environment which uses physical machines requires reinstalling the operating system after a malware infection. During this period the establishment cannot conduct business. The use of virtual machines minimizes the period of inactivity. Gold reported in 2007 of cyber attackers targeting virtualization (Steve 2007). Some malware is virtual machine aware which implies that the malware would not execute in the virtual machine environment (Zhu, Chin 2007). The malware writers added this feature to protect the malware against virtualised environments used by malware analysts. This could be beneficial to the establishments and reduce the infection rate due to the inactivity of the malware.

Users at establishments require access to the Internet. A customizable user management system would be required to control the sequence users follow to access the Internet and expose features of the SPSA system to the users. These features include the completion of a questionnaire and coverage and comprehension of the security awareness topics. The continuous exposure to security related content contributes the success of a security awareness program (Kruger, Kearney 2006). The SPSA system is designed to present security awareness content to the user before accessing the Internet thus reminding the user of safe practises against cyber attacks. Easyhotspot is an alternative solution for hotspot billing system released under the GNU general public license which implies that the software could be modified with the needed requirements of the SPSA system (The EasyHotspot team 2007). Easyhotspot consists of a user management system which allows users to access the Internet through the portal (See **Figure 1**). Modifications to the portal would presents users with access to the security awareness content or the questionnaire.

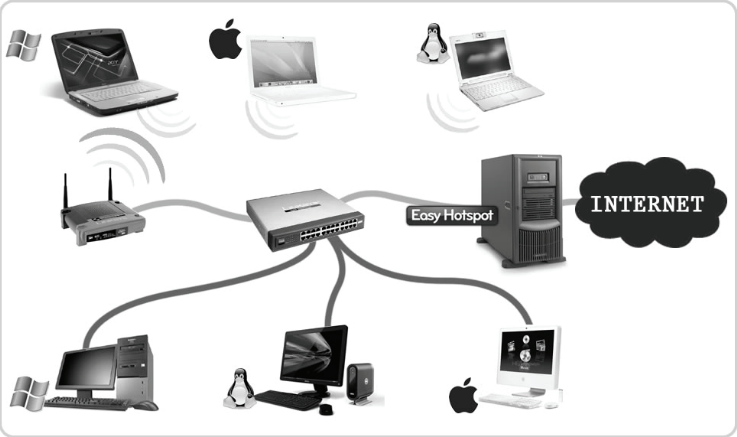


Figure 1: EasyHotspot management system

2.2 Threat collection and analysis

Abraham summarised an overview of social engineering malware which entices users to perform detrimental actions which could infect the computer system (Abraham, Chengalur-Smith 2010). The malware utilizes numerous avenues which include websites, social software and email for infection. Web browsers are used to access these avenues on the Internet. The inspection of the web sites visited is crucial in the identification of threats and determining the effectiveness of the security awareness program. Polychronakis proposed the design of a URL collection system used in exploring the life cycle of web based malware (Polychronakis, Mavrommatis & Provos 2008). The system analyzed the web pages for malicious content; this was achieved by visiting the URL and monitoring the system for new processes, file system changes and registry modifications. Provos also proposed a similar approach which consisted of identification of URL's, in-depth verification of maliciousness and aggregation of malicious URL's into site level ratings (Provos et al. 2007). These approaches are risky; a controlled approach is required by collecting the content from the URL and testing the content for maliciousness. Collection of the content from the web sites could be achieved with a web crawler. Mohhr discussed Heritrix which is an open source extensible, web scale, archival-quality web crawler (Mohr et al. 2004). Ikinci demonstrated the effectiveness of Heritrix as part of the MonkeySpider system used in the detection of malicious websites (Ikinci, Holz & Freiling 2008). The SPSA system follows a similar approach as demonstrated in the MonkeySpider system which includes the use of antivirus software in the identification of malicious content. These components discussed provide an automated and virtualized platform for the SPSA system.

The following section discusses the technical implementation of the components.

3. Shared public security awareness (SPSA) system architecture

The SPSA system consists of subsystems which as whole provide a virtualized automated platform to access the Internet, collect Internet behavioural data and delivery of a security awareness program at these establishments. These subsystems can operate independently of each other and thus are discussed separately. The automated virtualized environment is discussed in Section 3.1 and 3.2, followed by Section 3.3. and 3,4 which addresses the collection of data generated during browsing sessions and concluding with the elaboration of the security awareness program delivery mechanism in Section 3.5 and 3.6.

3.1 Internet access system

The Internet Access System is a modified user management system which based on configuration will direct users first to complete the security awareness questionnaire or direct users to the security awareness content before allowing access to the Internet (See **Figure 2**). The selection policy determines which functionality the user will interact with. The questionnaire functionality is used to assess the security knowledge of the user while the content functionality provides the user with an opportunity to learn about security related topics.

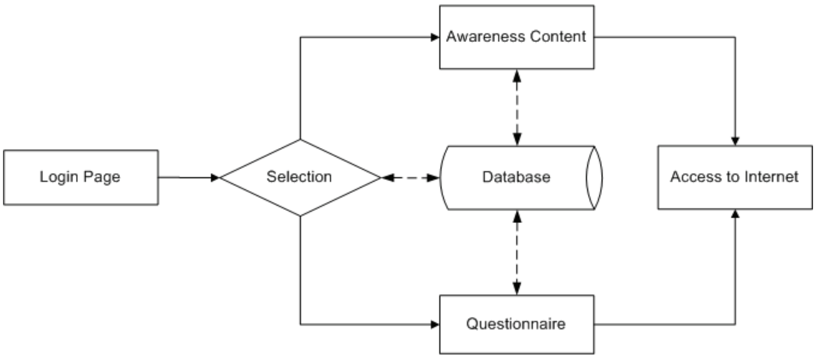


Figure 2: Internet access system

3.2 Virtual machine manager

The Virtual Machine (VM) Manager automates the operations of the SPSA system (See **Figure 3**). At the start of each day the VM manager loads a “clean” virtual machine for usage. A “clean” virtual machine represents a baseline installation of the operating system which has not been used by the users of these establishments. All components required to access the Internet are installed and configured. During the setup phase all software is tested for viruses and only reputable websites are visited to download software or update software. The task scheduler will initiate predefined scripts which will active the URL collection system to capture HTTP packet information into a file. Users will arrive at the workstations and start browsing websites. At the end of the day the task scheduler will initiate a script which will extract the data out of the file created and store the data in a database. The VM manager will shutdown the virtual machine which was used during the day, creates a backup of the virtual machine and assigns a date label to the virtual machine should forensics or malware analysis be required on the virtual machine.

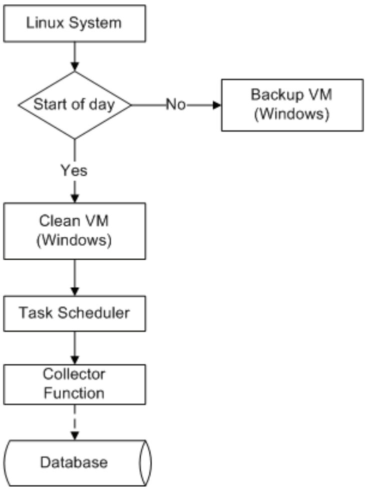


Figure 3: Daily virtual machine operations

3.3 URL collection system

The URL collection system is used in the collection of the web page address visited by the user and these include the web pages that are visited without the prior knowledge of the user. The URL collection is initialized during the start sequence of the user's virtual machine. TShark is a network protocol analyzer which provides the capability to capture packet data from a live network. Studies conducted by (Nascimento, Correia 2011) and (En-Najjary, Urvoy-Keller 2010) used TShark for the collection of specified network traffic. During the operation of the SPSA system a filter will be used to specify the required data to capture. Only outgoing HTTP traffic data is required which saves disk usage. The request line in the HTTP data packet contains the required data. The URL information is important to the work described here. According to (Forouzan 2003), "The URL is a standard for specifying any kind of information on the Internet. The URL defines four things: method, host computer, port, and the path." He states that host and path provide information on where the information is located. The URL provides a route to the content that was accessed by the user. TShark filter is configured only to collect the request line information encapsulated in the Hypertext Transfer Protocol (HTTP) header. An output file containing the captured data will be created when the time expires. This will contain the address of the webpage the user visited. Storage of the data is required and this is achieved by the URL transporter system which will analyze and extract the data from output file created by TShark. The URL transporter system is an application which will be executed at predetermined times during the day to poll a specified directory and extract the data from all the files within the directory and transport it to the external storage components for example a database server.

3.4 URL inspector

The URL Inspector component is designed to examine the URL's visited by the user. It consists of two components namely the URL Analyzer and the Malware Collection and Classification (MCC) system (See **Figure 4**). The URL Analyzer will examine each collected URL in the database against the Google Safe Browsing database, a service provided by Google, which enables applications to examine the location of the website against known phishing and malware websites (Google Code Lab 2008). This information is captured in a report. The MCC system also uses the URL captured in the database. The system consists of an Internet crawler called Heritrix which will be used to download the content of the URL and then use an anti-virus (AV) application called ClamAV to determine if the content is malicious. The list of malware found will be captured in a report. The report could assist in the identification of threats specific to these establishments and be used as a measure to determine the effectiveness of security awareness programs.

The data gathered about the browsing behaviour which include the destination address and the content of the web pages visited will be useful to determine the effectiveness of security awareness campaign by investigating the behavioural changes of the Internet users at these establishments

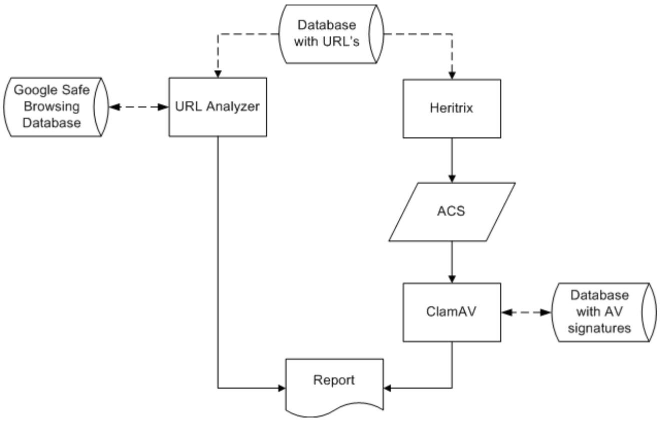


Figure 4: URL inspector

3.5 Awareness collection system

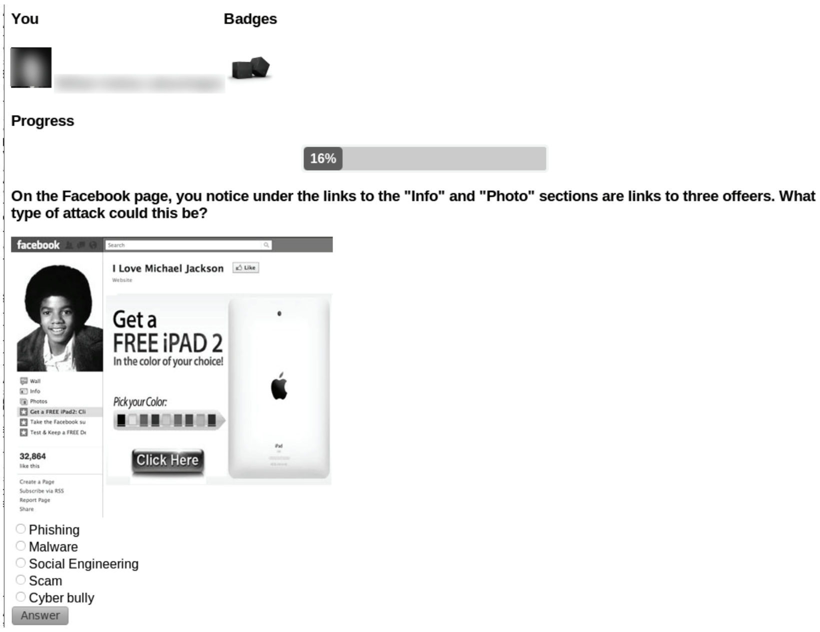
The security awareness levels of the users will be determined by completing a questionnaire. The users visiting these establishments are required to login. Thereafter the users will be presented with a set of questions which assesses the knowledge in security awareness related topics. Wilson reported on the best practises in the development of a security awareness program (Wilson, Hash 2003). One of the sections in the report discussed a comprehensive list of awareness topics some of these include but is not limited to:

- Password usage and management
- Spam
- Social Engineering
- Web usage
- Shoulder surfing
- Desktop security
- Unknown e-mail/attachments
- Incident response – contact whom? “What do I do?”

The Awareness Collection System was developed with requirements identified for the design of a security awareness game (See **Figure 5**). Game play encourages learning and with the use of game play components users are enticed to return to continue with the game. Using these principles would extend the contact time between the SPSA system and the user. Labuschagne recommended the use of Appointment, Influence and Status, and Progression dynamics (Labuschagne et al. 2011a). These dynamics are demonstrated visually with the use of badges. A badge is a visual indicator of an achievement. The appointment dynamic is represented with an image and is calculated with the consecutive logins over a period of three days. The user has to ensure that they continuing using the system after the badge have been obtained. The badge would be revoked should the user miss one day from using the system. The badge will be assigned again to the user after three consecutive day usage of the system. The status badge is provided when a user answers five questions correctly. The badge will be revoked in the event of an incorrect answer. Therefore the user is encouraged to provide the correct answers. The progression dynamic is represented with the progress bar which provides the user with a visual indicator on progress. The user is presented with randomized multiple choice questions. Labuschagne also identified security awareness topics which are applicable to establishments which allow resources to be shared amongst users accessing the Internet (Labuschagne et al. 2011b). These topics are more specific to the environment and include social media security awareness topics which is lacking in the work conducted by Wilson (Wilson, Hash 2003). The questions categories include but are not limited to the following:

- Spam
- Cyber bullying
- Malware
- Social Engineering

- Social Networking Sites
- Phishing



On the Facebook page, you notice under the links to the "Info" and "Photo" sections are links to three offeers. What type of attack could this be?

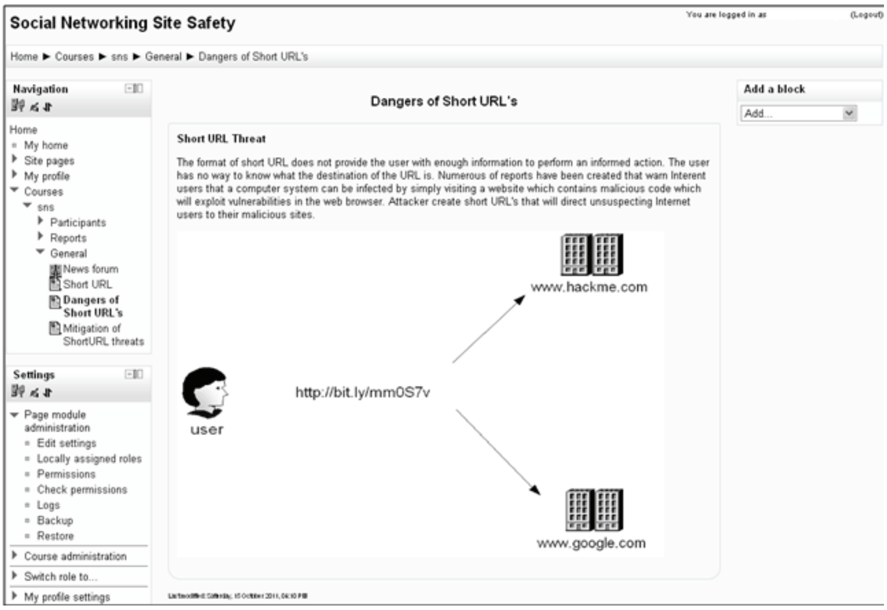
○ Phishing
○ Malware
○ Social Engineering
○ Scam
○ Cyber bully

Figure 5: Screenshot of security awareness questionnaire

A report will be generated upon the completion of the questionnaire. The report indicates areas of weakness for the user and provides the user access to resources which addresses the areas of concern. A comprehensive report could assist in the identification of security awareness topics specific to the establishment. These results could also be incorporated in to the E-Awareness Model (E-AM) proposed by Kritzinger and Von Solms. This model would not allow home users to access the Internet if their security awareness levels are not satisfactory. Also the users are required to complete remedial work to address the shortcomings before access to the Internet is granted (Kritzinger, von Solms 2010). The SPSA system is designed to determine the security awareness levels and provide users to opportunity to improve their security knowledge with topics specific to users at these establishments.

3.6 Awareness content system

The Awareness Content System makes use of a content management system (CMS) to deliver the material to the user. The CMS used for the study purpose is called Moodle. It is a software package for producing Internet-based courses and web sites (Dougiamas 1999). Some typical features of Moodle are assignment submission, discussion forum, files download, grading, instant messages, online calendar, online news and announcement, online quiz and a wiki. These features provide a platform that integrates into the requirements of the SPSA system in the delivery of security awareness content to the users and provide a mechanism for assessment. The CMS stores that material of the identified security awareness topics which the user can easily access. One of the topics addresses the dangers of short URL’s which could be encountered on social media platforms (See **Figure 6**). The user is provided with background information on the threat and suggests actions to perform once the threat is encountered. The CMS also provides functionality to assess the user’s knowledge on the topic that was accessed by the user. The material content is collected from different sources which include vendor specific security best practises provided to the community. For instance, McCarthy composed a guide to Facebook security which addresses safety topics relating to the social networking platform (McCarthy, Watson & Weldon-Siviy 2011). One of the topics in this guide provides readers the necessary steps required to protect their Facebook accounts. The material for the SPSA system is updated once new information has become available. The material on the SPSA needs to current to address the latest threats identified by security vendors. This is possible by following information security threat trends that affects the categories identified for the establishments.



Social Networking Site Safety

Dangers of Short URL's

Short URL Threat

The format of short URL does not provide the user with enough information to perform an informed action. The user has no way to know what the destination of the URL is. Numerous of reports have been created that warn Interent users that a computer system can be infected by simply visiting a website which contains malicious code which will exploit vulnerabilities in the web browser. Attacker create short URL's that will direct unsuspecting Internet users to their malicious sites.

Last modified: Saturday, 15 October 2011, 04:10 PM

Figure 6: Awareness content system

4. Conclusion

This paper describes the design of an automated and virtualized platform used to promote security awareness in rural areas where the community access the Internet through shared resources. The SPSA system is a collection of components identified in the body of knowledge which provides a singular tool to measure the proficiency of the community and promotes security awareness. The SPSA system resolves the problem of associated with conducting security awareness programs in rural areas; these include but are not limited to travelling to the destination, establishing trust with the community and the frequency of exposing the users at these establishments to security related content. It provides an automated and virtualized infrastructure which improves the availability of resources to access the Internet, collects data about the browsing behaviour of the users, the identification and classification of threats encountered by the users, and conducts a security awareness program. The SPSA system does however have limitations. Currently the SPSA system consists of two subsystems: The automated virtualized platform which delivers the security awareness program and a separate platform which is designed for the evaluation of content visited by the users during the browsing session. The process to transfer the data collected by the automated virtualized platform is not automated. The majority of these establishments do not have the infrastructure to provide enough bandwidth to harvest all the content from the web pages as this process requires the research team to collect the data from the establishments and complete the process at another location which provides high bandwidth infrastructure. Furthermore the identification of malicious sites and software is limited to the signatures identified by security vendors. The SPSA system does not provide a component to automatically update the security awareness content.

Future research will include an additional component to determine if the virtual machine used by the user resembles malware infection behaviour. This would improve the accuracy of malware infection identification. In addition, the SPSA system requires a mechanism to assess the factors affecting the behavioural change of the users at these establishments. This is required to evaluate the effectiveness of the SPSA system. The evaluation of the effectiveness of the SPSA system would be determined with the deployment of the system in identified rural areas.

References

Abraham, S. & Chengalur-Smith, I. 2010, "An overview of social engineering malware: Trends, tactics, and implications", *Technology in Society,* vol. 32, no. 3, pp. 183-196.

Agarwal, S., Rahman, S. & Errington, A. 2009, "Measuring the determinants of relative economic performance of rural areas", *Journal of Rural Studies,* vol. 25, no. 3, pp. 309-321.

Bell, M. & Lintumaa, K. 2011, *Virtual Machines: Added planning to the forensic acquisition process.*, InSecure.

Dougiamas, M. 1999, *Modular Object-Oriented Dynamic Learning Environment*, 2.1.2 edn, Moodle Pty Ltd.

England, P. & Manferdelli, J. 2006, "Virtual machines for enterprise desktop security", *Information Security Technical Report,* vol. 11, no. 4, pp. 193-202.

En-Najjary, T. & Urvoy-Keller, G. 2010, "A first look at traffic classification in enterprise networks", *Proceedings of the 6th International Wireless Communications and Mobile Computing Conference*ACM, , pp. 764.

Forouzan, B.A. 2003, "Hypertext Tansfer Protocol" in *TCP/IP Protocol Suite*, 2nd edn, McGrawHill, , pp. 649-663.

Google Code Lab 2008, *Google Safe Browsing*. [online], http://code.google.com/apis/safebrowsing/

Grobler, M., Flowerday, S., Von Solms, R. & Venter, V. 2011, "Cyber Awareness Initiatives in South Africa: A National Perspective", *Southern African Cyber Security Awareness Workshop*Defence, Peace, Safety and Security (CSIR), South Africa, 12 May 2011, pp. 32.

Harlan, C. 2005, "Malware analysis for windows administrators", *Digital Investigation,* vol. 2, no. 1, pp. 19-22.

Ikinci, A., Holz, T. & Freiling, F. 2008, "Monkey-spider: Detecting malicious websites with low-interaction honeyclients", *Proceedings of Sicherheit, Schutz und Zuverlässigkeit,* .

Kim, W., Jeong, O., Kim, C. & So, J. 2011, "The dark side of the Internet: Attacks, costs and responses", *Information Systems,* vol. 36, no. 3, pp. 675-705.

Kritzinger, E. & von Solms, S.H. 2010, "Cyber security for home users: A new way of protection through awareness enforcement", *Computers & Security,* vol. 29, no. 8, pp. 840-847.

Kruger, H.A. & Kearney, W.D. 2006, "A prototype for assessing information security awareness", *Computers & Security,* vol. 25, no. 4, pp. 289-296.

Labuschagne, W.A., Burke, I., Veerasmay, N. & Eloff, M.M. 2011a, "Design of cyber security awareness game utilizing a social media framework.", *Information Security South Africa*South Africa, 15 May 2011.

Labuschagne, W.A., Eloff, M.M., Veerasmay, N., Leenen, L. & Mujinga, M. 2011b, "Design of a Cyber Security Awareness Campaign for Internet Cafe Users in Rural Areas", *Southern African Cyber Security Awareness Workshop*Defence, Peace, Safety and Security (CSIR), South Africa, 12 May 2011, pp. 42.

Ligh, M.H., Adair, S., Hartstein, B. & Richard, M. 2010, *Malware Analyst's Cookbook and DVD: Tools and Techniques for Fighting Malicious Code,* 1st edn, Wiley.

McCarthy, L., Watson, K. & Weldon-Siviy, D. 2011, *A Guide to Facebook Security For Young Adults, Parents, and Educators*, Facebook.

Mohr, G., Kimpton, M., Stack, M. & Ranitovic, I. 2004, "Introduction to Heritrix an archival quality web crawler", *Proceedings of the 4th International Web Archiving Workshop (IWAW'04)*, sep.

Nascimento, G. & Correia, M. 2011, "Anomaly-based Intrusion Detection in Software as a Service", .

Polychronakis, M., Mavrommatis, P. & Provos, N. 2008, "Ghost turns zombie: exploring the life cycle of web-based malware", *Proceedings of the 1st Usenix Workshop on Large-Scale Exploits and Emergent Threats*USENIX Association, Berkeley, CA, USA, pp. 11:1.

Provos, N., McNamee, D., Mavrommatis, P., Wang, K. & Modadugu, N. 2007, "The ghost in the browser analysis of web-based malware", *Proceedings of the first conference on First Workshop on Hot Topics in Understanding Botnets*USENIX Association, Berkeley, CA, USA, pp. 4.

Steve, G. 2007, "Time to face virtualized realities", *Infosecurity,* vol. 4, no. 4, pp. 35-38.

The EasyHotspot team 2007, *EasyHotspot*., [online], http://easyhotspot.inov.asia/

Wilson, M. & Hash, J. 2003, *Building an Information Technology Security Awareness and Training Program*, National Institute of Standards and Technology, Gaithersburg.

Zhu, D. & Chin, E. 2007, "Detection of VM-Aware Malware", University of Berkeley, [online: http://radlab.cs.berkeley.edu/w/upload/3/3d/Detecting_VM_Aware_Malware.pdf

Information Security Model to Military Organizations in Environment of Information Warfare

José Martins[1], Henrique Santos[2], Paulo Nunes[3] and Rui Silva[4]
[1, 3]Military Academy – CINAMIL, Lisboa, Portugal
[2]University of Minho - Department of Information Systems, Guimarães, Portugal
[4]Lab UbiNET/IPBeja, INESC-ID, Lisboa, Portugal
jose.carloslm@gmail.com
hsantos@dsi.uminho.pt
pfvnunesam@gmail.com
rs.beja@gmail.com

Abstract: This article proposes a model to maximize the information security within military organizations, inserted in environment of Information Warfare. It attempts to answer three fundamental questions, *what to do, why and how*? to protect the information and Information Systems of possible incidents related to the information security that may affect confidentiality, integrity and availability of information. The main variables to be considered are defined and their possible values are proposed. These variables are obtained by means of an interpretative epistemological approach, through a literature review, the use of research methods of Contents Analysis, Focus Group and the General Morphologic Analysis method. To respond in an integrated manner to the three questions above, the model considers the possible incidents of information security in Information Systems, taking into account primarily the main components of the security risks of Information Systems that collect, store, process, transmit and disseminate the information. Its operation is guided by the military concepts of Information Warfare, Information Assurance, the most important principles of war applied to Defensive Operations and the military doctrine of Information Operations. Given the type of problem identified in the study, focusing primarily on the analysis of scenarios of information security incidents and interconnection with the planning and selection of security controls, the method used is the General Morphological Analysis. This method allows for the prediction of possible scenarios of incidents related to information security at the organizational level, which results in the selection of the most efficient solution of security controls, to maximize the security of information. Information security must guarantee confidentiality, integrity and availability of information and seeks to contribute, by means of the operational implementation of the military concept of Information Assurance, to achieve the information superiority.

Keywords: information security management, information assurance, information security model, general morphological analysis, information warfare

1. Introduction

Most organizations should consider as one of the most significant challenges to overcome, the increasing level of competition for obtaining information from competitors, suppliers and customers in an ethical and legal manner, i.e., through Competitive Intelligence (McCrohan, 1998).Yet many organizations still have not implemented a formal approach to information security management that enable them to ensure the safety of its own information or that that is of its responsibility(Barlette & Fomin, 2009; Fomin, Vries, & Barlette, 2008).

Information security management is a management process developed and implemented in an organization, in order to guarantee the main requirements to ensure information security, especially for information that is critical to support the business processes of an organization (Vermeulen & Von Solms, 2002). In this context, it is essential that in addition to protecting the information itself, is guaranteeing simultaneously the protection of the Information Systems that allow for the collection, processing, storage and transmission of information.

The international norm ISO/IEC 27001 (2005) considers information security as a structured management process which guarantees the main requirements of information security, providing a model for establishing, implementing, operating, monitoring, reviewing, maintaining and improving an Information Security Management System (ISMS).

Despite efforts to build tools that support a better application of these concepts to real cases, one cannot affirm that there is a universal model, fundamentally due to the specificity of the organizations.

Organizational information security can be considered a Wicked Problem, according to the criteria of acceptance of a Wicked Problem, which emphasizes the following five criteria: there is no definitive formulation of the problem; there is no immediate and final test to the encountered solution for the problem; each problem is unique and the causes of the problem can be explained in several way, hence the choice of explanation determines the nature of the problem's resolution (Ritchey, 2011).

Given the type of problem indentified in the study, focusing primarily on the analysis of scenarios on information security incidents, and interconnection with the planning and selection of security controls, the method of General Morphological Analysis (GMA) is used to obtain, validate and link the identified variables in the model of information security in military organizations, in addition to literature review and Content Analysis. This method allows developing the possible scenarios of information security incidents at the organizational level, which allow choosing the most efficient solution of security controls to maximize the security of information. The GMA allows structuring and investigating the total set of relationships contained in multidimensional complex problems, usually non-quantifiable (Ritchey, 2011; Zwicky, 1969).

The method of morphological analysis applied to the problem of information security at an organizational level, has the advantage of obtaining the possible combination of information security controls (outputs) to be applied in a military organization, faced with the choice of certain entries in the model parameters i.e. the possible methods of attack (inputs), or also permit through the indication of the possible inputs to obtain outputs.

To present the model, this article is divided into six sections. In the first section, the framework of the problem is carried out, the relevance of the topic is justified, and the principle motives and objectives are identified in order to achieve success. In section two there is an analysis of the general environment in which military organizations operate, essentially defining the concepts of Information Warfare (IW) and Information Operations (IO). Section three presents a literature review of the models and methods of information security currently available for conducting the management of information security at an organizational level. Section four consists of a brief summary that describes the research methodology used in the study, with emphasis on the use of the Focus Group research method. Section five describes the model of information security, making an analysis of its variable. Finally, conclusions of the study are presented in section seven, indicating limitations and possible open studies

2. Information warfare and information operations

In order to operationalize the security of information in a military organization because of its specificity, it is necessary to take into consideration military doctrine. In this context, one of the fundamental concepts to consider is the concept of IW, which have been studied and referred to by several authors (Alberts, 1996; Alberts, Garstka, Hayes, & Signori, 2001; Cronin & Crawford, 1999; Denning, 1999; Libicki, 1995; Waltz, 1998).

These authors have different interpretations of the concept of IW. This analysis falls outside the scope of this article. In essence, the concept of IW can be defined, as referred to in military doctrine of the United States of America (USA), as "*actions taken to achieve information superiority by affecting the adversary's information, information-based processes, Information Systems and networks based on computers of an adversary while defending our own information, information-based processes, Information Systems and computer-based networks*" (FM100-06, 1996, p. 224).

In an environment of IW, such actions may be conducted through IO, which consists of primarily, according to the USA doctrine, of a set of activities and skills used to affect adversary information and its Information Systems (IS), while defending our (FM3-13, 2003; JP3–13, 1998), which can be planned and used to obtain information superiority against an adversary.

To ensure that the operational capacity to "*collect, process and disseminate an uninterrupted flow of information while exploiting or denying an adversary the ability to do the same*"(JP-1-02, 2010, p. 225), which permits obtaining information superiority, fundamentally it is necessary to operationalize the concept of military Information Assurance, as a set of "*measures that protect and defend information and IS, ensuring their availability, integrity, authenticity, confidentiality and non-repudiation. This includes the ability to restore the operation of IS, incorporating capabilities for protection, detection and reaction*"(JP-1-02, 2010, p. 224).

In NATO, the IO consists of information activities, which "*are actions to affect the information and/or information systems. They can be performed by any actor and include protection measures*." (AJP-3.10, 2009, p. 1.3), that is, once again the main focus is to ensure the protection of information and IS of our forces, while searching to affect the information of IS of the opponent.

The offensive actions developed under IW and IO can be carries out and have an effect on primarily three levels or dimensions of performance, which are predominantly: a physical level, a level of information and a cognitive level (Alberts, et al., 2001; Andress & Winterfeld, 2011; Cronin & Crawford, 1999; Martins, Santos, & Nunes, 2009; Waltz, 1998).

The emergence of these concepts (i.e. the IW, IO and Information Assurance), developed specifically in this last decade and predominantly referenced in the military, where information is seen simultaneously as a weapon and a target (Hutchinson, 2003), leads to the need to devise new approaches to information security.

3. Literature review

Given the number of academic studies identified and analyzed in the literature review (Martins & Santos, 2010), with emphasis on management of information security at an organizational level, it appears that there are some difficulties in applying the existing methods and hence the management of information security, of which stress:

- First, considering an approach for risk analysis, it is difficult the calculate the probability of a threat exploring the vulnerability of an asset (Baskerville, 1993), in assessing the value of an asset and thus to prove that an investment in information security has an adequate return (Finne, 1998)**.**
- Also the multidimensionality of the problem of security of IS (e.g. threats, vulnerabilities, assets and impact) (Baskerville, 1993) increases the difficulty of the problem and possible solutions.
- The absence of a theory that covers all the management of information security through the integration of several theories of support (Hong, Chi, Chao, & Tang, 2003).
- Finally, the management of information security in organizations seems to have not yet reached a level of maturity that can make it a repeatable process of management (Nnolim & Steenkamp, 2008).

In the application of international norms for information security management, more importantly ISO 27001, there have been some difficulties that have resulted in a low adoption in organizations. Of these, one can refer to the nuclear reasons being: the high cost (i.e. time, resources and monetary value) associated with obtaining a certification in information security management; the difficulty of being understood by all employees of an organization due to its complexity; the contribution of certification to an increase in market share, for a short period of time, in comparison to other companies in the sector, and finally the very general guidelines of the rules in opposition to specificity of the organizations (Fomin, et al., 2008).

One can also identify some barriers to overcome that limit the adoption of this international norm of information security, complementary to the study by Fomin et al.(2008), of which stand out the need of managers to being sufficiently concerned and elucidated on the effectiveness of information security for one's organization in a cost/benefit perspective; managers must have knowledge of security controls applied in the organization, in perspective, "what? why?" of its adoption; the need to reduce the costs of implementing the norms and the need for highly qualified people for its implementation; the strategy and the security model implemented in an organization should consider the specificity of the organization, the cultural aspects and finally the human factor and the internal threats (Barlette & Fomin, 2009).

Recent studies show the main success factors for an efficient management of information security, the need for a commitment from the organization's top management, for security management not to be considered only as an aspect of IT, its management model be adapted to the organizational culture and possess mechanisms that allow timely update of security policies, as well as knowledge sharing (Barlette & Fomin, 2009).

These difficulties, barriers and success factors suggest some key clues or guidelines to consider in developing a model of information security for military organizations. The prospect of the military organization, in which the planning and military decision-making is focused on the analysis of possible courses of action (e.g. the most probable, the most dangerous, or other means) of the opponent, justifies the proposed model, focusing on an approach driven by scenarios.

4. Research methodology

In order to construct the model of information security, the following research approach was closer to the interpretive epistemological orientation. This is reflected accordingly in the used research methods, which consisted in the literature review and content analysis, and in a second phase in the use of Focus Group research method, in combination with the method of Analysis and Synthesis, used in General Morphological Analysis.

In the first phase of the study, focusing on the literature review and content analysis there is a search through the technique of triangulation for different sources of empirical material (e.g. international norms, military doctrine, industry recommendations, books and academic articles) and to obtain the key variables and their possible values or conditions.

However, due to the possible subjectivity of the results obtained, resulting from the interpretation of the researcher, in a second phase the Focus Group method was used, which is inserted in the application of the method of General Morphological Analysis. There is a search to ensure, in addition to the triangulation of sources, triangulation of methods of research and consequent rigorous in the research approach followed.

The choice of Focus Group method is due to the fact that this method allows one to control the critical aspects of the study, statistical projections are not necessary, the subject is technically dominated by all the participants and the objective is not measured, but to understand the phenomenon in order to interpret the interdependences between the different variables involved in the problem (Giovanazzo, 2001).

The ultimate goal of this method of investigation is to define a model of information security for military organization, which can later (i.e. in a second study) generate all the possible scenarios of information security incidents in the military, and identify possible security controls to implement with the object of maximizing information security in relation to each possible scenario. It is critical to the successful creation of the scenarios, the quality of the modeling problem, the rigor of the definition of variables and values or conditions, which are identified and analyzed in the next section.

In conclusion, the proposed model takes into account some of the key criteria for quality in interpretive research. It allows for one to understand the meaning of the parts (i.e. the variables) as a whole that was (principle of hermeneutic circle), according to the empirical data which was collected and contextualized (principle of contextualization). This model also complies with the principle of multi interpretations, as it takes into account the various sources of empirical data and narratives derived from the interaction between the researcher and participants in the Focus Groups.

5. Information security model

The security model of organizational information suggests the variables, or the principle parameters of the problem, followed by the identification of the range of values or conditions that each parameter expresses in the possible solutions for the problem of information security (and in a future study the possible relations between identified variables). All the parameters (i.e. variables in the method) and their values or conditions is a morphological field, which can be reduced to a number of settings in which only those that meet certain criteria remain in the end, or rather the model representation of the problem.

5.1 General description of the model

The model allows for one to analyze all of the possible information security incidents and IS, taking into account the main risks of the principle components of the IS, which collect, process, transmit and disseminate information. The model is driven by the concepts of IW and Information Assurance, by the most important principles of war applied to Defensive Operations and by military doctrine associated to IO. In addition to integrating taxonomy used by CERTS's for the description of computer security incidents (Howard & Longstaff, 1998).

The principle dimensions and types of information security control are also associated, resulting from the analysis of international standard ISO/IEC 27001 (2005) of the International Organization for Standardization, the special publication NIST-SP 800-53 (2007) of the National Institute of Standards and technology and NATO security, according to its reference model in the public domain.

The variables of this model identify the possible attackers, threats and attack methods (i.e. the actions, tools/weapons) and targets that can be achieved to affect the fundamental properties of information security (i.e. confidentiality, integrity and availability) directly or indirectly by exploiting the vulnerabilities of the major components of the IS. The specific nature of the military assets that support the critical systems of Command and Control are also taken into account, and responsible for collecting, storing, processing, transmitting and disseminating all the necessary information to decision-making of various levels of the military organization.

The model identifies the main targets that allow one to achieve directly or indirectly the fundamental properties of information security and consequently achieve the main objective of the action or set of actions realized. Therefore, only knowing with accuracy the course of action (i.e. the possible attack methods and vulnerabilities of the adversary's targets) is it possible to plan effectively the security of information. The planning, implementation and monitoring of information security controls, grouped according to the security dimensions proposed, makes it possible to remove from the opponent the opportunity to perform the method of attack that can impact the organization.

The obtained model takes into account the likely vectors of attack of the opponent, or rather their possible levels of actions (i.e. physical, information, cognitive). The minimization of the impact of actions from the attacker or the threats is key, through the implementation of the adequate group of security controls organized according to the principle dimension of the security of information i.e., the organizational, physical, human, and technological dimensions (Martins, et al., 2009). These controls seek to mitigate existing vulnerabilities in the main component of the IS likely to be exploited.

The planning of information security in military organizations, can take into consideration, with the appropriate adjustments to the specific problem, the following Principle of War: the principle of Economy of Forces (e.g. use of land, use of available time), the principle of Security (e.g. obtaining information from opponent), the principle of Maneuver (e.g. defense in depth, mutual support and flexibility), the Control Unit (e.g. cohesion) and the Offensive principle (e.g. offensive action) (Couto, 1988; RC80-5, 1991; RC130-1, 2005), to maximize information security.

The integration into a single model or information security, of the possible methods of attack of the opponent driven by the vectors of attack, with the types of security controls (e.g. prevent, detect, deter, deflect, recover and react) will allow the military decision maker to have a holistic view of the organization's information security and respond in an integrated and coordinated manner to information security incidents at all levels of the organization (i.e. the strategic level, operational level and tactical level).

5.2 Description of security model variables

The variables of the model presented in Figure 1 are briefly described, which define the space of solutions to the problem of information security for military organizations in IW environment. The model identifies the key variables (e.g. threats) and their possible values or conditions (e.g. interception, interruption), in which grey indicates some of the possible values already considered in the taxonomy of security incidents on computers and developed by Sandia Laboratories and CERT/CC (Howard & Longstaff, 1998).

In this model, the *Attacker* is fundamentally an individual or group of individuals who try to run one or more methods at attack to achieve the fundamental properties of information security (i.e. confidentiality, integrity and availability) in order to achieve a given objective (Howard & Longstaff, 1998; Mayer, 2009). It also included the concept of attacking natural disasters, which require a set of natural hazards on a particular component of components of organizational IS, which may impact on the physical structure of the organization and consequently on business processes.

The attacker directs its potential actions for possible *Threats*, which are not more than the potential causes of a security information incident, and can result in damage to the system or organization (Dhillon, 2007; ISO/IEC13335-1, 2004; NIST-SP800-53, 2007; Pfleeger & Pfleeger, 2007).

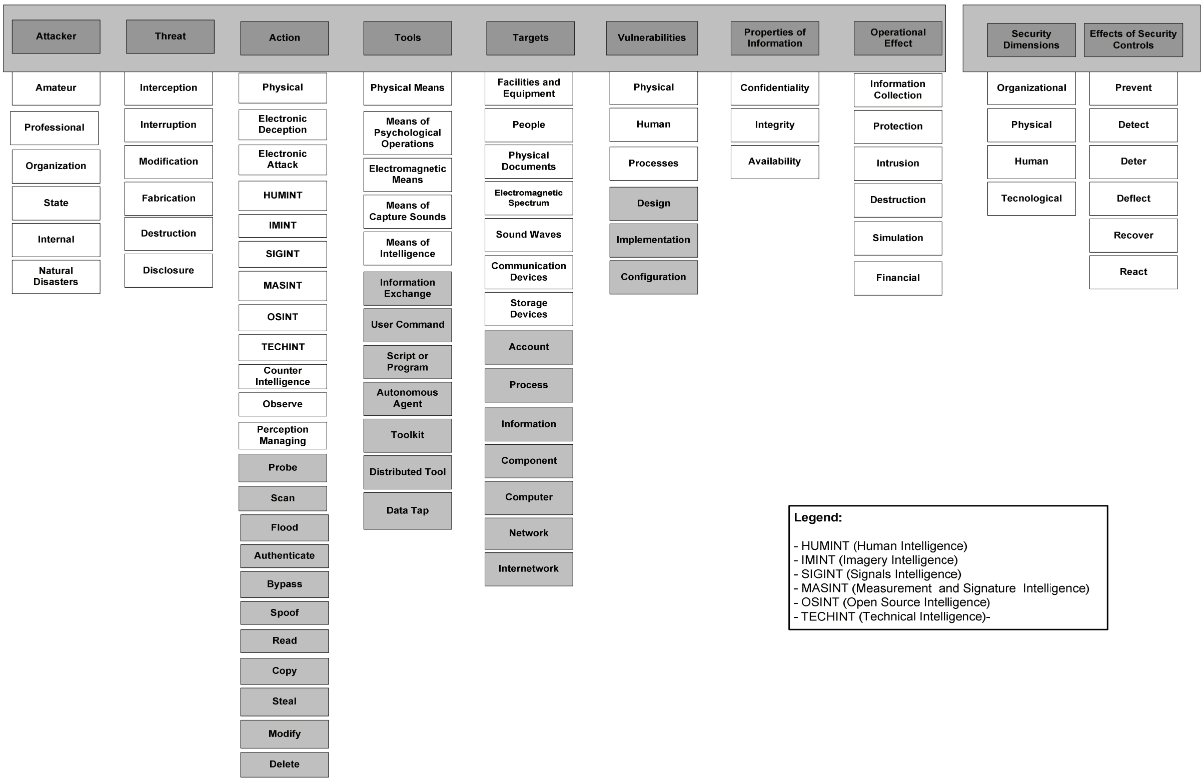


Figure 1: Information security model to military organizations

Therefore, the attacker will utilize methods of attack, or rather a series of actions destined to result in something that is not authorized to happen (FM3-13, 2003; Howard & Longstaff, 1998; Martins, et al., 2009; Mayer, 2009; Pereira & Santos, 2010).

These methods of attack are going to consist in an *Action* or group of actions, which are simply a group of steps that support specific *Tools* (i.e. weapons). These actions are aimed primarily at military organizations for possible IO (AJP-3.10, 2009; FM3-13, 2003), the activities of intelligence and counterintelligence (AJP2.0, 2003; FM2.0, 2010; JP2.0, 2007), and centered in the main components that support the IS system of the C2 organization.

In IO, take into account mainly the methods of attack used by adversaries to attack infrastructure and command and control systems (i.e. there IS), which according to military doctrine referenced in FM 3-13 (2003, p. 13), are classified as follows: unauthorized force or access, projection of malicious software, electronic deception, electronic attack, the computer network attack, physical destruction, management of perception. Activities of intelligence are also key activities which allow for more than one form of information gathering.

The actions proposed in the model looking to achieve the *Targets*, which may be logical entities (e.g. account, process, data) and physical entities (e.g. components, computer networks) (Howard & Longstaff, 1998), the organization’s human resources (e.g. decision-makers, managers of critical processes of the organization, experts), the means of transmitting information (e.g. electromagnetic radiation, sound waves), the physical infrastructure (e.g. facilities, data center, meeting rooms) that is all components that directly or indirectly help to achieve the fundamental properties of information security.

These actions taken by an attacker attempted to exploit the *Vulnerabilities* of the targets. These consist of certain weaknesses that allow themselves to be exploited by an unauthorized action or actions in the security policy of the organization. These vulnerabilities may come from the stakeholders in an organization, physical objects (e.g. hardware, cabling, data center, facilities), in how the organization’s processes are in place and functioning.

Regarding the technology, can be considered vulnerabilities project, introduced in the analysis phase of the design (e.g. software), implementation, when the installation in the organization and ultimately operational vulnerabilities, focused primarily on system settings (Correia & Sousa, 2010; Howard & Longstaff, 1998).

The effects produced on the *Property of Information* allow for the direct achievement of the objective or contribute to achieving them. The objectives of the attack methods can be focused primarily on gathering information (e.g. through the activities of Intelligence), the protection of information and systems that interact with it (e.g. by the use of Counter Intelligence activities, Physical Security, Human Security), in the intrusion or destruction of information of the means to support it (e.g. by the use of activities of Computer Network Operations, Psychological Operations or Social Engineering and Physical Destruction Attacks). One can also consider the purpose if financial gain and eventually the objective of the action or set of actions you can spend only a matter of personal challenge and recognition of professional skills of the attacker.

After analyzing all the possible information security incidents that may occur at the organizational level (i.e. the likely scenarios) can one then seek to maximize information security in accordance with the planning and implementing of security baselines. In this model, one has in mind a set of security controls suggested by industry norms such as ISO 27001 and NIST 800-53, by academic reference, primarily by the security of computer networks and security in software and by NATO military model of security.

Security controls, seek to reduce or remove the vulnerabilities of the assets (Pfleeger & Pfleeger, 2007) and thus reduce/avoid the effects of an incident (Pereira & Santos, 2010) or reduce the risk in information security IS (Mayer, 2009). Security checks can be various types, according to its purpose. It can be used to prevent, detect, deter, deflect, retrieve and respond to a security incident and it may have more technical characteristics (e.g. firewall), more formal characteristics (e.g. security policies) and informal characteristics (e.g. training and awareness of employees) (Dhillon, 2007; Pfleeger & Pfleeger, 2007).

6. Final considerations

This article proposes a model to maximize information security in military organizations, embedded in an environment of Information of Warfare. To protect information and Information Systems of the potential information security incidents that may affect the confidentiality, integrity and availability of information. This model, which attempts to answer three fundamental questions: What to do? Why? and How?

The key variables or parameters of the problem are identified, followed by identification of the range of values that each variable or conditions expressed in the possible solutions to the problem. Operationalizing the model according to the specific doctrine of military organization oriented planning and decision-making based on analysis of scenarios and carrying out war games.

However, the article only proposes and describes the main variables of the model of information security, as necessary in future work to identify relationships between identified variables in the conditions of the proposed model. Not all of the possible combinations of values of morphological parameters of the field are valid, requiring the total number of possible combinations, identify invalid combinations i.e. obtain the Cross – Consistency Matrix of the general method of morphological analysis.

In conclusion, this approach to information security for a military organization, through scenario analysis allow to support the creation of a descriptive model adjusted to the specificity of the military organization and its decision-making process.

References

AJP2.0. (2003). Allied Joint Intelligence: Counter Intelligence and Security Doctrine, NATO.

AJP-3.10. (2009). Allied Joint Doctrine for Information Operations, NATO.

Alberts, D. (1996). Defensive Information Warfare. *National Defense University Washington D.C. Institute for National Strategic Studies*.

Alberts, D., Garstka, J., Hayes, R., & Signori, D. (2001). *Understanding Information Age Warfare, CCRP Publication Series, Washington, United States of America*.

Andress, J., & Winterfeld, S. (2011). *Cyber Warfare: Techniques, Tactics and Tools for Security Practitioners*: Syngress Media Inc.

Barlette, Y., & Fomin, V. (2009). *The Adoption of Information Security Management Standards: A Literature Review, In Kenneth J. Knapp, Cyber Security and Global Information Assurance: Threat Analysis and Response Solutions (Chaper VI, pp. 119-140). Hershey - USA*: Information Science Reference.

Baskerville, R. (1993). Information Systems Security Design Methods: Implications for Information Systems Development. *ACM Computing Surveys (CSUR), 25*(4), 375-414.

Correia, M. P., & Sousa, P. J. (2010). *Segurança no Software*. Lisboa: FCA.

Couto, A. C. (1988). *Elementos de Estratégia* (Vol. I). Lisboa: Instituto de Altos Estudos Militares.

Cronin, B., & Crawford, H. (1999). Information Warfare: Its Application in Military and Civilian Contexts. *The Information Society, 15*(4), 257-263.

Denning, D. E. (1999). *Information Warfare and Security*. USA: Addison-Wesley.

Dhillon, G. (2007). *Principles of Information Systems Security - Text and Cases*: WILEY.

Finne, T. (1998). A Conceptual Framework for Information Security Management. *Computers & Security, 17*(4), 303-307.

FM2.0. (2010). *Intelligence*. USA: Headquarters Department of the Army.

FM3-13. (2003). *Information Operations: Doctrine, Tactics, Techniques, and Procedures*. Washington: Headquarters, Department of the Army.

FM100-06. (1996). *Information Operations*. Washington: Headquarters, Department of the Army.

Fomin, Vries, & Barlette. (2008). ISO / IEC 27001 Information System Security Management Standard: Exploring the Reasons for Low Adoption. Rotterdam School of Management, Erasmus University.

Giovanazzo, R. A. (2001). Focus group em pesquisa qualitativa-fundamentos e reflexões. *Administração on line, 2*(4), 1-13.

Hong, K., Chi, Y., Chao, L., & Tang, J. (2003). An integrated system theory of information security management. *Information Management and Computer Security, 11*, 243-248.

Howard, J., & Longstaff, T. (1998). *A common language for computer security incidents*: Citeseer.

Hutchinson, W. (2003). *The Changing Nature of Information Security.* Paper presented at the 1st Information Security Management Australian.

ISO/IEC13335-1. (2004). Information technology- Security techniques-Management of information and communications technology security. Part 1: Concepts and models for information and communication technology security management.

ISO/IEC27001. (2005). Information technology – Security techniques – Information Security Management Systems - Requirements.

JP2.0. (2007). Joint Intelligence: Joint Chiefs of Staff, USA.

JP3–13. (1998). *Joint Doctrine for Information Operation, United States of America*.

JP-1-02. (2010). Dictionary of Military and Associated Terms, *Washington D.C.*: Department of Defense.

Libicki, M. (1995). *What is Information Warfare?* Washington: National Defense University.

Martins, J., & Santos, H. (2010). *Methods of Organizational Information Security - A Literature Review.* Paper presented at the 6th International Conference On Global Security, Safety and Sustainability, Braga.

Martins, J., Santos, H., & Nunes, P. (2009). *Security Framework for Information Systems.* Paper presented at the 8th European Conference on Information Warfare and Security, Lisboa.

Mayer, N. (2009). *Model-based management of information system security risk.* University of Namur.

McCrohan, K. (1998). Competitive intelligence: Preparing for the information war. *Long Range Planning, 31*(4), 586-593.

NIST-SP800-53. (2007). Information Security. USA.

Nnolim, A., & Steenkamp, A. (2008). An Architectural and Process Model Approach to Information Security Management. *Information Systems Education Journal, 6*, 31.

Pereira, & Santos, H. (2010). *A Conceptual Model Approach to Manage and Audit Information Systems Security*. Paper presented at the 9th European Conference on Information Warfare and Security.

Pfleeger, C. P., & Pfleeger, S. L. (2007). *Securiy in Computing, Prentice Hall, 4ª ed, United States of America*.

RC80-5. (1991). *Brigada de Infantaria Independente*. Lisboa: Estado - Maior do Exército.

RC130-1. (2005). *Regulamento de Campanha - Operações*. Lisboa: Instituto de Estudos Superiores Militares.

Ritchey, T. (2011). *Wicked Problems-Social Messes: Decision Support Modelling With Morphological Analysis* (Vol. 17): Springer Verlag.

Vermeulen, C., & Von Solms, R. (2002). The information security management toolbox-taking the pain out of security management. *Information Management and Computer Security, 10*(2/3), 119-125.

Waltz, E. (1998). *Information Warfare: Principles and Operations, Artech House.*

Zwicky, F. (1969). *Discovery, Invention, Research - Through the Morphological Approach*. Toronto: The Macmillan Company.

Simulation Approach for Military Cyber Operations

Ben Morton[1], Sylvain Leblanc[1] and Melanie Bernier[2]
[1]Royal Military College of Canada, Computer Security Laboratory, Kingston, Canada
[2]Defence Research and Development Canada, Centre for Operational Research and Analysis, Ottawa, Canada
Ben.Morton@rmc.ca
Sylvain.Leblanc@rmc.ca
Melanie.Bernier@drdc-rddc.gc.ca

Abstract: Cyber operations are expected to become more important, and thus military commanders and staff will need to be trained in these operations. The aim of this paper is to describe an approach for simulating the effects of cyber operations in constructive simulations used for training by modern military forces. The paper argues that it is not currently possible to realistically simulate military cyber operations in a cost-effective manner, due to of the lack of existing data on the subject and the fact that it is not possible to validate available data from the civilian realm against military cyber operations. However, we argue that to educate senior military leaders, it is more important to simulate the effects of cyber attacks than to simulate the actual attacks themselves with a high degree of fidelity. The paper will discuss a set of cyber effects, and introduce an attack taxonomy that focuses on these effects. This taxonomy will discuss the effects of various attack types, along with the level of access to the target computing resource that is required to prosecute the attack. The effects of attacks will be described in terms of their impact on the computing network, computers or other devices. From this, we will derive impacts on mission capabilities, and discuss how these could be implemented inside constructive simulations. For example, to demonstrate the effects of a denial-of-service (DoS) attack, it is not necessary to carry out the attack itself; it may be sufficient to disconnect the server that is the target of the DoS attack. When prosecuting an attack, adversaries must always contend with limited resources and time. In order to integrate cyber operations in constructive simulations with a measure of realism, the paper will discuss a mechanism to limit the cyber attacks available to an attacker in terms of available resources and time. The approach will also introduce the concept of stochastic attack success by assigning probabilities of attack success against known defences. Finally, the paper will discuss avenues of future and related work, including the relationship of this work with the Metrics Framework for Cyber Command and Control paper, (Bernier et al. 2012) also presented at this conference.

Keywords: cyber operations, constructive simulation, education, cyber effects, military operations

1. Introduction

Modern armed forces increasingly depend on computer and network technologies as enablers for military capabilities. Computer and network technologies are inherently susceptible to 'cyber attack', the deliberate and malicious exploitation of vulnerabilities residing in computer networks. Cyber operations provide those who wish to challenge modern armed forces with a powerful and economical avenue to attack and thus impair or negate a crucial source of their opponents' strength, requiring defenders to expend vast resources. This asymmetric property of cyber operations, as well as its low cost relative to conventional military capabilities and the difficulty of attribution of cyber attacks, ensures future conflicts will occur not just in the kinetic realms of land, sea, air, and space, but also increasingly in cyberspace.

The potential for conflicts in cyberspace has driven the rapid development of cyber operations capabilities both by modern high-technology military forces, and their less sophisticated adversaries. Apart from a few isolated cases, we have yet to witness widespread or determined employment of cyber operations in a military context, beyond a few isolated cases (Clarke 2009). At the same time, events such as the advent of Stuxnet (Falliere et al. 2011) can lead one to confidently predict the prominent place cyber operations will take in future conflicts. There is therefore a lack of military-specific data on the subject of preparing military forces for cyber operations, but the need is real and urgent.

Unfortunately, military forces do not have the luxury of waiting for sufficient data to inform preparations for future conflict; the new challenges of this conflict will affect them whether they are ready or not. History is replete with the unforeseen disruptive impacts of technology, be it the defensive firepower of machine guns and indirect fire in the First World War, or the combination of mobility, firepower and communications seen in the Blitzkrieg in the Second World War; these help us appreciate the dangers of facing novel methods and technologies of warfare without adequate preparation. It is thus imperative to press on with urgency in developing military readiness for operations in cyber space. Simulation provides an alternative for training in the absence of large scale operational experience. This paper describes is an approach to the incorporation of cyber effects within simulation-based military training exercises, for the training and education of senior decision makers on the subject of military operations in cyber space.

2. Other cyber attack simulation efforts

Given the importance of understanding cyber attacks and their effects, a number of cyber attack simulations have been developed by both public and private sector researchers. Most of these focus on cyber attacks in the civilian realm, whether upon business or government networks (Futoransky et al 2003, Liljenstam & Liu, 2006, Cohen, 1999, Park et al. 2001), with a smaller proportion dealing with military scenarios (Carver et al. 2001, Geers, 2010). There are currently no unclassified simulations of cyber attacks in the military arena suitable for integration with tactical training simulations, a void which the approach detailed in this paper is seeking to fill.

3. Simulation priorities

Having identified the requirement for simulation of military cyber operations as a training and education tool, it is then necessary to identify the objectives of that simulation. This will guide the simulation design. If the emphasis is placed upon a realistic portrayal of real-world cyber operations, then the simulation should replicate not only the impact of a cyber attack upon military computer networks, but the actual process of attack and defence that causes those impacts.

While useful for providing validated lessons learned and security testing of actual network configurations (Bye et al. 2008), focusing on realism is problematic due to the above-mentioned extreme scarcity of available unclassified data. While it is widely acknowledged that numerous militaries have been developing offensive cyber operations capabilities, the exact nature of those capabilities is understandably not openly discussed.

On the other hand, instances of cyber attacks in the civilian realm are plentiful, and major examples occur with disturbing frequency. To simply apply these civilian cyber attack examples to military situations, however would ignore the unique character of the military realm. While some methods employed by cyber attackers may be similar, the motivations, resources, complementary physical capabilities, and the resolve of adversaries is another matter entirely. Modelling cyber operations simulations upon the day-to-day efforts of hackers, hacktivists and cyber-criminals may create a very false impression of things to come, by drawing its realism from the wrong context.

A more appropriate goal for this simulation is to accept the limitations in achievable realism and simply maximise its training and educational outcomes. Raising awareness of the potential operational impacts of cyber attack is one important achievable training/educational outcome. Such operational impacts may include degraded availability and performance of network enabled capabilities, as well as adversary informational advantages resulting from network penetration (Musman et al. 2009).

Preparing military decision makers for the psychological/behavioural impacts of cyber attack is another potential training benefit. Research on this subject has identified degraded decision making as a consequence of cyber attack (Stytz & Banks 2010). Repeatedly exposing decision makers to the effects of cyber attack as they attempt to conduct operations through simulation could reduce negative behavioural responses, in a similar manner to military combat drills using ‘operant conditioning’ techniques (Grossman 1996).

4. Simulation of effects

We have stated numerous complementary objectives for the use of simulation in military cyber operations:

- maximising training and educational outcomes in order to develop awareness of the threat of cyber operations among military decision makers;
- understanding what constitutes better defences against cyber attack; and
- Developing appropriate behaviours while operating in cyberspace.

We are now faced with the requirement to develop a simulation model to successfully meet these objectives. Though we argue for simple modelling of the attack, it must be said that a certain amount of palpable authenticity, or the *feeling* of realism, is required in order to elicit desirable behavioural responses. The scenario must also be plausible, in that it could represent the future of cyber operations well enough to justify and inform a coarse level of contingency and procurement planning. In this way, lessons learned by decision makers undergoing simulation-based training with incorporated cyber effects will remain generally transferrable to future real-world operations.

The result of this selective application of simulation realism is to free the simulation designer from the burden of slavishly recreating the entire cyber environment, attack process, and attack/defence interactions. In this context, simulation is not concerned with the actual technical details of cyber attack; instead, it approximates the attack process into a set of effects which can be artificially imposed upon the information systems used in the simulation and those military capabilities that depend on those information systems. Overlaying these cyber attack effects onto existing constructive simulation-based training will provide a simple, cost-effective method of cyber operations training and education for commanders.

5. The approach

Detailed here then is our proposed approach for integrating cyber operations effects in constructive simulation-based training for military operations. This approach is very simple in its application, yet captures and conveys the anticipated nature of cyber space in a military context. The approach described below is meant to be “simulation agnostic”, meaning that is it not tied to a particular simulation implementation. As this paper is written when we are in the initial stages of this research project, it is written more to seek comments from the research community than to describe a fully developed approach.

5.1 Scenario and adversary

The creation of this simulation necessarily begins with the development of a scenario involving a particular adversary. The scenarios upon which constructive simulations used in military training are generally built will give some background on the conflict setting, detail the nature of opposing forces (OPFOR), as well as the training subject’s forces (Blue Force - BLUFOR). From this exercise scenario, the likely capabilities of the adversary cyber operations forces may be deduced, and described in terms of both resources and sophistication. Cyber attack actions undertaken by the adversary should support their other actions in the simulation, and are therefore synchronised with the adversary scheme of manoeuvre. In this way, cyber operations activity will reflect the typical way in which any military capability is used; capabilities are of maximum utility when used in coordination with others. For example, infantry is coordinated with armour; artillery can be used in conjunction with reconnaissance etc. The adversary may therefore target command and control capabilities with cyber attack prior to commencement of an assault, or use cyber resources to target air defences prior to airstrikes (Fulgham et al. 2008)

5.2 Taxonomy of attacks

Once an adversary is defined, an array of cyber attacks will be selected by the OPFOR team. For this approach, the selection will be made from a taxonomy of attack (Partington et al. 2011) which categorizes a list of different attacks into three tiers. Each tier represents the level of access to a targeted computer network necessary to launch specific attacks: Tier 1 – No Network or Computer Access, including Phishing, distributed denial of service (DDoS) etc; Tier 2 – User Access with Limited Privileges, including Password Hacking, Sniffing etc; Tier 3 – Root Access, Administrative Privileges, including Backdoors, Rootkits etc; as well as an additional category of Physical Access, including Electronic Attack, Saboteur etc (shown below, figure 1). The taxonomy includes 19 different attack types in total. An attacker may escalate their access level through attacks conducted from lower Tiers, or alternatively, may be granted higher level access before the simulation commences to reflect preparatory cyber operations, depending on the nature of the scenario and the capabilities of the adversary. Physical Access requires specific physical capabilities to achieve. The most sophisticated attack types will be available to only to adversaries identified as having a high level of cyber operations capability.

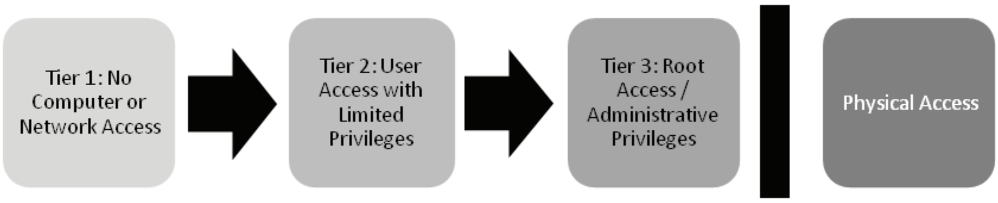

Tier 1: No Computer or Network Access
Tier 2: User Access with Limited Privileges
Tier 3: Root Access / Administrative Privileges
Physical Access

Figure 1: Access tiers

5.3 Costs

Each attack type in the taxonomy is then assigned a *cost* value. This value reflects the resources required to launch an instance of a given attack. The intention of these costs is to bound the attack activities of the OPFOR, in a manner that reflects their capabilities as defined in the scenario. OPFOR will thus be given a budget of resources from which to ‘purchase’ attacks. Assigning these costs with any level of validity is challenging, as many of the factors that will determine an adversary’s ability and requirements to carry out a given attack, such as talent, experience and labour, vary depending on the adversary modelled by the OPFOR. Another difficulty arises because of the scarcity and secrecy of data on this topic.

The premise on which our costs assignments are based is that more sophisticated attack types require more resources to execute. The 15 different cyber attack types discussed in Section 5.2 are therefore weighted in increasing order of sophistication. For example, Phishing ranks at 1, while Kernel-Level Rootkits are weighted at 15.

Recall that we are not proposing a fully developed approach; rather, we are suggesting a potential avenue for the inclusion of cyber effects in simulation based training. For this initial proposed approach, the probability of success of any given attack (discussed below) is fixed at 10%, or 0.1. The attack probability for each type of attack could be refined to differentiate between the cyber defensive posture of different BLUFOR elements as discussed in Section 5.6 below. With the probability of success fixed at 10%, one attack unit will buy the attacker a 0.1 probability of success for a given attack; a higher cost will deliver a higher probability of success. This figure is then multiplied by the sophistication ranking of the attack. Thus, achieving a 10% chance of success with a Phishing attack will cost 1 x 0.1 = 0.1, while achieving the same probability with a Kernel-Level RootKit will cost 15 x 0.1 = 1.5. In this way, the selection of more sophisticated attack types will incur a higher cost upon the allocated budget of the OPFOR, whereas instances of less sophisticated attacks will be more affordable. This linear relationship between cost and probability of success is one of many potential ways to bound OPFOR actions. A separate pool of resources is provided for physical attack types, which are more easily quantified based on available data.

5.4 Targeting

A deployed modern military force will take a variety of different computer networks with it into the field, depending on which capabilities are present. For example, Offensive Support (Artillery, Mortars, Rockets - OS), Command and Control (C2), Air Tasking, and Intelligence, Surveillance and Reconnaissance (ISR) functions may reside on separate networks. Furthermore, different services (Army, Air Force etc) and coalition partners will also likely deploy distinct networks (Wilson, 2007). As such, cyber attacks may target different capabilities via specific networks. For example, OPFOR may seek to deny Offensive Air Support (OAS) to BLUFOR during an offensive action. If such OAS is provided by a coalition partner, it is the partner’s networks, not BLUFORs’, which must be targeted. As the networks and the capabilities that they provide will vary, they must be specified in the simulation scenario and BLUFOR order of battle.

5.5 Duration and scheduling

Next, each attack type is assigned a *duration*, either of effect upon targeted networks, or of the time taken to achieve its intended goal, as well as a time of commencement allowing for synchronisation with the OPFOR scheme of manoeuvre. For example, a DDoS attack may begin two hours into the simulation, and will last 2.25 hours, or a password hacking effort launched at the beginning of the simulation will (if successful) provide an effect for 2.5 hours. Attacks may be repeated sequentially to increase their effect duration, but doing so will incur repeated costs.

To improve the probability of attack success (discussed below), several instances of an attack may also be launched simultaneously or sequentially. Doing so raises a complication, in that a successful attack at the beginning of a series may not need to be repeated. Resources assigned to the cost of unnecessary subsequent attacks should therefore be reallocated, requiring either ‘plan B’ choices for resource expenditure, or real-time adjustment by OPFOR of the cyber attack plan.

Valid durations are difficult to derive in a similar way as valid *costs* (as discussed in Section 5.3). While average durations for some attack types are available, these aggregate a diverse range of factors. For example, DDoS type attacks have been stated as having a global average duration of 0.5 hours (Anstee, 2011), however, a large proportion of these attacks may actually consist of practice attempts by novice attackers (iDefense, 2006). There are many other factors that will affect attack duration, such as the availability of resources and the expertise of the attacker (a botnet commanding millions of computers, or someone hiring their services, versus a handful of issue motivated but unskilled hacktivists working in concert), the determination of the attacker (how long they intend to prosecute the attack), and the defences of the target. This further demonstrates the trade-off between simulation validity and usefulness for our purpose.

In spite of these challenges, representing time is central to simulation efforts, and the approach must assign durations; whenever possible, this will be done based on data gained from a military context (for example, duration of a DDoS attack is pegged to the average of those directed against the government of Georgia during the 2008 Georgia-Russia conflict (Nazario, 2008)).

We must also address the fact that the duration of different attack types may exceed the timeframe of simulation-based training exercises, which tend to be of short duration (less than a week). In cases where the attack duration exceeds the simulation timeframe, the attack effects will be considered permanent for the purpose of simulation.

5.6 Probabilities

The probability of success denotes the chance that a given cyber attack will overcome computer-network defences and achieve its intended objective. For example, a phishing attack gaining the confidence of a user and generating a response, or a sniffing effort disclosing useful intelligence. As **discussed in section 5.3, the approach initially fixes probabilities and combines them with cost. We** understand that this has the result of effectively only assigning a cost to each type of attack. However, we believe that it is beneficial to identify the stochastic nature of cyber attacks in the model. Probabilities are currently not refined because of the paucity of data. We believe that probabilities can and should be refined as more data becomes available. This probability value will force OPFOR to weigh the cost of achieving higher degrees of certainty through combined lower cost attacks against the benefits of a single higher cost attack. As such, assigned probabilities of success will be useful in constraining OPFOR activities as well as educating simulation participants on the stochastic nature of specific attack types.

5.7 Effects

Contemporary military operations are highly sensitive to collateral damage. As such, modern military doctrine seeks to deliver a range of *effects* that further operational goals, which may be achieved through both kinetic and non-kinetic actions, rather than purely through destruction of enemy forces (Kelly & Kilcullen 2004). The proposed approach therefore describes the effect of cyber attacks on targeted computing resources with respect to processes and information. These effects, drawn from the paper *Computing the Impact of Cyber Attacks on Complex Missions* (Musman et al. 2011) are categorised as follows:

- **Interruption-** Targeted processes or information are unavailable for some time period and will not commence/ be accessible until the incident is recovered;
- **Modification-** Process characteristics have been altered in a way that can affect the output/result of the process; or Information has been altered, meaning that the processes that use it may fail, or produce incorrect results;
- **Degradation-** Speed of a targeted process/ rate of access to information on a targeted network is slowed by some multiple, or the quality or precision of information produced by an activity is decreased;
- **Fabrication-** A false mission instance has been inserted into the system, which may interfere with real mission instances; or false information has been entered into the system,
- **Interception-** The process (perhaps software, perhaps embodied in hardware), or information, has been captured by the attacker, and finally;
- **Unauthorized use-** Raises the potential for future effects on information, or unexpected outcomes on processes. In our approach, Unauthorised use may be further classified in terms of Escalation to Tier 2 access, or Escalation to (or maintenance of) Tier 3 access. Each attack type may deliver a combination of these different effects upon success.

5.8 Imposition of effects

The final element of this approach is the imposition of the computer/network effect of a specific cyber attack upon the user. This imposition should mimic the actual real-world effect of the cyber attack, but with minimal unintended disruption to the simulation itself. Simplicity is therefore a priority in choosing how to impose effects. The effects may be achieved by either physical or electronic interference with the simulation participants' ability to interact with the simulation, by providing information to OPFOR, or by imposing rules that restrict BLUFOR actions. The imposition of an *Interruption* effect for example may involve simply switching off a targeted user's monitor, or unplugging their mouse and keyboard to prevent interaction with forces under their command, for the duration of the attack. *Modification* and *Fabrication* are more complicated, and require input from the exercise controller (EXCON), for example the temporary addition, removal, or relocation of displayed OPFOR or BLUFOR units from the simulation in order to confuse the BLUFOR intelligence or C2 picture and represent the manipulation or *fabrication* of this information by a cyber attacker. *Degradation* can be imposed through the restriction of the BLUFOR targeted user to a lower level of simulation resolution (enforcing a zoomed out view for example), or by running other resource intensive applications on the user computer to slow it down. *Interception* may be imposed relatively easily, by handing OPFOR screen shots or a real time feed of the BLUFOR C2 and intelligence picture. *Unauthorised Use* can simply enable the OPFOR cyber attacker to carry out other previously unavailable attack types. Effects targeting specific capabilities other than C2 (ISR, OS, OAS etc) will see those capabilities either limited or denied for use by BLUFOR.

6. Approach in application

The application of this approach to a simulation-based exercise may then unfold as follows:

6.1 Selection of scenario including OPFOR capabilities

The exercise is based on a scenario that includes a specified adversary (OPFOR). The nature of this adversary will dictate the resources available for cyber attack. For example, a high technology nation-state military adversary with a strong cyber operations capability maybe apportioned a high propensity for cyber attack and physical attack on computing resources. They could potentially begin the exercise with *Tier 3* access, representing prior penetration of BLUFOR networks. Alternatively, an irregular adversary albeit with some medium level cyber operations capability would have a lower level of cyber attack resources and limited potential for physical attack on computing resources. Such an adversary could commence the exercise with *Tier 1* or *2* access. Where the adversary is a low technology irregular force with no cyber operations capabilities, cyber attacks may originate from sympathetic "hacktivists" of limited ability.

6.2 Resource allocation

Once the adversary is identified, a budget of resources can be apportioned and appropriate restrictions can be placed upon the OPFOR. Prior to the beginning of the simulation, the OPFOR can spend their budget on a selection of attacks chosen and sequenced to support their scheme of manoeuvre, choosing from the list of attacks discussed in Section 5.2 as though choosing a meal from an à la carte menu. They may for example opt to undertake attacks with an *Interception* effect initially to help build a detailed intelligence picture of BLUFOR. They may then launch DoS attacks that will disrupt BLUFOR C2 or OS synchronised with an offensive kinetic manoeuvre, or alternately keep resources for such attacks in reserve to use as a spoiling measure when BLUFOR commences offensive kinetic manoeuvres.

6.3 OPFOR adjustments during the simulation

As the simulation progresses, OPFOR may be able to modify the selection and sequencing of attacks to adapt to the evolving operational picture; for example, as discussed in Section 5.5, some successful attacks will not need to be repeated. Where cyber attacks originate from sympathetic ‘hacktivist’ from outside of OPFOR, this synchronisation of cyber with kinetic activities would be restricted to the point where the ‘hacktivist’ attacks could be randomised.

6.4 Effects calculation

While the simulation is under way, the success of each attack is computed using its *probability* value. As attacks succeed, their effects can then be appropriately imposed upon the BLUFOR. Successful attacks providing higher *Tier* access will enable new attack options, which may require fine tuning of the cyber operation plan, and will lead to an escalating imposition of effects against BLUFOR commanders.

7. Further work

This paper describes a rough outline of a proposed approach, and seeks comments from the research community. Much works remains to be done. As the approach is still in development, and due to the previously discussed shortage of military-specific cyber attack data, many improvements remain to be made. Values for *cost*, *probability* and *duration* may be refined as more data becomes available, this would allow for the fine-tuning of the economics of attack as a means of OPFOR attack selection and the severity of effects for appropriateness in training and education. The taxonomy of attacks can also be updated as novel attacks come to light. It is noteworthy that the approach to replicate the effects without attempting to model the attacks themselves can have the benefit of incorporating undisclosed attacks in the simulation, without alerting potential adversaries to the specifics of the attack. We also believe that the metrics and measures described in the paper “Metrics Framework of Cyber Operations on Command and Control” (Bernier et al. 2012), also presented at this conference, will further assist in the understanding of the effects of cyber operations on C2 capability.

8. Other applications

Although we have focussed on an application to the education of senior military commanders, this approach could easily be adapted for business use through a series of simple substitutions. The simulation of a military operation over which this approach is proposed could be substituted for a simulation of normal business operations, while the adversary could take the form of competing companies, hacktivists or cyber-criminals. Military capabilities could be substituted for business processes. The replication of the cyber attack effects could also then begin to benefit senior business managers.

9. Conclusion

The expected growth in military cyber operations poses a serious threat to information technology-enabled modern military forces who depend on computer networks. As such, a pro-active training and education effort is required to prepare senior military decision makers for future conflicts. Simulation offers a means to achieve this, although data from which to develop a model of military cyber operations is scarce. However, to provide a general military cyber operations education to senior military leaders it is more important to simulate the effects of cyber attacks than to replicate the actual mechanics of the attacks themselves. Therefore the data-intensive and highly complex task of replicating entire attack processes with a high level of fidelity can be avoided, while still providing the needed educational benefits.

The approach proposed here provides a simple and cost-effective method for the incorporation of military cyber effects into constructive simulation-based training exercises, which enables the integration of cyber operations with kinetic operations. Within simulations, adversary actions may be controlled by restricting resources and sophistication available to the OPFOR. This approach could potentially be adapted for business use, where more data is available. It is anticipated that this approach will be developed and refined as more data becomes available through use and other research efforts in the field of cyber effect simulation.

References

Anstee, D. (2011) “DDoS Attack Trends Through 2010: Infrastructure Security Report & ATLAS Initiative” [online], Arbor Networks, http://ripe62.ripe.net/presentations/88-Darren-Anstee-AA-RIPE-2011-DDoS_Trends.ppt.pdf (accessed 01/02/2012)

Bernier, M., Leblanc, S. and Morton, B (2012) “Metrics Framework of Cyber Operations on Command and Control” Defence Research and Development Canada – Canada Operational Research and Analysis, Ottawa.

Broad, J., Markoff, J. and Sanger, D. (2011) “Israeli Test on Worm Called Crucial in Iran Nuclear Delay”, *The New York Times*.

Bye, R., Schmidt, S., Luther, K. and Albyrak, S. (2008) “Application-Level Simulation for Network Security”, in *Proceedings of the First International Conference on Simulation Tools and Techniques for Communications, Networks and Systems*.

Carver, C., Surdu, J.,Hill, J., Ragsdale, D., Lathrop, S., & Presby, T., (2001) “Military Academy Attack Defense Network (MAADNET)” [online] http://www.bucksurdu.com/professional/documents/maadnet.pdf. (accessed 19/03/12)

Geers, K., (2010) “Live Fire Exercise: Preparing for Cyber War, in Journal of Homeland Security and Emergency Management”, 7(1) [online] http://www.bepress.com/jhsem/vol7/iss1/74 (accessed 19/03/12)

Cohen, F. (1999). Simulating cyber attacks, defences, and consequences. *Computers & Security* (pp. 479-518). Elsevier Science Ltd.

Clarke, R. (2009) “War From Cyberspace”, *The National Interest,* November-December Issue, p 32.

Falliere, N., Murchu, L., & Chien, E. (2011) “W32.Stuxnet Dossier Version 1.4.” [online], Symantec, http://www.symantec.com/en/ca/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf. (accessed 01/02/2012)

Fulghum, D. A., Wall, R., & Butler, A. (2007) “Cyber-Combat"s First Shot” *Aviation Week & Space Technology,* Vol. 167 No. 21, p. 28.

Futoransky, L., Notarfrancesco, L., Richarte, G. & Sarraute, C. (2003) “Building Computer Network Attacks” [online]http://www.coresecurity.com/files/attachments/Futoransky_Notarfrancesco_Richarte_Sarrute_NetworkAttacks_2003.pdf. (accessed 19/03/12)

Grossman, Dave (1996) “On Killing: The Psychological Cost of Learning to Kill in War and Society”, Back Bay Books, New York.

iDefense Security Report (2006) “Distributed Denial of Service (DDoS) and Botnet Attacks” [online] http://complianceandprivacy.com/WhitePapers/iDefense_DDoS_20060428.pdf. (accessed 01/02/2012)

Kelly, J and Kilcullen,D. (2004) “Chaos Versus Predictability: A Critique of Effects-Based Operations,” *Australian Army Journal*, 2 (Winter 2004), p. 90.

Liljenstam, M., & Liu, J. (2006) “Rinse: the real-time immersive network simulation environment for network security exercises (extended version)” *SIMULATION, 82*(1), 43-59.

Musman, S., Tanner, M., Temin, A., Elsaesser, E., and Loren, L. (2011) “Computing the Impact of Cyber Attacks on Complex Missions”, The MITRE Corporation, in *Systems Conference (SysCon), 2011 IEEE International.*

Musman, S., Temin, A., Tanner, M., Fox, D., & Pridemore, B. (2009) “Evaluating the Impact of Cyber Attacks on Missions”, in *Proceedings of the 5th International Conference on Information Warfare and Security*, pp. 446-456.

Nazario, J (2008) “Georgia DDoS Attacks – A Quick Summary of Observations” [online], Arbor Networks, http://ddos.arbornetworks.com/2008/08/georgia-ddos-attacks-a-quick-summary-of-observations/ (accessed 01/02/2012)

Park, J. S., Lee, J., K, H. K., Jeong, J., Yeom, D., & Chi S. (2001) “Secusim: a tool for the cyber-attack simulation”. *Information and Communications Security* (pp. 471-475). Heidelberg: Springer Berlin.

Partington, A., Leblanc, S. & Morton, B. (2011) “A Taxonomy of Cyber Attacks for Use in Computer Network Attack Modelling and Simulation-Version 2.0”, Royal Military College of Canada Computer Security Laboratory, Kingston.

Stytz, M. & Banks, S. (2010) “Addressing Simulation Issues Posed by Cyber Warfare Technologies” [online] *SCS M&S Magazine*, http://www.scs.org/magazines/2010-07/index_file/Files/Article_Stytz.pdf (accessed 01/02/2012)

Wilson, C. (2007) “Network Centric Operations: Background and Oversight Issues for Congress”, Technical Report ADA466624, Library of Congress Washington DC Congressional Research Service, Washington DC.

A Taxonomy of Technical Attribution Techniques for Cyber Attacks

Andrew Nicholson, Tim Watson, Peter Norris, Alistair Duffy and Roy Isbell
De Montfort University, Leicester, UK
abn@dmu.ac.uk
tw@dmu.ac.uk
pdn@dmu.ac.uk
apd@dmu.ac.uk
risbell@dmu.ac.uk

Abstract: In recent years the number of cyber-attacks has dramatically increased, affecting military, government, business and home users. For example, the UK Ministry of Defence claims to have blocked and investigated over 1000 serious cyber-attacks in 2010 while in 2011 Detica reported that the cost of cybercrime in the UK is estimated to be £27 billion per annum. In cyber-attacks numerous methods exist that can be used to discover information about the attacking entity, otherwise termed as attribution. Attribution is a desired quality to counter a variety of attackers. Cyber-crime attribution can aid police investigations in identifying cyber criminals. In cyber warfare and conflict an attribution capability is desired to enhance decision making of Computer Network Operations (CNO). Attribution of terrorist cyber-attacks may help to prevent future attacks. Highly publicised attacks such as Stuxnet and Night Dragon have been subject to intense analysis, yet published attribution of these attacks has been minimal. The complexity of reliable attribution is increased by an attacker's ability to route attacks through compromised systems, anonymised networks, proxy servers and various jurisdictional boundaries. There are numerous technical attribution techniques ranging from traceback, malware inspection and honeypot deployment. In this paper we present a taxonomy to classify these techniques, using five different classes: acquired attributes, proposed/in use, external party involvement, sabotage opportunity and prepositioning depth. The novelty of this paper is its scope; classifying the landscape of technical attribution techniques.

Keywords: cyber, attribution, profiling, traceback, honey-pots, taxonomy

1. Introduction

In the context of a cyber-attack, attribution has been defined as "determining the identity or location of an attacker or an attacker's intermediary" (Wheeler, 2003). The identity may be of digital form (e.g. Internet Protocol (IP) address, account name) or physical form (e.g. name, geographical address) (Guan and Zhang, 2010: 197).

Many cyber defensive technologies exist, such as anti-virus, firewalls and intrusion detection systems (IDS). However attackers continue to succeed and do so with minimal apprehension. An attribution capability is desirable for the following reasons (Hunker et al., 2008: 5):

- The prospect of an attacker being identified can serve as a deterrent to future attacks.
- Knowing the identity of an attacker, and information gained in the process of attribution, can be used to improve defensive techniques.
- Attribution, even partial attribution, can provide the basis for interrupting attacks in progress.

The history of attributing cyber attacks can be traced back to the 1980's when Stoll (1989) profiled and learnt about intruders who penetrated the computer networks of Lawrence Berkeley National Laboratory. His techniques formed the basis of Honeypot research, which flourished in the early 2000's and is now well founded and used in business and academia. Also at the turn of the millennium, a flurry of traceback techniques were proposed which, in most cases, called for changes to the core Internet infrastructure. Shortly after, Internet redesigns were considered for attribution purposes, and researchers now consider how non-technical and technical means can be used together to attribute effectively.

In the field a number of technical difficulties have been identified. The 'traceback problem' is concerned with the idea that no authentication of the packet header source address takes place as packets travel across routers in the Internet infrastructure. Therefore attackers are able to forge packets and evade attribution. The ‘stepping stone problem’ is concerned with the idea that attackers may use multiple hosts to launch attacks. Clark and Landau (2010: 26) define this as ‘multi-stage attacks’ and note that routing through multiple jurisdictional boundaries can cause additional difficulties and further evade attribution. Wheeler (2003) provides a further discussion of technical difficulties. At a non-technical level there are also difficulties. Should a ‘perfect’ technical solution that overcomes technical problems ever be realised, privacy advocates highlight that governments or intelligence agencies could use techniques for nefarious purposes. Some argue that attribution solutions such as an Internet redesign would erode the anonymity that society values (Clark and Landau, 2010: 26).

Clark and Landau (2010: 25) state that "There are many types of attribution, and different types of attribution are useful in different contexts". Technical examples include traceback, malware inspection and honeypots. Non-technical examples include tracking money flows, intelligence efforts and cui bono (who benefits) analysis. We limit the scope of this paper to classifying technical attribution techniques. Attribution already takes place in applications, such as banking, where customers are positively identified through authentication mechanisms (something you have/know/are). We rule these scenarios out of the scope of this paper and focus on proposals that have been designed for attribution of attacks. Finally, rather than classify every single technique (for example, in traceback there are many), we categorise and classify the most promising techniques.

The remainder of this paper is structured as follows: Section 2 describes related work, Section 3 defines the objective and approach, Section 4 describes the taxonomy criteria, Section 5 is the taxonomy, Section 6 concludes and Section 7 considers future work.

2. Related work

Similar taxonomies covering the spectrum of technical attribution techniques are limited. Santhanam et al (2006) presented a taxonomy on traceback techniques classifying by reactive and proactive approaches. Seifert et al (2006) presented a taxonomy on honeypots, creating classes of common characteristics. Wheeler (2003) presented a survey of attribution techniques, focusing on traceback techniques and provided a simplistic taxonomy of the field (see Figure 1). The taxonomy provides a basic flat structure of available technical attribution techniques, although does not assign characteristic classes.

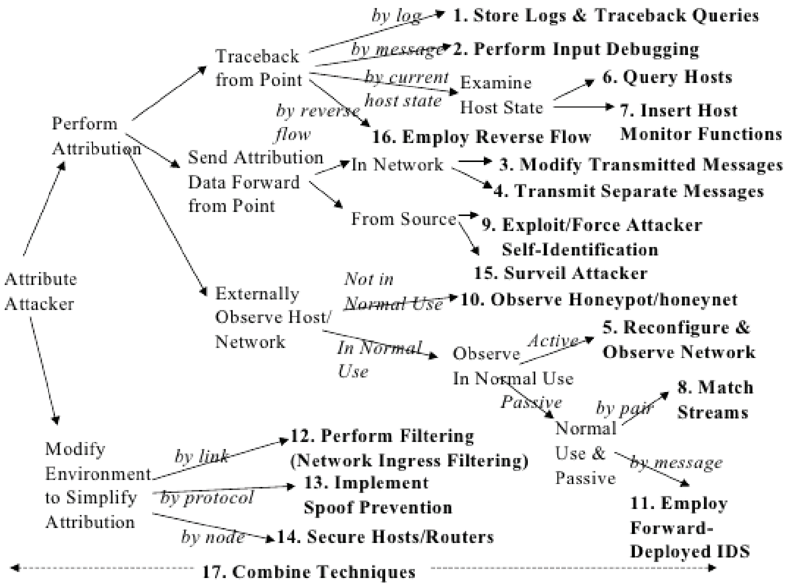

Attribute Attacker
Perform Attribution
Traceback from Point
by log
1. Store Logs & Traceback Queries
by message
2. Perform Input Debugging
by current host state
Examine Host State
6. Query Hosts
7. Insert Host Monitor Functions
by reverse flow
16. Employ Reverse Flow
Send Attribution Data Forward from Point
In Network
3. Modify Transmitted Messages
4. Transmit Separate Messages
From Source
9. Exploit/Force Attacker Self-Identification
15. Surveil Attacker
Externally Observe Host/ Network
Not in Normal Use
10. Observe Honeypot/honeynet
In Normal Use
Observe In Normal Use
Active
5. Reconfigure & Observe Network
Passive
by pair
8. Match Streams
Normal Use & Passive
by message
11. Employ Forward-Deployed IDS
Modify Environment to Simplify Attribution
by link
12. Perform Filtering (Network Ingress Filtering)
by protocol
13. Implement Spoof Prevention
by node
14. Secure Hosts/Routers
17. Combine Techniques

Figure 1: Wheeler's (2003) taxonomy

3. Objective and approach

The objective of this work is to create a new taxonomy that covers the landscape of technical attribution techniques. Previous taxonomies have covered distinct fields such as traceback, honeypots and IDS, or have provided a basic, flat taxonomy. In this paper we create a taxonomy of technical attribution techniques and combine these distinct fields based on characteristics.

A taxonomy is concerned with identifying, naming and classifying objects. Our taxonomy complies with characteristics compiled by Seifert et al (2006); comprehensible, conforming, useful, determinism, objectivity, repeatability, exhaustive, mutually exclusive and specificity. Our approach involved systematically analysing the literature of technical attribution techniques and assigning these techniques to a set of classes that we define in Section 4.

4. Taxonomy criteria

We now describe each class, it's possible values and the chosen classification method. In Figure 2 we present the taxonomy in graphical form, complete with a key to show the results of applying classes to objects.

Class 1: Acquired Attributes

Description: Guan and Zhang (2010: 197) identify two types of identity that may be acquired as a result of technical attribution; physical and digital. This classification identifies whether a given technique can reveal physical, digital or both.

Values: Physical | Digital

Classification Method: Values are not exclusive and both may be selected, in fact both is preferable.

Class 2: Proposed / In Use

Description: Due to the high level view of the taxonomy we find that a number of technical attribution techniques have been in use for some time, while others have reached the proposal stage and not proceeded further. We define In Use as having a generally known use in the wide scale Internet.

Values: Proposed | In Use

Classification Method: Values are binary and exclusive and thus only one value may be selected.

Class 3: External Party Involvement

Description: External party involvement, such as ISPs or governments are sometimes required by technical attribution techniques, which impacts their applicability.

Values: Required | Optional | Not Required

Classification Method: Values are exclusive and thus only one may be selected.

Class 4: Sabotage Opportunity

Description: Researchers have identified numerous ways to subvert or bypass technical attribution techniques. The attribution data that these techniques produce is of little relevance if there is the possibility that an attacker may have subverted or modified the data. Our scale is based on our interpretation of the literature.

Values: High | Medium | Low

Classification Method: Values are exclusive and thus only one may be selected.

Class 5: Prepositioning Depth

Description: Wheeler (2003) discusses depth of prepositioning as being preparations that must take place in order for the technique to be used. Our scale is based on our interpretation of the literature.

Values: High | Medium | Low

Classification Method: Values are exclusive and thus only one may be selected.

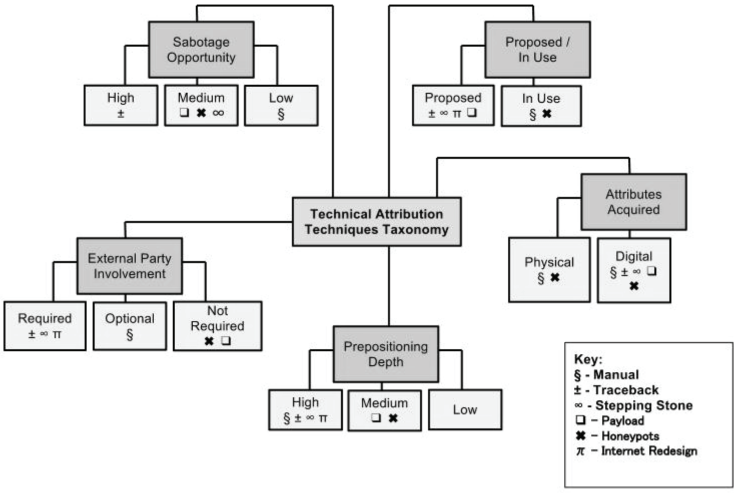


Figure 2: Taxonomy of attribution techniques

5. Taxonomy of attribution techniques

In this section we classify the spectrum of technical attribution techniques in accordance with the classes outlined in Section 4. Techniques are categorised into the following sub-sections; manual, traceback, stepping stone, payload, honeypots, Internet redesign and finally frameworks.

5.1 Manual attribution

While a number of attribution techniques involve manual efforts, we define manual attribution as those which primarily consist of manual efforts. Manual attribution requires human involvement, most often technical specialists, and can be classed as network-based or host-based. Usage varies, for example, in network-based manual attribution specialists may use command line tools such as tcpdump and traceroute to learn more about a source IP address in a suspicious packet. Internet Service Providers (ISP) may be contacted to further investigations. However, this technique requires manpower, technical knowledge and a network of contacts. The duration for which ISPs maintain logs is also a critical factor. Additionally, ISPs may not speak the same language, working hours may clash and geographical politics may erode communication.

In a host-based method, one approach may be for a specialist to examine malware by reverse engineering. Linguistic properties such as distinctive programming practice may be identified. Fuzzy hash analysis may be used to identify code in malware that appears in other malware/software (Apel et al., 2009). This technique may reveal the author of the malware in question, but not necessarily the instigator of the attack as malware is sold on the Internet.

5.1.1 Classification

- Acquired Attributes - Physical and Digital - A variety of both attribute types may be identified, however, this is on a case by case basis, offering little consistency.
- Proposed/In Use - In Use
- External Party Involvement - Optional - However is more successful with external party involvement and less intrusive for ISPs than traceback techniques.
- Sabotage Opportunity - Low - A point of concern here is that investigator mistakes may be made, as is entailed in any manual work. Also intelligent attackers may introduce deceptive techniques, such as removing evidence of their intrusion.
- Prepositioning Depth - High - Manual attribution requires human involvement, meaning that it is expensive in both time and cost. This means that less critical attacks, or less-able victims, may not have the skills, time or money to perform manual attribution.

5.2 Traceback techniques

Traceback techniques aim to address the traceback problem and are defined as follows: “Any attribution technique that begins with the defending computer and recursively steps backward in the attack path towards the attacker.” (Hunker et al., 2008: 5). Traceback techniques account for a large proportion of academic attribution proposals. The process results in the creation of an attack graph (see Figure 3) containing at least the originating IP address and possibly intermediate router IP addresses. Traceback has been approached in a number of ways, the first we address is marking packets.

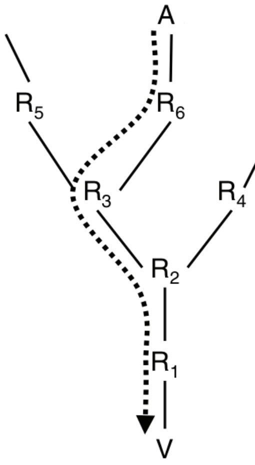


Figure 3: A theoretical attack graph (Savage et al., 2000)

5.2.1 Packet marking

Packet marking involves the modification of packet header fields by core routers so that they contain information that describes the path that the packet has taken. Savage et al (2000) proposed probabilistic packet marking (PPM) in order to counter the limited storage space that packet header fields offer. This technique probabilistically marks path data, therefore a victim must collect multiple packets to generate an attack graph. Marking information is stored in the 16-bit Identification field, which means that legitimate fragmented packets are lost. The authors suggest 75 packets would be sufficient when the path length is 10 and the number of attackers is small. When the number of attackers is large, this technique becomes ineffective; thousands of packets are required and convergence time increases.

Song and Perrig (2002) proposed Advanced and Authenticated Marking Schemes (AMS and AMSII). AMS advances upon the work of Savage et al by compressing entire traceback data into the Identification field. AMS II introduces authentication so that each router uses a unique secret key to mark packets. Despite these modifications, the technique still remained weak against distributed denial of service attacks (DDoS) and spoofing. Goodrich (2002) proposed Randomize-and-Link which aimed to counter these weaknesses. This technique uses large checksums to link packets across a wide spectrum meaning an attacker’s chance of spoofing is minimised. Finally Belenky and Ansari (2003) proposed Deterministic Packet Marking (DPM), which aimed to stop spoofing and allow low packet quantity.

5.2.2 Packet logging

In packet logging routers store information about packets that have passed through them. During or after an attack, data stored at each router may be analysed to determine if the router was involved in the attack path. Snoeren et al (2001) proposed the first feasible logging technique, Source Path Isolation Engine (SPIE). SPIE-enabled routers store a hash of the 32-bit IP packet in hierarchical bloom filters, therefore alleviating privacy and storage concerns. SPIE prompted a number of extensions ranging from high-speed Internet (Li et al., 2004) to IPv6 (Strayer et al., 2004). A problem with Snoeren et al's proposal is that single packet attacks, such as Ping of Death, could not be traced. Zhang and Guan (2007) proposed a single-packet traceback scheme. A victim is able to send a single packet signature to an enabled router which confirms whether or not the packet has been stored. One problem in logging is determining who should incur costs for infrastructure changes. Haeberlen et al (2011) proposed Packet Attestation in which responsibility is placed upon the ISP. ISPs must hash all upstream packets which can then be verified by a victim. Uniquely, the authors suggest that ISP customers should pay for the service, calculating costs to be roughly $0.68 per month per customer.

5.2.3 ICMP traceback

In this technique, first proposed by Bellovin et al (2003), routers generate custom ICMP packets that contain path data. Each custom ICMP packet contains the IP address of the router, previous and next hop routers. The ICMP packet is probabilistically sent to the source or destination IP address. Packets are created with low probability; hence a victim collects multiple ICMP packets in order to create an attack graph. Unlike packet marking, ICMP packets are sent out-of-band, meaning that the original packets are not modified and are forensically sound. However, treatment of ICMP packets in the Internet is diverse. ICMP network attacks have resulted in ISPs and organisations dropping ICMP packets, meaning that traceback packets could be lost.

5.2.4 Hybrid traceback

Hybrid traceback techniques combine two or more types of traceback. A hybrid should theoretically enable the best features of each technique, such as single packet traceback (logging) and reduced storage and access times at routers (marking). Gong and Sarac (2005) proposed a hybrid of marking and logging techniques. In this proposal routers always mark packets and probabilistically log hashes of packets using Snoeren et al’s (2001) approach. Results find that this hybrid technique reduces storage overhead at routers by half when compared to other logging proposals.

5.2.5 Classification

- Acquired Attributes - Digital - Traceback approaches provide an attack graph containing an attacker's IP address, and optionally, intermediate routers. However, the stepping stone problem explains why this data may be of little use. Compromised hosts are only useful if the owner is cooperative, allowing forensic investigation of their machine(s).
- Proposed/In Use - Proposed - Despite research dating back to 1998, traceback proposals have remained firmly in a proposed state, apart from within local networks.
- External Party Involvement - Required - ISPs around the globe would need to be involved. The onus of who should pay for and manage these techniques is unclear, although proposals have been made (Haeberlen et al, 2011).
- Sabotage Opportunity - High - Researchers have highlighted numerous circumvention techniques. Traceback techniques may introduce new attack vectors. For example, packet logging produces additional traffic and could cause a DDoS attack in itself, also affecting legitimate network traffic.
- Prepositioning Depth - High - Traceback is intrusive, requiring ISP infrastructure changes for deployment and packet/router modifications, additional traffic, storage and processing requirements. Implementing and maintaining such changes would be both costly and timely.

5.3 Stepping stone attribution

An attacker may route their attack through numerous entities, such as compromised hosts or proxy servers, which results in the victim viewing a source address that does not represent the attacker. Techniques have been proposed for the stepping stone problem, first discussed by Staniford-Chen and Heberlein (1995). They defined the first proposal for stepping stone attribution, termed ‘thumbprints’. This solution involves watermarking traffic and then using intermediate routers to report whether or not the watermarked traffic has passed through. To be effective this requires a globally configured router and thus a high prepositioning depth. Wheeler (2003) defined another class of solutions to the stepping stone problem as stream matching, which involves matching characteristics between two streams to assert a relationship, such as timing, packet size and content. In Zhang and Paxson’s (2000) ON/OFF technique, two characteristics are used to identify consistencies; timing and packet size. Since the packet content is not used, this technique works against encrypted traffic. Wheeler (2003) identifies that timing stream matching is defeated by zombies, in which a zombies response may occur much longer after the initial request was made, rather than symmetrically.

5.3.1 Classification

- Acquired Attributes - Digital
- Proposed/In Use - Proposed - Similar challenges to traceback.
- External Party Involvement - Required - Similar challenges to traceback.
- Sabotage Opportunity - Med - Researchers have highlighted ways to circumvent this technique, such as employing high numbers of stepping stones (Wheeler, 2003).
- Prepositioning Depth - High - ISPs need to make costly changes to their infrastructure.

5.4 Payload attribution

Investigators sometimes do not have access to packet headers and therefore traceback techniques are ineffective. They may however, have access to a packet payload excerpt. An investigator can query the system using the excerpt to find out more information, such as Source/Destination IP address and port numbers involved. In payload attribution, researchers find that it is not acceptable to store full packet capture data and therefore approach the problem with compression and hashing techniques such as bloom filters (Shanmugasundaram et al., 2004). This technique may identify infected machines within the network and is therefore useful for remediation.

5.4.1 Classification

- Acquired Attributes - Digital
- Proposed/In Use - Proposed - Prototypes have been designed and tested.
- External Party Involvement - Not Required - External parties do not need to be involved, the technique takes place in an Intranet rather than Internet.
- Sabotage Opportunity - Med - Researchers have highlighted that this technique is foiled by high volume, repetitive, network traffic e.g. DDoS packets (Shanmugasundaram et al., 2004).
- Prepositioning Depth - Med - Efficient network monitoring needs to take place in order to capture and archive full content data.

5.5 Honeypots

Honeypots are specially crafted systems that attempt to draw attackers towards them. This is accomplished by imitating vulnerable systems, services and software. They are monitored with tools such as Sebek so that interaction between the attacker and the honeypot is stored and can be analysed by an investigator. Within this interaction information, attribution artefacts may be found. Honeypots combine an automated data capture process with a manual attribution process involving the analysis of captured data. Advanced proposals may also partially automate the analysis process. Raynal et al (2004a) presented a methodology for honeypot forensics. In this methodology both host-based and network-based honeypot data is correlated to identify patterns and highlight unexplained results. In further work, Raynal et al (2004b) considered 'black hat profiling' in which passwords, email addresses and URLs found in their honeypot data was examined. The authors used deductive reasoning to draw conclusions, such as: did the attacker destroy anything? If so then the motivations of the attacker may be predatory, so the analyst might focus on discovering who has a personal intent against the target, such as an irate customer or inside attacker.

Ramsbrock et al (2007) created numerous honeypots in which attackers were monitored after successful SSH compromises. Authentication details were deliberately simple to guess and were successfully brute-forced by low-skill attackers. The authors defined five post-compromise activities; configuration checking, password changing, file downloading, installing/running rogue code and changing system configuration. All of which revealed attribution artefacts such as exploit packs chosen, keyboard typing mistakes, command line preferences and a perceived skill level. Wagener et al (2009) proposed adaptive honeypots that engage with attackers using probabilistic events. Programs unexpectedly fail or insults are directed at the attacker. This proves to be a novel way to glean attribute data from an entity, by actively engaging with them and exploiting their tolerance threshold. The persistence of an attacker could also be measured using such a technique. Similarly, Bilar and Saltaformaggio (2011) discuss the use of baits to encourage attackers to engage with the honeypot.

5.5.1 Classification

- Acquired Attributes - Physical and Digital - Honeypots offered the most varied selection of attributes, however, they are restricted to capturing data only when an intruder enters the honeypot system.
- Proposed/In Use - In Use
- External Party Involvement - Not Required
- Sabotage Opportunity - Med - Researchers have identified numerous methods for detecting honeypots such as analysing clock skews (Kohno et al., 2005). Therefore it could be claimed that honeypots are only useful for catching lesser skilled attackers.
- Prepositioning Depth - Med - Honeypots require maintenance and creating convincing honeypots is a challenge.

5.6 Internet redesign

Finally researchers have considered a complete redesign of the Internet, disregarding the current design. Mike McConell (2010), once Director at the National Security Agency, stated: “We need to re-engineer the Internet to make attribution, geolocation, intelligence analysis and impact assessment, who did it, from where, why and what was the result”. However, this is a complex task, one that has been explored by a small number of researchers. Efforts such as the National Science Foundation’s Future Internet Network Design and the Department of Defence’s Global Information Grid are already underway. Passport systems have been devised, in which anonymity on the Internet is almost extinct (Liu et al., 2008). Anderson (2008) proposes Accountable Internet Protocol (AIP) as a replacement for IP which uses self-certifying addresses. These techniques have a low readiness level and are costly.

5.6.1 Classification

- Acquired Attributes - It is not currently possible to assign this class.
- Proposed/In Use - Proposed
- External Party Involvement - Required - Global cooperation would be required from ISPs.
- Sabotage Opportunity - It is not currently possible to assign this class.
- Prepositioning Depth - High

5.7 Attribution frameworks

Finally, we define an additional technique, attribution frameworks, as more than one attribution technique used together in order to perform attribution. For example, the WOMBAT Project, is a collaboration between commercial organisations, academic groups and worldwide CERTs (Dacier et al., 2009). An API communicates with honeypots, darknets and relevant shared security databases to streamline incident response. One project goal is root cause analysis, which aims to identify involved groups, organisations, machines and attack methods. Since this technique combines numerous attribution techniques that have been previously discussed, it has not been possible to appropriately classify based on the given classes. However, it has been included for completeness.

Table 1: Summary of taxonomy

Name Category	Manual	Traceback	Stepping Stone	Payload	Honeypot	Re-design
Acquired Attributes	Physical and Digital	Digital	Digital	Digital	Physical and Digital	NA
Proposed / In Use	In Use	Proposed	Proposed	Proposed	In Use	Proposed
External Party Involvement	Optional	Required	Required	Not Required	Not Required	Required

Name Category	Manual	Traceback	Stepping Stone	Payload	Honeypot	Re-design
Sabotage Opportunity	Low	High	Med	Med	Med	NA
Prepositioning Depth	High	High	High	Med	Med	High

6. Summary

In this taxonomy we have introduced cyber attribution and described the key problems. We then presented a taxonomy criteria which defined five suitable classes for the high level scope of technical attribution techniques. We then analysed the attribution literature and assigned our set of attribution objects to classes. Figure 2 and Table 1 summarise the findings of the taxonomy.

We found that each attribution technique provides its own merits and, as was suggested by Wheeler (2003), a combination of technical attribution techniques is likely to be most successful. Further, researchers have commented that a combination of technical and non-technical attribution is likely to provide an ideal solution with real-world uses (Hunker et al., 2008).

One possible causality that warrants further investigation is between Prepositioning Depth, External Party Involvement and Proposed/In Use. We suggest that High and Required values limit In Use values, respectively. This is true in traceback techniques, where despite numerous years of research, techniques have not seen widespread deployment, due in part to the changes required in the Internet infrastructure.

7. Future work

Non-technical attribution techniques were outside of the scope of this paper and therefore it may be in the interest of the research community to classify these techniques. Other classes might extend the taxonomy, such as skill level of attackers attributed by techniques and reliability of attribution technique. Additionally, Use/Proposed, could employ a scheme such as the Technology Readiness Level (TRL) scale. Furthermore, researchers could validate and extend our classes with practical experiments, testing the technical attribution techniques in laboratory conditions.

References

Andersen, D., Balakrishnan, H., Feamster, N., Koponen, T., Moon. and D., Shenker, S. (2008) Accountable Internet Protocol (AIP), Proceedings of ACM SIGCOMM Conference on Data Communication, pp 339-350.

Apel, M., Bockermann, C. and Meier, M. (2009) Measuring Similarity of Malware Behavior, SICK IEEE Explorer, pp 891-898.

Belenky, A. and Ansari, N. (2003) IP Traceback with Deterministic Packet Marking, IEEE Communications Letters, Vol.7, No.4, pp 162-164.

Bellovin, S., Leech, M. and Taylor, T. (2003) ICMP Traceback Messages, Internet Draft.

Bilar, D. and Saltaformaggio, B. (2011) Using a Novel Behavioral Stimuli-Response Framework to Defend Against Adversarial Cyberspace Participants, 3rd International Conference on Cyber Conflict (ICCC), pp 1-16.

Clark, D. and Landau, S. (2010) Untangling Attribution, Proceedings of Workshop on Deterring CyberAttacks: Informing Strategies and Developing Options for U.S Policy.

Dacier, M., Pham, V. and Thonnard, O. (2009) The WOMBAT Attack Attribution Method: Some Results, Proceedings of 5th International Conference on Information Systems Security, pp 19-37.

Gong, C. and Sarac, K. (2005) IP Traceback Based on Packet Marking and Logging, Proceedings of IEEE International Conference on Communications, pp 1043-1047.

Goodrich, M. (2002) Efficient Packet Marking for Large-scale IP Traceback, Proceedings of 9th ACM Conference on Computer and Communications Security, pp 117-126.

Haeberlen, A., Fonseca, P., Rodrigues, R. and Druschel, P. (2011) Fighting Cybercrime with Packet Attestation, Proceedings of 9th ACM Conference on Computer and Communications Security.

Hunker, J., Hutchison, B. and Margulies, J. (2008) Role and Challenges for Sufficient Cyber Attack Attribution, Dartmouth College: The Institute for Information Infrastructure Protection.

Kohno, T., Broido, A. and Claffy, K. (2005) Remote Physical Device Fingerprinting, IEEE Transactions on Dependable and Secure Computing, Vol.2, No.2, pp 93-108.

Li, J., Sung, M., Xu, J. and Li, L. (2004) Large-scale IP Traceback in High-speed Internet: Practical Techniques and Theoretical Foundation, Proceedings of IEEE Symposium on Security and Privacy, pp 115-129.

Liu, X., Li, A., Yang, X. and Wetherall, D. (2008) Passport: Secure and Adoptable Source Authentication, Proceedings of 5th USENIX Symposium on Networked Systems Design and Implementation, pp 365-378.

McConnell, M. (2010) How to Win the Cyber-War We're Losing, [online], Washington Post, www.washingtonpost.com/wp-dyn/content/article/2010/02/25/AR2010022502493_pf.html.

Ramsbrock, D., Berthier, R. and Cukier, M. (2007) Profiling Attacker Behavior Following SSH Compromises, Conference on Dependable Systems and Networks, pp 119-124 .

Raynal, F., Berthier, Y., Biondi, P. and Kaminsky, D. (2004b) Honeypot Forensics, Part II: Analyzing the Compromised Host, IEEE Security and Privacy, Vol.2, No.5, pp 77-80.

Raynal, F., Berthier, Y., Biondi, P. and Kaminsky, D. (2004a) Honeypot Forensics, Proceedings of 5th Annual IEEE SMC Information Assurance Workshop, pp 22-29.

Santhanam, L., Kumar, A. and Agrawal, D. (2006) Taxonomy of IP Traceback, Journal of Information Assurance and Security, Vol.1, No.1, pp 79-94.

Savage, S., Wetherall, D., Karlin, A. and Anderson, T. (2000) Practical Network Support for IP Traceback, Proceedings of ACM SIGCOMM, pp 295-306.

Seifert, C., Welch, I. and Komisarczuk, P. (2006) Taxonomy of Honeypots, [online], www.mcs.vuw.ac.nz/comp/Publications/archive/CS-TR-06/CS-TR-06-12.pdf.

Shanmugasundaram, K., Brönnimann, H. and Memon, N. (2004) Payload Attribution via Hierarchical Bloom Filters, Proceedings of 11th ACM Conference on Computer and Communications Security, pp 31-41.

Snoeren, A. (2001) Hash-based IP Traceback. Proceedings of Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, pp 3-14.

Song, D. and Perrig, A. (2002) Advanced and Authenticated Marking Schemes for IP Traceback, Proceedings of 20th Annual Joint Conference of IEEE Computer and Communications Societies, pp 878-886.

Staniford-Chen, S. and Heberlein, L. (1995) Holding Intruders Accountable on the Internet, IEEE Proceedings of Security and Privacy, pp 39-49.

Stoll, C. (1989) The Cuckoo’s Egg. UK: Doubleday.

Strayer, W., Jones, C., Tchakountio, F. and Hain, R. (2004) SPIE-IPv6: Single IPv6 Packet Traceback, 29th Annual IEEE International Conference on Local Computer Networks, pp 118-125.

Guan, Y. and Zhang, L. (2010) Managing Information Security, Syngress Media Inc.

Wagener, G., State, R., Dulaunoy, A. and Engel, T. (2009) Self Adaptive High Interaction Honeypots Driven by Game Theory, Stabilization, Safety and Security of Distributed Systems, Vol.11, pp 741-755.

Wheeler, D. and Larsen, G. (2003) Techniques for Cyber Attack Attribution, Defense Technical Information Center.

Zhang, L. and Guan, Y. (2007) TOPO: A Topology-aware Single Packet Attack Traceback Scheme, IEEE Securecomm and Workshops, pp 1-10.

Zhang, Y. and Paxson, V. (2000) Detecting Stepping Stones. Proceedings of 9th USENIX Security Symposium.

A Vulnerability-Based Model of Cyber Weapons and its Implications for Cyber Conflict

Karlis Podins and Christian Czosseck
Cooperative Cyber Defence Centre of Excellence, Tallinn, Estonia
karlis.podins@ccdcoe.org
Christian.czosseck@ccdcoe.org

Abstract: Throughout history, mankind has developed and employed novel weapons systems and equally novel countermeasures. Naturally, both offensive and defensive systems are limited by the laws of nature. Consequently, military concepts and doctrines were designed by implicitly taking into account those same limitations. The digital age has introduced a new class of weaponry that poses an initial challenge to our common understanding of conflict and warfare as for their different characteristics: cyber weapons. Cyber weapons and other terms like hacking are used frequently, commonly without giving clear definitions in the given context. We propose a restricted definition of cyber weapons as consisting primarily of data and knowledge, presenting themselves in the form of prepared and executed computer codes on or a sequence of user interactions with a vulnerable system. This article explores the crucial differences between the conventional weapons and cyber weapons domains, starting a debate on to which extent classical concepts and doctrines are applicable to cyber space and cyber conflict. This motivates a discussion on the role of vulnerabilities in IT systems and their impact to IT security and cyber attacks. The authors describe a vulnerability-based model for cyber weapons and for cyber defense. This model is then applied to describe the relationship between cyber-capable actors (e.g. nation-states). The proposed model clarifies important implications for cyber coalition-building, and disarmament. Furthermore, it presents a general solution for the problem of the destruction of cyber weapons, i.e. in the context of cyber arms control.

Keywords: cyber weapons, cyber defense, disarmament, coalition, vulnerabilities

1. Introduction

As conflicts have moved into cyberspace (and vice versa), a clear understanding of cyber weapon becomes a necessity. The development of weapons was always part of mankind's history. Weapons evolved to suit the tactics, but from time to time new weapons revolutionized the tactics and strategies of warfare. The developments of artillery, gunpowder, aviation or weapons of mass destruction are just some examples from recent history. They all caused dramatic changes on the face of the battlefield. But all those weapons developed so far have similar kinetic and/or thermal properties due to the shared physical domain.

In the cyber attacks on Estonia in 2007, a campaign of massive distributed denial of service (DDoS) attacks against government websites, paired with hacking attempts against valuable targets like ISP backbone routers, a new type of conflict became a reality (Evron, 2008). Many more politically motivated DDoS attacks followed (Nazario, 2009). It drew strong and inconsistent media attention, up to being termed the first cyber war in history, as discussed by (Farivar, 2009). While Article 5 of the North Atlantic Treaty, governing mutual support among NATO nations in case of an armed attack was not invoked, national security leadership around the world received a wakeup call. However, several years later we still have not seen a commonly agreed definition of cyber weapons or cyber warfare (Ottis & Lorents, 2010).

In first section we are going to define basic terms and give short overview about special properties of cyber domain and cyber weapons. The concepts and terminology initially used were based on our understanding of conventional weapons or weapons of mass destruction (Sharma, 2009). Unfortunately, cyber weapons, being data and knowledge, follow other rules than their conventional counterparts. Thus the effects of cyber weapons on their target are different. In 2007 a direct, destructive cyber attack on a power generator was proven possible in a real life experiment (Herold, 2007). Latest prominent case, the Stuxnet virus, which is assumed to have sabotaged the Iran nuclear program starting from 2009, shows how powerful a cyber weapon in a real world setting can be (Falliere, Murchu, & Chien, 2010; Langner, 2011). Some research states that conventional weapons framework is ineffective, is not applicable or has left a big gap between our assumptions about cyber weapons and reality (Sulek & Moran, 2009). We do not want to get involved in this controversial discussion, but rather explicitly show some surprising aspects of cyber domain.

Section 2 of the paper introduces a vulnerability-based model of cyber conflict. We argue that knowledge about vulnerabilities is key element in understanding both cyber offense and cyber defense, and build a simplified model of cyber conflict around this idea.

In section 3 we will apply the vulnerability-based model on selected aspects of cyber conflicts. By using the proposed model to describe relationship and interactions between cyber-capable actors we demonstrate applications of proposed approach.

2. The cyber domain and its weapons

In recent years, the prefix „Cyber“ has become quite common and was added to many existing terms with the intention of extending its meaning into cyberspace (Ottis & Lorents, 2010). While terms like cyber war (warfare), cyber defense or cyber weapons are widely used, there is a lack of commonly accepted definition, and authors often use terms without precisely describing their meaning.

In the context of this paper we would like to offer the following definitions.

Definition 1: Cyber Weapon and Cyber Attack.

A cyber weapon is data and knowledge that is capable of, designed to and executed with the intention to affect the integrity, availability and/or confidentiality of an IT system (target) without its owner's approval. The target's defense is overcome by abusing existing vulnerabilities in the target.

Application of cyber weapon shall be named a Cyber Attack.

Definition 2: Vulnerability.

A vulnerability is an exploitable flaw in an IT system, which allows an attacker to usurp privileges or trust, access data or execute commands, he normally would not be allowed or expected to. This includes mis- or known default configuration but excludes social engineering means. Possible examples are: taking control of a system, reading or modifying information stored or processed or adding functionality.

When starting discussions on conflicts in cyberspace, one should clearly understand the cyber domain and its special properties. Cyberspace differs from conventional weapons domains, some differences being minor while some are of the utmost importance. We believe that ignoring these differences and straight-forward application of established concepts especially on weapons has caused some of the confusion around the cyber domain. In the following an overview of key differences between the cyber domain and the domain of conventional arms is provided. As for space reasons, an analogy to the domains of nuclear, biological and chemical (NBC) weapons is not made.

The confusion can be easily explained just by looking at the definitions. While established concept assumes weapon to be a physical object, cyber weapon under proposed definition is knowledge and data derived from it.

The extent and severity of cyber attacks can vary as recent history has shown, from local (loss of email confidentiality due to loss of password) to nation-wide (Ottis, 2008). As for the time being, *cyber attacks do not directly kill* living beings, but can cause abuse of, malfunctions or the destruction of equipment. That, as a 2nd or higher order result, can lead to a lethal effect or further destruction (Ziolkowski, 2010).

The distance between attacker and target is irrelevant for conducting the attack as long as there is connectivity between them. By connectivity we mean possibility to communicate between the attacker and the target,including not only wired or wireless media, but alsomethods to transfer information between air-gapped networks (see e.g. Falliere et al., 2010; Langner, 2011). Considering the Internet as the global network, all connected computers, smart phones, cars, industrial control systems (SCADA), or Internet enabled household electronics are just some of the possible targets. As of the very nature of the Internet a cyber attack can be initiated from any (connected) place on this planet to reach almost everywhere.

Launched attacks hit their target nearly instantly. While preparing a cyber attack might demand time consuming preparations, some of the effects could manifest itself in the blink of an eye. With this the defensive aspect of time, which can be used in conventional warfare to start a countermeasure against an attack, becomes less relevant.

There is no border or neutral area. National territory or borders on the Internet are only considered by legal departments, e.g. (Kelsey, 2008), but so far those concepts have almost no everyday practical meaning. While it is possible to cut or limit nation's connectivity to Internet, the consequences for any modern nation state are too bizarre to consider it a long term solution.

The attribution of an attack to a cyber attacker is, with technical means, nearly impossible (Hunker & Hutchinson, 2008). While attribution might be possible by also considering information from other sources and by analyzing the behavior and beneficial outcome for other parties (Ottis, 2008), this is not guaranteed. The Conficker worm so far has not been attributed to any party, although considerable public and security community attention was focused to it, as well as USD 250,000 bounty was offered by Microsoft("Microsoft Collaborates With Industry to Disrupt Conficker Worm," 2009). If an attacker plots his attacks with enough care, he could even maintain a steady attack on his adversary without being identified (Lemay, Fernandeza, & Knight, 2010). But having, to a reasonable extent, a confirmed source is a common requirement for the attacked party to take active countermeasures, especially if they are of violent nature (Ziolkowski, 2010).

Cyber weapons do not have a physical nature, they are knowledge and data as defined above. To conduct a cyber attack, the attacker has to send data to the victim. Most can be made persistent, e.g. by writing a script or program and storing this on an IT system. Like ordinary files, those cyber weapons can then be copied without noteworthy costs so the number of copies of a cyber weapon is usually irrelevant. Storage and transportation of cyber weapons also seem easier than for physical counterparts, one party could even decide to keep a copy or even the whole cyber weapon arsenal outside its own borders, hiding it in different places all over the Internet.

Cyber weapons production costs are mainly human resources related. While we recognize the initial investment in R&D as necessary, success of lone hackers and loose groups (Anonymous, LulzSec) shows that a lot can be achieved with limited funds. We believe that especially for a nation-state it is a minor expense compared to development of advanced conventional weapons. The production of conventional weapons depends mostly on the costs of fabrication, which often is a combination of labor costs, materials, machines and infrastructure. In contradiction to this, cyber weapons are constructed mainly by human knowledge on relatively cheap IT equipment. Their skills and knowledge in areas like software development, exploit development and penetration testing are the essence of the attack. Hiring the right personnel and keeping or extending their skills is the major investment needed to maintain or extend a cyber weapons arsenal (Miller, 2010).

The knowledge of a target's vulnerabilities and how to exploit them is of central importance to the attacker. On the other side, this knowledge is also the most important piece of information needed by the target of the attack to set up an efficient defense.

This leads us to the 1st paradox of cyber weapons: they are subject to time-decay. If a cyber weapon is exploiting a vulnerability, it is effective as long as the target IT system has this vulnerability. Vulnerabilities get discovered and patched, and target systems can change their software/hardware at their will, so the period of effectiveness for a cyber weapon is undefined, but likely to decrease the longer the vulnerability is known. This aspect is also discussed in (Moore, Friedman, & Procaccia, 2010).

The *2nd paradox is that of cyber weapons usage may lead to the target enhancing its defense* in a very short time. As soon as a cyber attack is executed, the target might have means in place to detect that an attack occurred and how it was conducted. While the initial attack might have been successful, the likelihood that a proper defense can be built after it is reasonably high, rendering the used cyber weapon useless against the same target and even against others in cases of existing cooperation or disclosure. An example is specially crafted malware. As soon as samples of a particular new malware get collected and analyzed, appropriate antivirus, IDS and software patching could be done rather fast. Still this is ultimately decided by the capabilities for the actors.

This has consequences for one's ability to test the effectiveness of cyber weapons. As soon as a certain cyber weapon is tested in the wild or even against the target itself, the likelihood is reasonably high to believe that security-aware targets will find a way to enhance their defense against the attack. (It is recognized that techniques are available to the attacker to re-launch the same attack using changed code, and that the target can only learn from an attack if he recognizes it in the first place.) But testing those attacks in a closed test environment will not always guarantee their successful later deployment, as it will be hard for the attacker to build a reasonably complete testing environment with the true attributes of the real target. Nevertheless, it has to be pointed out, that it is quite common to use commercial off-the-shelf hardware and software, which can also be easily acquired by the attacker in order to test his defenses.

The given unique attributes and examples support our opinion, that conventional arms and the ways we use them, are different from cyber weapons. This is why in the following section we propose a model to describe cyber domain and conflicts within.

3. A vulnerability based view on cyber weapons

The actors (parties) of cyber incidents might be individuals, groups of individuals, companies or nation-states. All possible combinations of parties attacking each other with cyber weapons reflect different scenarios, which, depending on effects, we would call criminal acts, industrial espionage, terrorism, conflicts or even (cyber) war. While a discussion on mapping these terms in the cyber landscape might be reasonable, we would like to state that the proposed model is generic enough to be applicable to all of them, because it focuses on the underlying mechanics. Considering the scope of this paper, we will use the term cyber conflict for all those mentioned combinations.

3.1 Vulnerabilities

Vulnerabilities, the knowledge about and the skills to exploit them, are of utter importance in any form of cyber conflict. It is quite clear that without exploitable vulnerabilities, cyber attacks would be pretty much limited to (D)DoS attacks (as they do not rely on a flaw in the target, but limit their targets' accessibility by exhausting bandwidth, CPU or other limited resource)or attempts to manipulate users through social engineering to get access to the target's system. To reflect that, we propose to take vulnerability as a basic entity and analyze cyber weapons and conflict by looking at them from a vulnerability point of view.

Vulnerabilities differ in their severity. Some of them enable an adversary to conduct a cyber attack against a target (e.g. root access), while some are likely not to be of any use to an attacker. The internationally recognized Common Vulnerability Scoring System introduces an open framework enabling one to score and compare the severity of a vulnerability based on its intrinsic qualities, its behavior over time and its environmental uniqueness (Mell, Scarfone, & Romanosky, 2007). While comparing individual vulnerabilities, it makes a difference how severe every single one of them is. However, we will save a detailed examination of this issue for future research. For the moment, it might be safe to only consider those potentially leading to a complete violation of availability, confidentiality or integrity of an attacked IT system.

The dynamic nature of IT implies that the set of vulnerabilities changes over time. When new code gets installed, new vulnerabilities are added and/or other vulnerabilities get removed. There is a considerable amount of vulnerabilities publicly disclosed every day (e.g. at the National Vulnerability Database (Security, 2010)) but some vulnerabilities remain unpatched although known to the public for decades (Oiaga, 2010).

Naturally, there exists a set of all vulnerabilities that are in the software installed somewhere. This is not constant and changes every time new software is installed, removed or configured. Each party has knowledge about some of the vulnerabilities. There are vulnerabilities unknown to all parties. As such the union of all parties known vulnerabilities still might not give all existing vulnerabilities.

3.2 Cyber defense

Apparently a system without vulnerabilities is well defended against majority of cyber attack(apart from DoS based attacks). The defender of a target system has three options to render a vulnerability-based cyber weapon *w* useless.

By knowing about vulnerabilities exploited by *w* and having own infrastructure immune against w by patching its own systems accordingly;

By putting in place means to effectively make an existing, unpatched vulnerability not exposed to the attacker (e.g. by using firewalls hiding a service to the Internet or signature-based attack detection). This would also apply to vulnerabilities unknown to the defender but coincidently covered;

By putting in place means to detect and mitigate cyber attacks before their effect manifests itself.

One could assume that cyber security aware parties have full control over and knowledge about their own IT assets, making sure not to be vulnerable to attacks they known by themselves (as they ultimately relay on vulnerabilities). Realty shows us that this does not hold true all the time, but nevertheless we keep this assumption for the sake of simplicity, assuming the target to be capable and willing for defending his cyber frontline.

3.3 Cyber weapons development

Cyber Weapons are produced based on knowledge of target‘s vulnerabilities. It should be clear that before developing any weapon *w*, a party must have knowledge about the vulnerabilities the weapon will exploit and for any given weapon *w* it is possible to find respective vulnerabilities.

More formally speaking, there exists a natural mapping *e* from cyber weapons to vulnerabilities they exploit. For any given weapon *w*, the set *e(w)* is either empty (in the case of a DoS attack or cases of social engineering) or it contains at least one vulnerability.

Sometimes, successfully attacked IT Systems can be used for manual or automated creation of (new) cyber weapons for further attacks, e.g. sending infected emails to all persons in contact list or probing the local network for other vulnerable machines.

To effectively enlarge one's weapons arsenal with new weapons, a party is required to find an additional vulnerability.

As time passes, other parties discover vulnerabilities in a non-deterministic manner. When a party discovers a vulnerability v, we can assume that their defenses are upgraded, making weapons exploiting vulnerability v ineffective against party p. Research by Moore et al. (Moore et al., 2010) propose a similar vulnerability-based look but assumes the opposite, that own infrastructure is not patched not to warn the opposing parties of discovered vulnerabilities, coming to noteworthy conclusions.

Additionally, it seems reasonable to assume that the target party *t* has adequate logging and attack detection systems in place and will be capable of identifying exploited vulnerabilities at least after they have been used by a cyber weapon, leading to improved cyber defense.

This stresses again the overlap between cyber attacks and defenses and vulnerability discovery plays a central role both for attackers and defenders.

4. Implications on party relationships

In following section we will test vulnerability-based thinking by applying it to model cyber disarmament, coalitions, collective defense, pacifism and arms control.

4.1 Cyber disarmament

In contrast to vague cyber disarmament discussion (Gjelten, 2010) the proposed model offers an effective method of cyber weapon disarmament, as asked for by (Cahill & Rozinov, 2003). Disarmament of cyberspace could be achieved by public vulnerability disclosure. If a party wants to verifiably dismantle some of its cyber weapons, it could publicly disclose all the vulnerabilities used by them. Unlike in real-world disarmament, which affects only the disarming party, cyber disarmament is global. After other parties have adequately reacted to disclosed information, all cyber weapons using disclosed vulnerabilities are rendered useless. This is regardless if they are in the possession of the disarming party or by another party.

But there are further peculiarities. Unlike in the conventional weapons domain, if a party *p* dismantles a set of cyber weapons by disclosing the respective vulnerabilities, defense of *p* is not affected. Following our assumption that a party always tries to be protected against all known vulnerabilities, *p* does not disclose a new attack vector against itself, leaving its adversaries without an advantage by this action.

P's offensive capabilities are decreased by an unknown factor, depending on how many other parties had the disclosed vulnerabilities in their IT systems.

4.2 Cyber coalitions

If two or more parties would mutually agree on cyber disarmament, public vulnerability disclosure can be used with the consequences as described before. But an interesting approach would be a mutual vulnerability disclosure; it effectively means that cyber weapons involved in the mutual disarmament would be ineffective against other parties involved in the disarmament agreement, but still potentially effective against 3rd parties.

By exchanging information on vulnerabilities between each other one enables the other to build new weapons based on the newly learned vulnerabilities. But at the same moment both become immune to cyber attacks based on the exchanged vulnerabilities, as soon as they have implemented the necessary changes in their defense. Their offensive capabilities to attack each other decreases or remains the same.

Let *c* be any other 3rd party not part of the treaty or coalition between *a* and *b*. With the newly won knowledge both parties *a* and *b* might have gain additional knowledge to build a cyber weapon *c* is vulnerable to, enhancing their offensive capabilities against *c* or at least remaining the same.

4.3 Cooperative cyber defense

Based on the definition of cyber defense given above, cooperation in cyber defense between two or more parties can manifest by exchanging of vulnerability information and/or attack detection procedures, if it provides tangible benefits to each party.

But in the case of vulnerability information exchange, there is no difference from the aforementioned cyber coalition. This leads to the observation that it is possible for cooperative cyber defense to be indistinguishable from a cyber coalition.

To arrange a cooperative cyber defense that is purely defensive in nature, the cooperation between the parties would be limited to the exchange of signatures and other attack detection information without sharing vulnerabilities per se. If knowledge about vulnerabilities shall also be shared without giving an offensive advantage to the parties, it must be done by public disclose with the consequences as described under cyber disarmament.

4.4 Cyber pacifism

Pacifistic movement is usually limited to protest when in conventional weapons domain But this could change in the cyber space.

Apart from voicing concerns about cyber weapons as discussed in (Rowe, 2007), a more active strategy could be to actively search for vulnerabilities and publicly disclose them as soon as possible. As in the case of cyber disarmament, it would universally destroy the corresponding cyber weapons, hitting the weapons arsenals of many parties simultaneously.

Assuming enough resources are present, the cyber pacifist can have a significant impact on the cyber weapons arsenals of other nation-states without committing any aggressive actions. This again shows how different the cyber weapons world is, leading to new possibilities, but also new limitations.

4.5 Cyber arms control

A problem of what to do with cyber weapons found during an inspection as illegal under an Cyber Arms Control treaty is stated by Dorothy Denning (Denning, 2001). Within our proposed model, the same public disclosure procedure as for *cyber disarmament* could be applied, after the found cyber weapon got analyzed to identify the vulnerability exploited by it . Wherever other copies of discovered cyber weapon are stored, they all rely on the same vulnerability (or same set of vulnerabilities), by disclosing those vulnerabilities one can effectively neutralize each and every of those copies at the same time.

5. Conclusion

Cyber weapons are a new type of armament that follows different rules than the established rules of conventional weapons or weapons of mass destruction. The revolutionary nature of cyber weapons as a means and method of warfare demands the creation of a new conflict model.

This article proposes a highly-simplified, vulnerabilities-based model that defines and describes cyber weapons and cyber defense. The authors analyzed the relationships between hypothetical parties who would develop and employ cyber weapons. The model revealed a number of important findings, such as the fact that a defensive alliance, in which the parties exchanged their knowledge of cyber vulnerabilities, would lead to an enhanced offensive capability by every member of the alliance. Further, the authors proposed a strategy for cyber attack deterrence based on the introduced model. The restricted definition of cyber weapons together with vulnerability-based model has been shown useful for solving cyber arms control problem, and we hope that the model will show useful in other problems as well.

The examples used in this article primarily highlight the typical roles expected of nation-states, but the authors believe that the model's principles will apply in the corporate world as well.

Disclaimer: The opinions expressed here are those of the authors and should not beconsidered as the official policy of the NATO Cooperative Cyber DefenceCentre of Excellence or NATO.

Acknowledgements

Both authors contributed equally. The opinions shared in this article are those of the authors and do not represent the views of NATO or its member states. The authors would like to thank Mr. Kenneth Geers and Prof. Peeter Lorents who gave us valuable feedback while writing this article.

References

Cahill, T., & Rozinov, K. (2003). Cyber warfare peacekeeping. *Assurance Workshop, 2003.*, (June).

Denning, D. (2001). Obstacles and Options for Cyber Arms Control. *Arms Control in Cyberspace, Heinrich Böll Foundation, Berlin, Germany*, 1-13. Retrieved from http://faculty.nps.edu/dedennin/publications/berlin.pdf

Evron, G. (2008). Battling botnets and online mobs: Estonia's defense efforts during the internet war. *Georgetown Journal of International Affairs*, *9*(1), 121–126.

Falliere, N., Murchu, L. O., & Chien, E. (2010). W32. Stuxnet Dossier. *Symantec Security Response*, *3*(November), 1-64. Symantec.

Farivar, C. (2009). A Brief Examination of Media Coverage of Cyberattacks (2007 - Present). In C. Czosseck & K. Geers (Eds.), *The Virtual Battlefield: Perspectives on Cyber Warfare* (p. 183). Amsterdam: IOS Press.

Gjelten, T. (2010). Debating Cyber Disarmament. *World Affairs*. Retrieved March 1, 2012, from http://www.worldaffairsjournal.org/article/shadow-wars-debating-cyber-disarmament

Herold, R. (2007). DHS Exploding Generator Shows Dire Need For Better Computer Security&nbsp;-&nbsp;Realtime IT Compliance. *www.realtime-itcompliance.com*. Retrieved August 17, 2010, from http://www.realtime-itcompliance.com/information_security/2007/09/dhs_exploding_generator_shows.htm

Hunker, J., & Hutchinson, B. (2008). Role and Challenges for Sufficient Cyber-Attack Attribution. Institute for Information Infrastructure Protection. Retrieved from http://www.thei3p.org/docs/publications/whitepaper-attribution.pdf

Kelsey, J. T. G. (2008). Hacking into International Humanitarian Law□: The Principles of Distinction and Neutrality in the Age of Cyber Warfare. *Michigan Law Review*, (May), 1427-1452.

Langner, R. (2011). Keynote speech on Stuxnet @ ICCC 2011. Tallinn: CCD COE Publications. Retrieved January 1, 2012, from http://www.ccdcoe.org/280.html

Lemay, A., Fernandeza, J. M., & Knight, S. (2010). Pinprick attacks, a lesser included case? In C. Czosseck & K. Podins (Eds.), *Conference on Cyber Conflict Proceedings* (pp. 183 - 194). Tallinn: CCD COE Publications.

Mell, P., Scarfone, K., & Romanosky, S. (2007). A complete guide to the common vulnerability scoring system version 2.0. *Published by FIRST-Forum of Incident Response and Security Teams* (pp. 1–23). Retrieved from http://www.first.org/cvss/cvss-guide.pdf

Microsoft Collaborates With Industry to Disrupt Conficker Worm. (2009). Retrieved January 3, 2012, from http://www.microsoft.com/presspass/press/2009/feb09/02-12confickerpr.mspx

Miller, C. (2010). Kim Jong-il and me: How to build a cyber army to attack the U.S. Tallinn: CCD COE Publications. Retrieved from http://www.ccdcoe.org/conference2010/agenda.html

Moore, T., Friedman, A., & Procaccia, A. D. (2010). Would a'cyber warrior'protect us: exploring trade-offs between attack and defense of information systems. *Proceedings of the 2010 workshop on New security paradigms* (pp. 85–94). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1900559

Nazario, J. (2009). Politically Motivated Denial of Service Attacks. In C. Czosseck & K. Geers (Eds.), *The Virtual Battlefield: Perspectives on Cyber Warfare* (pp. 163-181). 163-181: IOS Press.

Oiaga, M. (2010). Windows Blue Screens of Death After Patch for 17-Year Old Vulnerability Is Applied. Retrieved August 17, 2010, from http://news.softpedia.com/news/Windows-Blue-Screens-of-Death-after-Patch-for-17-Year-Old-Vulnerability-Is-Applied-134808.shtml

Ottis, R. (2008). Analysis of the 2007 Cyber Attacks Against Estonia from the Information Warfare Perspective. *Proceedings of the 7th European Conference on Information Warfare* (p. 163). Academic Conferences Limited.

Ottis, R., & Lorents, P. (2010). Cyberspace: Definition and Implications. *Proceedings of the 5th International Conference on Information Warfare and Security* (pp. 267-270). Academic Publishing Limited.

Rowe, N. C. (2007). War Crimes from Cyber-weapons. *Journal of Information Warfare*, *6*(3), 15–25. Retrieved from http://academic-conferences.org/pdfs/JIW6_3.pdf#page=21

Security, D. O. H. (2010). National Vulnerability Database (NVD) Search Vulnerabilities. Retrieved August 17, 2010, from http://web.nvd.nist.gov/view/vuln/search?execution=e2s1

Sharma, A. (2009). Cyber Wars: A Paradigm Shift from Means to Ends. In C. Czosseck & K. Geers (Eds.), *The Virtual Battlefield: Perspectives on Cyber Warfare* (Vol. 34, pp. 3-17). Amsterdam: IOS Press.

Sulek, D., & Moran, N. (2009). What Analogies Can Tell Us About the Future of Cybersecurity. In C. Czosseck & K. Geers (Eds.), *The Virtual Battlefield: Perspectives on Cyber Warfare* (pp. 118-131). Amsterdam: IOS Press.

Ziolkowski, K. (2010). Computer Network Operations and the Law of Armed Conflict. *Military Law and the Law of War Review*, (49), 47-94.

Modelling Emergency Response Communication Systems

Graeme Pye and Matthew Warren
School of Information Systems, Faculty of Business and Law, Deakin University, Geelong, Australia
graeme@deakin.edu.au
mwarren@deakin.edu.au

Abstract: Subsequent to the Australian 'Black Saturday' bushfires there were a number of issues arising from investigations with regard to the functional stability and resilience of communications systems and the flow of information between emergency response organisations, and their ability to provide relevant information to the general public. In some cases, the transference of information failed or was late or ineffective with regard to decisions, advice and information broadcasting during the crisis. This was particularly evident in terms of managing emergency organisational information requests and field situational advice both to and from emergency response management teams and the delivery of informative advice to the public. This paper analyses one such case study with a view of applying a systems modelling technique to determine the viability of the communication systems and information exchange structures associated with an emergency response agency.

Keywords: communication, resilience, emergency response, systems modelling

1. Introduction

The Australian Federal Government in 2010 launched the *Critical Infrastructure Resilience Strategy*. The aim of this new strategy is the continued operation of critical infrastructure in the face of all hazards as these infrastructure systems support Australia's national defence and national security and underpins economic prosperity and social wellbeing. Therefore, a more resilient critical infrastructure will help to achieve the continued provision of essential services to the community (Australian Government, 2010). This strategy also deals with associated areas such as disaster protection and disaster resilience and this shift in policy is going to have a major impact upon Australia, as this now places disaster management under the security domain of critical infrastructure protection (Warren and Leitch, 2011).

On or around Saturday the 7th of February 2009 a series of bushfires were burning across the southern Australian state of Victoria creating a crisis situation. As many as 400 individual fires occurred in various locations across the state and in their aftermath the day became known as 'Black Saturday'. These fires occurred towards the end of summer during extremely hot, dry and windy weather conditions with temperatures reaching 46.6^0 Celsius (115.9^0 Fahrenheit) and winds in excess of 100kph. This, in conjunction with an intense heat wave and little or no rain for the previous two months resulting tinder-dry fuel loads, exacerbated the ferocity of the bushfires resulting in Australia's highest loss of life from bushfires of 173 deaths and 414 injuries. Despite a total of 3582 fire fighting personnel from the Country Fire Authority (CFA) and the Department of Sustainability and Environment (DSE) being placed on standby in preparation for the extreme weather conditions (Teague 2009; State Library Victoria 2012).

Initially, this paper provides a brief contextual overview of what to expect when experiencing a bushfire and discusses the response options and actions that affected individuals can take related to their circumstances. Then a discussion is undertaken to outline and identify some of the information communication and warning system shortfalls experienced during the 'Black Saturday' disaster as identified by the 2009 Bushfire Royal Commission. This is followed by identifying the primary response organisations involved, before proposing the Viable Systems Model (VSM) approach to assess organisational structure and communication from a *systems approach* perspective. The VSM approach is then applied to a specific bushfire event case study of the Bendigo Fire that occurred on 'Black Saturday' and is discussed with regard to understanding the situation and the resilience of critical information communication infrastructures.

2. What to expect in a bushfire

Experiencing a bushfire of any size can be a very stressful and frightening event. While no one willing wants to be anywhere near an out of control bushfire, the best way to survive such a threat is to relocate away from the direct threat as early as possible. Typically, when a bushfire is approaching a location its presence is characterised by small spot fires igniting ahead of the main fire front, smoke, heat, noise and possibly darkness due to the smoke plume. The fire front itself can be difficult to pinpoint as it can approach locations from multiple directions depending on environment factors and geographical circumstances, with lots of smoke and burning embers landing ahead of the fire, which can continue for hours after the fire has passed (CFA 2012).

Furthermore, there is an expectation that the various infrastructure systems will be disrupted or fail completely, including: power, mains water supplies and telecommunications. Mains water supplies will be affected with other residents and fire trucks accessing water, telephone lines may be cut by falling trees and mobile phone coverage can become congested and the loss of electricity supply will further compound the situation. Moreover, road travel at this time is fraught with danger due to poor visibility, fallen trees and branches, other traffic including many emergency vehicles and escaping livestock and native fauna. Then there is the personal human aspects of confusion, fear, being unable to breathe properly, the possibility of heat stress, dehydration and tiredness that all impact on rational decision making during the crisis itself (CFA 2012).

These are the obvious factors that will occur when a bushfire is threatening individual residents and communities and therefore having a predetermined plan to survive a bushfire is a recommendation made to all rural and regional residents in Victoria. Having a well written and rehearsed plan of action will assist those affected, with what needs to done in preparation if the choice is to stay and defend property rather than evacuating. However if the plan is to evacuate, then authorities recommend enacting this early rather than at the last moment when the threat is imminent (CFA 2012).

3. Stay and defend or evacuate?

In Australia emergency bushfire response falls within the jurisdiction of the States and Territories, the powers granted to the emergency service organisations are in some cases different in particular circumstances enabling emergency services organisations to implement forced mandatory evacuation of people from their homes (Loh 2007).

The policy in the State of Victoria regarding community response to bushfire, directs residents to Prepare, Stay and Defend or Leave, which is locally known as the '*stay or go*' policy. This approach to bushfire preparation has been developed over many years based on research and experience of past bushfire incidents. It stipulates that with proper planning and preparation, most building structures can be defended during a bushfire, alternatively the plan is to voluntarily depart the location early (Teague 2009).

Obviously, to enact any plan successfully there is a need for timely warnings and information; in fact the community expects and depends on the broadcasting of high quality and detailed information before, during and after a bushfire. Furthermore, there is an expectation that timely warnings and information will be communicated to emergency response agencies during a crisis so they can allocate resources to respond appropriately and proportionately to the situation. However, in terms of the 'Black Saturday' bushfires there was issues with the effectiveness of communicating information and warnings (Jackson 2009; McDermott 2009; Teague 2009).

4. Information communication and warnings?

The subsequent state government sponsored Royal Commission into the 'Black Saturday' bushfires noted a number of weaknesses and failures with the Victorian warning systems where warnings were often delayed, meaning many people were not at all aware of the amount of time they had to respond or the magnitude and circumstance of the bushfire threats. This information shortfall was not only the case for people in the affected communities, but also for the localised bushfire emergency response services and teams in the area (Teague 2009; Jackson 2009).

Prior to the 'Black Saturday' disaster the State Government had devoted significant resources towards a campaign of informing the broader Victorian community of the potential fire risks in relation to the weather conditions and environmental circumstances. While this was beneficial, it could not on its own construe that the information provided could be universally translated into direct awareness and preparedness for what would eventuate. It is the quality of the information and the multiple modes of dissemination, linked with the capacity of people to listen, comprehend and act on the information provided that is the basis of a shared responsibility between government and the community (Teague 2009).

It is the ongoing provision of bushfire information that prepares communities, the Royal Commission's findings were that the delivery of warnings was inadequate, with techniques to capture people's attention (i.e. sirens) when warnings were broadcast not being used. Similarly, the use of local community sirens, commercial radio and television was not encouraged either. Furthermore during 7 February, the emergency telephone call services (Telstra's Triple Zero service) were overwhelmed by demand resulting in calls going unanswered and callers not being connected to the relevant authorities resulting in a large number of abandoned calls. Other sources of information and warnings did not cope with the level of demand either with people attempting to access emergency service websites and about eighty percent of calls to the Victorian Bushfire Information Line left unanswered, which was further compounded with incomplete or out of date information for those that did get through (Teague 2009).

Recommendations of the Royal Commission investigation broadly suggested that improvements to the dissemination and provision of information and warnings to the public where as follows (Teague 2009):

- improving the quality of bushfire information and warning messages by adopting standard language already developed for national usage;
- simplifying the format of bushfire warnings;
- reintroducing the Standard Emergency Warning Signal to draw attention to broadcast warnings about life threatening fires;
- extending the broadcasting of official warnings to commercial radio and television;
- allowing the reintroduction of sirens in local communities where there is demand for them;
- supporting the acceleration of the full introduction of a nationally developed telephone based automatic warning system;
- pursuing research into the development of improved fire danger index systems;
- enhancing the role of the Bureau of Meteorology in issuing daily information on bushfire risk;
- improving technology and processes to accelerate the updating of common bushfire information on agency websites;
- increasing the capacity of the bushfire emergency networks, the Victorian Bushfire Information Line, Telstra's Triple Zero service and the Emergency Services Telecommunications Authority to better handle peak demands, and to work more collaboratively during severe fire risk days.

Many of these changes will need to be accompanied by an education campaign so that people understand the changes and how to interpret the information that is provided. Much of this work has been undertaken and is now integrated and in use presently, although these information communication and warning systems have not been tested to the magnitude and extent of the 'Black Saturday' bushfire emergency.

5. Managing resource allocation and coordination

Within Victoria there are principally two bushfire response organisations, namely the CFA and DSE that are tasked with responding to bushfire emergencies within rural and regional Victoria. The CFA is largely staffed by trained local community volunteers and the DSE is a state government department consisting of paid personal. The jurisdiction of the Metropolitan Fire and Emergency Services Board (MFB) is largely dedicated to fire and rescue response within the urban built environment of Melbourne and is generally only deployed in bushfire emergencies when requested by the CFA (CFA 2012; DSE 2012; MFB 2012).

However, from this research paper's perspective it is the CFA organisational response, communication systems and the information flows during a bushfire incident that is of interest, as applied in a specific case study situation. The VSM approach to modelling organisational communication should provide insight into the communication structures and information flows that occurred within the CFA organisation during the specific bushfire event. Although, initially it is necessary to provide some background and an overview explanation of the VSM approach and how it is applied, as the following example outlines.

6. Viable system modelling

A viable system is one that is deemed to be adaptable and is capable of continuing to meet the demands and challenges of a changing environment. The Viable System Modelling (VSM) approach can be utilised to express an abstract model of a viable system and afford an applicable description of an organisational system that is viable and capable of functional autonomy (Beer 1985).

This system modelling approach extends on the earlier work of Beer (1985) in defining the laws that govern a viable system and their application for comprehending and determining the viability of an enterprise system and for coping with system complexity. The purpose of the Viable Systems Model (VSM) approach is to diagnose or design organisational structure and communication by distinguishing between five management functions and a number of vertical and horizontal communication channels. Through the application of a recursive process, VSM enables the description and comparison of various functions at differing organisational levels from community to international, and is most beneficial when applied via a shared agreement regarding the focal entity for modelling, its system boundaries and fundamental goals (Beer 1985; Leonard & Beer 1994).

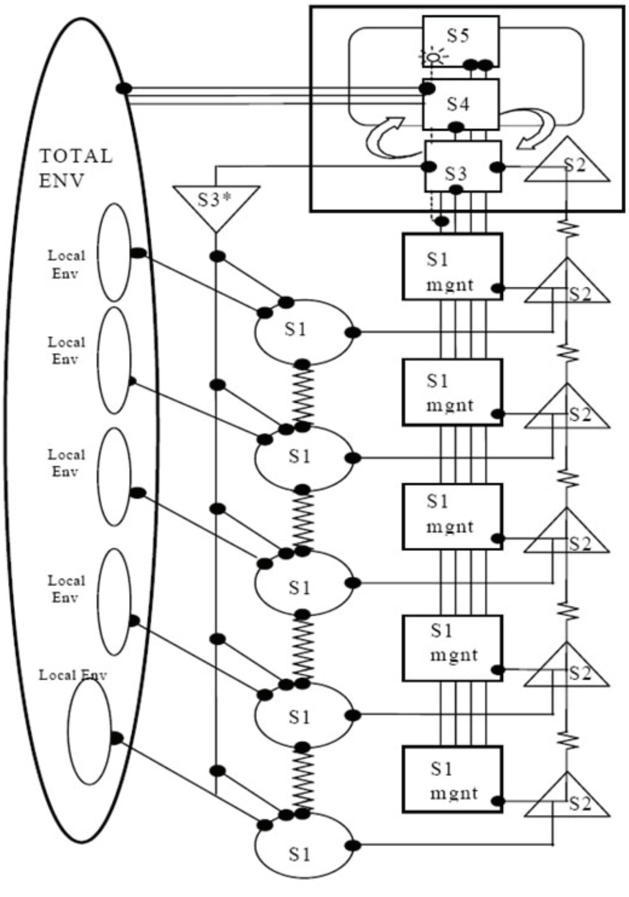
The research of Hutchinson and Warren (2002) applied the VSM framework and its principles to managing the security of large multi-level organisational information systems to detect, check and identify system security threats and vulnerabilities as they appear. By applying VSM and utilising its local subsystem monitoring it is possible to distinguish between threatening and non-threatening behaviours and adjust the whole system in response. From the research perspective of Hutchinson and Warren (2002), VSM provides a framework to manage cooperatively, the normal organisational information system functions and security concerns as a single overall information system to provide continuous system security that takes into consideration all levels within the greater system view.

In order to utilise VSM with this system approach, it is essential to understand the dynamics of its applicability as illustrated in Figure 1. VSM consists of five subsystem component structures or functions as follows (Warren & Hutchinson 2005):

- Implementation (S1): this function consists of semi-autonomous units, which carry out the operational tasks in the system and are fundamental to the existence or purpose(s) of the system by interacting with their local environment and each other. Each unit has its own local management function and connects to wider management by the vertical information flows. This function is the ‘doing’ part of an organisation and with each recursive element of VSM; each S1 has another VSM embedded.
- Co-ordination (S2): this function coordinates the S1 units to ensure that each S1 unit acts in the best interest of the whole system, rather than its own. This could be represented by something as simple as a timetable, or as subtle as morale among the workforce.
- Internal Control (S3): the function that interprets policy information from ‘higher’ functions (S4), and ‘lower’ functions. It controls the operational levels but its function is not to create policy, only to implement it. The periodic auditing of information arriving from the S1 function ensures its quality and correctness and denoted as the S3* audit function in the model.
- Intelligence and Development (S4): the function that acts as a filter of information from the S3 function and the overall outside environment. Its purpose is to ensure that policy-making function (S5) information is current before transmitting decisions to S3.
- Strategy and Policy (S5): the function responsible for the direction of the whole system, which must balance the requirements of both internal and external factors.

In this example, the data flows between S1 and S5 and the environment as shown in Figure 1 indicates the potential points of vulnerability to a 'computer-based attack'. With this conceptual viable system model an organisation can prepare strategies and tactics to make the system 'non-viable' or dysfunctional. The logic behind applying VSM is that through investigating functional shortcomings in this manner it can improve organisation preparedness and identify its weaknesses and vulnerabilities by highlighting possible avenues for attack (Warren & Hutchinson 2005).

It is the disruption or destruction of information systems that can cause serious loss of service to customers and increasingly information systems are under threat from both internal and external sources, and there is a need to establish a robust and dynamic response to protect information assets (Gokhale & Banks 2004). In view of the structure of information flows and their reliance upon information systems, there is obviously a need to establish ways to protect such systems and the VSM may offer benefits from this perspective.



(Note: Env = Environment, mgnt = management)

Figure 1: The viable system model (Warren & Hutchinson 2005)

It is with this in mind that VSM framework is applied in the following case study context as a means of modelling the viability of the communication systems supporting the information flows for the CFA during a specific bushfire incident. The research intention is to apply the VSM approach to the case study description to assess the effectiveness and viability of the communication systems and resilience of the supporting communication infrastructures utilised by the CFA during the Bendigo bushfire incident on ‘Black Saturday’.

7. Case study – The Bendigo Fire

The following case study was based upon evidence submitted, to the Royal Commission reviewing the Bushfire of 2009 (Cooke, 2009).

The incident control centre managing and coordinating the fire fighting activities at Bendigo on ‘Black Saturday’ was almost entirely cut off in its communications with firefighters and could not get detailed updates of the blaze as it burned on the edge of the town. Country Fire Authority (CFA) incident controller Peter Rogasch told the Royal Commission into the February bushfires that the centre he was managing on Adam Street in Bendigo had virtually no radio communication with firefighters in the field and could only tune in to what was being said about how the fire was developing on the central radio system.

The centre, which was set up for a less complex level-two fire, was equipped with two computers but he could not log on to the CFA's network and only about one in 10 attempts to contact people on their

mobile phones succeeded, he said. Heavy smoke played havoc with radio channels and a back-up telephone system limited the people that could be contacted.
Mr Rogasch said due to these circumstances decisions about where to send resources and set up road blocks were made by those on the fire ground. He sent observers out to gather information but, "once they got out there … we really couldn't talk to them either".

The centre had no information until 6pm and Mr Rogasch received a detailed briefing at 9.30pm, hours after the fire's main impact. The fire, also known as the Bracewell Street or Long Gully fire, began around 4.35pm and it burned about 330 hectares, killed one person and destroyed 58 homes.

The following is an applied assessment of the case study using the VSM criteria. The assessment is:

S1 Implementation

The impact upon the S1 is major; the localized fire fighting units would be able to fight the fires, but only in an ad-hoc fashion. Without intelligence or coordination their applied strategic approach to fight fighting may be sub-optimal.

S2 Co-ordination

Co-ordination of the localised control of the S1 (Bendigo) would be impacted. Ad-hoc strategies and approaches would to be developed in real time to try and implement coordination. There would also a lack of upwards coordination with the CFA Headquarters (Burwood). The backup communication devices also failed.

S3 Internal Control

Internal control would have failed. Strategic direction from the CFA Headquarters (Burwood) would not be passed to Regional Control (Bendigo). Ad-hoc measures would need to be developed in real time, but the 170 kilometre distance between the CFA headquarters (Burwood a suburb of Melbourne) and Regional Control Centre (Bendigo) would be an issue and the rapid occurrences of other fires in separate locations across Victoria is a major issue..

S4 Intelligence and Development

The role of S4/S5 is to try and determine the reason for the attack, for example, possible security weakness

S5 Strategy and Policy

The organisation (CFA) would need to (as they did) develop a more robust and resilient communication strategy and robust infrastructure.

In summary, this VSM assessment highlights some major communication shortcomings that became apparent as the situation and response circumstances of the Bendigo fire changed, which had a notable impact on the resilience and ongoing viability of the CFA communication systems.

It is evident that management and emergency response coordination issues were significantly and adversely affected due to the inability of the organisation to communicate advice both upward to CFA headquarters (Burwood) and downward to the Regional Control Centre (Bendigo) and the CFA fire-fighters on the ground at the fire site. From a wider perspective of the events of ‘Black Saturday’, this also indicates some of the shortcomings with communication infrastructures during such catastrophic circumstances that were evidenced by the broader activities of coordinating and managing numerous major bushfire responses in a number of differing state-wide locations simultaneously.

8. Discussion

Although it is impractical to consider that any emergency management organisational communication system could adequately cope with every situation effectively and simultaneously under such a load as occurred during the events of ‘Black Saturday’. There is nevertheless an expectation that emergency service communications can be maintained to some rudimentary extent though the redundant communication system structures in place. However, the events of and circumstances of ‘Black Saturday’ could not be predicted to have the effect they did on the communication systems and information flows of the CFA. Furthermore, as discussed previously from the public perspective, those citizens directly affected were also unable to access relevant up-to-date information regarding the fire pertaining to their situation and nor was the CFA able to effectively deliver early warnings either.

In this paper we have presented a vision of the future of emergency management that requires more robust support structures for the inclusion of activities and information from members of the public during disasters and mass emergency events. Such a vision relies on integration of multiple subfields of critical infrastructure systems resilience, steadfast information technology and communication systems, including in-built redundancy and a commitment to an understanding of the domain of application.

Emergency management is more than service personnel responding quickly to a situation; it is also about providing, capturing and distributing relevant and up-to-date information across a number of infrastructure enabled communication systems. Emergency management communication for coordinated responses and for public awareness need to consider adopting approaches that go beyond ‘pull’ technologies and actually ‘push’ information out to those located within the vicinity of the event. Also consideration should be given to incorporating the public as a source of situational intelligence that can utilise information communication technology infrastructures to provide information, which can potentially play a transformational role in a crisis (Palen *et al* 2010).

Although there are implications regarding the quantity, quality and trustworthiness the information provided by the public, if applied constructively there is great potential to utilise mobile phone networks, social media and alike to further enhance the flow of information. This requires consideration of how to utilise and filter both the formal and informal information communication mechanisms within emergency management organisations, their response services, information accuracy and relevance, including warnings to the public to ensure effective distribution to those that need it.

9. Conclusion

Coordination and conveying up-to-date information is critical to managing incidents such as bushfires and the knowledge generated from information flows provides the opportunity to forewarn the public and better manage and coordinate emergency responses. If the lines of communication are not resilient, too narrow or collapse due to infrastructure overload stress or failure, then the ability to generate good decisions from remote locations is adversely affected. This situation was evident during the ‘Black Saturday’ bushfires where the organisational and infrastructure communication systems were overloaded and became either ineffective or the quality of information was poor due to latency and distribution issues.

Additionally, the issue of ‘information overload’ arose where emergency response people were inundated with information that was not relevant to their current situation; to the point where the communication devices were either ignored or turned off. Furthermore, the impact of communication infrastructure loss during the emergency also raises issues regarding the level redundancy and the applicability of alternative communication mediums within these emergency response communication systems.

The reliance of established communication infrastructures was found to be capacity inadequate and less resilient than otherwise expected during the emerging bushfire emergency. In this instance, this research utilises the ‘*systems approach’* of the VSM modelling technique to exemplify the viability or otherwise of established communication and information exchange structures. In this case study the information communication structures did not function as expected requiring a rethink of approach to better manage information, to distil and disseminate important details quickly. The ‘Black Saturday’ Royal Commission findings suggest a review to investigate the underlying communication structures of the emergency response agencies to determine if structural and capacity issues need to be addressed. At this stage progress has been made towards this end although the pace of change is somewhat slower than expected.

References

Australian Government (2010), Critical Infrastructure Resilience Strategy, Attorney General's Department, Commonwealth of Australia, ISBN: 978-1-921725-25-8.

Beer S. (1985), Diagnosing the System for Organizations, Wiley, Chichester UK.

CFA (nd), Fire Ready Kit, Country Fire Authority (CFA). Online: http://www.cfa.vic.gov.au/firesafety/bushfire/firereadykit.htm Accessed: 20 January 2012.

Cooke D. (2009), Control Centre "couldn't talk' to firefighters, The Age Newspaper, 29th October, Fairfax media.

Gokhale G.B. & Banks D.A. (2004), 'Organisational Information Security: A Viable System Perspective', in 2nd Australian Information Security Management Conference, School of Computing and Information Science, Edith Cowan University, Perth, WA, pp. 178-184.

Hutchinson W. & Warren M. (2002), 'Information Warfare: Using the viable system model as a framework to attack organisations', Australian Journal of Information Systems, Vol.9, No.2, pp. 67-74.

Jackson L. (2009), Transcript and Verbatim: Interview Glen Fiske, ABC. Online: http://www.abc.net.au/4corners/content/2009/s2553476.htm Accessed: 12 January 2012.

Leonard A. & Beer S. (1994), The Systems Perspective: Methods and Models for the Future, [Online], URL: <http://www.agri-peri.ir/AKHBAR/cd1/FORESIGHT%20METHODOLOGY%20&%20FORECASTING/FORESIGHT%20METHODOLOGY/related%20articles/books/Future%20Research%20Methodology/6-sysmeth.pdf> Accessed: September 2008.

Loh E. (2007), 'Evacuation powers of emergency workers and emergency-service organisations in Australia', The Australian Journal of Emergency Management, Vol.22, No.4, pp. 3-7.

McDermott Q. (2009), Transcript from "Two Days In Hell", ABC - Four Corners. Online: http://www.abc.net.au/4corners/content/2008/s2492832.htm Accessed: 12 January 2012.

Palen L. Anderson K.M. Mark G. Martin J. Sicker D. Palmer M. Grunwald D. (2010), 'A Vision for technology-Mediated Support for Public Participation & Assistance in Mass Emergencies & Disasters', in ACM-BCS Visions of Computer Science 2010, British Informatics Society Ltd, Edinburgh.

State Library Victoria (2012), 2009 Bushfires in Victoria, State Library of Victoria. Online: http://guides.slv.vic.gov.au/content.php?pid=36394&sid=267991 Accessed: 12 January 2012.

Teague B. McLeod R. Pascoe S. (2009), 2009 Victorian Bushfires Royal Commission Interim Report, Parliament of Victoria, Melbourne, Interim Report.

Warren M. & Hutchinson W. (2005), 'A Systems Approach to Security', in 11th ANZSYS Conference Managing the Complex V, ISCE Publishing NZ, Christchurch, NZ.

Warren M. & Leitch S. (2011), "Australian National Critical Infrastructure Protection: A Case Study", Conference Proceedings of the 10th European and Information Warfare Conference, Tallin, Estonia.

Digital Finland: Life at the Screen

Jari Rantapelkonen[1] and Saara Jantunen[2]
[1]Department of Tactics and Operations Art, National Defence University, Helsinki, Finland
[2]Department of Leadership and Military Pedagogy, National Defence University, Helsinki, Finland
jari.rantapelkonen@mil.fi,
sijantunen@gmail.com

Abstract: This article aims to critically engage the Finnish cyber narrative from the perspective of human life. It discusses the representation and the narrative of future 'cyber life', as presented by the Finnish authorities through the *Digital Finland* document written in 2010. *Digital Finland* is a report on future prospects of the Finnish society by the Ministry of Transport and Communications. It proposes three alternative government programs for managing Finnish communication services and labels them as "progressive", "dynamic" and "decisive". These proposals make rhetorical claims and narrative assumptions about what is noteworthy in cyberspace, providing empirical data for analysis. The analysis discusses the actors and the actions reported in the document in the spirit of critical discourse analysis: Who are the participants in cyber life, and what are their actions and responsibilities? What do these narratives reveal of the authorities' perceptions of the cyberspace and its actors? The results show that the alternative programs are not only positive metanarratives of imagined future, but arbitrary and vague descriptions of the use of political power: without any reference to what is meant by "digitalizing", it remains unclear what makes the program labeled as "Decisive Finland" decisive. By choosing the terms "progressive", "dynamic" and "decisive" to describe the options, the authorities impose a desired image that corresponds with the promotion of the Nordic welfare model. What is listed under these definitions remains political. The "decisive" program proposes most funding and government control, and presents the most concrete and detailed plan for future cyber life. In contrast, the "dynamic" program is discussed in highly abstract terms and lacks indication of the roles and responsibilities, making it an unattractive option. In terms of content, *Digital Finland* fails to recognize the interrelationship between information technology and social life as complex questions about the nature of society, and ignores how networks, technology and society co-constitute each other.

Keywords: cyber strategy, cyber policy

1. Introduction

Finland is a keen supporter and an avid actor in developing the Nordic welfare model. The model itself is, according to the latest (June 2011) program of the Finnish Government, based on a high employment rate, competitive economy, and accessible services and care for all. This is what is considered the best social system -- a model that combines social cohesion and competitiveness. The Finnish government is willing to take a determined approach to improving the basic structures of the welfare society. Information and technology are defining the structures of welfare society. Finland has been a pioneer in the field of broadband and mobile telephone technology, which have been the key key innovations of the information society for the everyday lives of Finns. In 2003, there were some 300 000 broadband connections in Finland, compared to 2,4 million households. In 2006 the number of broadband connections had increased to 1,4 million, meaning 50 % of households were connected. According to the Global Information Technology reports, in 2006-2007 Finland ranked number 4 in the world in the networked readiness index. In 2007-2008, Finland fell to the sixth position, staying there until 2009-2010. The latest Global Information Technology Report 2010-2011 confirms that the leadership of the Nordic countries and the Asian Tiger economies in adapting and implementing information and communication technology advances both growth and development.

2. Government: Finland leading country in the development of cyber security

The Finnish government has recognized that the reliability of information networks is vital to the operation of modern information societies. The newly elected government in 2011 has declared cyber issues as a matter of “security and defence policy”. Preparing a cyber strategy for the national information security is an important objective. Also important is to actively participate in international cooperation in the information security field. “Finland’s goal is to become a leading country in the development of cyber security” (Finnish Government, 2011). Actually the aim of “leading light” was already mentioned in the Government Resolution on National Information Security Strategy called “Everyday security in the information society - a matter of skills, not of luck.” The strategy was adopted by the Finnish Government on the 4th of December 2008. The aim of the National

Information Security Strategy aim is to make everyday life in the information society safe and secure for everyone in Finland. Strategy defines everyone as individuals and businesses, administrative authorities, and all other actors in society. Strategy defines vision for the Finland as "people and businesses will be able to trust that their information is secure when it is processed in information and communications networks and related services." According to the National Information Security Strategy "by 2015 Finland will be the leading country in the world in terms of information security." The strategy focuses on three priority areas: 1) Basic skills in the ubiquitous information society, 2) Information risk management and process reliability, and 3) Competitiveness and international network cooperation." The term 'information security' mentioned in the 2008 Government document has been replaced by 'cyber security' in 2011. One may ask whether these are synonymous concepts or what the difference is, if the aim is has remained the same.

Defining the "transport and communication policy", the Finnish government sets a role for the information and communication technology: "The importance of information and communications technology for the improvement of growth and productivity is decisive." The government refers to the economical growth and the productivity of industry arguing that clear goals will be set for improved productivity (Finnish Government 2011). More particularly, the government declares what kind of aims Finland has with communications technology: "The provision and use of high-speed broadband connections will be promoted to make Finland the leading European country in terms of broadband access. The introduction of high-speed broadband connections will be promoted throughout the country and the expansion of the freely-available wireless network will be accelerated." One of the projects to enhance this is *Broadband for All by 2015* (Finnish Government 2011). For citizens, the government promises that "[a]ll citizens will be guaranteed barrier-free participation in the information society and the digital world regardless of their income level, health, financial status or place of residence." The government clearly is against digital exclusion and would like to see all citizens to be members of digital Finland. "The goal is to make digital data materials managed by the public sector available to citizens, companies, enterprises and organisations, authorities, and for research and education purposes in an easily reusable format via information networks.", announces the Finnish government (Finnish Government 2011).

3. From connectedness to productivity and well-being - Digital Finland

Digital Finland is a 2010 report on future prospects by the Ministry of Transport and Communications. It discusses themes relating to the Finnish society that require guidelines in the Government Programme in 2011-2015. The report puts forward three "suggestive" government programs, labeled as "Progressive Finland", "Dynamic Finland" and "Decisive Finland". The report states that it does not take a stance for any of the suggestive programs but that "based on them, it is also possible to put together different entities with varying emphasis relating to transport and communications policies." (LVM 2010, p.3).

"Progressive Finland" is the most moderate option of the suggested three. The aim is to increase the productivity of businesses and the public administration. The general aim is to advance democracy and Finland's competitiveness. (LVM 2010, p.3) "Dynamic Finland" aims to promote digital Finland as part of digital Europe. It, as well, aims to improve productivity and therefore the well-being of the citizens. (LVM 2010, p.3) "Decisive Finland" is the most radical with its aims. According to this program suggest, Finland is aiming for a remarkable competitive edge in international markets and economy (LVM 2010, p.3).

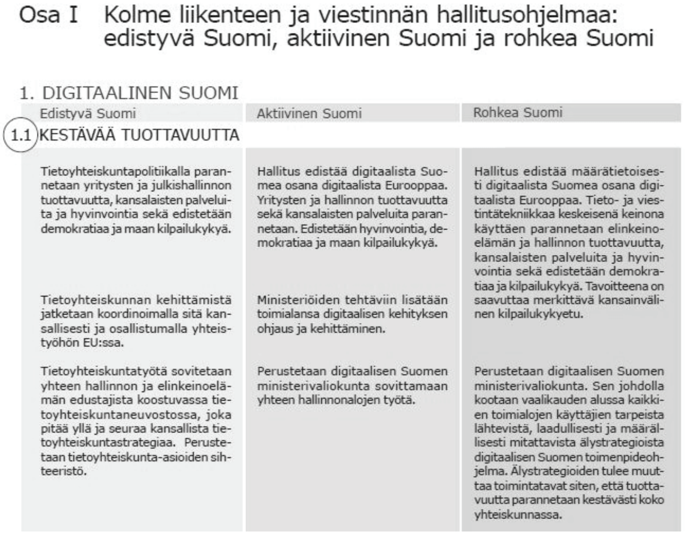
In addition to the theme of productivity, *Digital Finland* discusses the improvement of networks. The basic idea is that the Ministry of Transport and Communications seeks to promote an efficiently functioning society and national well-being by making sure that people and businesses have access to high-quality, safe and reasonably priced communication networks. This is ensured by competition between communications companies.

4. Analysis and discussion

This analysis focuses on three sections of the report, which discuss the digitalization process from the perspective of information networks. First, the references to the interdependence of productivity, communication and connectedness are discussed. The second part of the analysis focuses on the security discussion of the report. Third, the analysis is concluded by discussing the labels of each program: What makes the progressive program progressive, or the decisive program decisive?

4.1 Who is who in digital Finland?

This section discusses the first part of the *Digital Finland* report, which focuses on productivity.



Osa I Kolme liikenteen ja viestinnän hallitusohjelmaa: edistyvä Suomi, aktiivinen Suomi ja rohkea Suomi

1. DIGITAALINEN SUOMI

Edistyvä Suomi	Aktiivinen Suomi	Rohkea Suomi
1.1 KESTÄVÄÄ TUOTTAVUUTTA		
Tietoyhteiskuntapolitiikalla parannetaan yritysten ja julkishallinnon tuottavuutta, kansalaisten palveluita ja hyvinvointia sekä edistetään demokratiaa ja maan kilpailukykyä.	Hallitus edistää digitaalista Suomea osana digitaalista Eurooppaa. Yritysten ja hallinnon tuottavuutta sekä kansalaisten palveluita parannetaan. Edistetään hyvinvointia, demokratiaa ja maan kilpailukykyä.	Hallitus edistää määrätietoisesti digitaalista Suomea osana digitaalista Eurooppaa. Tieto- ja viestintätekniikkaa keskeisenä keinona käyttäen parannetaan elinkeinoelämän ja hallinnon tuottavuutta, kansalaisten palveluita ja hyvinvointia sekä edistetään demokratiaa ja kilpailukykyä. Tavoitteena on saavuttaa merkittävä kansainvälinen kilpailukykyetu.
Tietoyhteiskunnan kehittämistä jatketaan koordinoimalla sitä kansallisesti ja osallistumalla yhteistyöhön EU:ssa.	Ministeriöiden tehtäviin lisätään toimialansa digitaalisen kehityksen ohjaus ja kehittäminen.	
Tietoyhteiskuntatyötä sovitetaan yhteen hallinnon ja elinkeinoelämän edustajista koostuvassa tietoyhteiskuntaneuvostossa, joka pitää yllä ja seuraa kansallista tietoyhteiskuntastrategiaa. Perustetaan tietoyhteiskunta-asioiden sihteeristö.	Perustetaan digitaalisen Suomen ministerivaliokunta sovittamaan yhteen hallinnonalojen työtä.	Perustetaan digitaalisen Suomen ministerivaliokunta. Sen johdolla kootaan vaalikauden alussa kaikkien toimialojen käyttäjien tarpeista lähtevistä, laadullisesti ja määrällisesti mitattavista älystrategioista digitaalisen Suomen toimenpideohjelma. Älystrategioiden tulee muuttaa toimintatavat siten, että tuottavuutta parannetaan kestävästi koko yhteiskunnassa.

Figure 1: "Sustainable productivity", part 1

The three program alternatives are presented side by side, as can be seen above. The first phrases set the aims, and are then followed by suggestions on how to actualize them. The following chart lists these elements: first the aim and then the actors that are mentioned, and finally the proposed actions. The chart below presents the contents:

Table 1: Aims, actors and action descriptions in Digital Finland

Progressive Finland ("Edistyvä Suomi")	Improving the productivity of the private and public sector, improve public services and welfare, and advance democracy and competitiveness through information society policies.
Actors	Government administration Businesses Information Society Council Universities, research organizations
Actions	The development of information society is continued by coordinating it nationally and by participating in international cooperation Information society work is coordinated between government and business representatives in the information society council Maintain and observe national information society strategy work Found a secretariat for the matters of information society The "smart strategies" of all branches of administration are composed in the supervision of ministries and in the coordination of information society council Public sector information administration is developed Universities and research organizations are encouraged to participate in information society research and development work
Active Finland ("Aktiivinen Suomi")	The Government advances digital Finland as part of digital Europe. The productivity of businesses and the public sector are being improved. Welfare, democracy and competitiveness are being improved.
Actors	Ministries Cabinet committees Government
Actions	The management and improvement of digital progress are included in the tasks of the ministries A cabinet committee is founded to incorporate the work of different branches of administration In the supervision of the cabinet committee, Ministries compose their "smart strategies" that are based on the needs of the branch user The government decides on the Digital Finland operational programme which is used to improve sustainable productivity in the society Public information reserves are used in the development of electrical services The public sector information administration is developed and electrical public services

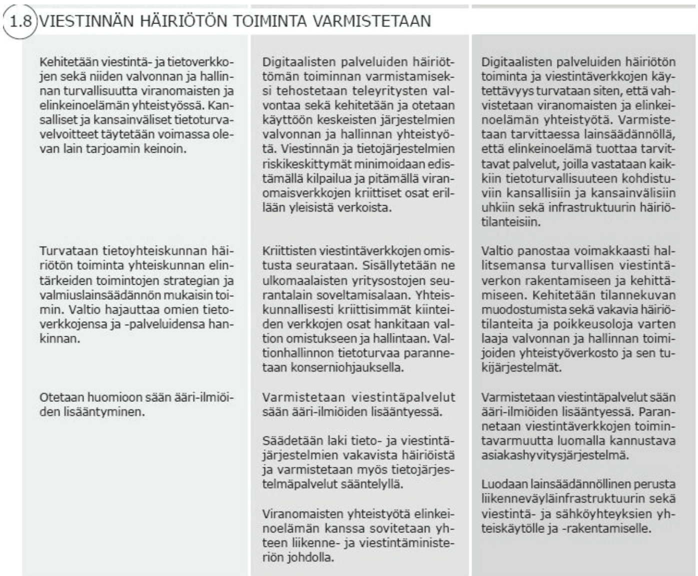
	advanced through corporate guidance Steer public research funding to the basic social research of digital society
Decisive Finland (“Rohkea Suomi”)	The government determinedly advances digital Finland as part of digital Europe. Through the use of information- and communications as central strategies, the productivity of businesses, the services and welfare of the citizens are being improved, and democracy is advanced. The aim is to achieve a significant, international competitive advantage.
Actors	Ministries Cabinet committee
Actions	A cabinet committee is founded. In its supervision a Digital Finland operational program, which takes into account the needs of the users in each branch, and the qualitatively and quantitatively measurable smart strategies. The smart strategies must alter actions towards the improvement of sustainable productivity in the entire society. The Cabinet committee prepares the procedures of the ministries in the matters of communication policy, the advancement of information and communication, communication technology, innovations, the copyrights of digital society, privacy, electrical trade, and other matters of advancing digital Finland. The public information reserves are widely opened for the development of electrical services The Cabinet Committee also supervises the public sector information administration, which is developed through strong corporate guidance in the government and municipal information administration. Steer significant amounts of public research funding to the basic social research of digital society

The observations that can be drawn from the section of the report are that

- *Progressive Finland* is characterized by ambiguous action descriptions, such as “improve sustainable productivity”. *Active Finland* and *Decisive Finland* present more descriptive actions and responsibilities, *Decisive Finland* being the most detailed about who does what. All programs are described in the passive form, which make them more or less ambiguous.
- The program titled as *Progressive Finland* brings up cooperation that would involve both the businesses and the government administration. The role of government increases in *Active Finland* and even more so in *Decisive Finland.*
- Competitive and productivity are assumed to rise according to the increase of government control.
- Only *Decisive Finland* explicitly states its aim (“competitive advantage”).

Even though the authors of *Digital Finland* claim they do not favor any single option of the three, the ambiguity of *Progressive Finland* and to a degree of *Active Finland* makes them the less attractive options, whereas *Decisive Finland* is presented as the most comprehensive and thorough option of the three.

4.2 Digital security



1.8 VIESTINNÄN HÄIRIÖTÖN TOIMINTA VARMISTETAAN

Kehitetään viestintä- ja tietoverkkojen sekä niiden valvonnan ja hallinnan turvallisuutta viranomaisten ja elinkeinoelämän yhteistyössä. Kansalliset ja kansainväliset tietoturvavelvoitteet täytetään voimassa olevan lain tarjoamin keinoin.	Digitaalisten palveluiden häiriöttömän toiminnan varmistamiseksi tehostetaan teleyritysten valvontaa sekä kehitetään ja otetaan käyttöön keskeisten järjestelmien valvonnan ja hallinnan yhteistyötä. Viestinnän ja tietojärjestelmien riskikeskittymät minimoidaan edistämällä kilpailua ja pitämällä viranomaisverkkojen kriittiset osat erillään yleisistä verkoista.	Digitaalisten palveluiden häiriötön toiminta ja viestintäverkkojen käytettävyys turvataan siten, että vahvistetaan viranomaisten ja elinkeinoelämän yhteistyötä. Varmistetaan tarvittaessa lainsäädännöllä, että elinkeinoelämä tuottaa tarvittavat palvelut, joilla vastataan kaikkiin tietoturvallisuuteen kohdistuviin kansallisiin ja kansainvälisiin uhkiin sekä infrastruktuurin häiriötilanteisiin.
Turvataan tietoyhteiskunnan häiriötön toiminta yhteiskunnan elintärkeiden toimintojen strategian ja valmiuslainsäädännön mukaisin toimin. Valtio hajauttaa omien tietoverkkojensa ja -palveluidensa hankinnan.	Kriittisten viestintäverkkojen omistusta seurataan. Sisällytetään ne ulkomaalaisten yritysostojen seurantalain soveltamisalaan. Yhteiskunnallisesti kriittisimmät kiinteiden verkkojen osat hankitaan valtion omistukseen ja hallintaan. Valtionhallinnon tietoturvaa parannetaan konserniohjauksella.	Valtio panostaa voimakkaasti hallitsemansa turvallisen viestintäverkon rakentamiseen ja kehittämiseen. Kehitetään tilannekuvan muodostumista sekä vakavia häiriötilanteita ja poikkeusoloja varten laaja valvonnan ja hallinnan toimijoiden yhteistyöverkosto ja sen tukijärjestelmät.
Otetaan huomioon sään ääri-ilmiöiden lisääntyminen.	Varmistetaan viestintäpalvelut sään ääri-ilmiöiden lisääntyessä. Säädetään laki tieto- ja viestintäjärjestelmien vakavista häiriöistä ja varmistetaan myös tietojärjestelmäpalvelut sääntelyllä. Viranomaisten yhteistyötä elinkeinoelämän kanssa sovitetaan yhteen liikenne- ja viestintäministeriön johdolla.	Varmistetaan viestintäpalvelut sään ääri-ilmiöiden lisääntyessä. Parannetaan viestintäverkkojen toimintavarmuutta luomalla kannustava asiakashyvitysjärjestelmä. Luodaan lainsäädännöllinen perusta liikenneväyläinfrastruktuurin sekä viestintä- ja sähköyhteyksien yhteiskäytölle ja -rakentamiselle.

Figure 2: “The undisrupted operating of communications is ensured”

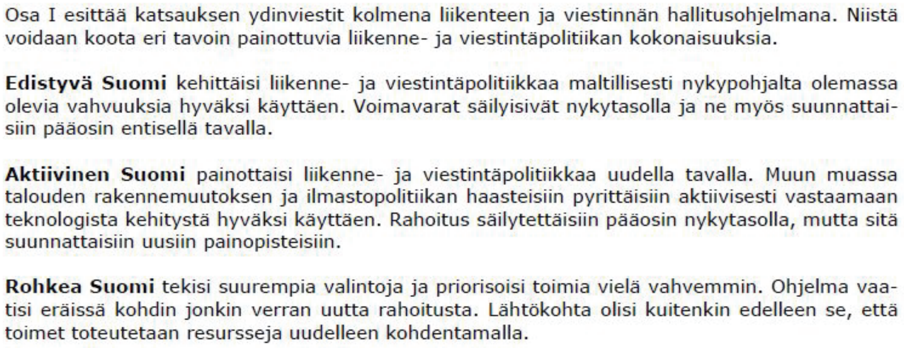
Table 2: Actors and actions in Digital Finland

	Action	Target	Actor
Progressive Finland	Develop	communication and information networks and the security of their control in cooperation between government and businesses	(passive form)
	Fulfill	national and international information security responsibilities	(passive form)
	Ensure	the undisrupted operationality of the information society	(passive form)
	Decentralize	procurement of information networks and services	Government
	Take into account	extreme weather conditions	(passive form)
Active Finland	Strengthen	the supervision of telecommunication businesses	(passive form)
	Deploy	cooperation in the supervision and management of central infrastructure	(passive form)
	Minimize	risk concentration of communication and information systems	(passive form)
	Promote	competition	(passive form)
	Keep	the critical systems of government networks separated from public networks	(passive form)
	Observe	ownership of critical communication networks	(passive form)
	Obtain	the most critical parts of the networks into the ownership and control of the government	(passive form)
	Improve	government information security through corporate management	(passive form)
	Ensure	communication services in case of the increase of extreme weather conditions	(passive form)
	Legislate	about critical problems of information and communication networks.	(passive form)
	Ensure	services through management	(passive form)
	Incorporate	the cooperation of government officials and businesses in the supervision of the Ministry of Communications	(passive form)
Decisive Finland	Secure	the undisrupted operability and usability of digital services and communication networks	(passive form)
	Strengthen	the cooperation between government officials and businesses	(passive form)
	Ensure through legislation	that businesses produce necessary services to respond to all national and international threats and infrastructure disruptions	(passive form)
	Invests heavily	in the development of the safe communication network it manages	Government
	Develop	the formation of situation awareness and, in case of a serious disruption or state of emergency, a comprehensive cooperation network for surveillance and control	(passive form)
	Ensure	communication services in case of the increase of extreme weather conditions	(passive form)
	Improve	the operability of communication networks	(passive form)
	Create	a legislative foundation for the joint use and construction of traffic routes and communication and electricity connections	(passive form)

In security, the government is presented as the only actor. All three programs mention cooperation between government and the private sector, but again the trend is that *Decisive Finland* is the most comprehensive and concrete version of the three options. Again, the role of the state and government is most prominent in Decisive Finland, although strongly present in *Active Finland* as well.

4.3 Progressive, active or decisive - but what would make Finland so?

The prologue of the report describes the programs as moderate policy-making with the emphasis on already existing assets (*Progressive Finland*), as “active” policy making with new points of emphasis (*Active Finland*) and as heavily prioritizing policy-making (*Decisive Finland*). The authors also estimate, that the policy introduced in *Progressive Finland* would not be likely to demand more resources, whereas *Decisive Finland* would. In other words, the amount of government control correlates positively with the need for funding, whereas public/private cooperation is seen as moderate and low cost.



Osa I esittää katsauksen ydinviestit kolmena liikenteen ja viestinnän hallitusohjelmana. Niistä voidaan koota eri tavoin painottuvia liikenne- ja viestintäpolitiikan kokonaisuuksia.

Edistyvä Suomi kehittäisi liikenne- ja viestintäpolitiikkaa maltillisesti nykypohjalta olemassa olevia vahvuuksia hyväksi käyttäen. Voimavarat säilyisivät nykytasolla ja ne myös suunnattaisiin pääosin entisellä tavalla.

Aktiivinen Suomi painottaisi liikenne- ja viestintäpolitiikkaa uudella tavalla. Muun muassa talouden rakennemuutoksen ja ilmastopolitiikan haasteisiin pyrittäisiin aktiivisesti vastaamaan teknologista kehitystä hyväksi käyttäen. Rahoitus säilytettäisiin pääosin nykytasolla, mutta sitä suunnattaisiin uusiin painopisteisiin.

Rohkea Suomi tekisi suurempia valintoja ja priorisoisi toimia vielä vahvemmin. Ohjelma vaatisi eräissä kohdin jonkin verran uutta rahoitusta. Lähtökohta olisi kuitenkin edelleen se, että toimet toteutetaan resursseja uudelleen kohdentamalla.

Figure 3: Progressive, active and decisive Finland as presented in the report summary

The cyber domain seems to be more or less understood as a technical tool for improving the economy a society where productivity is a high aim. The vision is that a global, computerized, networked economy is the one Finland is aspiring to. Some experts have even argued that profits for the Finnish society can be in the billions. Without question, *Digital Finland* sees information and communications technology in a positive light, whereas dreams of “ubiquitous”, “fast” networks and “connectivity” are a path to a better economy and life.

On the other hand, the document sees the information society vulnerable for the interferences and interruptions.

5. Conclusion

Digital Finland is a grand narrative of imagined cyber life that unveils the future without negative aspects. The future, according to the narrative, will be better if government is put in the frontlines of cyber security development and responsibility.

Cyber life, according to the “decisive” program (which is obviously the preferred option), is very much an instrument necessary to achieve the best productivity and economic growth for Finland. *Digital Finland* contains the idea of a close connection between the pursuit of economic efficiency and digitality.

However, *Digital Finland*’s metanarrative does not discuss the consequences of possible fragmentation of other narratives that have been present in the shadows of globalization. *Digital Finland* is silent about the disappearance of the Subject. Therefore the future *Digital Finland* proposes is very much about emphasizing both power and politics, where digitality is the part of technology that actually is a mode of revealing what Heidegger has written about technology. The future of *Digital Finland* reveals only one form of a computer culture. This takes us back to Sherry Turkle’s findings on cyber cultures and how “[i]n the 1970s through the mid-1980s, the ideology that there was only one right way to “do” computers nearly masked the diversity of styles in the computer culture. In those days top-down thinkers didn’t simply share a style; they constituted an epistemological elite” (Turkle 1995, p.54).

The nature of technology includes the paradox that claims “to bring the future under the rule of instrumentality, while it both largely devoid of historical awareness and the primary source of the disruptions that will falsify any expectations of the future” (Ross 1990, p.261). *Digital Finland,* in the name of productivity and efficiency, denies how the past and future of technology are disruptive. The "decisive" program contains the idea of the inseparability of power, desire, and truth. Future digital Finland with the “decisive” narrative is actually the desired life, which can not be an explicitly presented as a utopian dream, nor an implicit presentation of the invisible, such as Neuromancer’s cyberspace: “A consensual hallucination experienced daily by billions…” (Gibson 1984, p.51). Denying the disruptiveness of the future is typical for high-tech narratives, and in that sense *Digital Finland* is not different from any other plans of technology for the future cyber life.

Politics is a struggle between narratives. *Digital Finland* recognizes the production of services, such as broadband for every Finn and online services as its outcome. This 'revolutionizes' and turns human life into digital life in order to create economic productivity. This is the key to better life, a cyber life in digital Finland. It should be noticed that the concept of cyber or digitality are not clearly defined - or, in fact, defined at all. This is what should make us approach *Digital Finland* critically. According to the report, the citizens of digital Finland should trust the government that determines the policy on communication and its organization, but also decides to introduce the society to the 'improved cyber life'. Digital Finland does not discuss how cyberspace or “the digital environment is foremost a culture of rapid changes and adaptability: it is a cultural phenomenon driven by social adaptations of technological innovations, and thus it calls for a dual inquiry into its inner mechanisms and structures” (Doueihi 2011, p.xvii). This takes us to think about the competitive edge, “customer is always right”, and whether this aproach suits the aims of digital Finland. However, in the future, life will be experienced through the computers and lived at the screens - which is already very much the state of hybrid virtual life of many who reshape cyberspace. This is no longer a future challenge for the political narratives -- it is our present.

References

Dutta, Soumitra; Mia, Irene (2011) The Global Information Technology Report 2010-2011, Transformations 2.0. World Economic Forum, Geneva.

Finnish Government (2011) Programme of the Finnish Government, 22 June 2011. Available at http://www.valtioneuvosto.fi/hallitus/hallitusohjelma/pdf332889/220611hallitusohjelma_en.pdf

Doueihi, Milad (2011) Digital Cultures. Harvard University Press, Cambridge, Massachusetts, and London, England.

Gibson, William (1984) Neuromancer. Ace books, New York.

LVM (2011) "Suomi tietoturvan suunnannäyttäjäksi. Suomalaisen tietoturvaosaamisen levittäminen ja aktiivinen osallistuminen standardien kansainväliseen kehittämistyöhön", The Ministry of Transport and Communications. Julkaisuja-sarja 17 / 2011. Available at http://www.lvm.fi/c/document_library/get_file?folderId=1551284&name=DLFE-11972.pdf&title=Julkaisuja%2017-2011

LVM (2010) "Digitaalinen Suomi, uusi liikennepolitiikka. Liikenne- ja viestintäministerön tulevaisuuskatsaus puolueille 10.9.2010", The Ministry of Transport and Communication. Julkaisuja-sarja 33 / 2010. Available at http://www.lvm.fi/c/document_library/get_file?folderId=964900&name=DLFE-10937.pdf&title=Julkaisuja%2033-2010%20LVM%20Tulevaisuuskatsaus%2010092010

Ross, Stephen David (1990) "Power, Discourse, and Technology: The Presence of the Future". In Shapiro, Gary (ed.). 1990. After the Future. Postmodern Times and Places. State University of New York, Albany, NY, pp.255-272.

Turkle, Sherry (1997) Life on the Screen: Identity in the Age of the Internet. Simon & Schuster.

Finding Suspicious Activity on Computer Systems

Neil Rowe and Simson Garfinkel
U.S. Naval Postgraduate School, Monterey, California, USA
ncrowe@nps.edu

Abstract: When computer systems are found during law enforcement, peacekeeping, counter-insurgency or similar operations, a key problem for forensic investigators is to identify useful subject-specific information in a sea of routine and uninteresting data. For instance, when a computer is obtained during a search of a criminal organization, investigators are not as much interested in the machines used for surfing the Internet as the machines used for accounting of drug deals and emailing to co-conspirators. We are doing research on tools to enable investigators to more quickly find such relevant information. We focus on the directory metadata of a computer drive, the listing of the stored files and directories and their properties, since examining it requires much less time than examining file contents. We discuss first what ways people try to hide things on drives. We then discuss clues that suggest concealment or atypical usage of a drive, including encryption, oddities in file names, clusters of deletions, atypical average values, and atypical clusters of files. We report on experiments we have conducted with a corpus of drives purchased from a range of countries. Processing extracted the directory metadata, classified each file, and calculated suspiciousness metrics on the drives. Experimental results showed we could identify some suspicious drives within our corpus but with a certain number of false alarms.

Keywords: digital forensics, law enforcement, drive classification, metadata, suspicion

1. Introduction

An increasingly important aspect of modern warfare is control of criminal and terrorist activities in occupied territories. This includes searches during law-enforcement, peacekeeping, and counter terrorism operations, and these increasingly encounter computers and other digital devices (Pearson, 2010). The data in the secondary storage or “drives” of these devices can provide considerable information about illegal activities. Drives can also provide information of the development and deployment of cyber arms that we are attempting to control (O'Neill, 2010). We are thus developing methods to automatically aid in assessing drives.

Usually investigators search for files on a drive that contain certain keywords called “selectors.” These can be names of particular people and organizations such as drug producers or human traffickers. They might occur in electronic mail, Web-page downloads, and documents. However, inspecting all the contents on a drive takes a good deal of time, as many files are of no interest such as software and operating-system bookkeeping files. So a key challenge for digital forensics is to quickly assess a drive’s investigatory value.

We propose a first step to examine just the metadata (directory information) of the file system on the drive (Buchholz and Spafford, 2004). Metadata is typically 1000 times smaller in size than the contents of a drive, and it alone provides sufficient descriptions and statistics about files to give a good picture of what users are doing (Agrawal et al, 2007). Often drives captured during raids were captured by surprise and did not allow owners to conceal or destroy much. So we ought to be able to make discoveries with them.

2. Clues that a drive is of interest

Criminals who are aware that their computers and digital devices could be seized bylaw enforcement may try to conceal information with anti-forensics techniques (Garfinkel, 2007). Anti-forensics is rare since most users have little need to hide anything besides passwords, so any evidence of itis a clue that a drive is worth inspecting further. (Jahankhani and Beqiri, 2010) enumerates nineteen categories of anti-forensic techniques, some of which can be detected in the metadata alone:

- Media wiping, or erasing of the contents of a drive: Detectable in metadata. It is useful to know even if thorough because it makes other related drives more interesting.
- Steganography (hidden messages): Not detectable in metadata since affects only the contents of the files.
- Anonymization of data: Not detectable, similarly.
- Rootkits: Not detectable.
- Tools for analysis of the hardware: Detectable in metadata.

- Homographic naming, using a name similar to a well-known one: Detectable in metadata.
- Signature modification on files: Ineffective when forensics is used since forensics will create its own signatures.
- Encryption: Detectable in metadata from the file extension or from encryption markers.
- Modified metadata: Not detectable.
- Hiding files in the marked-as-unusable ("slack") space of a disk or in other odd places in the file system (Huebner, Bem, and Wee, 2006): Detectable in metadata if we have it for those parts of the drive, but this requires specialized techniques.
- Files constructed to create deliberate hash collisions with known files: Detectable in metadata.
- Hiding data within memory: Not detectable, but irrelevant if drive can be analyzed.
- Planted misleading evidence: Detectable in metadata if it makes the drive atypical among similar drives.
- Non-encryption encoding of files: Detectable in metadata from the file extension.
- Exploiting of forensic-tool vulnerabilities: Possible but depends on the vulnerability.
- Creation of large numbers of dummy files to confuse analysis: Detectable in metadata.
- Internet anonymization techniques: Not detectable.
- Use of anti-forensics hardware: Not detectable.

Applying these ideas, the following clues can be sought in the metadata of a file system to detect possible concealment and anti-forensics:

- *Encrypted files and directories*
- *Extensions indicating encryption*
- *Encrypted directories*
- *Encryption programs*
- Suspicious file extensions
- *Nonstandard ones, especially long ones and integers*
- *Double extensions, e.g. "setup.exe.txt"*
- *Rare extensions*
- Deceptive paths (the directory sequence to a file)
- *Rare characters such as "{" or the higher-numbered HTML-encoded characters.*
- *Unnecessarily atypical ways of specifying characters, e.g. "&#65;" for the letter "A"*
- *Misspellings (e.g. "little figher", "vewing", "addresss"), which can be camouflage*
- *Engineered hash collisions with very different paths and file names*
- Malicious software
- *Malware executables, as per extension or filename*
- *Tools known to be associated with malware or anti-forensics*
- *Source code of malicious software*
- Deletions of files
- *Many deletions*
- *Clusters of deletions around the same time, suggesting attempts to destroy evidence*
- Many small similar files or files not accessed after being created, suggesting they were created as decoys.
- Atypical counts of certain file types of interest to investigators. This depends on the investigation but could include email, photographs, video, or program source code.

However, there are also legitimate reasons for the appearance of these clues:

- Encryption is needed to hide sensitive or private information, passwords, and keys.
- Pathnames may use unusual characters if they are created by software.
- Double extensions are used by certain kinds of software.
- Rare extensions may represent rare but legitimate use.
- Malicious software may infect a system without the user being aware of it.
- Some users are atypical in their software use.
- Users delete files in clusters whenever they feel a need to cleanup their systems.

We can be more accurate in identifying suspicious drives if we look for multiple clues on the same drive, much as analysts look for multiple clues to confirm hypotheses (Hollywood et al, 2004). We can add suspiciousness of individual clues to get an overall suspiciousness of a drive.

3. The experimental test bed

We have been experimenting on the Real Drive Corpus (RDC) (Garfinkel et al, 2009), a corpus of drives purchased as used equipment from 22 different countries. As of March 2012 we had 2310 drive images containing 40.0 million files comprising 13.7 million unique paths. Most were disks from desktops and laptops, but some were from mobile devices and some were from storage devices. They represent a variety of usage including business users, home users, and servers of several kinds, and ranged from 0 to 25 years old. For these experiments the corpus was augmented with a few drive images created specifically for testing that included deliberate suspicious behavior. Most used NTFS file systems, but some had the older Microsoft FAT system. The great majority of the drives appeared to have normal usage without any criminal or terrorist activity.

We used the Fiwalk program(Garfinkel, 2009) to extract metadata for these files including file path and name, file size, times, NTFS flags (allocated, empty, compressed, and encrypted), fragmentation status, as well as each file’s MD5 hash values. 27.7% of the filesin the corpus are “unallocated” or deleted.

FAT file systems modify the directory entries of deleted file by changing the first character to hexadecimal 0xE5. In addition, metadata for deleted files on all systems may also be missing its directory information. Since deleted files are of special forensic interest, we try to reconstruct the original file path in these cases when a correction is unambiguous in our corpus. Our methods found corrections for 15.9% of the relevant files in the current corpus. Additional methods could recover more of the remaining deleted file paths (Naiqi, Zhongshan, and Yujie, 2008).

3.1 File classification based on extension and directory

Quickly assessing a drive requires some statistics as to what kinds of files are on it. This requires classifying the files into semantically meaningful groups like pictures, spreadsheets, and word-processing documents. Three kinds of groups were defined based on the 22,565file extensions (like “htm”), the 8,373top-level directories in which the files occurred (like “WINDOWS”), and the 5,159immediate directories in which the files occurred (like “photos”). When immediate directories are ambiguous or just arbitrary codes, we search their parent directories in succession to find one with an unambiguous group assignment. An example where this is important is Documents and Settings/Administrator/Application Data/Microsoft/Internet Explorer/Quick Launch/Show Desktop.scf where the immediate directory that explains the purpose of this file is “Application Data”. Currently we assign all extensions and directories that occur at least 200 times in our corpus, and others are assigned to the category of “miscellaneous”.

For grouping file extensions, we used Wikipedia's list of common extensions and the lists of www.file-extensions.org. For grouping directory names, we used common-sense knowledge for known items and we ran Google queries for unknown items. For instance, the directory names of “videos”, “movies”, “peliculas”, “multimedia”, and “my videos” all map to the "video" category of immediate directory, but “clips” does not because it could also mean an image. For the European-language words in the RDC (mostly Spanish and German) we used a dictionary (like for "peliculas"). For Asian languages we used Google Translate, but often their file paths included English words that we could exploit. Currently 8,102 extensions and directories are mapped to 77 categories. Table 1 shows the major file groups in the corpus, with percentages both before and after the filtering described in the next section.

Table 1: Percentages of major file groupsin our corpus, before and after filtering for known files (of total of 40.0 million files)

Extension		Graphics	28.8%, 34.1%	None	14.8%, 14.5%	Executable	12.5%, 11.8%
Web	5.6%, 4.7%	Microsoft OS	5.3%, 4.2%	Camera image	5.1%, 6.8%	Audio	3.6%, 1.7%
Config-urations	3.1%, 2.1%	Game	2.6%, 1.2%	Non-MS document	2.1%, 2.0%	Multiple use	1.7%, 1.5%
Temporary	1.5%, 1.8%	Links	1.2%, 1.5%	Help	1.1%, 1.2%	XML	1.0%, 0.9%
Low frequency	0.9%, 1.2%	Log	0.8%, 1.1%	Program source	0.8%, 0.6%	Microsoft Word	0.7%, 0.9%
Query	0.6%, 0.9%	Spread-sheet	0.6%, 0.3%	Encoded	0.5%, 0.6%	Copy	0.5%, 0.5%
Database	0.4%, 0.4%	Integer	0.3%, 0.4%	Video	0.3%, 0.2%	Security	0.3%, 0.2%
Disk image	0.3%, 0.2%	Present-ation	0.3%, 0.3%	Geogra-phic	0.2%, 0.1%	All other	1.1%, 1.2%
Top directory		Deleted file	27.7%, 40.0%	Program	23.4%. 15.7%	Microsoft OS	19.6%, 20.5%
Document	13.6%, 14.0%	Temporary	4.4%, 4.1%	Unix and Mac	3.8%, 2.0%	Game	3.5%, 1.8%
Hardware	0.7%, 0.7%	Root	0.3%, 0.2%	Microsoft Office	0.1%, 0.0%	Docs. and Settings	0.1%, 0.1%
Immediate directory		Root (mostly default)	25.7%, 36.0%	Temporary	15.3%, 17.3%	Operating system	13.7%, 12.3%
Application	10.1%, 8.1%	Visual images	9.8%, 7.7%	Documents	4.6%, 3.3%	Hardware	3.3%, 2.3%
Audio	3.1%, 1.1%	Games	2.0%, 1.2%	Installation	1.5%, 1.1%	Data	1.4%, 1.1%
Help	1.4%, 1.3%	Web	1.3%, 1.1%	Logs	1.2%, 1.1%	Program-ming	1.1%, 0.8%
Security	1.1%, 0.9%	Sharing	0.9%, 0.9%	Video	0.3%, 0.2%	All other	0.6%, 0.7%

3.2 Filtering out known files

Forensic investigators are primarily interested in user-created files. So it is useful to exclude files of the operating system, applications software, and hardware since they do not say much about the distinctive characteristics of the user. We can do this by searching for the hash codes that Fiwalk computes on the files in the set of known hash values of the National Software Reference Library Reference Data Set (NSRL, from the U.S. organization NIST at www.nist.org/nsrl). This is an extensive collection of hash values on published software and its accompanying files. 12.2 million of the files in our corpus,or 30.4%, had hash values in NSRL, though not always under the name listed by NSRL.

A weakness of the NSRL is that it currently provides hash values only from the static files supplied with software. Some important files are created once software is installed and starts running, such as default documents. We, however, can exploit our large corpus to guess likely additions to the NSRL hash values from those files that occur on more than a certain minimum number of disks in our corpus. A minimum of five occurrences worked well in our tests. It is also reasonable to eliminate files having the same name and path as other files in the corpus that do have an NSRL hash code, since these are likely to be different versions of the same file, This eliminated an additional1.1 million of the original corpus files as being uninteresting for further analysis, giving a total reduction of 33.26%. The second percentage given in Table 1 is for after this filtering.

4. Experiments

We implemented software to test our corpus for the clues to suspicious behavior mentioned in section 3. These tools for preprocessing the metadata are part of the Dirim system first reported in (Rowe and Garfinkel, 2011). Dirim currently follows 59 steps to produce 180 analysis files.

4.1 Encryption

Encryption is an overt clue to concealment. NTFS metadata allocates bits to indicate that a file or directory is encrypted. We did not see these bits set in any of the files of our corpus. However, we did see files whose encryption was indicated by their file extension. There were 32,806 of these in the corpus after filtering known files. Drives with a significant number of encrypted files were suspicious. We also looked for encryption software that was not part of the operating system since it is not normally installed except by people with something to hide; we counted each occurrence of such software as equivalent to 20 encrypted files in the overall total as a quick way to credit it.

4.2 Suspicious file extensions

Clues to suspicious files occur in their file extensions. Unrecognized extensions longer than 4 characters are suspicious since they are generally nonstandard and an easy way to hide data and programs. An example is avgxpl.dll.prepare where extension “prepare” is nonstandard. There were 7,215 occurrences of these in the corpus after excluding accepted known ones. Double extensions can also be suspicious since the outer extension may serve to conceal the inner extension. We found 25,718 suspicious double extensions on the corpus after excluding some judged as legitimate. Links, copies, and compression extensions like "lnk", "bak", "zip", and “manifest” have legitimate double extensions to represent the object of the action, files ofInternet addresses often use the periods of the address, and some legitimate periods are associated with abbreviations. A suspicious example is ActSup.dll.tag, where “tag” conceals an executable extension “dll”. Drives high on the number of suspicious extensions were judged suspicious.

Rare extensions are suspicious since they are unusual use. Rarity should not be defined by the overall count in a corpus, however, because many rare extensions occur numerous times on the drives on which they are found. Wethus focus on the number of drives on which an extension occur, which we define as m_j for extension j of M extensions. Then for each disk *i*, average rarity of its extensions can be calculated as $r_i = \sum_{j=0}^{M-1} (o_{ij} / m_j) / \sum_{j=0}^{M-1} o_{ij}$ where o_{ij} is 1 if extension *j* occurred at least once on disk *i*. We got a mean of 0.0272 and a standard deviation of 0.0599 with this metric, but some values were much higher, like one drive that had a value 0.893 on 33,017 files, indicating nonstandard usage.

4.3 Suspicious paths

Files can also be suspicious if they have apparent obfuscation in their paths in the form of significant numbers of punctuation marks and, to a lesser extent, digits. Examples are “program files/!$!$!$!$.mp2” which has too many punctuation marks to be honest, and “windows/{15d372b6-e470-11da-bb68-00105a10e007}.dat” which fails to indicate what kind of data it holds unlike most Windows operating-system files. In addition, names of files and directories that start with a punctuation mark are suspicious because this is not standard English and it is an easy way to obfuscate, though there are important exceptions such as “#” and “$” (standard program prefixes) and “&#” (HTML character codes). We found 29,002instances of this kind of apparent obfuscation in the NSRL-filtered corpus. We used the identity of the group of the immediate directory to exclude those that were frequently seen as legitimate use of automated naming: temporaries, encodings, installation files, logs, data, and security information.

Certain characters alone are inherently suspicious, such as hexadecimal codes for characters rather than standard UTF-8 or UTF-16 encoded code points, HTML-encoded code points less than U+007F(since they can be written in UTF-8 with a single byte), and code points larger than U+1000. We found 426,142, 1,765, and 786,770 instances respectively of these characters in the corpus, so the first and third are not strong clues.

We also sought directory and file names that were misspellings of common names, another way to obfuscate, and found 4,194 occurrences in the file names of the corpus. This required a 172,173-item list of common words in the corpus languages, as well as software and hardware terms, that we compiled from a range of sources. False alarms were reduced by only counting misspellings differing by one alphabetic letter that were at least 10 times less common than their properly spelled counterpart in names at least 5 characters long.

4.4 Malicious software

The presence of malicious software may indicate an attempt to distribute it. Known malicious software can be detected by running antivirus software on a disk image. Clam AntiVirus was run on a sample of our Windows drive images. It found 6874 files on a 48-disk subset of our corpus whose contents matched virus signatures. Correlation with the other suspiciousness factors was weak.

As a shortcut to signature checking, sources like www.fileextensions.org list extensions generally associated with malware like “pid”, “blf”, and “gbd3”. We found 5,559instances of these in our corpus, all in software directories. But many appear to be legitimate uses that either unwittingly use a malware extension or that predate the occurrence of the malware. Some file names are specifically associated with malicious software, but most use well-known or random names for camouflage.

As for development of malicious software, the developers may have model software that will have recognized signatures. If not, the occurrence of specific software associated with malware development such as Metasploit is a clue, as is the weaker clue of file extensions and directories known to be associated with software development.

4.5 Deliberate hash collisions

A clever way to conceal a file from detailed forensic inspection would be to cause it to have the same hash code as a known innocent file. This would be useful because inspectors often use hash values from NIST or other vendors to rule out uninteresting files from further analysis. This is quite difficult because of the high computational cost to find hash collisions with the standard algorithms of SHA and MD. But it is at least worth looking for such sophisticated attacks.

A benefit of our checking files against the NSRL database is that we can assemble lists of file names of files with the same hash value; a name different than the predominant name is suspicious. We counted 340,739 such files on the corpus, where a hash value occurred at least 20 times, the predominant name occurred at least 50% of the time, but the file name in question occurred only once. Drives with large numbers of such files are more suspicious. There are legitimate reasons to rename files with unique names as when copying them, but a large amount of copying can be suspicious too.

4.6 Clusters of deletions

We can seek clusters of activity at suspicious times, such as just before the drive was captured from an insurgent. To find deletion clusters, Dirim counts the deleted files (marked by the "unallocated" flag) by day of modification for each drive and subtracts the number of files created on that day. Drives that have an unusually large number of days where this number exceeds a threshold (currently 100) are suspicious. We found 5,753 instances of such days in the corpus. (Rowe and Garfinkel, 2010) discusses more of what can be detected in analysis of file times. The total number of deletions on a drive can also be a suspiciousness clue, as people engaged in clandestine enterprises have more reason to delete files than ordinary users.

4.7 A typical drive averages

Dirim computes averages for each drive on a number of parameters obtainable from metadata, as well as counts on the file groups of Table 1. Drives atypically high or low on these statistics may be suspicious depending on the investigation goals. For instance, the following automated summary of a student-created drive shows an unusually large number of small files created in a narrow time period, indicators of suspiciousness.

Summary of drive 1457 summer11_scenario4.xml:
Temporal characterization: little-used
low_standard deviation of modification-creation
low_standard deviation of access-creation
high_standard deviation of log of length of filename
low_average filename alphabeticality
low_average filename commonality
low_standard deviation of filename commonality
high_fraction_of_Windows_OS_topdir
high_fraction_of_logs_and_backup_botdir
high_suspicious_extensions
high_suspicious_path_characters
high_rare_extensions

4.8 A typical clusters

More detailed differences between drives can be seen by comparing their file clusters. Table 2 gives the 34 properties we found after experiments to be the most useful for clustering. The first ten are normalized by mapping onto ranges of 0 to 1 by functions of the form $f(x) = \Phi((x - \mu_x) / \sigma_x)$ for ordinary properties, or $f(x) = \Phi((\log(x) - \mu_{\log(x)}) / \sigma_{\log(x)})$ for widely varying properties like file and directory size, where Φ is the integral of the normal distribution with mean of 0 and standard deviation of 1, μ_x is the mean of the property over the entire corpus, and σ_x is the standard deviation. This transformation maps the value to its fractional rank order in the sorting set of all values assuming it has a normal distribution, and most of the properties were close to normal; it provides a quick estimate of rank order without sorting. The remaining 24 properties are unnormalized and assigned by feature vectors provided for each group in Table 1; a file's values are the weighted average of 55% of the feature vector of its extension group, 10% of the feature vector of its top-directory group, and 35% of the feature vector of its immediate-directory group.

Table 2: Properties of files used in clustering them

Log of size	Modification-creation	Access-creation	Access-modification
Log of depth	Log of name length	Alphabetic fraction	Log of count of foreign characters
Log of frequency in corpus	Log of size of containing directory	Degree of frequent update	Degree of being user-owned
Degree to which relates to operating system	Degree to which relates to hardware	Whether is an executable	Degree to which relates to executable support
Degree to which relates to application support	Whether at root	Whether has no extension	Whether is temporary
Whether is encoded	Whether is a disk image	Degree to which is a document	Degree to which relates to mail
Whether is a presentation	Whether is a spreadsheet	Degree to which relates to the Web	Whether is a visual image
Whether is audio	Whether is video	Degree to which relates to programming	Degree to which relates to specialized applications
Degree to which relates to games	Degree to which relates to security	Whether is data	

Rowe and Garfinkel, 2011) describes a clustering algorithm based on K-Means clustering the files of each drive, including iterative splitting and merging of clusters, and then clustering the clusters. We have since improved performance by taking a large random sample of the entire corpus, clustering it, mapping the entire corpus to the cluster centers found, and then clustering the residual files insufficiently close to any cluster center to provide additional cluster centers.

Figure 1 summarizes the clustering found for the 837 Windows drives in our corpus by plotting the clusters by the first two principal components, where size of the circle represents the size of the cluster. The big clusters are for caches, operating-system files, and applications files. Suspiciousness is related to the size of the cluster, not its position in this display, since there are many legitimate reasons for files to have anomalous principal components. We measure suspiciousness of a drive's clusters by the average of the reciprocal of the total number of drives with at least one representative of a cluster that has a representative on the drive.

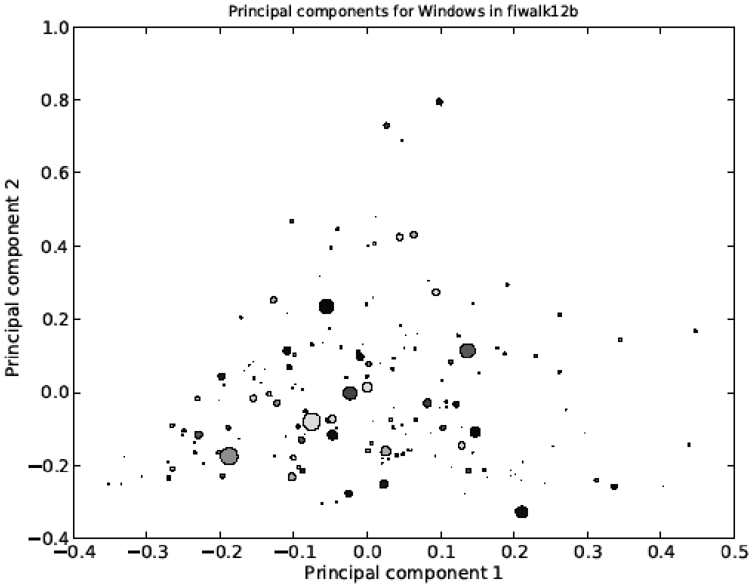


Figure 1: First two principal components of the files of the 837 Windows drives in our corpus

4.9 Rating overall drive suspiciousness

To rate the overall suspiciousness of a drive, we can combine the abovementioned clues by taking a weighted average of their suspiciousness measures. We did experiments using the unweighted average of 15measures on the files after filtering out known files: number of bad extensions, number of bad paths, extension rarity metric, number of misspellings, number of hexadecimal characters, number of low HTML code numbers, number of high HTML code numbers, number of files with unique names for their hash code provide another name occurred at least 10 times, number of encrypted files, number of deletion clusters, fraction of files on drive that were deleted, fraction that were email files, drive-cluster uniqueness, average file size (a negative factor), and variance in access time minus creation time (a negative factor). We took logarithms of one plus the value for the first nine since there values varied considerably between drives. We normalized the measures using the formula of the last section, and then took their average.

Figure 2 shows the histogram of overall suspiciousness for our 837 Windows drives. The mean was 0.47 with a standard deviation of 0.10. Of the test drives, the one with repeated deletions rated 0.54; the one with many encrypted messages rated 0.51; and three used in the “M57” experiments simulating more subtle malicious activity rated 0.56, 0.53, and 0.55. These values were above the mean, suggesting the drives were worth investigating. But there are too many factors here which have legitimate explanations that interfere with obtaining clearer suspiciousness ratings. The drives rated above 0.6 all had intriguing features and justify further study.

5. Conclusions

In investigations of criminal activity, several clues can quickly distinguish a suspicious drive from an uninteresting drive using just its metadata. Certainly we can look for keywords representing targets of interest, but we can also look for general evidence of concealment and deception just in the file system. These clues may save us valuable time in directing our attention for more detailed analysis of file contents.

The views expressed are those of the author and do not represent those of any part of the U.S. Government.

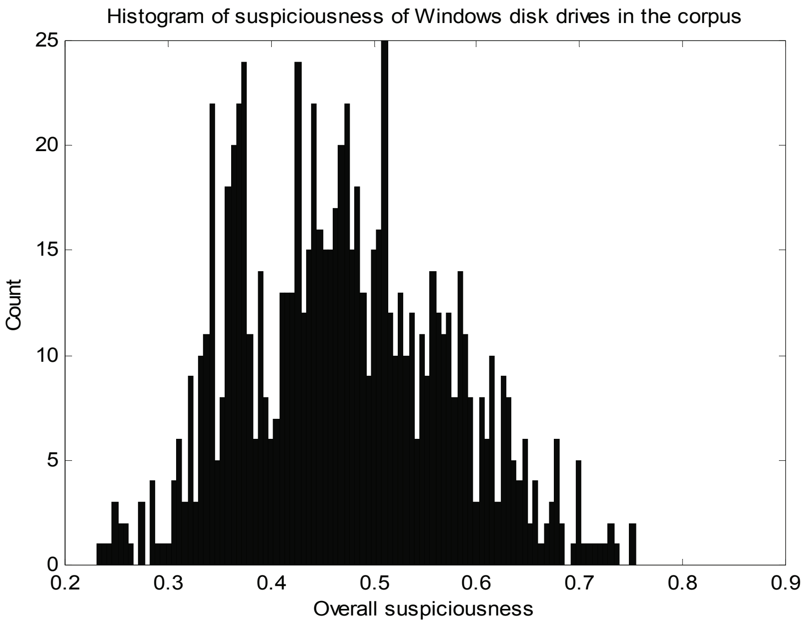


Figure 2: Histogram of the suspiciousness of the 837 Windows drives in the corpus

Acknowledgements

Thanks to Hector Guerrero and Jose Ruiz. Our software and results on our corpus are freely available for further research.

References

Agrawal, N., Bolosky, W., Douceur, J., and Lorch, J. (2007)“A Five-Year Study of File-System Metadata”,*ACM Transactions on Storage*, Vol. 3, No. 3, October, pp. 9:1-9:32.

Buchholz, F., and Spafford, E. (2004)“On the Role of File System Metadata in Digital Forensics”,*Digital Investigation*, Vol. 1, pp. 298-309.

Garfinkel, S. (2007)“Anti-Forensics: Techniques, Detection and Countermeasures,” 2nd International Conference on I-Warfare and Security (ICIW), Naval Postgraduate School, Monterey, CA, March 8-9.

Garfinkel, S. (2009)“Automating Disk Forensic Processing with SleuthKit, XML and Python”, in Proc. Systematic Approaches to Digital Forensics Engineering, Oakland, CA, USA.

Garfinkel, S., Farrell, P., Roussev, V., and Dinolt, G. (2009)“Bringing Science to Digital Forensics with Standardized Forensic Corpora”,*Digital Investigation*, Vol. 6, pp. S2-S11.

Hollywood, J., Snyder, D., McKay, K., and Boone, J. (2004)*Out of the Ordinary: Finding Hidden Threats by Analyzing Unusual Behavior*, Rand Corporation, Santa Monica, CA, USA.

Huebner, E., Bem, D., and Wee, C. (2006)“Data Hiding in the NTFS File System”,*Digital Investigation*, Vol. 3, pp. 211-226.

Jahankhani, H., and Beqiri, E. (2010) “Digital evidence manipulation using anti-forensic tools and techniques”, Chapter 2 in *Handbook of Electronic Security and Digital Forensics*, World Scientific, Singapore, pp. 411-425.

Naiqi, L., Zhongshan, W., and Yujie, H. (2008)“QuiKe: Computer Forensics Research and Implementation Based on NTFS File System”, in Proc. Intl. Colloquium on Computing, Communication, Control, and Management, Guangzhou, China, August, pp. 519-523.

O'Neill, P. (2010)*Verification in an Age of Uncertainty: The Future of Arms Control Compliance,* Oxford University Press, New York.

Pearson, S. (2010)*Digital Triage Forensics: Processing the Digital Crime Scene,* Syngress, New York.

Rowe, N., and Garfinkel, S. (2011)“Finding Anomalous and Suspicious Files from Directory Metadata on a Large Corpus”,3rd International ICST Conference on Digital Forensics and Cyber Crime, Dublin, Ireland, October.

Rowe, N., and Garfinkel, S. (2010) “Global Analysis of Disk File Times”,Fifth International Workshop on Systematic Approaches to Digital Forensic Engineering, Oakland CA, USA, May.

The Comprehensive Approach as a Strategic Design to run the Military-Industrial Complex in Operations

Mirva Salminen[1] and Aki-Mauri Huhtinen[2]
[1]Department of Social Sciences, University of Lapland, Finland
[2]Department of Leadership and Military Pedagogy, National Defence University, Finland
msalmine@ulapland.fi
aki.huhtinen@mil.fi

Abstract: How to steer the 21st century military actions that are concurrent both in the real world and in virtual networks? The 20th century saw immense struggles between nation states that contained a level of unprecedented material and human destruction. After the great wars, the Cold War and nuclear weapons paralysed the application of military plans in the organisation. The new wars and global threats that emerged towards the change of the millennium required the re-thinking of military organisation, planning and conduct. Simultaneously, the information revolution penetrated the battle space. These developments have lead to an increasingly complex security space, in which the *military-industrial complex* influences both virtual and material aspects of warfare, politics and economy. The 21st century introduced the general public with so far undetected actors in conflict zones, that is, with private military and security contractors (PMSCs). The policy of outsourcing warfare and security related functions that were seen primarily, at times even solely, as state functions raised heated discussion; especially, when an increasing number of scandals related to PMSCs' conduct as well as to the governmental contracting practices was discovered. Despite the strong rhetorical opposition to PMSCs, they have become codified as a steady part of the military-industrial complex. This paper scrutinises the basic principle and key concepts of the new western politico-strategic level military planning model called the *Comprehensive Approach (CA)*. It is a Wikileaks type of open door policy: everyone operating in the same real space can participate in the shared virtual planning space in order to fill in the comprehensive picture of the parallax and the narrative gap. The model's usability in a new military atmosphere in which private contractors operate alongside public soldiers and attend the planning is under scrutiny. Challenges that the open planning creates to the military organisation are highlighted.

Keywords: comprehensive approach, strategic planning, military-industrial complex, private military and security contractors

1. Introduction

In his farewell address, given in 1961, President Dwight D. Eisenhower famously warned about the rise of what he named as the military-industrial complex. Due to the great wars, the United States had "[....] been compelled to create a permanent armaments industry of vast proportions." However, "[i]n the councils of government, we must guard against the acquisition of unwarranted influence, whether sought or unsought, by the military-industrial complex. The potential for the disastrous rise of misplaced power exists and will persist. We must never let the weight of this combination endanger our liberties or democratic processes". (Eisenhower 1961.)

Were the President's warnings heard or not, the military-industrial complex was never without influence during the Cold War (in this paper, understood as a commonly shared cognitive constellation of international relations rather than as a particular period of time [Patomäki 2011]). Even if no major conventional war was fought between the main ideological opponents, the military-industrial complex did not find itself redundant. Preparedness for war in the western world was at an unprecedented level and the market for standardised, industrially produced military products reached a global scale. Towards the end of the Cold War, and especially after it, the complex has gradually reformed itself and expanded into new sectors. The traditional production sector has been supplemented with what can be categorised as the service sector; the sectors have enmeshed, allied with different sectors in the civil society and hence created new relationships of power. Virtualisation of the social space has played an important facilitating role in this process by creating new potential for service development and by increasing the speed at which information is globally disseminated. Therefore, following James Der Derian's (2001) theorising, it might be more appropriate to discuss the military-industrial-media-entertainment network than mere military-industrial complex. However, even if this paper acknowledges the importance of public media and entertainment aspects, it concentrates on discussing the role of PMSCs in military planning and organisation.

The mainstream historical narrative told about PMSCs states that the reformation of the military-industrial complex has taken place within the possibility conditions created during the Cold War and at its end. The large supply of surplus armament and skilled personnel to the global market combined with an increased demand for capabilities that had been run down established favourable conditions for the enlargement of the complex (Avant 2005, 30–38; Singer 2003, 49–55). Nevertheless, this is only part of the reasoning and cannot explain the entire phenomenon (Rosén 2008). At the same time, the information revolution changed the possibility conditions of almost every aspect of life – including the spheres of security and warfare. Thus currently, it is next to impossible to give estimations about the scale or influence of the reformed (and globally linked) military-industrial complex. What can be said, however, is that PMSCs have become an important component of military and security operations that are executed in the same manner as they are viewed: on real-time – actually and virtually (Der Derian 2001).

With regard to military planning, historically, states with strong economic and military power base have created the way of war for fighting the next one. According to Raitasalo and Sipilä (2006), the Cold War conceptualisations of war were based largely on traditional Clausewitzian definitions. Political and/or military power was to be used for compelling the enemy to consent to our will. After the Cold War, the key conceptualisation in the US was “the Revolution in Military Affairs”, which based on unidealistic or nonnormative way of war. Instead, it used war as a test framework for the western technological revolution. Recently, the US has followed and contributed to the general development of the art of war through her doctrines and war experiences (for example, the War in Vietnam and at the Persian Gulf) as well as by emphasising the new threats; by defining terrorism/extremism, restricted or denied access to the global commons, and proliferation of weapons of mass destruction as the main military threats (the National Security Strategy [NSS] 2010; the National Military Strategy [NMS] 2011); and, especially, through the diversified technological development.

The worldwide role of the US armed forces has been emphasised in the post-9/11 world, in which the World Trade Centre and global economy became, in a strong symbolic manner, tied to war and warfare. War no longer seemed to be about international politics and trade; instead, it became perceived as a struggle in the omnipresent networks which nobody was able to escape. The different networks have enabled multiple actors – from states to non-governmental organisations (NGOs), from individuals to mafia – to organise and cooperate. Hence, the malevolent side of the networks has been defined as a threat to the economic stability and wellbeing in the US (which according to both NSS 2010 and NMS 2011 is the basis of the US security). The existence and operation of these networks have compelled the US to extend the doctrine of pre-emptive strike to the so-called conventional sphere of warfare. The (potential) execution of the doctrine requires a comprehensive planning model and it is believed necessary in order to safeguard both homeland security and foreign trade. The globalisation of economy and the multiplication of information networks have also enabled and increased the appeal of cyber-attacks. Encountering these kinds of problems demands comprehensive planning.

In the same vein, demands for 24/7 situational awareness, too low levels of potentially needed military capabilities, and changes in the principles of military planning have required the development of a more comprehensive approach; not only to crisis management, but to the overall provision of security. This has called for coordinating the benevolent side of the networks; not only to carry out a whole-of-government effort, but also to harness the multiple semi-governmental and private actors to serve the common purpose. For example, in order to obtain air lift capability when needed the armed forces has turned towards the US aircraft industry – which has tackled with its own problems after the 9/11 – and contributed to its protection and preservation. Science and research in general (for example, in the fields of space and nanotechnology) have become strongly intertwined with defence industries (including the development of robotics).It is currently believed that the goals and practices of various actors can be accommodated in the CA so that the maximum efficiency of the overall effort can be exploited.

Next, the study’s design will be introduced. After establishing the theoretical framework, the policy of privatising and outsourcing the state’s warfare and security related functions will be discussed further. The paper’s main and conclusive chapters will concentrate on the inclusion of PMSCs into the CA based planning.

2. Research design

As noted, the paper examines the CA's usability in the military atmosphere in which private contractors operate alongside public soldiers and attend the military planning. It comments on the privatisation and outsourcing of the state's warfare and security related service functions in the 21st century and lessons (potentially) learnt from the experiences in conflict zones. The CA is seen as a holistic attempt to tie the contractors, firstly, into the military planning and conduct and, secondly, into the political and diplomatic efforts as well as into the chains of responsibility and accountability that are perceived as preconditions to a functioning democracy. The Approach can be seen as an attempt to shake the power position that the industry has occupied, and to balance the powers within the military-industrial complex as well as between the state administration and the complex. In the process, outsourcing of the state's warfare and security functions has become a better controlled policy (nevertheless, a lot remains to be done as, for example, the Government Accountability Office's reports have reminded throughout the 21st century). On the other hand, the inclusion has normalised PMSCs as actors in contemporary (and future) conflict zones and is thus seeking wider legitimacy to the complex.

The paper is a theoretical examination of the opportunities and challenges that the inclusion of PMSCs in military planning creates to the armed forces. It is based on three disciplinary traditions and comments on the relevant aspects of each tradition. Firstly, following the President Eisenhower's line of argument, the strengthening of the military-industrial complex is a political concern. However, at the peak of outsourcing and privatisation, the issue was not approached so much as a political concern than as a matter of economic efficiency and business logic. Majority of the state's warfare and security related functions were then evaluated as subject to outsourcing and thus, a new potential field for the creation of economic activity in the private sector. The CA, finally, brings the contractors into the field of military planning. It focuses on the inclusion of a variety of actors in the planning, and addresses the issue of PMSCs in a way that crosses administrational boundaries and disciplinary divisions. The question addressed in the paper is the connection between the development of the CA and of the military-industrial complex. We argue that there are parallel development trends and inter-dependence.

3. Privatisation and outsourcing of the state's warfare and security related functions

"Private military and security contractors" has become a binge category into which all companies, NGOs and other entities (to whom the state's warfare and security related functions in the international context have been contracted to) have been thrown. Differentiating between actors or their fields of service provision is without a purpose for this paper. Some examples of the functions that PMSCs take care of are training and consultation; interpretation and translation; development, maintenance and operation of weapons systems; intelligence and surveillance; and logistics (Avant 2005, 16–22; Singer 2003, 88–100). What the contractors have in common, however, is that their relationship to the US state administration has been formed through contract. The contract holds the service providers as outsiders in relation to the administration (Resteigne & Soeters 2009, 309–311), which in the military sphere has meant that the contractors have operated outside the command and control structure of the armed forces. The contract has also served as the main tool of control for the administration.

During the main military operation in Iraq, it was estimated that there was a contractor working in the battle space for every ten service members (Avant 2005, 1). The Department of Defence (DoD) began to keep records on its contractors in 2007. The first USCENTCOM Quarterly Contractor Census Report to be found on DoD's Program Support website is dated for August 2008 and it declares 242,558 PMSC employees working in its area of operations. In October 2011, the same report declared 175,045 employees, which represented approximately a half of the total contribution in the area. Contractors have thus become an important part of the US contingency operations; moreover, it has been claimed that the state is dependent on them. (It needs to be kept in mind that all figures presented are mere estimates and do not include contractors working for other US state agencies, international organisations, NGOs, media and so forth.) This kind of intrusion of company interests into the core of the state, that is, into the realm of warfare and national security decision making has raised many concerns. One of the main ones has been that the state in itself is no longer able to command and control all entities using violent means in its name.

PMSCs and their employees do not fall under the same legislation and regulation as the armed forces and soldiers. Contractors working for different state agencies have had different operational rules and guidelines, and there has been confusion about the legislation applicable to them. In addition, when, for example, security services have been sub-contracted, the main contract between the state and the prime contractor has not directly established guidelines or a chain of responsibility for the sub-contractor. Next to that, PMSC employees may, in principle, choose whether or not to undertake the task assigned to them. This creates a factor of uncertainty in the operation. PMSC employees have not automatically been entitled to military protection (but may have received this) or to the status of legal combatant; neither have they had an obligation to inform the armed forces about their operations, which has occasionally caused confusion and even violent confrontations. The aforementioned factors, in addition to the negative impact that the PMSC-related scandals have had on the US foreign policy as well as on the potentiality to achieve the political and military goals, have given an incentive to enhance coordination and to include contractors better in the military planning.

Operations of the armed security companies have often been perceived as counterproductive to the overall military and security effort due to lacking cultural awareness, exaggerated use of violent means, neglect towards the interests and aspirations of local population, wasteful use of resources and fraud, insufficient local employee background checks which facilitates infiltration, and unwillingness and inability to share information with the armed forces. The state agencies using contractor services, on the other hand, have been blamed for not having planned, monitored and supervised the contractor work sufficiently, for not having enough skilled personnel, resources and the right attitude to cut off the contractor wrongdoings, and for not having been able to subject the contractors to adequate regulation. This, again, has deepened the legitimacy crisis that the state has already been seen to be in. Eventually, plans for (re)in-sourcing some of the outsourced functions has been made (Gansler 2010), and research agencies have been asked to produce reports on which of the functions are to be evaluated as inherently governmental, that is, as unsuitable to outsourcing (see CRS R41209; R40641; and GAO-11-192). (Re)In-sourcing and cross-state agency agreements have been steps towards a better-thought, structural inclusion of contractors in the military planning and organisation. (It is not suggested that PMSCs would not have been included in the planning before. However, this has usually been done on an ad hoc basis.)

The centralisation of contractor activities under the military command has confronted opposition from different sides. The strongest opponents seem to have been the US state agencies other than DoD, who also use PMSCs' services. The agencies have been worried about the potential militarisation of, for example, foreign policy and foreign aid, as well as about the loss of control over their own policies and practices in conflict zones. The serving military personnel, on the other hand, have opposed the contractor penetration, because they do not wish to be associated with PMSC employees who have the reputation of gun-carrying cowboys; because they doubt the trustworthiness of contractors; and because they question the morality and ethics of PMSC personnel. (Salminen 2010.) In addition, the centralisation of contractor oversight is likely to increase the demand for personnel as well as to change the power relations within both DoD and the armed forces as well as between the different state agencies. Nevertheless, intense demands for the centralisation have also been presented. It is believed that this would enable the establishment of more comprehensive situational awareness and give some control to the armed forces over PMSC operations. The inclusion of PMSCs in military planning has also been presumed to improve the contractor oversight and accountability. Enhanced oversight would establish a clearer chain of responsibility and point out an administrative body that is responsible for the contractor conduct. In this way, the broken chain of democratic control could be re-established. The CA is ought to be the planning framework into which PMSCs for the aforementioned reasons could and should be accommodated.

4. The comprehensive approach

In general, there is no unanimous definition of a comprehensive approach (Johnson 2010, 8). It has been identified, for example, as a philosophy, a network of actors, a means to an end (conflict resolution) or a network of key drivers. Moreover, it has been designated as the model for, for example, crisis management, counter-insurgency or prevention of conflicts. In this paper, it is understood as a planning model used before, during and after conflict operations. In addition, this paper comments only on the US/NATO planning model (which emphasises the role of the armed forces as the leading actor in operations) and leaves the comprehensive approaches of, for example, the UN, the EU or other national governments aside.

The essential foundation of any comprehensive approach is the recognition of a need to combine the efforts of multiple actors (both civilian and military) operating in the contemporary, complex conflict environments in order to produce a coherent and effective response to the situation (Johnson 2009, 8–9). This combining is ought to take place on several, overlapping levels, for example, at the inter-organisation, intra-organisation and national levels or at the politico-strategic, military-strategic, operational and tactical levels (Rintakoski & Autti 2008, 25; Simon & Duzenli 2009). In this way, the comprehensive approach could be viewed as an ideal to strive for. Reaching this ideal would require the successful coordination of the conduct of numerous actors (for example*, PMSCs*) that operate with different mandates (*assigned in the contract*) and value bases (*vary from patriotism to pure profit motivation*), resources (*usually well financed with plentiful material resources or open channels for obtaining what is required*) and ideologies (*vary from the desire to turn the world into a better place to war profiteering*), and for different purposes and aims (*usually to fulfil the contract requirements*).

Trust building through dialogue is perceived as crucial starting point in the approach; especially, when different actors may not share the view of the desired end state and may not necessarily desire to cooperate with one another (as the discussion briefed in the previous chapter showed) (Molnar & Smith-Windsor 2008, 1–2). Information technology, again, is believed to be an important enabling tool in the creation of a network based structure that the successful application of the CA requires. PMSCs often serve as the developers and maintainers of these information technology based solutions as well.

In the planning of the use of force, the CA is executed by the coordination of the efforts of diplomacy; production, dissemination and control of information; economy; and military. If the structure of the armed force was previously based on concrete threat models (the military organisation was symmetric in relation to the potential adversary), the current comprehensiveness based organisation is built on operational performance requirements. In other words, if civil crisis management capacity is required, it has to be possessed; if air fuelling capacity is demanded, it needs to be obtainable (from either public or private sector). The organisation is no longer symmetrical in relation to one’s potential opponent, but in relation to politically and economically defined readiness. However, it is not only the armed forces that has to modify its structure in order to meet the current demands, but “for a truly comprehensive approach to emerge, each part of each [attending] government needs to transform [...] (Rotmann 2010, 4)”. For PMSCs this could mean, for example, opening up their planning and information bases which thus far have been perceived as business secrets and assets that guarantee their competitive edge on the market. NATO’s organisational form, again, seems to be converging closer to the form of a multi-national corporation.

The CA thinking thus moves actors from the “need-to-know” to “need-to-share” world. In the approach, information must be made accessible to all relevant actors by all relevant actors. Information serves as a basis for knowledge development, which is the key to understanding of the complex engagement space and the actors’ roles and capabilities in it. (Engagement space can be defined as the space in which NATO decides to engage and in which the interaction of different actors creates conditions that may be perceived as acceptable or unacceptable in terms of the end state which NATO desires to reach [Simon & Duzenli 2009, 16].) Actors attending the knowledge development (including PMSCs) are to explicate the frame of intentionality, in other words, to analyse the factors and relations that influence the situation. The matter to be taken into consideration is how one’s frame of intentionality condition his/her perception, choices and actions in the situation. The knowledge development is to take place in the domains of politics, military, economy, social, infrastructure and information (Simon & Duzenli 2009, 16).

In the CA thinking, actors in the engagement space are to be seen as systems; and as parts of bigger systems. Systems analysis (which is closely connected to resource needs [Johnson 2010]) is used as a methodology for developing situational knowledge, which, again, is to serve as the basis for developing effects based solutions for turning the unacceptable conditions into acceptable ones. First, the current situation is to be described. Then, the desired kind of situation is to be envisioned. Finally, actions that are believed to produce effects that transform the current situation towards the desired kind are explicated. (Simon & Duzenli 2009, 16–17.) This is how the CA process is envisioned to work in practice. Its successful application requires continuous assessment of the previous, current and even planned actions. The assessment’s core questions are: Are operations progressing according to the plan, and are they having the anticipated kind of outcomes? If not, what needs to be adjusted? Traditionally, these kinds of questions have been addressed by reporting through the chain of command. Moreover, reporting has often been qualitative in nature and dependent on subordinate commanders' skills and experience. Yet, in NATO and, especially in the US, the culture is that of quantitative measures. Difficulties have arisen in the linking of the abundance of data produced by evidence-based quantitative methods with the politico-strategic goals. The question is to find a balance between experience-based (leadership) and evidence-based (management) methodologies. A proper command system should be able to set goals for itself and strive for reaching those goals – in spite of realising that some things will go wrong, and with confidence that, when they do go wrong, the system will be able to overcome the obstacles (van Creveld 2003, 194–195).

Within the CA framework, the underregulated conduct of PMSCs and the effects it has caused in conflict zones may be perceived as an unacceptable condition. The Approach itself, again, could be seen as the design envisioned turning the condition into an acceptable one. By including PMSCs better in the military planning and organisation, the US could involve them in the efforts to reach the overall political and military goals and hence, make PMSCs more accountable for their conduct. This conclusion is to be drawn in the assessment of how the outsourcing and privatisation policies have materialised in the 21st century conflict zones. Unfortunately, the assessment was not to take place within the organisation before the public and social media highlighted the apparent problems.

5. Conclusion

Are contemporary warfare and the essence of military organisation more a question of justifying warfare as one of society's functions or of the impact of warfare and military organisation on society? The aim of good government is to guarantee the functions of the so-called invisible power, so that the masses under control feel "less controlled". Involving and empowering citizens as well as increasing their interaction with the administration are leading trends in the information age. This point of departure needs new models for planning, such as the Comprehensive Approach. In the CA, control is based on the interaction between actors in systemic networks. The CA strives for supporting the actors' self-government when they participate in the common process. Everyone can have his/her voice heard without having to follow an unfamiliar process. This does not mean, however, that everyone's voice will be listened equally.

If the 20th century idea was to turn the battle into an industrial type of operation (which was impossible due to the lack of suitable communication system), the aim of the 21st century command is based more on General Bernard Montgomery's idea of a "Phantom" system of liaison officers. These officers used to visit the battle space by car or plane and report directly back to the headquarters; today computers and social media are in use (van Creveld 2003, 191).

In this paper, we have claimed that the inclusion of PMSCs in military planning is not an easy exercise. Friction arises for different reasons. Firstly, PMSCs may or may not share the military's rules, guidelines, goals and concepts, thus, making coordination more complicated. Next to that, the practice of contracting has augmented the demand for specialists and legal advisors, when the contract has turned commands into matters of bargaining. This changes the relative power positions inside the armed forces and more widely within the state administration – mainly, but not restricted to, in the fields of defence and security. In addition, economisation of warfare and problems in consolidating economic logic and good governance are reflected in the planning. Outsourcing has enlarged the space of political calculations in military planning by enabling intervention without prior public legitimisation (this has kept the decisions to engage outside the political struggles by turning them into administrative decisions), but also by drawing attention to ethical and politically sustainable conduct. The current development is likely to lead to a greater number of public-private partnerships, in which public and private can hardly be differentiated from one another. In this context, the US armed forces is likely to recede from carrying out tasks that do not fall under its core functions (unlike it has done, for example, in Afghanistan), and the question of whether PMSCs should be managed as insiders or outsiders becomes less reasonable (albeit it is likely to colour the intra- and inter-administrative struggles for the time being).

References

Avant, D.D. (2005) The Market for Force. The Consequences of Privatizing Security, Cambridge University Press, Cambridge.

van Creveld, M. (2003) Command in War, Harvard University Press, London.

Der Derian, J. (2001) Virtuous War: Mapping the Military-Industrial-Media-Entertainment Network, Westview Press, Boulder (CO).

Eisenhower, D.D. (1961) The Farewell Address, http://www.americanrhetoric.com/speeches/dwightdeisenhowerfarewell.html, accessed 29.12.2011.

Gansler, J.S. (2010) “The Dangers of Over Insourcing” PeaceOps.com – Journal of International Peace Operations, Vol. 6, No. 3, http://web.peaceops.com/archives/1058, accessed 30.12.2011.

Hill, C. (2010) Grand Strategies. Literature, Statecraft, and World Order, Yale University Press, London.

Johnson, T.F. (2010) “The Comprehensive Approach and the Death of the Term ‘EBAO’” The Three Swords Magazine, Issue No. 17, pp 9–13, http://www.jwc.nato.int/files/17_10_Magazine.pdf, accessed 11.1.2012.

Molnar, F. and Smith-Windsor, B. (2009) 10 Things You Should Know About a Comprehensive Approach. NATO Defence College, Research Division, Seminar Report, http://natolibguides.info/nato-eu, accessed 9.1.2012.

Patomäki, H. (2011) “On the Complexities of Time and Temporality: Implications for World History and Global Futures” Australian Journal of Politics and History, Vol. 57, No. 3, pp 339–352.

Raitasalo, J. and Sipilä, J. (2006) “Reconstructing war after the Cold War”. In Jeppsson, T. and Mikkola, E. (eds.) Perspectives on the Evolving Nature of Military Power, Finnish National Defence University, Department of Strategic and Defence Studies, Helsinki. Series 2, Research Reports No. 36, pp 1–24.

Resteigne, D. and Soeters, J. (2009) “Managing Militarily” Armed Forces & Society, Vol. 36, No. 2, pp 307–332.

Rintakoski, K. and Autti, M. (eds.) (2008) Comprehensive Approach. Trends, Challenges and Possibilities for Cooperation in Crisis Prevention and Management. Seminar Publication. Crisis Management Initiative, Helsinki.

Rosén, F. (2008) “Commercial Security: Conditions of Growth” Security Dialogue, Vol. 39, No. 1, pp 77–97.

Rotmann, P. (2010) Build on Shaky Ground: the Comprehensive Approach in Practice. NATO Defence College, Research Division, Research Paper No. 63, http://natolibguides.info/nato-eu, accessed 9.1.2012.

Salminen, M. (2010) Struggle over outsourcing of the security functions of the state: The case of September 16, 2007 shooting in Baghdad, University of Tampere, Tampere. M.Soc.Sc. thesis.

Simon, G. and Duzenli, M. (2009) “The Comprehensive Operations Planning Directive” NRDC-ITA Magazine, Issue No. 14, pp 16–19, http://www.nato.int/nrdc-it/magazine/home.html, accessed 12.1.2012.

Singer, P.W. (2003) Corporate Warriors. The Rise of the Privatized Military Industry. Cornell University Press, Ithaca N.Y.

The United States National Security Strategy 2010

The United States National Military Strategy 2011

USCENTCOM Quarterly Contractor Census Report for the 3rd quarter of FY2008 and for the 4th quarter of FY2011, http://www.acq.osd.mil/log/PS/CENTCOM_reports.html, accessed 30.12.2011.

User-Side Password Authentication: A Study

Libor Sarga and Roman Jašek
Tomas Bata University in Zlín, Zlín, Czech Republic
sarga@fame.utb.cz
jasek@fai.utb.cz

Abstract: Researchers have for a time been struggling to change inert mindset of users regarding passwords as a response to advances in processing power, emergence of highly-scalable computing models, and attackers prioritizing human element for attacks. Recommendations regarding security are ignored as documented by recent corporate database breaches and releases of unencrypted password caches which corroborated lacking security awareness in vast majority of Internet users. In order to educate users about computer security, terms such as hashing, cipher systems and their weaknesses, brute-force attacks, social engineering, multi-factor authentication, and balance between usability and ease of use must be clearly explained. However, academia tend to focus on areas requiring deep mathematical or programmatic background, clear communication of these security elements while minimizing scientific rigor thus remains challenging. The article aims to provide a concise, comprehensive research overview and outline of authentication, including information entropy, hashing algorithms, reverse password engineering, importance of complexity and length in passwords, general-purpose attacks such as brute-force and social engineering as well as specialized ones, namely side-channel interception. Novel ways of increasing security by utilizing two- and multi-factor authentication, visual passwords, pass phrases, mnemonic-based strings will be considered as well along with their advantages over the traditional textual password model and pitfalls for their widespread propagation. In particular, we hypothesize that technological developments allow vendors to offer solutions which limit unauthorized third parties from gaining windows of opportunity to exploit weaknesses in the authentication schemes. However, as infrastructure becomes more resilient, attackers shift their focus towards human-based attacks (social engineering, social networking). Due to largely unchanging short-term behavior patterns, institutions need to lecture employees over extended periods about being vigilant to leaks of procedural and organizational information which may help attackers bypass perimeter-level security measures. We conclude the article by listing emerging threats in the field, specifically social networks-distributed malware and mobile devices targeting.

Keywords: authentication, security, hash, password, mnemonic, visual, multi, factor, social, brute-force, attack, engineering, passphrase, side-channel

1. Introduction

Authentication techniques were devised to prove one’s identity using exclusively-possessed identifiers. Passwords were used in fortified cities where the sole way of entering was to procure a word/phrase to “pass” the gate-keeping authority. Security management and engineering had to be taken into consideration even then: how to distribute passwords securely in case of a change while avoiding leaks during physical transport (robbers etc.)? How often is it viable to change passwords to balance security and costs involved with “user” comfort? How to treat forgotten or deprecated passwords? Maintaining a system secure while relying on human element’s discretion and rationality has remained a challenge up to now.

Dealing with consequences of irresponsible users’ behavior is much more pronounced, however, that it has been in the past. Password leaks involve serious repercussions as global society steadily shifts towards intangible virtualized cyber system in which personal identity gives way to electronic presence.

The article provides high-level overview of electronic authentication principles, schemes and vectors the attackers utilize to illicit confidential data. Social engineering, side-channel attacks along with multi-factor authentication will be discussed, too. It is structured as follows: the second part discusses underlying principles of textual passwords. The third part provides overview of alternatives to the traditional password authentication model. The fourth part comments on two types of attacks aimed at obtaining users’ passwords via unconventional means: social engineering and side-channel interception. Conclusion summarizes the article, briefly hypothesizing about the future of authentication.

2. Password authentication principles

Recent cases of security breaches prove that any authentication system has to meet certain requirements. Passwords must be able to withstand a worst-case scenario – offline cracking where attacker gains access to a list of strings obfuscated by a cryptographic function, and proceeds to systematically enumerate all probable and possible combinations. Moreover, sensitive data must be stored in separately encrypted databases from any identifying tokens to minimize exposure in case of a perimeter security breach. The requirements are, unfortunately, not enforced despite legislative attempts to bolster server-side security when dealing with storage and handling of sensitive data, such as Payment Card Industry Data Security Standard (PCI DSS) initiative.

Algorithms called one-way cryptographic hash functions, or hashes, are utilized to store such data. A hash is a "...function, mathematical or otherwise, that takes a variable-length input string (called a pre-image) and converts it to a fixed-length (generally smaller) output string (called a hash value)" (Schneier 1996). Two characteristics of a hash are: different pre-images produce different outputs; and the hash itself is theoretically irreversible as it is both resource- and time-intensive to obtain the pre-image from the hash. A bit difference in the pre-image changes at least 50 % of output bits (i.e., avalanche effect). Additionally, modern cipher systems employ multiple rounds, the effect proliferating through every iteration which makes predicting effects of the initial change on the hash intractable. When wishing to impersonate the user using their credentials the attacker may resort to hash collision, generating identical hash output from a different pre-image via exhaustive search, thus forging his own valid password. Since hashes are of fixed lengths, collisions are bound to exist but computational requirements to find them were initially deemed sufficient.

Suppose the attacker obtains database which utilizes SHA-1 and MD5 cipher systems. Many vendors support these solutions by default, database administrators are therefore more likely to use them. Both SHA-1 and MD5 have in recent years been proven unsecure.

SHA-1, published by the United States National Security Agency (NSA) is a 160-bit digest function iterated for 80 rounds (Eastlake 2001). The cipher was broken when a researcher shown it is possible to produce a custom hash colliding with any hash to be attacked, thus bypassing user authentication with a forged pre-image (Manuel 2008). The SHA-1 standard has been since updated to SHA-2, SHA-3 is tentatively scheduled for release in 2012. However, due to all three cipher variants being based on identical algorithmic operations, optimized attack on SHA-1 may prove effective.

The MD5 Message-Digest Algorithm is a 128-bit, 4-rounds function proposed by Ronald Rivest (Rivest 1992). Widely distributed and used for generating electronic certificates, MD5 was found to be flawed in 2004. Five years later, Chinese researchers succeeded in finding a hash collision (Xie 2009). Furthermore, given a known hash a method was devised to extract the original pre-image (Sasaki 2009). Validating rogue electronic certificates using hash collisions was also demonstrated, allowing attacker to intercept encrypted communication (Sotirov et al 2008).

Publicly accessible databases allow users to enter an MD5 hash and get a plaintext pre-image, extracting commonly-used passwords in a matter of seconds. While the services are limited in maximum number of queries per unit of time, offline cracking tools provide no such safeguards. Commercially available parallelized GPGPUs (General Purpose computation on Graphics Processing Units) and FPGAs (Field-Programmable Gate Arrays) greatly reduce time required to obtain pre-images in bulk. Due to architecture variations between CPU (Central Processing Units) and GPGPU/FPGA, substantial differences have been observed in their performance (Bakker & van der Jagt 2011).

The often-mentioned password requirement, sufficient length, is tied to the concept of information (alternatively Shannon) entropy (Shannon 1948). Claude E. Shannon introduced the measure to quantify expected value of information in bits contained within a message (password). Entropy represents uncertainty about the message content based on previously known (intercepted, decoded) parts. Suppose the user chooses her password to be 20 characters long, consisting of a single repeating symbol. If the password's encrypted version will be analyzed, the interloper doesn't have any prior information or assumptions regarding its structure or length. If brute-force attack is utilized, he will have to try every combination (lowercase/uppercase letters, numbers, special symbols, non-Latin characters) from an arbitrary value up to twenty to succeed. The password thus ensures high level of security.

Suppose, however, the attacker still doesn't have information as to its strength but is aware of all correctly guessed characters. When the password's first character is decoded it is insufficient to determine whether it will occur again. After decoding the second symbol, probability that the third character will be the same increases. Getting 10 identical symbols, there is high probability the next position will be occupied by the same character, too. As the attacker reinforces his belief the password consists of a predictable pattern, he is able to adjust search algorithms accordingly. Shannon entropy envelopes the second situation, considering the message is discovered bitwise, not as a whole entity. This is advantageous for users as resources needed for reverse password discovery in the first case are comparatively higher.

This was corroborated by a research stating that "[d]ue to the all-or-nothing nature of guessing a password, it is meaningless to calculate the entropy of a password based on its composing characters. Without the knowledge of the characteristics of the password, the only assumption we can make, apart from trying our luck with a big dictionary, is that every character has the same probability" (Ma 2010). Shannon entropy holds true only for stream-processed messages with applicability in the fields of error-correcting codes, data compression, and recovery.

It has been also stated that entropy is not a suitable metric of password strength as "…passwords of the same length have the same amount of entropy. As the length becomes longer, the entropy of the password goes up, and the average entropy per character stays the same" (Ma 2010). NIST (National Institute of Standards and Technology) provides comprehensive overview of how entropy changes based on input character space and length variations (Burr et al 2011).

User has 93 printable characters at their disposal when creating passwords: 26 lowercase letters, 26 uppercase letters, 10 digits, and 31 special characters. There are drawbacks to using the full input set, though, namely inability of human brain to memorize complex long strings and application-side restrictions disallowing some symbols when creating passwords. Due to language's natural tendency to group letters into strings (words) of predictable composition, frequency analysis was devised providing a link between cryptanalysis and linguistics. Research on the subject exists with a classical treatise "Codes & Secret Writing" in which the most common letters and double letter pairs in English were analyzed (Zim 1978). These properties help reduce search space during exhaustive searches when dealing with a known, English-speaking subset of users. Markov chains, a lattice containing a finite number of states and probabilities associated with transitions between any two of them, in particular benefit from frequency tables as they decrease occurrences of less probable doublets such as "qz" while increasing occurrences of "th", "he" and others found commonly in English words. Probabilistic password cracking automates the process further: a computer generates matches based on Markov chaining rules (Weir 2009). Academic coverage of the topic remains scarce.

Sufficient password length reduces effectiveness of brute-force attacks by increasing time factor involved by incorporating wider range of characters which broaden the search space. Choosing strings not included in word lists to prevent attacks in which "…we assume the attacker has already built a database of possible passwords, called a dictionary," (Chakrabarti & Singhal 2007) is strongly advised. Techniques to thwart password-guessing attempts (substitution, padding, prepending, appending characters) operate on assumption that automated tools are unable to produce word permutations. Yet if Markov chains are part of the search routines, every dictionary word may be converted in this way. Such recommendations on passwords, most of which are considered more 'secure' than basic textual inputs, only slow down the attempts with marginal performance degradation. Additionally, complex character sequences are regularly being written on paper in plaintext format for better memorability. Frequent corporate password policy changes thus have adverse effect of lowering security while opening additional attack vector to be exploited.

Mitigating server-side procedure involve per-user generated salts, (pseudo)random strings concatenated with plaintext passwords which are subsequently hashed. If the salt is unique and sufficiently long, brute-force attack is effectively precluded by rapid expansion of the search space.

3. Alternatives to textual passwords

To address the shortcomings of textual input alone, schemes have been proposed to substitute or complement the 'traditional' password approach. While several increases security with little impact to overall complexity, two-factor authentication (2FA) is currently the only technology seeing its adoption rising due to decreasing costs of producing tokens or harnessing existing infrastructures without additional costs incurred. Support from major hardware and software vendors together with ease of use for end-users are a factor, too.

3.1 Multi-factor authentication (MFA)

MFA adds additional layer of security by requiring near-simultaneous input from independent sources. A conventionally implemented measure is to use unique OTPs (One Time Passwords) sent to

a mobile device, a 2FA approach based on "something the user knows, i.e., a static password, and something the user has, i.e., an OTP" (Eldefrawy et al 2011). Cellular phones capable of displaying SMS (Short Message Service) were chosen primarily due to their availability and wide penetration rates. Tokens generating (pseudo)random strings came into prominence recently, too. A proposal has been also made to gradually phase in smart phones as an automated authentication vector, sidelining textual passwords in favor of exchanging identifiers between devices wirelessly (Umezawa et al 2011). Despite substituting human element, it is nevertheless prone to passive reconnaissance where attacker intercepts the signal as a man-in-the-middle. Suitable encryption scheme is therefore necessary to be put in place.

The system requires inputting password in combination with an OTP sent to a mobile device or generated on a token, usually during limited timeframe. If the information are not entered in time the request may be repeated with a new OTP, after single/several attempts a lockdown protocol should be initiated. By typing the OTP string, user proves knowledge of the password as well as having access to a device which the OTP was sent to or on which the OTP was generated.

MFA expands the 2FA federation process onto more than two separate layers. Biometrics is cited as being a promising development in conjunction with cryptographic primitives (Sarier 2010). Some challenges preclude broader commercial adoption, particularly technology and deployment costs, setting optimal fault thresholds, and pre-images' (irises, fingerprints, faces) encryption and storage according to legal statures on handling personally identifiable data. Non-compliance may result in financial penalties or lawsuits which for many corporations outweighs its benefits. Biometrics also faces the threat of permanent feature compromise. As pointed out, "[b]iometrics, unlike passwords or PINs, cannot be changed during the course of an individual's life" (Schreier & Boult 2009). The fact led to a proposal of revocable, biometric data-based tokens which may be invalidated and released anew in case a security breach occurs.

3.2 Passphrases, mnemonic phrases, visual passwords

Also proposed have been techniques which supplement and reinforce passwords to withstand brute-force and dictionary attacks. They exploit properties such as length, complexity, and inclusive, orders of magnitude harder to guess knowledge unique to every user.

Passphrases are rooted in Shannon entropy described above. First introduced in 1982, they suggest expanding search space the attacker must take into consideration when reverse engineering passwords (Porter 1982). As it is assumed only the password's digest with unknown length and complexity of the pre-image is available, a viable course of action would be to employ common words dictionary, probabilistic grammar rules, or brute-force attack. Passphrases consist of several compounded words, sometimes with respect to language punctuation rules and grammar. For instance, both "This is a passphrase" and "Thisisasshprase" may be used in place of a one-word password, either providing higher level of security. The user may generate such strings using dictionary words as the output is able to withstand dictionary attacks due to the fact extensive out-of-string permutations (combining multiple words) would slow the cracking down considerably, and are thus frequently omitted. The disadvantage is that the user is often limited in the password length allowed, i.e., 8 characters and no spaces which provides the attacker with a clue regarding search space's upper bound. Another, more serious downside is that users choose quotes from popular movies, lyrics, and books which are included in dictionary lists. In this case, the advantage of passphrases is diminished as increased complexity is irrelevant when dictionary phrases can reveal the pre-image.

Mnemonic passwords are closely related and based on human ability to "…use many mnemonic techniques to successfully memorize apparently random sequences" (Yan et al 2004). Instead of incorporating words into a passphrase, the user chooses a sentence and extracts characters from every word using an individualized scheme. In the example of "This is a passphrase", the user may choose to extract first letter of each word, substitute 'i' by '1' and spaces by dots. The result, "T.1.a.p" despite being too short demonstrates structure of a mnemonic-based string. Low probability of being included in dictionaries as well as nonexistent dedicated mnemonic phrases lists make them preferable to passphrases. Brute-force attacks have to operate on much wider search space than in case of simple passwords. The longer the source phrase, the longer the resulting string which coupled with substitution rules provide users with a tool disobeying perceived negative correlation between secure passwords and low memorability, i.e., it is possible to create high-entropy output while avoiding improper storage, such as writing on paper.

Visual (or graphic) passwords were introduced as "…a possible alternative to text-based schemes, motivated partially by the fact that humans can remember pictures better than text..." (Suo et al 2005). Instead of keyboard, graphic passwords utilize mouse interactions which hampers some reconnaissance agents (keyloggers). Dictionary and brute-force attacks are severely limited due to amount of combinations growing with the underlying graphics' complexity. Visual passwords are represented in different ways, either as a set of pictograms out of which users establishes a pattern, a picture/map into which user clicks to a single/multiple locations, a grid consisting of human faces valid order of which must be established, a color palette out of which colors must be selected in a specific way, or an interactive board onto which user reproduces predetermined sequence of strokes (signature, picture, text) to authenticate. Visual passwords are, nevertheless, vulnerable to shoulder surfing where the attacker observes the user during password inputting. As textual passwords are entered in form of stars or dots, shoulder surfing may at worst reveal password length, not its contents if additional keystroke-intercepting mechanism is not employed. No dedicated technology for graphic passwords is necessary except for storage capacities, in case of a high-resolution graphics several megabytes per user. Adoption momentum is not accelerating, though, due to general's public inertia and limited support. Fault thresholds, especially when selecting locations in a photograph/map, must be tweaked to minimize false negatives while keeping false positives on an acceptable level.

4. Social engineering, side-channel attack

Two types of attacks with which the attacker may choose to extract passwords from users without resolving to brute-force or dictionary techniques will be briefly mentioned. The first one focuses on the human element in the authentication process, considered to be the most likely point of failure; the second one on the inherent weaknesses of electronic devices. Attempts are also commonly made server-side to extract contents of databases via SQL injections, buffer overflows, directory traversals, null byte injections, uncontrolled format strings, and others. These will not be discussed in the paper.

4.1 Social engineering

Social engineering is defined as "…the act of manipulating a person to take an action that may or may not be in the 'target's' best interest. This may include obtaining information, gaining access, or getting the target to take certain action" (Hadnagy 2010). The technique stems from analysis of human behavior and interaction psychology. Social engineer approaches the target either indirectly by electronic means (email, instant messaging, social networks), or directly face-to-face. Social engineering is closely associated with phishing, vishing, baiting, tailgating, pretexting, and shoulder surfing, each exploiting particular technology, gaining physical access to facilities, inventing scenarios into which the victim is injected, or intercepting information by direct observation.

The process is divided into several stages, usually consisting of information reconnaissance (target's position in corporate hierarchy, names of colleagues, location, phone number) and subsequently establishing contact on the basis of offering a solution to a problem either perceived or real which may be instigated by the attacker himself. In the next stage, the victim is exposed to psychological duress under which she chooses to contact the social engineer. Trust model is thus established with the target providing sensitive data to get the problem fixed as soon as possible. The last stage involves social engineer making an exit without alerting the victim. In certain situations, she may be exploited over longer time.

Human psychology plays a vital role in success of social engineering with countermeasures dealing primarily with prevention and simulated or live trainings. Traits such as fear of authority, excitement, fear of job loss, laziness, ego, financial stimuli are inherent to any individual, ubiquitous in one's personality, relatively stable and modifiable only by long-term exposure. Forcing employees to change routine behavior patterns within a short timeframe may trigger psychological blocks, rejecting any further external information contradicting one's own opinions. Social engineering is considered "…the most dangerous form of security attack due to its nature. Because [it] utilizes the basic qualities of human nature such as trust, it is impossible to defend against it with just hardware or software alone" (Nyamsuren & Choi 2007).

4.2 Side-channel attack

Side-channel attack is a type of vulnerability found in every unprotected electronic system. Rather than attempting to extract information as bits, the system is analyzed by means other than direct interaction, such as power consumption fluctuations and radiation emissions which are subsequently interpreted as discrete states out of which source information may be reconstructed. First detailed in 1985 as proof-of-concept work focusing on CRT (Cathode Ray Tube) emanations, side-channel attack evolved into a series of techniques specifically designed to eavesdrop on sensors, keystrokes and other hardware (van Eck 1985). As LCD (Liquid Crystal Display) panels proliferated, it was demonstrated the vector continues to be relevant with the emissions readable at 10 meters distance (Kuhn 2004). Attacks against encryption systems were also proposed in the form of timing attacks during which the attacker observes time periods of individual cryptographic operations (hashing) and works backwards to determine the input data. Another technique is called cold boot and exploits inherent properties of materials out of which hardware components are produced, namely magnetic remanence. If the attacker has physical access to a running machine, he is able to make a copy of all decrypted plaintext passwords stored in memory via physical RAM (Random-Access Memory) module extraction and forensic analysis. No operating system contains a safeguard against such action, for example by unloading sensitive data off memory when not used as well as decrypting password ad hoc without storing them in RAM unless absolutely necessary (Halderman 2008). Keylogging may also be considered a side-channel attack as data are reconstructed ex-post instead of obtained in a batch from a database. Hardware keyloggers presuppose physical access to the machine while software alternatives run a hidden or obfuscated process on the host. It may periodically connect to external sources, exfiltrating keystrokes, screenshots, and statistics about the machine – running processes, software installed along with version numbers to enumerate existing vulnerabilities. Advent of smart phones brought about novel exploitation vectors and mobile side-channel attacks will continue to spread. A research demonstrated keylogging using iPhone solely through analyzing vibrations when the device is positioned next to a keyboard on which the user is typing the password (Marquardt 2011).

4.3 Conclusion

Electronic presence protection has become a serious issue as cyber criminals rely on lackluster approach to security. Social networks together with mobile devices will be targeted due to massive congregations of users, their inappropriate security habits, and widespread adoption rates. Polyfunctional malware strains will be deployed to extract identity data of users, including passwords, logins, banking details, list of visited sites along with associated credentials. Sophisticated measures are employed to detect and thwart such attempts but ultimate responsibility still resides with rational behavior of end-users both online and offline. While trusted computing may alleviate current hardware/software challenges, human psychology cannot be expected to change so rapidly. Training and education must hold pace with technological developments and prepare users for situations beyond their control: what to do when infected with a virus, where to seek help or advice regarding cyber security, how to effectively protect personal information, are simple but longer passwords better than complex but shorter ones, what capabilities do my mobile phone allow the attacker to exploit? Until answers to these questions are clearly communicated and understood by majority of Internet-enabled users, paradigm shift will not take place. Especially mobile devices may provide a convenient means of authentication but traditional, one-factor textual password model will prevail. Two-factor authentication, while already supported by major vendors, is expected to rise only slowly. Multi-factor authentication such as biometrics is dependent on additional layers of technology and deployment costs, a disadvantage and a dissuading argument for organizations. Substitutes (graphic-/sound-based passwords) will have to go a long way towards providing a positive cost/benefit ratio, however, new operating systems promise to provide graphic passwords as a login option, albeit with on-demand fallback to textual input. While increased complexity is disadvantageous to attackers, fault thresholds must be set adequately and inert mindset of users modified for its adoption to rise.

References

Bakker, Marcus, and van der Jagt, Roel. (2011) "GPU-based Password Cracking", [online], University of Amsterdam, Faculty of Science, Informatics Institute, System and Network Engineering, http://hgpu.org/?p=6060 (accessed on 03/01/2012).

Burr, William E., Dodson, Donna F., Newton, Elaine M., Perlner, Ray A., Polk, Timothy W. et al. (2011) "NIST Special Publication 800-63-1 Electronic Authentication Guideline: Information Security", [online], National Institute of Standards and Technology, Information Technology Laboratory, Computer Security Division, http://www.nist.gov/itl/csd/sp80063-121311.cfm (accessed on 03/01/2012).

Chakrabarti, Saikat, and Singhal, Mukesh. (2007) "*Password-Based Authentication: Preventing Dictionary Attacks*", **Computer**, June 2007, Vol 40, No. 6, pp. 68—74.

Eastlake, Donald E. 3rd, Jones, Peter. (2001) "US Secure Hash Algorithm 1 (SHA1)", [online], Internet Engineering Task Force, http://tools.ietf.org/html/rfc3174 (accessed on 08/12/2011).

Eldefrawy, Mohamed Hamdy, Alghathbar, Khaled, and Khan, Muhammad Khurram. (2011) "*OTP-Based Two-Factor Authentication Using Mobile Phones*", Proceedings 2011 Eighth International Conference on Information Technology: New Generations (ITNG), Riyadh, Saudi Arabia, April 11—13, 2011, pp 327—331.

Hadnagy, Christopher. (2010) **Social Engineering: The Art of Human Hacking**, Wiley, New Jersey.

Halderman, Alex J., Schoen, Seth D., Heninger, Nadia, Clarkson, William, Paul, William, Calandrino, Joseph A. et al. (2008) "Lest We Remember: Cold Boot Attacks on Encryption Keys", [online], 17th Usenix Security Symposium, San Jose, California, July 28—August 1, 2008, http://www.usenix.org/events/sec08/tech/halderman.html (accessed on 05/01/2012).

Kuhn, Markus G. (2004) "*Electromagnetic Eavesdropping Risks of Flat-Panel Displays*", Proceedings 4th International Workshop on Privacy Enhancing Technologies (PET 2004), Toronto, Canada, May 26—28, 2004, pp 88—107.

Ma, Wanli, Campbell, John, Tran, Dat and Kleeman, Dale. (2010) "*Password Entropy and Password Quality*", Proceedings 2010 Fourth International Conference on Network and System Security, Melbourne, Australia, September 1—3, 2010, pp 583—587.

Manuel, Stéphan. (2008) "*Classification and Degeneration of Disturbance Vectors for Collision Attacks against SHA-1*", Designs, Codes and Cryptography, April, Vol 59, pp 247—263.

Marquardt, Philip, Verma, Arunabh, Carter, Henry, and Traynor, Patrick. (2011) "*(sp)iPhone: Decoding Vibrations From Nearby Keyboards Using Mobile Phone Accelerometers*", Proceedings 18th ACM Conference on Computer and Communications Security (CCS 2011), October 17—21, 2011, pp 551—562.

Nyamsuren, Enkhbold, and Choi, Ho-Jin. (2007) "*Preventing Social Engineering in Ubiquitous Environment*", Proceedings 2007 International Conference on Future Generation Communication and Networking (FGCN 2007), Jeju-Island, Korea, December 6—8, 2007, pp 573—577.

Porter, Sigmund N. (1982) "*A password extension for improved human factors*", **Computers & Security**, January 1982, Vol 1, No. 1, pp 54—56.

Rivest, Ronald. (1992) "The MD5 Message Digest Algorithm", [online], Internet Engineering Task Force, http://tools.ietf.org/html/rfc1321 (accessed on 08/12/2011).

Sarier, Neyire Deniz. (2010) "*Practical Multi-factor Biometric Remote Authentication*", Proceedings 2010 Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS), Bonn, Germany, September 27—29, 2010, pp 1—6.

Sasaki, Yu and Aoki, Kazumaro. (2009) "*Finding Preimages in Full MD5 Faster Than Exhaustive Search*", Lecture Notes in Computer Science 2009, Vol 5479/2009, pp 134—152.

Schneier, Bruce. (1996) **Applied Cryptography, Second Edition: Protocols, Algorithms, and Source Code in C**, Wiley, New Jersey.

Schreier, W.J., and Boult, T.E. (2009) "*Bipartite Biotokens: Definition, Implementation, and Analysis*", Lecture Notes in Computer Science 2009, Vol 5558/2009, pp 775—785.

Shannon, Claude E. (1948) "*A Mathematical Theory of Communication*", The Bell System Technical Journal, July, October, Vol 27, pp 379—423, 623—656.

Sotirov, Alexander, Stevens, Marc, Appelbaum, Jacob, Lenstra, Arjen, Molnar, David, Osvik, Dag Arne, and de Weger, Benne. (2008) "MD5 considered harmful today", [online], Technische Universiteit Eindhoven, http://www.win.tue.nl/hashclash/rogue-ca/ (accessed on 08/12/2011).

Suo, Xiaoyuan, Zhu, Ying, and Owen, G. Scott. (2005) "*Graphical passwords: a survey*", Proceedings 21st Annual Computer Security Applications Conference (ACSAC 2005), Tucson, Arizona, December 5—9, 2005, pp 462—472.

Umezawa, Katsuyuki, Tezuka, Satoru, and Hirasawa, Shigeichi. (2011) "*An Authentication System using Smart Phones as Secure Storage*", Proceedings 2011 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Yokohama, Japan, October 9—12, 2011, pp 415—420.

Weir, Matt, Aggarwal, Sudhir, Medeiros, Breno de, and Glodek, Bill. (2009) "Password Cracking Using Probabilistic Context-Free Grammars", Proceedings 2009 30th IEEE Symposium on Security and Privacy, Oakland, California, May 17—20, 2009, pp 391—405.

Xie, Tao, and Feng, Dengguo. (2009) "How To Find Weak Input Differences For MD5 Collision Attacks", [online], International Association for Cryptologic Research, http://eprint.iacr.org/2009/223 (accessed on 08/12/2011).

Yan, Jeff, Blackwell, Alan, Anderson, Ross, and Grant, Alasdair. (2004) "*Password memorability and security: empirical results*", **IEEE Security & Privacy**, September—October 2004, Vol 2, No. 5, pp 25—31.

van Eck, Wim. (1985). "*Electromagnetic Radiation from Video Display Units: An Eavesdropping Risk?*", **Computers & Security**, December 1985, Vol 4, No. 4, pp 269—286.

Zim, Herbers Spencer. (1978) **Codes & Secret Writing**, Scholastic, New York.

Multi-Level Security Cannot Realise NEC Objectives

Harm Schotanus, Tim Hartog and Cor Verkoelen
Information Security Dept., TNO Information and Communication Technology, Delft, The Netherlands
Harm.schotanus@tno.nl
Tim.hartog@tno.nl
Cor.verkoelen@tno.nl

Abstract: Multi-Level Security (MLS) is often viewed as the holy grail of information security, especially in those environments where information of different classifications is being processed. In this paper we argue that MLS cannot facilitate the right balance between need-to-protect and duty-to-share as required for a Network Enabled Capability (NEC) based military operations. This is due to the fact that MLS is deemed rigid in its restrictions; it obstructs the flow of information towards lower classifications by definition and thus influences duty-to-share; furthermore MLS results in a set of rigid preconditions for the physical environment to guarantee the required need-to-protect. The focus of a security solution instead should be on flexibility towards information sharing and reducing risks to be useful in a NEC environment. This can be achieved by firstly reducing the size (and complexity) of the systems that contain the classified information systems, using Multiple Independent Levels of Security (MILS) to create these smaller, separated compartments; and secondly controlling the information flow between the (different) classified compartments by dynamic policies. Moreover, the realignment of classification provisions can make management of information much more flexible and efficient. Hence, we can finally forget MLS.

Keywords: MLS, MILS, information security, classified information, policies

1. Introduction

Ever since the 1970s Multi-Level Security (MLS), defined by the National Security Institute (1985), is often viewed as the holy grail of information security for working with information of different classifications. Despite these high expectations, it has never become reality on a large scale and most of the time a system-high approach is taken instead. MLS is still an idea for the future. Despite that MLS is mainly concerned about regulating access to classified information and the actions that are allowed, it is often expected that MLS will be able to solve any security problem with regard to classified information.

The question nevertheless is whether MLS would solve the actual problems at hand. In other words, can MLS find the right balance between need-to-protect and duty-to-share that is required for realizing the Network Enabled Capability (NEC)?

2. Objectives of military information systems

Military operations are increasingly carried out with partners from different nations and in partnership with non-military organisations in order to reach the operational objectives. Current military communication infrastructures are deployed as stand-alone networked information systems operating in the system-high mode. However, effective and efficient cooperation among the different coalition partners requires information sharing – with the motto “the right information at the right place at the right time” (Buckman 2010). It is not acceptable in many cases to outright share all information among all partners. Hence, it is necessary to be able to determine which information is suitable for sharing with other partners, and enforcing these decisions.

MLS is often positioned to solve the problem of “the right information at the right place at the right time”, without unnecessarily violating information security rules. The main question is whether that is correct. In this paper we attempt to debase the unrealistically high expectations on MLS and provide alternative focal areas. We argue that MLS will not provide the right balance between need-to-protect and duty-to-share and we recommend that the development of military information systems focuses on alternatives to realise this goal.

These alternatives should provide flexibility towards implementing security restrictions and permissions. Rigid rules regarding the treatment of classified information, as is the case in MLS, inhibits the sharing of information as they only define limitations but do not capture the need-to-share principle. The need-to-share principle implies that the mandate for information sharing decision making is delegated to the authorised users of the information. Thereby increasing the risks involved proportionally to the amount of information accessible to a user. It is therefore fundamental to reduce risk without impeding the capability to share information.

3. Multi-level security

3.1 Origin

MLS has been created to enforce access control in military and government information systems, especially aiming at the separation of information of different classifications within a single system. This implies a focus on confidentiality protection. To provide this protection, the MLS concept was created, for which the primitives are formulated in the Bell-LaPadula Model by Bell and La Padula (1973). The BIBA model by Biba (1977) is an alternative MLS model. However in most military domains, confidentiality is still the primary concern, whereas BIBA primarily addresses integrity.

The Bell-LaPadula access control model states that based on the user's clearance and their need-to-know the user may access a subset of the classified information within the system. Access is defined by two rules of mandatory access control (MAC), as shown in Figure 1:

- No read up (NRU) – no read access is permitted to an object with a higher classification than the clearance of the user.
- No write down (NWD) – no write access is permitted to an object with a lower classification than the clearance of the user.

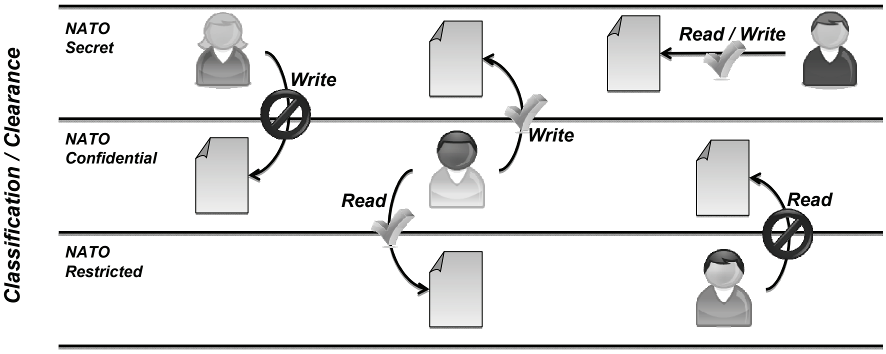


Figure 1: Access rules within Bell-LaPadula

This ascertains that information cannot leak to users that do not have the proper clearance for that information and therefore fulfils the 'duty-to-protect' requirement.

3.2 Consequences

Fahs (2004), Bell (2005), Schaefer (2004) and Levin (2007) state that the application of the MLS model has a number of significant drawbacks with respect to the military information system objectives, aside from its technical feasibility. First, information can only flow upwards by definition – i.e. information with a certain classification can only be accessed by users with a higher or equal clearance. MLS does not provide any means for a user to share information with another user having a clearance that is lower than the information requires. All access control in a MLS system is only restrictive, and not permissive. The two MAC rules cannot be extenuated by a user, but only reinforced by discretionary access rules (DAC). The classification is always leading. This is in direct conflict with the objective of actually sharing information, need-to-share, as in many cases information is shared with users having a different or lower clearance. E.g. NATO Secret cannot be shared with NATO personnel with a NATO Confidential clearance, as the clearance does not permit this.

Second, whether information can be processed is not merely dependent on the clearance of the user and the classification of the information, but also depends on the environment in which the information is accessed. E.g. in an office cleared to process NATO Confidential information, an MLS terminal can be placed. Subsequently, a user with a NATO Secret clearance should be able to access NATO Secret information on the terminal assuming the user has a need-to-know. However, this would not be permitted because the environment is only cleared to process NATO Confidential information. The terminal is typically not capable of determining its location, and the security properties of this location.

This means that the highest classification of information that can be processed is determined by a combination of the clearance of the user and the physical location of the MLS terminal. Consequently the user cannot always access all the information for which he is cleared. Hence, this means that the configuration of an MLS system should address the physical location security properties and thereby making the configuration very static and inflexible. This inflexibility inhibits the possibility to access important information when it is needed and thus is in contradiction with the "right information at the right place, at the right time" as needed for an effective and efficient cooperation among the different coalition partners.

In addition the result of MLS is one system where all types of information can be processed and any user satisfying the access control requirements can access the system. However creating such a large system also increases the risk associated with processing classified information. This means that any possible flaw – technical or procedural – can escalate quickly into serious incidents. In such a large, and complex system as this, flaws are nearly a certainty.

In other words, an MLS solution is a very restrictive and inflexible model that enforces a single, rigid policy. It also inhibits any option for a user to make an alternative decision regarding the access to information or sharing of information. Access control and possible information sharing is determined by the MLS model itself, intelligent decisions made by a user are in direct conflict with the MLS concept. Hence it can be concluded that MLS does not facilitate sharing of information but instead will actually hinder it.

4. Alternatives to MLS

We need an alternative to MLS that provides a better balance between information sharing and information protection. This implies a shift in focus on solely protecting the confidentiality of the information to incorporate the assurance of the availability of the information. All the aspects that need to be addressed are briefly described in the following paragraphs. This includes technical components but also requires a change of attitude.

4.1 Classification management

An important aspect of working with classified information, is actually properly classifying the information. In traditional system high environments the incentive is to classify anything in the environment to the highest possible level, as stated by Neugent (2005). E.g. information in a NATO Secret domain is by default and implicitly classified as NATO secret. As a result the information pyramid is top-heavy as shown in Figure 2. In an optimal situation, most information will be classified as low as possible. Only a small amount information will have a high classification. In system high and MLS solutions the opposite is much closer to the truth. E.g. for a map with coordinates of artillery the map is not classified, but only the coordinates are. In a system high or MLS environment the whole map is classified.

As a result, the required security measures to protect all the information are costly, redundant and impede sharing information – the higher the classifications are, the more difficult it is to share. The value of classifying information is eroded. The noise of all the over-classified data will result in an increase of the risks by the incapability to distinguish the actual value of the information.

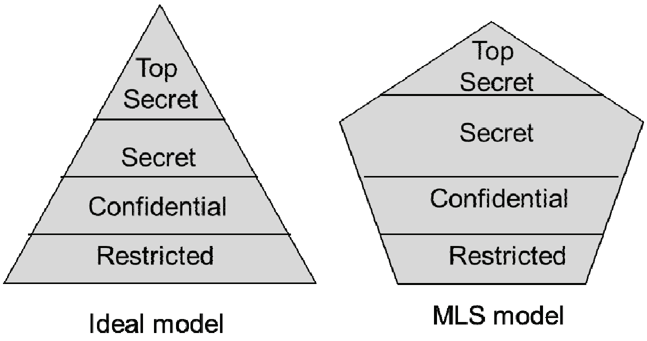


Figure 2: Information amounts per classification

Therefore in an environment such as envisioned by Buckman (1973) and proposed in by Hartog (2011) it is necessary to realign the way information is classified, with two important characteristics:

1. Information is classified as low as possible – implying that only that part of the information is classified that can be justified as such. E.g. for a map with coordinates of artillery, only these coordinates are classified and not the entire map.

2. Information is only classified for as long as the classification is actually needed and declassified when the classification is no longer reasonable.

The information system should provide the user the means to de- or reclassify information and as such make it easier to share information. However, in any case the user must be made responsible for all such actions.

4.2 Compartmentalization of information

Multiple Independent Levels of Security (MILS) is often cited as an alternative to MLS by Rushby (1981). Although the acronym is quite similar, the model is different. The MILS concept creates different information compartments on one system or computer, as shown in Figure 3. Each compartment can be used e.g. for a different classification. Vanfleet (2005) states that a so-called separation kernel ensures that information cannot flow between different classified compartments. And because each compartment behaves like an independent system, access control to, authorisation within these compartments can be specific for each compartment.

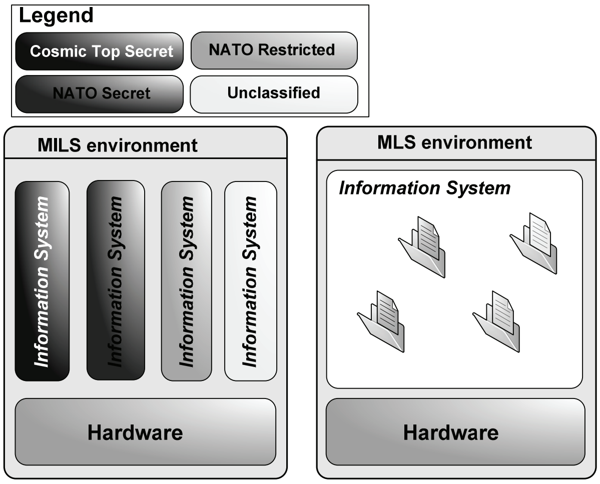


Figure 3: MILS versus MLS

In case a user needs access to information in a certain compartment it is his own decision to access that compartment given in which environment he is. In an MLS system this decision cannot be made by a user because the relation between information and the location is rigidly defined in the configuration of the MLS system.

4.3 Reduction of information per system

Processing classified information is always associated with residual risks that cannot be mitigated fully. The more information a system contains and the more users have access to that system, the higher these risks are. Hence, less information in a compartment means that the risks can be mitigated and managed better. As the amount of information needed by a user will not decrease as a whole, this implies that the compartments must be made smaller¸ but users may have to access more compartments. Because the residual risks are smaller, it is more acceptable for the organisation to delegate the responsibility to the user to make intelligent decisions about information sharing.

To reduce the size of the systems into smaller compartments and to support the decision making by a user, we must disentangle the infrastructure and the information and shift the security of the information towards the individual computer systems or eventually application as stated by Verkoelen (2010) and Hallingstad (2007)(2008). Having several compartments available for different types of information allows the reduction of the amount of information in each compartment.

4.4 Controlled information exchange

The information exchange between compartments of different classifications should not take place unrestrictedly. Hence, we need mechanisms to establish the controlled information exchange based on user decisions. The mechanisms should reflect the agreements regarding the conditions of information sharing.

The design of a plane of interconnection between different compartments can be very complex. The elements of trust and threats should be leading in the design. Thereby the security functionality and policies are geared to the actual situation instead of forming rigid default configurations. Schotanus (2011) and Boonstra (2011) describe a methodology to analyse interconnections based on the trust assumptions and applicable threats for these interconnections. The methodology determines the security requirements and associated assurance levels.

The focus should be on creating standardised modules to realise a plane of interconnection that can provide means to control the information exchange between different compartments. The methodology can be used to select the modules for a specific situation.

4.5 Policies

The fixed security policies that many current systems and the MLS concept have, incur the inflexibility of the system. This inflexibility impedes the capability to have the user in control of which information can be shared within all the rules and agreements associated with modern military operations. To provide a flexible model for security, the applicable security policies should be separated from the security systems and enforced by the security mechanisms. The policy can be aligned with the present risks. This enables the adjustment of policies when needed, and the enforcement by security systems of the correct policies.

A security policy describes a set of rules that determines how the information must be processed in order to ensure the protection of the information sufficiently. We need a methodology to translate the abstract rules as defined in agreements into policies that can be enforced by the security mechanisms. The methodology should ensure the completeness and consistency of the applicable rules. Neugent (2005) states that the security mechanisms should be aligned with the present risks and that we need means to manage the security policies.

The rules should reflect a user action or in some cases an automatic action that has been applied to the information. Based on the act it is determined whether the permission to share the information can be granted. A crucial part of a policy is the accountability and therefore we need to register the information flow. An example of an act can be secure labelling of information as described by Oudkerk (2010), Hartog (2011) and Eggen (2010).

5. Conclusions

We have argued that MLS - the oft-cited all-in solution for processing classified information - cannot facilitate efficient sharing of a subset of information while guaranteeing the security requirements on the entire set of information. The reasons behind this are that in an MLS system information can flow only towards higher classifications and for information sharing, we actually need to facilitate the controlled flow of information to lower classifications. MLS cannot take the environment into account hence it has to on a static configuration whereas contrariwise flexibility is needed. Furthermore MLS moulds a huge system of all information which will only increase the risk in case of flaws, and for information sharing we actually need to reduce risks of flaws. MLS is a very strict and static model that enforces a fixed policy.

Hence it is necessary to shift the focus to other solutions instead of clinging to an unrealistic and essentially undesirable concept such as MLS. We have identified five elements that should be integrated into system designs. These are:

1. Improvement of classification management to reduce the amount of information that is classified and reduce the time information is classified.

2. Compartmentalisation of information by separating information of different classifications using Multiple Independent Levels of Security.

3. Reduction of the amount of information per compartment to limit the risk of flaws.

4. Controlled information exchange by using control mechanisms between different compartments based on user decisions.

5. Policy management to translate and manage abstract rules and agreements to machine interpretable rule sets that can be implemented in control mechanisms that validate user actions against a policy to facilitate the sharing of information between different compartments.

In this manner we can actually realise the objectives stated by the Network Enabled Capability. This primarily reflects the need for flexibility towards determining what information can be shared and with whom. However the need-to-share principal needs to be properly balanced by the need-to-protect. It is necessary to reduce the risk of information sharing that can be attained by dividing the information in smaller separate compartments. In addition, for disentangling the information and infrastructure, it is necessary to relinquish the implementation of security in the infrastructure and work towards applying policies on information flows. And then the concept of MLS can finally be shelved indefinitely.

References

Bell, D.E. (2005) "Looking Back at the Bell-La Padula Model," *Proceedings of the 21st Annual Computer Security Applications Conference* (ACSAC), pp 337–351, 2005.

Biba, K.J. (1977) "Integrity Considerations for Secure Computer Systems", MTR-3153, The Mitre Corporation, http://seclab.cs.ucdavis.edu/projects/history/papers/biba75.pdf, April 1977.

Boonstra, D. and Schotanus, H.A. and Verkoelen, C.A.A. (2011) "A Methodology for the Structured Security Analysis of Interconnections", http://ieeexplore.ieee.org/iel5/6111660/6127424/06127476.pdf,
Milcom 2011

Buckman, T. (2010) "Nato Network Enabled Capability Feasibility Study – Executive Summary", version 2.0, NC3A

Eggen, A., et al. (2010) "Binding of Metadata to Data Objects – a proposal for a NATO specification",
Norwegian Defence Research Establishment (FFI) & NC3A

Eggen, A., et al. (2010) "XML Confidentiality Label Syntax – a proposal for a NATO specification", Norwegian Defence Research Establishment (FFI) & NC3A, 22 april 2010.

Fahs, R and Wiseman, S.R. (2004) "Re-Floating the Titanic: Multi Level Security in Contemporary Environments" Defence Research Agency. March. 2004.

Hallingstad, G. and Oudkerk, S. (2007) "Protected Core Networking – Initial concept description"

Hallingstad, G. and Oudkerk, S. (2008) "Selected aspects of Protected Core Networking"

Hartog, T., et al. (2011) "Labelling: Security in Information Management and Sharing" 6th International Conference on Information Warfare and Security

La Padula, L.J. and Bell, D.E. (1973) "Secure Computer Systems: A Mathematical Model," MTR–2547, Vol. II, The MITRE Corporation, Bedford, MA,31 May 1973. (ESD–TR–73–278–II)

Levin, T.E., et al. (2007) Analysis of three multilevel security architectures; Proceedings Of Computer Security Architecture Workshop.

National Security Institute (1985) "5200.28-STD Trusted Computer System Evaluation Criteria", http://csrc.nist.gov/publications/secpubs/rainbow/std001.txt, December 1985.

Neugent, W., (2005) "Assured Information Sharing," The Edge, The MITRE Corporation, Fall 2005

Oudkerk, S., et al. (2010) "A proposal for an XML Confidentiality Label Syntax and Binding of Metadata to Data Objects", IST 091, 2010.

Rushby, J. (1981) "Design and Verification of Secure Systems", 8th ACM Symposium on Operating System Principles; pp. 12-21; Asiloma, CA; December 1981; (ACM Operating Systems Review, Vol. 15, No. 5).

Schaefer, M. (2004) "If A1 is the answer, what was the question? an edgy naif's retrospective on promulgating the trusted computer systems evaluation criteria", In Annual Computer Security Applications Conference, pages 204–228. IEEE Press, 2004.

Schotanus, H.A., et al. (2011) "Decomposition of the Security Requirements for Connected Information Domains", Military Communications and Information Systems Conference, Amsterdam

Vanfleet, W.M., et al. (2005) "MILS: Architecture for High-Assurance Embedded Computing," STCS CrossTalk

Smulders, A.C.M. (2010) "Classification as a bottleneck for information sharing", VOVKLICT January 2010.

Verkoelen, C.A.A., et al. (2010), "Security shift in future network architectures," Information assurance and cyber defence, NATO IST 091, 2010.

From Kinetic Warfare to Strategic Communications as a Proactive and Mind-Centric Paradigm of the art of war

Torsti Sirén
National Defence University, Finland
torsti.siren@mil.fi

Abstract: Traditionally, the purpose of war has been to influence the Other's behaviour, or even to destroy the existence of the Other. In this article, war has been referred to as the violent, or killing-prone, state of the human mind. The art of war, for its part, involves an understanding of ideational possibilities and restrictions along with future-oriented open-mindedness to reach grand-strategic ends set by political authorities. If the fundamental maxim of the art of war is understood as the proactive avoidance of kinetic wars without losing peace as it has been understood in this article, then individual and collective minds should be recognised as the essential targets to be influenced, not by kinetics, but by incentives for cooperation. Methodologically, this article focuses on both Social Constructivism (habituation and reification theses) and Critical Realism (emancipation thesis) as social theories and abductive content analysis as a method. While we may be habituated into the past contexts of war and the art of war, which may even be reified in their nature, there is still space for emancipation – the power of new ideas. This article argues that during the past decade the mentioned maxim of the art of war has found prominence in the comprehensive, proactive and mind-centric paradigm of the art of war, Strategic Communications (*StratCom*), which is based on the theory of positive recognition. The theory of positive recognition assumes that human societies may reconstruct their domestic structures (identities, interests and social systems) in order to earn 'universal recognition' in the eyes of other human societies. The challenge of liberal democratic human communities and societies is then to seduce intolerant human communities and societies to reconstruct their constitutive identity structures by 'being wonderful and acting accordingly'.

Keywords: war, the art of war, paradigm, strategic communications, emancipation

1. Introduction

War and the art of war are probably among the most researched human phenomena. However, a uniform definition of at least for the concept of war cannot be found due to its political nature. Generally speaking, war may be understood as a military crisis or militarily culminated conflict. We may, for example, speak about the conflict between Israel and the Palestinians which in time may develop into a military crisis or war. The *Stockholm International Peace Research Institute* (SIPRI) argues that a crisis may be called a wide-scale armed conflict or war if there are at least 1,000 casualties as a result of battles during one year (SIPRI 2005: 135). Kalevi Holsti, on the other hand, argues that 1,000 casualties is only a guiding variable. There may be 'only' 0–750 casualties and even then we may call the crisis a war (Holsti 1995: 10). Combining both of the above mentioned arguments, *war* has been understood in this article simply as a state of human mind and as a time-bound episode of a conflict cycle which may be called a war only if it suits the political agenda of the involved or third parties, even if a single casualty cannot be found during the episode. This definition allows us, for e.g., to speak intellectually about such phenomena as media war, information war, Phoney War, etc., even if there was not one human casualty during such an episode.

There are numerous books, researches and authors on the concept of the art of war and principles of warfare (e.g. Sun Tzu, Flavius Vegetius, Niccolò Machiavelli, Miyamoto Musashi, Sir Basil Liddell Hart, etc.), but most of them have concentrated on the use of the military in traditional kinetic war, which is not exactly the point of this article. The point of this article is to promote the idea of *proactively avoiding wars without losing peace as a maxim of the art of war*. This idea necessitates a global perspective as well as a grand-strategic, proactive and comprehensive approach. In this article, *the art of war* has been understood as an affective and cognitive skill in combining cumulated historical 'knowledge' with today's materiel and ideational possibilities and restrictions as well as with the future-oriented open-mindedness in a way that facilitates the attaining of the grand-strategic ends set by the legitimate political authorities by minimizing the use of kinetic force as far as it is possible. *The paradigm of the art of war* has been understood here in a Kuhnian way, as a general ideational orientation which may be updated if the paradigm ceases to function satisfactorily (Kuhn 1996: 176; Pocock 1989: 13).

Strategic Communications has been understood here as a proactive and comprehensive paradigm of the art of war. Strategic Communications stresses the need for exerting an enduring psychological effect on intolerant human identity structures by offering incentives for reconstructing more tolerant ones. Strategic Communications is based on the *Theory of Positive Recognition*, which assumes that human societies may reconstruct their domestic structures (identities, interests and social systems) in order to earn 'universal recognition' in the eyes of other human societies (Others). The challenge then is to seduce human societies not only to behave in tolerant ways, but also to have them internalise liberal values as an integral part of their identity structures. The additional challenge is that the already tolerant human societies should act according to their 'seductive messages' in order to change the world into a more tolerant and better place to live in. To put it shortly, liberal democracies should show by the power of example what 'being wonderful and acting accordingly' means, and not go 'crusading' in the name of liberal democracy.

2. Theory and method

The aim of this article is to enhance our understanding of how we could reach the above mentioned maxim of the art of war. The article takes a reflective approach and emphasises the human potential for emancipation. Reflectivism has not been considered here as a synonym for normativism. It is not then that post-modern Western societies should entice the Others to reconstruct their intolerant identity constructions into more tolerant ones, but they could do so.

Social Constructivism (habituation and reification theses) and Critical Realism (emancipation thesis) constitute the social theoretical framework for this article. In this context, the constructivist social theory (Constructivism) emphasises our habituation of thinking of war and warfare in traditional terms (*habituation thesis*). Constructivism's *reification thesis* means that this habituation tends to get stronger generation by generation. (See, for example, Sirén 2009: 19)

Critical Realism supplements Constructivism by emphasising emancipation and the reflective power of the human mind (See, for example, Sirén 2009: 23–52). This means, for e.g., that if our understanding of the paradigms of the art of war were not relevant any more, we could then maybe emancipate ourselves from our habituation and reified modes of thinking and start to respond to the call of the future by using novel ideas. The call of the future is understood here as *the need to avoid wars without losing peace*. This call cannot be answered merely by preparing ourselves for the past or next wars (*peace by force*) and threats, but by trying to use our human potential (i.e. ideas) in proactively constructing new *silos of human tolerance*, as well as enhancing the existing ones, nationally, regionally and globally.

Qualitative content analysis, as the method used here, is basically nothing but the thematisation of the source material. In this article, thematising has been conducted by leaning loosely (i.e. abductively) on the themes of Constructivism and Critical Realism as social theories.

3. Paradigms of the art of war

There are at least four paradigms of the art of war:

- Target-centric paradigm
- Effects-centric paradigm
- Comprehensive paradigm of crisis management and
- Strategic Communications

Each paradigm promotes different ends and targets. *Target-centricism* emphasises direct kinetic influence on the Other's military, population and infrastructure. Target-centricism is based on a traditional thought according to which it is possible to win a war and suppress the Others in one crucial battle and conquer their territory by military means; one wins, another looses (*win–loose-situation*).

Effects-centricism, introduced by the USAF in the 1990s, emphasised the significance of U.S. air power, much in the spirit of *Douhetism* (Douhet 1983 [1921]; Cerasini 2003: 124). According to effects-centricism, or EBO as it was starting to be known, the Other's will, understanding and capabilities should be attacked so fast and accurately that the Other would not have time to react to the striking tempo (e.g. *Shock and Awe* phase prior to the 2003 Second Gulf War). John Warden's (1995) model of 'five circles' (see Figure 1), as a theoretic motive of EBO, argues that the fastest way to win a crucial battle is to affect directly the adversary's decision-making systems and key-functions by using developed information technology, precision kinetics and invisibility. The Gulf War in 1991 saw EBO's integrated use of satellites and new information technological innovations (e.g. computers and fibre optics) with the wider usability of the electromagnetic spectrum that changed the four-dimensional battlefield (land, maritime, sea, space) into a five-dimensional battle space (land, maritime, sea, space, information and electromagnetic spectrum). (Organisation and Instruction Authority, Defence Staff, Norwegian Armed Forces 2007)

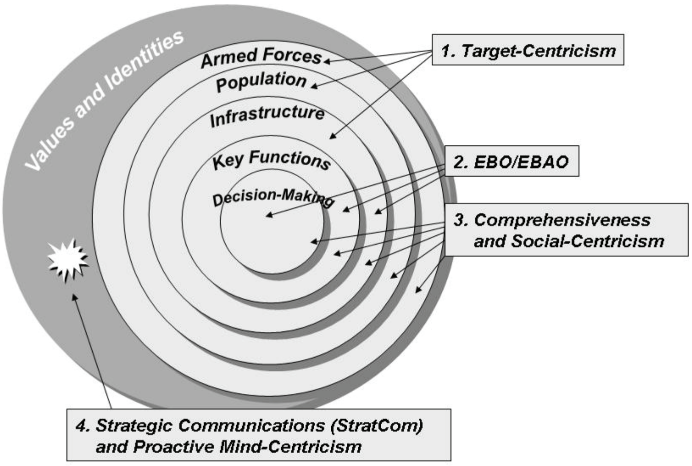


Figure 1: Targets and paradigms of the art of war

EBO's European version is called the *Effects-Based Approach to Operations (EBAO)*. The main difference between EBO and EBAO is that EBAO prefers soft ways and means when affecting the Others. The technology- and science-driven orientation of EBO has been criticised even in the USA (Mattis 2008: 105–108). Both EBO and EBAO take war as a science, based on systemic thinking according to which states and human societies are regarded as unitary systems of their own (*Internal System of Systems*), consisting of political, military, economic, social, infrastructure and information systems (*PMESII*) (See, for example, Vego 2006: 51–57). Furthermore, EBO and EBAO emphasise that states and human societies, as layered systems, are linked internationally to other equivalent state- and non-state systems (*International System of Systems*) (Organisation and Instruction Authority, Defence Staff, Norwegian Armed Forces 2007: 82; NATO 2009b: 1–5). However, EBAO could also be called an evolutionary continuum of EBO, because it stresses the significance and use of soft ways and means when affecting the Others instead of kinetics or precision kinetics (Paul 2008: 1).

According to social-centric comprehensiveness, states as well as civilian and military organisations may act in accordance with shared normative principles. The comprehensive paradigm of crisis management also prefers soft means and ways to the kinetic or precision kinetic ones when affecting the Others, *but the main focus of comprehensiveness is on the efforts of coordination and integrated communication between the actors prior to the forthcoming international crisis-management operation* (Social-Centricism; See Figure 1). Coordination and integrated communication includes such issues as political coordination; human rights in an area of operation; the need to maintain a legal judiciary, societal order as well as justice (*Rule of Law*) in an area of operation; perspectives on human security (See, for example, UNDP 1994: 22), and other possible humanitarian issues (Crisis management Initiative, Rintakoski, Autti (eds.) 2008: 11). The comprehensive paradigm of crisis management also includes an ambition, at least implicitly, to avoid crises and wars, without conducting kinetic battles, based on preventive diplomacy (See, for example, United Nations, General Assembly 2001).

Strategic Communications is explicitly focused on the proactive and continuous affecting of values and identities in order to prevent inter- and intra-state armed conflicts before they even occur. Strategic Communications complements John Warden's "model of five circles" (see Figure 1) by arguing that the 'sixth circle', values and identities, would be the more correct target to be affected than, e.g., the armed forces, or the decision-making systems of the Other. Due to the Internet revolution, StratCom brings forth the new dimension of warfare, namely the mind; *the mind is a weapon and target of the modern 6-dimensional global mind space*.

Strategic Communications comprises a strategic narrative defined by a national political leadership or international society (e.g., the UN), as well as a set of ways and means by which the narrative is to be mediated. Defined narrative and acts should correspond; it is not possible to say something and act differently, if the aim is to have a positive effect on intolerant identity structures. Strategic Communications does not necessarily necessitate the existence of any specific target audience to be affected, at least if thinking globally. However, StratCom may have various target audiences, if thinking operationally, but then StratCom easily slips into the domain of propaganda, cultural imperialism, the systemic thinking of natural science and strategic psychological operations (Psyops). In other words, we tend to trace causal relationships in human societies without recognising that human societies are open and blurred systems of reduced predictability that do not operate in laboratories (See Figure 2). One may try to trace causal relations and social networks, say in an operational area of some crisis management operation (CMO), and by this way try to detect the local key leaders or human communities to be influenced. But if thinking globally, systemic thinking has its limits – there are unknown causalities ('unknown unknowns'; See Figure 2) that are not detectable. What we have to do then is to seduce intolerant human individuals, communities and societies (*silos of intolerance*) to behave and think like 'Us'. But this necessitates that 'We' act in accordance with our narratives of tolerance and 'wonderfulness'.

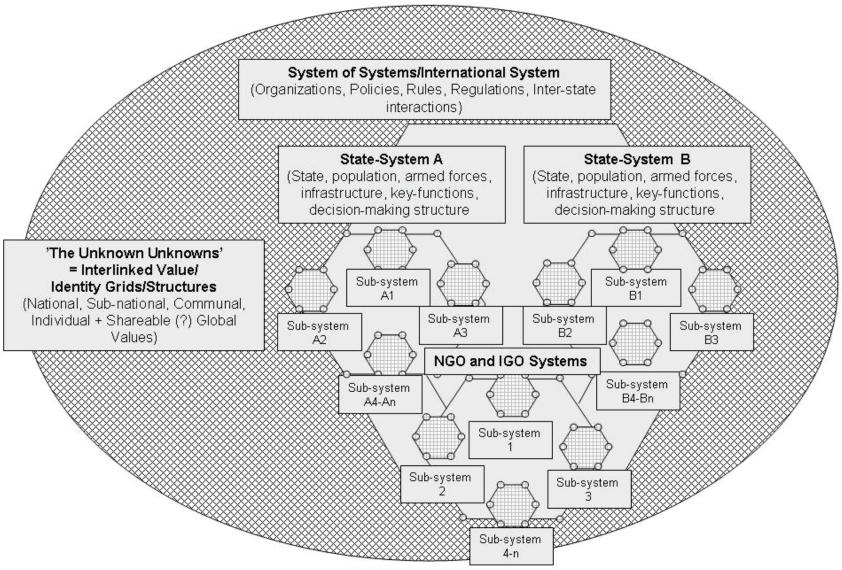

System of Systems/International System
(Organizations, Policies, Rules, Regulations, Inter-state interactions)
State-System A
(State, population, armed forces, infrastructure, key-functions, decision-making structure
State-System B
(State, population, armed forces, infrastructure, key-functions, decision-making structure
'The Unknown Unknowns'
= Interlinked Value/ Identity Grids/Structures
(National, Sub-national, Communal, Individual + Shareable (?) Global Values)
Sub-system A1
Sub-system A2
Sub-system A3
Sub-system A4-An
Sub-system B1
Sub-system B2
Sub-system B3
Sub-system B4-Bn
NGO and IGO Systems
Sub-system 1
Sub-system 2
Sub-system 3
Sub-system 4-n

Figure 2: From systemic thinking to open and blurred human systems of reduced predictability

For it to be Strategic Communications instead of communication (without an *s*), measures of effectiveness (MOE) should be included into the set of means and ways by which we mediate our narrative. This means that we should be able to evaluate the reactions our narratives and actions have caused among non-targeted or targeted audiences so that we are able to change the content of our narrative, or the used means and ways, if necessary. This article will not, however, deal further with MOEs.

4. Strategic communications in proactive action

Target-centric and the effects-centric paradigms of the art of war share three basic elements: the main body, the manoeuvrable striking element and the 'follow-on-forces' attack' element. The task of the main body (e.g. the Macedonian phalanx) is to challenge and bind the Other's main forces, while the striking element (e.g. Macedonian cavalry) conducts deep operations and strikes against the Other's rear echelons. The follow-on-forces' attack element (e.g. the Macedonian light infantry (Hypaspists)) supports the breakthroughs of the striking element. In addition, the main body offers shelter for the other elements if their strikes are not successful.

Figuratively, StratCom also shares the mentioned three basic elements (See Figure 3). The main body of StratCom consists of the *values and freedoms of liberal-democracy*. The striking element of StratCom, '*the unleashed cavalry of the free world*', consists of the free media, such as Facebook and Wikileaks, as well as of NGOs such as the Nobel Committee. It charges, challenges and outflanks even yourself if you do not act according to your narrative. But if you act according to your narrative, 'the cavalry' probably won't charge your sheltering main body and you may even be able to support the charges of the cavalry by the follow-on-forces' attack element of StratCom, namely '*the mobile troops of liberal-democracy*' (public diplomacy, official statements, peace-mediations, crisis management operations and so on). This entity has been considered here as *StratCom's six-dimensional mind-space strategy* (See Figure 3).

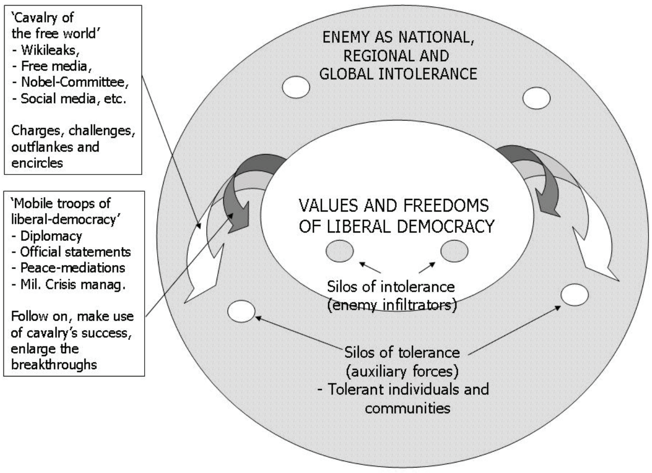

'Cavalry of
the free world'
- Wikileaks,
- Free media,
- Nobel-Committee,
- Social media, etc.
Charges, challenges, outflankes and encircles
'Mobile troops of liberal-democracy'
- Diplomacy
- Official statements
- Peace-mediations
- Mil. Crisis manag.
Follow on, make use of cavalry's success, enlarge the breakthroughs
ENEMY AS NATIONAL, REGIONAL AND GLOBAL INTOLERANCE
VALUES AND FREEDOMS OF LIBERAL DEMOCRACY
Silos of intolerance (enemy infiltrators)
Silos of tolerance (auxiliary forces)
- Tolerant individuals and communities

Figure 3: Six-dimensional warfare: the basic principle of the mind-space strategy

4.1 The main body of stratcom – values and freedoms of liberal democracy in action

The strategic narrative of a state, human community or society, should include one's own national or societal mission (who are we?). We have to know and critically recognise our own identity- and culture-related weaknesses first to be able to credibly seduce intolerant human individuals, communities and societies to reconstruct their values and identity structures into more tolerant ones. The strategic narrative should also include one's vision of the future world which could be based, e.g., on the UN Charter. The narrative should be understandable and palatable, otherwise we may be driven into *cognitive dissonance*. Cognitive dissonance describes the disharmony between unpalatable and palatable information: "unpalatable information falls on barren ground because people cannot see where it fits into their way of seeing and believing" (Taylor (2006: 11).

Most of the existing nations and states have not explicitly defined a strategic narrative of their own, including Finland. The following strategic narrative, proposed by the author, would probably be acceptable to all of the legislative, judiciary and executive authorities of Finland and probably all of their counterparts in other existing liberal democracies:

> *Finland is a liberal democracy that in all of her activities stresses and strives to promote norms, values and goals based on the equality of nations, freedom of speech and religion, the equality of the sexes as well as on a comprehensive security concept shared by the UN, the OSCE (Organization for Security and Co-Operation in Europe) and the EU.*

The mentioned narrative defines Finland's mission or self-image (liberal democracy) and her values and vision (promoting norms, values and goals based on…). The narrative also recognises that no-one is infallible, but at least Finland is trying to act according to her narrative (… strives to promote…), which increases the sheltering power of the main body if Finland sometimes does not happen to act according to her narrative. Liberal democracy, as a key concept of the narrative, is a mental state based on liberal values and identity structure, civil (non-political) rights (e.g. freedom of expression, press, religion/non-religion, sexual orientation and gender identity) and political rights (e.g. right to vote in free elections). Liberal-democracy also means transparency – the opportunity for citizens to see inside the power structures and the outcomes produced by the power structure (see, e.g., Fukuyama 1992: 43–44). Democracy, for its part, can be understood merely as a right to vote in elections. Thus a country is democratic if it grants people the right to choose their government through some kind of elections, but it is not a liberal democracy if it does not guarantee freedom of speech or religion. Islamic Iran, e.g., may be called a democracy, but not a liberal democracy, since it does not guarantee freedom of speech or religion, which consequently makes it vulnerable to the charges of 'the unleashed cavalry of the free world'.

4.2 The striking element of stratcom – 'the unleashed cavalry of the free world' in action

There is no escape for today's repressive governments unless they realise that all the governments are for the people, not vice versa. In order to maintain power, today's repressive and intolerant governments have to kill the Internet, or the Internet will kill intolerant governments in the long run. There are over 2,000 million Internet users today and the amount is growing. People share information and meet each other in virtual meeting places (social media) such as Facebook, Twitter, LinkedIn and so on. In other words, people are globally connected, which opens new possibilities to affect them. Simultaneously, the societies and militaries of the world have become more vulnerable since their key functions are operated, more or less, with Internet-based applications. The most interesting issue concerning social media is that when people become active in one or many social media networks, they purposefully use those networks to spread their individual or cultural ideas (Rainie, Purcell and Smith 2011). In other words, the Internet and social media have emancipated people to express themselves even in societies which do not support freedom of speech. Consequently, repressive governments try to set Internet restrictions (firewalls) and people try to bypass those. This could be prescribed as a cyberspace arms race, but it is also about the universal need for individual freedom, and the Internet offers a tool for satisfying that need. For example, in January 2011 the so-called *Jasmin Revolution* in Tunisia was, at least partially, ignited by Wikileaks, which confirmed the view of the Tunisian people about their government being corrupt (Payne 2011). Tunisia's revolution inspired protests at least in Egypt, Libya, Yemen, Jordan, Bahrain and Syria. Even China set firewalls that block social media, fearing her people's claims for political freedom.

In addition to the Internet, the Nobel Peace Prize Committee may be prescribed as a representative of the cavalry of the free world which strikes against intolerance and repression when they are detected. By 2011 the Nobel Peace Prize has been awarded 92 times between 1901 and 2010 and ever since 1901 the Prize has caused tension between the Nobel Committee and repressive governments (*Nobelprize.org* 2011a and *Nobelprize.org* 2011b). It is precisely the Nobel Committee's non-political role (even though the Committee is chosen by the Parliament of Norway) that makes its role as a 'cavalry charger' so significant. Many Nobel Peace Prize laureates may be named as good examples of successful 'cavalry charges' of the Nobel Committee against intolerance and suppression of intolerant governments. The German Carl von Ossietzky, for e.g., was awarded the peace prize in 1935 due to his liberal and pacifist thoughts. The then German government declared that no German could accept any Nobel Prize (including Nobel Peace Prize) in the future (*Nobelprize.org* (2011c). Another illuminating case is the Chinese Liu Xiaobo, who in 2008 was the first citizen of the People's Republic of China to ever receive the Peace Prize. According to the Nobel Committee, Liu deserved his prize, because of "his long and non-violent struggle for fundamental human rights in China." As Carl von Ossietzky, Liu Xiaobo was also arrested at the time he was awarded the peace prize, because he was a criminal, according to the Chinese government. In the West he was interpreted as being arrested for criticising the religious and political intolerance of the Chinese government. China's government made a statement that the award should have been given to someone who has "focused on promoting international friendship and disarmament", but the Nobel Committee does, after all, have the right to interpret the rules of the award itself. (Pomfret 2010)

In addition to social media and the Nobel Committee, nation rankings by, e.g., education, human development, level of (liberal) democracy and transparency, quality of life as well as gender equality and gay rights and so on, are tools of the unleashed cavalry of the free world. Nation rankings conducted by the UN, international magazines, newspapers, etc., fit well into the content of the theory of positive recognition when exposing the ranking of states among other states of the world through education level, for example. (United Nations Development Programme 2011 and Crabtree 2005+) In August 2010, *Newsweek* evaluated Finland as the best country of the world to live in, which was positively recognised in Finland at least, but was also discussed on Facebook by the representatives of other nations. However, it is unclear whether this caused any public debate in countries which were ranked as not being such good countries to live in. And even if Finland were the best country to live in, Finland might not be the true emancipatory power in world politics, since there is still work to do, e.g., with the habituated and reified attitudes Finns have towards Russia — according to which the only threat to Finland – in addition to economic and environmental threats – is Russia (however publicly called a challenge and not a threat).

4.3 Follow-on-forces' attack element of stratcom - 'the mobile troops of liberal democracy' in action

The strategic narrative and StratCom's message of 'wonderfulness' could be mediated, e.g., by repeating the defined strategic narrative in official statements whenever and wherever it is possible. The narrative may be mediated through so-called DIMEX (Diplomacy, Information systems, Military Policy, Economic Policy and other policies (X), such as environmental and social policy). The challenge is to act according to the narrative; otherwise a government's narrative causes more damage to the government itself and to the image of the whole nation or society. In this regard, StratCom is close to the concept of perception management. What distinguishes StratCom from perception management is that StratCom's mind-set is global, whilst perception management is basically interested in egoistically enhancing one's own national image and reputation, which sometimes may necessitate lying and the overt appeasement of intolerant, albeit economically important, governments.

If a government or society acts in accordance with its narrative and exposes itself seductively as a wonderful example to be followed, the unleashed cavalry of the free world will not charge this wonderful example's sheltering main body. The wonderful government or society may then even support the charges of the cavalry by the follow-on-forces' attack element, namely '*the mobile troops of liberal-democracy*' (public diplomacy, official statements, peace-mediations, crisis management operations and so on). A good example of this is Finland's active role in international crisis management and peace mediations, albeit without a publicly defined strategic narrative. Recognising this, the Nobel Committee awarded Finland's former President Martti Ahtisaari the Nobel Peace Prize in 2008. He was awarded the peace prize for his "important efforts, on several continents and over more than three decades, to resolve international conflicts" (*Nobelprize.org* (2011d).

5. Conclusions

War has always been part of human interaction. Traditionally war has been defined by the terms of target-centricism, but this has been set under critical evaluation in this article. This article defines war as being a mental state, which means that war may be considered as peace, if we could change our narrow national and egoistic mind-sets into the wider and global one, based on liberal values without hate, agony and repression. Consequently, the art of war has traditionally focused on the kinetic means and ways of target-centricism as guarantors of achieving the set political ends. In this article the maxim of the art of war has been understood as proactively avoiding wars without losing peace. The challenge to us as individuals and human societies is then to emancipate ourselves from our habituated and reified concepts and approaches of the past and start to think and act globally as well as be future-oriented.

Strategic Communications may be considered an evolutionary maxim of all the previous paradigms of the art of war. Strategic Communications is not about cultural imperialism or propaganda, but about an indirect grand strategic approach, based essentially on uniformity of the message (strategic narrative) and acts. The strategic narrative should inform us and the Others about our mission, values and vision of the future world. Strategic Communications includes the contents of the previous paradigms of the art of war (target-centricism, effects-centricism and a comprehensive paradigm of crisis management), but it stresses that instead of leaning mostly on the systems of preparedness (e.g. kinetic and precision kinetic defence or cyber defence) we should proactively and non-kinetically influence intolerant value and identity structures worldwide by being wonderful and acting wonderfully ourselves. In other words, liberal democracies should offer incentives for intolerant and repressive governments and societies to reconstruct their intolerant value and identity structures into more tolerant ones.

References

Cerasini, Marc (2003) The Future of War – The Face of 21st-Century Warfare, Alpha, United States of America.

Crabtree, Vexen (2005+) Which Countries Set the Best Examples? www.vexen.co.uk/countries/best.html.

Crisis Management Initiative (CMI), Rintakoski, Kristiina; Autti, Mikko (eds.) (2008) Comprehensive Approach – Trends, Challenges and Possibilities for Cooperation in Crisis Prevention and Management, Seminar Publication based on Comprehensive Approach Seminar 17 June 2008, Helsinki, Ministry for Foreign Affairs, Finland, Helsinki.

Douhet, Giulio (1983 [1921]) The Command of the Air. USAF Warrior Studies, Office of Air Force History, Washington D.C.

Fukuyama, Francis (1992) The End of History and the Last Man, Penguin Books, England.

Holsti, Kalevi (1995) Peace and War: Armed Conflicts and International Order 1648-1989, Cambridge University Press, Cambridge.

Kuhn, Thomas, S. (1996) The Structure of Scientific Revolutions (Third Edition), The University of Chicago Press, United States of America.

Mattis, James N. (2008) "USJFCOM Commander's Guidande for Effects-Based Operations", JFQ, issue 51, 4th quarter 2008, pp 105–108.

NATO (2009) Allied Joint Doctrine for Information Operations (AJP-3.10) (November 2009), Promulgated Version. NATO/PfP Unclassified. Material at the possession of author

Newsweek (2010) "The World's Best Countries – A Newsweek Study of Health, Education, Economy, and Politics Ranks the Globe's Top Nations", www.thedailybeast.com/topics/the-world-s-best-countries.html.

Nobelprize.org (2011a) "The Nobel Peace Prize", www.nobelprize.org/nobel_prizes/peace.

Nobelprize.org (2011b) "All Nobel Peace Prizes", www.nobelprize.org/nobel_prizes/peace/laureates.

Nobelprize.org (2011c) "The Nobel Peace Prize 1935 − Carl von Ossietzky",www.nobelprize.org/nobel_prizes/peace/laureates/1935.

Nobelprize.org (2011d) "The Nobel Peace Prize 2008 − Martti Ahtisaari",www.nobelprize.org/nobel_prizes/peace/laureates/2008.

Organisation and Instruction Authority, Defence Staff, Norwegian Armed Forces (2007) Norwegian Armed Forces Joint Operational Doctrine (FFOD), 15 June 2007, www.mil.no/multimedia/ archive/00106/ FFOD_English_106143a.pdf.

Paul, Christopher (2008) Information Operations – Doctrine and Practice: a Reference Handbook,
Praeger Security International, United States of America.

Payne, Matthew (2011) Bradley Manning and the Jasmine Revolution, www.thepaltrysapien.com/2011/01/bradley-manning-and-the-jasmine-revolution.

Pocock, J.G.A. (1989) Politics, Language and Time – Essays on Political Thought and History, The University of Chicago Press, United States of America.

Pomfret, John (2010) "China's Liu Xiaobo wins Nobel Peace Prize", Washington Post, October 8, 2010, www.washingtonpost.com/wp-dyn/content/article/2010/10/08/AR2010100801502.html.

Rainie, Lee; Purcell, Kristen; Smith, Aaron (2011) The Social Side of the Internet. Pew Research Center, www.pewInternet.org/Reports/2011/%20The-Social-Side-of-the-Internet.aspx.

Sirén, Torsti (2009) State Agent, Identity and the "New World Order" – Reconstructing Polish Defence Identity after the Cold War Era (Academic Dissertation). Edita Prima Oy, Helsinki.

Stockholm International Peace Research Institute (SIPRI) (2005) SIPRI Yearbook 2005, Oxford University Press, Oxford.

United Nations Development Programme (1994) Human Development Report, Oxford University Press, New York.

United Nations Development Programme (2011) Human Development Report 2011 – Sustainability and Equality: A Better Future for All, http://hdr.undp.org/en/reports/global/hdr2011/download.

United Nations, General Assembly (2001) Prevention of Armed Conflict – Report of the Secretary-
General (A/55/985-S/2001/574), www.reliefweb.int/library/documents/2001/un-conflprev-07jun.htm.

Vego, Milan N. (2006) "Effects-Based Operations: A Critique", JFQ, issue 21, 2nd quarter 2006, pp 51–57.

Warden, John (1995) "Air Theory for the Twenty-first Century", in Battlefield of the Future, www.airpower.maxwell.af.mil/ airchronicles/battle/chp4.html.

Cablegate Analysis of Likely Espionage of Nokia by the United States

Daniel Strmecki[1], Wilke Schwiedop[2], Emmanuel Oyo-Ita , Brigitte Kaagman[4], Pierre Leandre[5], Enrique Santos-Brihuega[6], Lateef Kadiri[3] and Jessica Dufmats[7]
[1]University of Zagreb - Faculty of Organization and Informatics, Croatia
[2]University of Applied Sciences Bonn-Rhein-Sieg, Germany
[3]University of Salford, UK
[4]Hogeschool van Amsterdam, Netherlands
[5]ESIEA, France
[6]Universidad de Alcalá, Spain
[7]Mid Sweden University, Sweden
leandre@et.esiea-ouest.fr

Abstract: In recent years computing has shown an increasing shift towards mobile devices. Smartphones and similar devices such as tablets are becoming more powerful and less expensive every day and as such are becoming more widespread not only in developed, but also in developing countries. Alongside the development of the mobile devices, the internet offers an increasing amount of services for these devices. This evolution of mobile devices from simple telephones to portable computers as well as their increased interconnectivity however also made them more prone to security issues. As such secret services around the world are given more possibilities and opportunities to use these mobile devices for espionage and widespread surveillance. The recent leak of cables sent by US embassies around the world also known as the 'cablegate' gives us an opportunity to get a better understanding of this issue. In the light of these events we tried to measure to which extent US have spied on European companies, especially one of the leading manufacturers of mobile devices around the world, Nokia. We set up a database and preprocessed the embassy messages to allow us to search through the huge amount of data in short time. We then investigated Nokia's fields of business to find possible contact points to special agencies. Additionally we looked for other reasons why special agencies might have a specific interest in Nokia. With this information we analyzed the data. The analysis clarifies two major key points:
Firstly it validated the assumption that mobile devices, even civil ones, play an important role in modern warfare. They are used not only by US special agents, but also by guerrilla forces to coordinate military operations. Secondly Nokia is a main competitor to American companies in the fastest growing markets worldwide, such as India and China. This paper intends to present the results and the main conclusion of our analysis.

Keywords: Wikileaks, cablegate, Nokia, espionage, mobile devices, smartphones

1. Context

This Intelligence Report has been created as part of the European IP eDiscovery 2012 in Manchester. On behalf of the EU-Commission for Trade and Industry an international group of students from different branches of study analyzed the United States diplomatic cables leaked between 8th February 2010 and 1st September 2011 for industrial espionage and give the EU-Commission for Trade and Industry a recommendation on further actions.

During this leak, widely known as “cablegate”, over 250,000 diplomatic cables sent by US embassies, consulates and diplomatic missions around the world were released to the public by wikileaks. The not-for-profit organization Wikileaks launched in 2006 and provides an international and anonymous platform for news leaks and whistleblowers. Wikileaks is best known for their publication of classified information about the wars in Afghanistan and Iraq, Guantanamo and internet censorship.

The published cables date from 28th December 1966 to 28th February 2010 and contain information on US External political relations, internal government affairs, Human rights, Economic conditions, Terrorists and terrorism and the UN Security Council, including confidential and secret material.

The investigation's foremost purpose was to prove or disprove the suspicion that United States embassy employees conducted or were involved in industrial spying on european companies. The ten investigated companies were:

- Alcatel
- Cobham

- Datong
- Infineon
- Nokia
- Phillips
- Siemens
- Telefonica
- Thales
- Unilever

Of these companies only Nokia will be discussed in this paper.

Nokia is a multinational communications company based in Espoo, Finland. Nokia is most famous for its mobile phones but also offers various other electronic mobile devices and internet services. Nokia has about 120,000 employees in 120 different countries around the world.

In 2011 Nokia was the leading manufacturer of mobile phones.

2. The electronic discovery reference model

During our research we tried to measure the extent Nokia has been spied upon by the United States and why US secret services could have a special interest in Nokia and its business.

The model we used for our research is the Electronic Discovery Reference Model (EDRM) which provides a common, flexible and extensible framework of guidelines and standards for an eDiscovery process. The EDRM was created in May 2005 to address the lack of standards in the eDiscovery market and was placed in the public domain in May 2006.

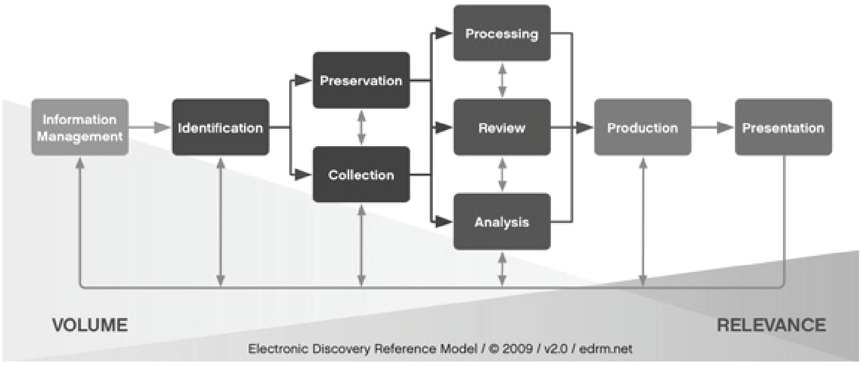


Figure 1: Electronic discovery reference model

The EDRM describes the following stages:

Information Management: The information management track describes how to manage all sources of ESI during the complete information lifecycle, from creation through usage to archival or deletion to ease future eDiscovery processes.

Identification: During the Identification track all potentially relevant sources are identified including the scope, breadth and depth of the required ESI. We searched information about Nokia and it’s link with the United States. Basically the first thing we needed to do was to find the most relevant haystacks and the keywords for all of the haystacks. After many changes in this part we determined that the most relevant haystacks are United States and China.
Relevant keywords for the United States haystack are:

- Companies: Apple, Motorola, Google and Microsoft.
- People: Steve Jobs, Stephen Elop, Bill Gates and Sanjay Jha.
- Other: RINOA, MOSAID, Iraq, Iran, espionage, spyware, patent, war, military, 3GPP…

Relevant keywords for the China haystack are:

- United States-China Economic and Security Review Commission,
- China Communications Standards Association,
- Other: Huawei, trade agreement, standard, market, lawful interference…

Preservation/Collection: The preservation track ensures that the required ESI can be extracted without being interrupted by running businesses processes and is protected against modification and deletion. All the ESI we required for our project were stored on a DVD, the cable files in csv and sqlformat, the spyfiles as a single rar archive. The first thing that we did is that we took the hash values of the given data so our research could be used later even in a court or law.

The internet was our second main source for keywords. We used Google search to get a clearer picture of Nokia and the United States possible interest in the company. By searching the Internet we found some possible evidence about a patent war. It included articles about the patent companies such as Mosaid, Sterling and WiLan. We also found more information that explained the meaning of some of the words and some more clues about what to search for in the cables. The main time period is from the year 2006 up to about 2010.

During the Collection track the required ESI is extracted from the identified sources in a way that the integrity and authenticity can be verified. For our everyday work every member of our made copies of the files for him or herself. To ensure the data had not been unintentionally modified we took md5 checksums of each copy and compared them to the original.

Processing/Review/Analysis: The goal of the Processing track is to speed up the process of reviewing the collected ESI by unifying the data's format and data deduplication. In our case we inserted the cables into a postgresql database. To speed up searching and to give us better results we preprocessed the cable contents by converting them to tsvectors, a specialized data type designed for full text searches in the postgresqldatabase.We removed all cables containing none of our keywords which massively reduced the amount of data we had to search through.

For the actual searching we developed a small search application for the full cables database and only Nokia related cables. The application is written in PHP web programming language, using PostgreSQL database and ran on an XAMP web server.

During the Review track the gathered data is checked for relevance and privileges. Privileged information that may not be used in court is flagged and not considered for the analysis track.

The analysis track is the final evaluation of the relevant information. The information is reviewed to draw conclusions and build a case strategy.

Production: Deliver the selected ESI to third parties in a form established beforehand.

Presentation: Final presentation of key findings and conclusions.

3. Main findings

Initially our hope was that the cables just flat out said “we spied on Nokia”. Unfortunately this was not the case. We did however find several clues and reasons for possible espionage of Nokia by the USA in the cables.

Shia militants using Nokia equipment (REFID: 09STATE45504):

A very interesting cable that we have found in the cablegates is that Nokia’s equipment (N-95 smartphone’s) is being used by the Shia militants in Iraq. Basically the militants use the GPS functions of the Nokia smartphone’s to coordinate attacks from various bases near Baghdad. Here is a list of the most important information inside the cable:

- “GPS-equipped Nokia N-95 smartphone’s were used to collect coordinates throughout the installation as well.”
- “Shi'a militants and sources throughout Iraq have used the GPS function on the N-95 to pinpoint point-of-impact coordinates on various forward operating bases near Baghdad and inside the IZ.

Iranian-backed Shi'a militias in southern and eastern Iraq possess the weapons, training, manpower, capability, and intent for IDF attacks against BIA."

The fact that the Iraq militants are using Nokia's equipment could be a very good reason for the United States to spy on Nokia because Nokia is not allowed to sell this kind of equipment to Iran, Iraq etc..., due to Wassenaar arrangement.

Nokia undercover surveillance equipment (REDIF's: 10SANJOSE130, 09TIRANA255, 09HELSINK13):

This cable proves that the United States are using Nokia's equipment for surveillance. Here are the most important parts of the cables:

- "Undercover Surveillance Equipment: In CY 03, CY 04 and CY 06 NAS funded the purchase of equipment to provide video and audio surveillance for PCD operations. This equipment was highly versatile, essential for officer/agent safety, and has been used for documentation of undercover narcotics purchases, providing valuable documentary evidence used in criminal judicial proceedings. However, some of this equipment, such as the Nokia cell phone audio transmitter, is outdated and easily recognized during operations and requires an oversize shirt for PCD's officials. All this equipment is well maintained and is being used for its intended purpose."
- "II. STATUS-COMMODITIES - 2008 1A. Type of Commodity and Project - Surveillance Equipment - Nokia spy phone and Night Vision Goggles. 1B. Location - Organized Crime Directorate of ASP 1C. Use & Condition - The OC Directorate maintains accountability for these items and continues to employ them during investigations of organized crime and narcotics trafficking cases. 1D. Disposal of Commodities - N/A 1E."

If the United States rely on Nokia equipment and in the same time Nokia is selling the same / similar equipment to United States enemies (1st cable) this could be a very good reason for United States to spy on Nokia so that the United States could find out where exactly is the equipment going and how.

This was not that surprising as communication as exchange of information is obviously an expected target for any secret service.

United States interest in Nokia's market shares in India, Iran and China (REDIF's: 07NEWDELHI5230, 07CHENNAI289):

We have found several cables mentioning Nokia's market shares and investment. The search using keywords "market" and "share" gave us 30 results (cables). Here are a few of the most important sentences from those cables:

- "Seeking to take advantage of Tamil Nadu's revised industrial policy, both Dell and Nokia announced expansion plans in the state this week. A senior Dell official announced on December 4 that the company will invest USD 30 million in its existing facility in Chennai, which will allow it to produce an additional 500,000 desktop computers per year. Not to be outdone, Nokia announced on December 6 an additional investment in its Chennai manufacturing facility of USD 75 million, along with a planned increase of its employees, who will total 30,000."
- "At Maran's urging, mobile phone manufacturer Nokia in April 2005 invested $150 million in a manufacturing facility at Sriperumbudur (50 kilometers from Chennai). Nokia's investment also attracted seven of its component suppliers to invest in the region. Nokia's printed circuit board supplier, Aspocomp Group Oyj, invested $70 million. At the same time, Nokia's handset mechanics supplier Perlos Corporation invested $12 million. Nokia and its suppliers are located in a special economic zone spread over 210 acres, which provides fiscal incentives and allows for considerable flexibility in the hiring of personnel. By December 2006, the facility had shipped 25 million handsets for the booming Indian mobile phone market. Motorola followed in Nokia's wake with an announcement in June 2006 of its plans to invest $30 million in a facility to produce handsets. The facility will produce low cost handsets for the Indian market."
- "What is the nature of investments (and names, if known) that host country businesses have in Cuba? What host country businesses participated in the Havana Trade Fair (November 3)? Major Finnish companies such as Wartsila and Nokia have investments and sell products and services in Cuba."

From this cables we can see that the United States are very interested in Nokia market shares and investments all over the world. Nokia is a competitor to lot of huge United States companies including Dell and Apple. There is no evidence that this data from the cables is gained illegal, but again, this is one more big reason for the United States to spy on Nokia.

Nokia 3GPP Project:

In the spyfiles, Wikileaks had also alleged that NOKIA was involved with the production of surveillance technologies that intruded on the privacy of individuals. With release of the files that would later come to be referred to as spyfiles, wikileaks leaked a "*company confidential*" memo detailing the "*3GPP Project*". This document is a presentation speaking about a new standard of telecommunication. 3GPP mean "*3rd Generation Partnership Project*", it's collaboration between telecommunications group from the entire world in order to make a new globally applicable third generation (3G) of mobile phone system.

3.1 Additional information

We found some additional information on the internet as a reason for the US to spy upon Nokia.

Commercial rivalry between US companies and Nokia (patent war):

- A commercial rivalry between Microsoft and Nokia : "Mr. Lindgren couldn't talk about was a deal he was about to strike with Microsoft Corp. and Nokia Corp. that would give Mosaid control of 2,000 advanced cellular technology patents originally filed by the Finnish phone giant."
- A commercial rivalry with Apple: "There's no question Apple lost the legal battle that pitted its significant intellectual property holdings against Nokia's even deeper patent portfolio"
- Some article about the deal between Nokia and Motorola: "Huawei is pleased that the court continues to recognize the merits of our claim that Motorola must abide by its contractual obligations to protect Huawei's trade secrets and intellectual property" said Hauwei in a statement.

Backdoor hidden in Nokia cell phones for the Indian Government:

- "The news of this day is that Nokia, RIM and Apple have provided to the Indian government, particularly its military, a backdoor that allows the ability to monitor each mobile device."

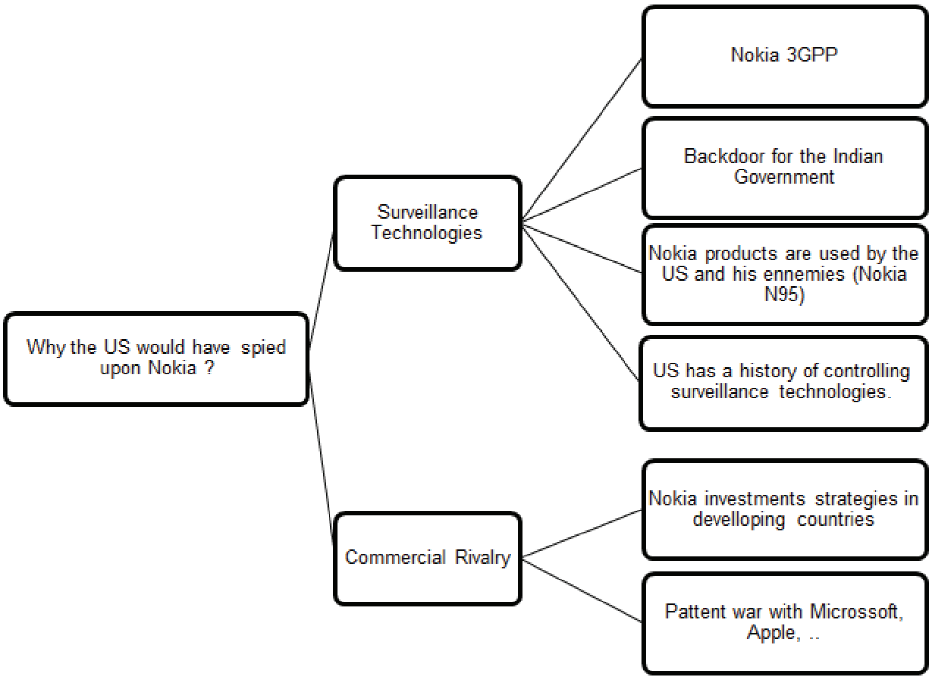

Why the US would have spied upon Nokia ?
Surveillance Technologies
Commercial Rivalry
Nokia 3GPP
Backdoor for the Indian Government
Nokia products are used by the US and his ennemies (Nokia N95)
US has a history of controlling surveillance technologies.
Nokia investments strategies in develloping countries
Pattent war with Micrososft, Apple, ..

Figure 2: Diagram

4. Conclusion

As earlier stated our client the EU Commission for Trade and Industry requested an e-discovery work process into the Wikileaks-Cablegate-Spyfiles concerning Nokia, the EDRM process was undertaken in the process of e-discovery.

In this process we were guided by several things such as hypothesis, questions, allegations and eventual determination of that which is true and that which is false.We founded 2 main allegations:

- Wikileaks had alleged that NOKIA was involved with the production of surveillance technologies that intruded on the privacy of individuals. With release of the files that would later come to be referred to as spyfiles, wikileaks leaked a "*company confidential*" memo detailing the "*3GPP Project*" that was planned to be undertaken by NOKIA SIEMENS in partnership with 6 other organizational partners.
- The other allegation was that the mention of NOKIA in the US DIPLOMATIC CABLES; where an indication of US GOVERNMENT involvement of industrial espionage towards a Finnish company. The allegation here was that the US was spying on NOKIA in other to attain for its own companies, an advantage over EU companies in the field of surveillance.

These were the two allegations which we worked with; and so the E-Discovery work process was geared towards what was fact, what was fiction.

First we looked at the allegations by Wikileaks that Nokia amongst other telecommunications companies, were developing surveillance technology the issue that Wikileaks had with this was that it was not regulated. With the release of the "*company confidential*" document from NOKIA, we concluded that Wikileaks allegation against NOKIA was indeed fact.

Finally we looked at the allegation that the US Government was involved in industrial espionage; here we noticed a convergence of fact and fiction and for this we shall use "partially fact, partially fiction" to describe this circumstance.

"Partially fact", yes the US government had mentioned NOKIA in its diplomatic cables especially in the same sentence as China, the US government was actively spying on Nokia. There are several sources to strengthen this view; one was from the US Government Cable, the fact that Nokia is mention in those cables infers that there is a file on Nokia. Nokia is involved in the development of several technologies that have intersected with the US context for example the N-95 which was used by insurgents to pinpoint concentrated mortar fire.

"Partially fiction", from our E-Discovery work process, we do not see the intention of the US government to give US companies undue advantage. The US government has a sort of monopoly over surveillance technologies; this is achieved through different means, the most plausible explanation reached was that they were interested in Nokia as a manufacturer and dealer in surveillance technologies, who the buyers were and whatever weakness might have been present in the system.

5. Legal aspect

Since the use of internet increases we have to be more careful with the accessibility of information. As said before one of the products Nokia is manufacturing and selling listening devices. The American military is a client from these listening devices. In the US embassy's secret files we found a report which shows that two years ago the use of Nokia N95 equipment in the Iraqis military's.

The European Union was not that found about these findings because they were illegal. The following paragraphs the legal aspects about exporting sensitive products will be defined.

Export of sensitive products

The Wassenaar arrangement is an agreement between different countries about exporting military products in the participating countries. According to the Wassenaar arrangement exporting the Nokia equipment as mentioned before, are according to Category 5- part 1 of the Wassenaar arrangement not allowed without permission. This is on the "*List of Dual-Use Goods and Technologies*"

Embargo

It’s probated for the participating countries to export sensitive product to countries on the embargo list. These rules two of the countries on the embargo list are China and Iran. Although they’re on this list and it was not allowed to export sensitive products to Iran. Nokia did export listening devices and to Iran. The export of these Nokia listening devises to Iran was contrary to this regulation.

Trades EU and Iran

In October 2010 the European Union made a regulation, which contains guides for exporting to Iran. It’s a list with the entire product which is not allowed to transport from the EU to Iran. At 27 of October this regulation is expanded with additional principles. The reason for expanding this regulation was because Iran denied discontinuing his nuclear activities.

Trading of listening devices and other important telecommunication equipment which could be a risk for the privacy of people and companies within the European Union should be denied according to this regulation.

After this regulation Nokia changed his export controls for some of the product they export to the countries at the list. On the Nokia site we discovered it is still possible to order product which may be prohibit sending these countries. Nokia excludes the responsibility of the export of this forbidden product through the general terms of business.

Possible problems with the regulation

In our opinion there could be a few problems with the recent regulation on trading.

With increasing the use of internet and the amount of purchases on the internet it’s harder to check to who get this product in his possession. Therefore the EU can’t check if countries on the embargo lists could get sensitive products.

The second thing that could be a problem for the regulation is that the export of technologies, like Nokia did, has to be reported at the after they are exported already.

References

Berkow Jameson (2011), “Mosaid signs ‘transformational’ deal for Nokia patents; may scuttle hostile takeover” – Financial Post. [online] http://business.financialpost.com/2011/09/01/mosaid-signs-transformational-deal-for-nokia-patents-may-scuttle-hostile-takeover/

Elmer-DeWitt Philip (2011), “Nokia patent settlement a 'sweet defeat' for Apple” – CNN Money. [online] http://tech.fortune.cnn.com/2011/06/14/nokia-patent-settlement-a-sweet-defeat-for-apple/

Electronic Discovery Reference Model (EDRM), [online] http://www.edrm.net/

LoekEssers (2012), “Kabinetkoketteert met exportboycotafluistertech” – webwereld.nl [online] http://webwereld.nl/nieuws/109238/-kabinet-koketteert-met-exportboycot-afluistertech-.html

Nokia Siemens - press Statement (2009), “Provision of Lawful Intercept capability in Iran”. [online] http://www.nokiasiemensnetworks.com/news-events/press-room/press-releases/provision-of-lawful-intercept-capability-in-iran

Nokia Terms And Conditions [online] http://www.nokia.com/us-en/about-nokia/terms/site/terms-conditions/

PierluigiPaganini (2012), “Rim, Nokia and Apple Providing Government Back Doors” – Infosec Island. [online] http://infosecisland.com/blogview/19253-Rim-Nokia-and-Apple-Providing-Government-Back-Doors.html

Press release(2011) - “Controlling dual-use exports” – European Parliament [online] http://www.europarl.europa.eu/news/nl/pressroom/content/20110927IPR27586/html/Controlling-dual-use-exports

Thomson Iain (2011), “Huawei decision clears Motorola-Nokia Siemens Networks deal” – v3.co.uk [online] http://www.v3.co.uk/v3-uk/news/2030677/huawei-decision-clears-motorola-nokia-siemens-networks-deal

René Schoemaker(2011), “Europa gedoogtict-export naardictaturen” – webwereld.nl [online] http://webwereld.nl/nieuws/108055/europa-gedoogt-ict-export-naar-dictaturen.html

Wikileaks, Cables Gates and Spy Files. [online] http://wikileaks.org/

Practical Application of Open Source Frameworks to Achieve Anti-Virus Avoidance

Ignus Swart
CSIR, Pretoria, South Africa
ISwart@Csir.co.za

Abstract: A common aim of malware creators is to have the ability to spread their software undetected through various networks until the required goal is completed. In response to this, anti-virus vendors have implemented various strategies to detect viruses as they attempt to execute and propagate from one target to the next. Some of the anti-virus vendors claim to achieve impressive success rates as high as 98.7% that indicates the problem of spreading viruses and malware is well taken care of. Yet, despite the impressive detection rates, a proliferation of open source tools, frameworks and utilities are being introduced that claim to have the ability to avoid anti-virus detection. As an example, the very popular Metasploit framework has several encoders available that can alter the virus signature in such a way that it will avoid the anti-virus engine and allow the malicious code to be executed. This approach has been implemented and simplified in the Social Engineering Toolkit (SET) as part of a menu driven approach that is accessible to people with a relatively low skill level. The SET framework, implemented in Metasploit, is only one such framework and several more specialised open source tools exist, that does not only focus on encoding but on other common anti-virus avoidance techniques such as binary editing, packing and encryption. Open source packages such as UPX compress the data in the selected virus executable to such an extent that it will most likely completely circumvent the anti-virus and similarly so for a program that is encrypted with a common encryption product such as TrueCrypt. Should the anti-virus still detect the offending executable after either packing or encryption a combination of the two applications might yield superior results. The aim of this paper is to experiment on a common executable that is classified as malware e.g. the meterpreter module of Metasploit, and make use of the various open source frameworks and utilities to document the techniques and success rate of anti-virus avoidance. By presenting the results of this research, it will contribute to the understanding of security personnel / researchers on what can be achieved with open source frameworks and how to better protect against the virus threat. Paper Relevance: While great strides have been made in anti virus detection it is not nearly perfect and many open source tools can be used to effectively hide even old executables flagged as malicious. The question is how difficult is it to use these tools and how effective are they?

Keywords: antivirus avoidance, open source packer, protector, binder

1. Introduction

The current worldwide strategy for dealing with malicious software has not really changed much in the past five years and significant challenges still remain. (McAfee 2011) in their third quarter threat report states that well over 70 million malware types has been identified and the complexity of the malware is increasing. Malware detection happens if a sufficient number of machines on the internet have been infected and a sample of the malware reached one of the antivirus vendors. Detailed analysis is performed and an identification signature is determined for the new malware that is then added to the vendors antivirus database. The new detection signature reaches the clients when a software update is performed and the malware infection starts to recede until it is finally almost negligibly encountered in the wild. After this, the malware creators restart the creation process to maintain the foothold they have over currently infected machines that does not yet have the newly created Anti Virus (AV) signatures and to maintain the ability to infect new machines.

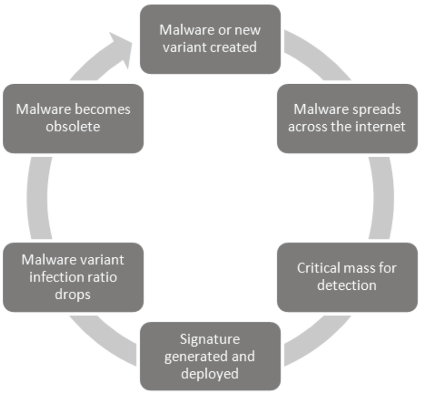

Malware or new variant created
Malware spreads across the internet
Critical mass for detection
Signature generated and deployed
Malware variant infection ratio drops
Malware becomes obsolete

Figure 1: Current malware detection and perpetuation cycle

Research conducted by (Kanich et al. 2011) found that although on average 95.5% of all PC users had an antivirus installed on their system, only 42.9% on average has kept the installation up to date with the latest virus definitions. This leaves a huge potential for malware perpetuation even though antivirus companies has already found an effective detection mechanism. In a report by (Verizon 2011) they stated that only 1% of victims of a successful attack that they investigated was notified by their antivirus of the attack on their system.

Table 1: Antivirus installation and updated percentage breakdown

	A.V Installed (%)	Up to date (%)
US	98.7	22.8
India	92.7	68.7
Other	95.2	37.3
Average	95.5	42.9

It should be noted that although the malware detection cycle has remained the same, antivirus vendors has made significant inroads to simplify and speed up the submission of potential new malware. Additionally tremendous effort has been put in place to automate the detection and classification of malware by automated reverse engineering since the sheer number and complexity of malware would make manual inspection nearly impossible. (Debrey & Patel 2010) in a publication estimate that between 79%-92% of all malware employs packers to increase the difficulty of reverse engineering the malware binary, effectively eliminating the option of only employing humans in the analysis process.

This paper will examine several open source antivirus avoidance tools and frameworks that require little or no programming experience to gauge their effectiveness at avoiding antivirus detection. Section two will describe techniques that antivirus vendors use to detect malware while section three will describe methods used by malware creators to hide malware. In section four the various selected tools and frameworks are described along with the experimentation procedure with the results and conclusion presented in section five and six respectfully.

2. Antivirus methodologies used for detection

According to (Daoud et al. 2008) antivirus detection strategies fall predominantly into two methods namely static and dynamic analysis. The main distinction between static and dynamic analysis are that static detection does not actively execute the code but employs several scanning techniques to detect a potential threat. According to (Brand 2011) some examples of static analysis are control flow graphs, dataflow analysis, string extraction and target architecture determination. Dynamic analysis on the other hand executes the code and monitors it for known suspicious behaviour such as opening an executable file in read and write mode, or attempting to write to the boot sector. Both methods have significant problems as of late and so far no easy solution has been found.

Static analysis can only be effective if the suspicious binary can be matched to a already identified signature and with the introduction of tools such as packers and cryptors, this analysis can be severely hampered. Packers are already widely used as previously noted and (Guo et al. 2008) explains that currently not all packers can be detected and unpacked by antivirus vendors. Packers such as Ultimate Packer for eXecutables (UPX) is already widely know and can be easily unpacked, but in the case of proprietary tools such as Themida it might take up to six months to write a unpacker program that will work effectively.

Dynamic analysis allows the executable to execute in a virtual environment where the full calling sequence, file requests and input output operations are mapped. In theory it should present a clear picture of what the malware's intended purpose is regardless of the encryption or packing features used to defeat static analysis. The significant problem with dynamic analysis according to (Nataraj et al. 2011) however, is the time that is required for the dynamic analysis and the fact that avoidance techniques such as virtual machine detection and debugger attachment detection is also available to malware creators. This will not only prolong the dynamic analysis to an inconvenience, but as the author mentioned could possibly take 254 year of machine time to test against the full 2010 Symantec malware library. Various researchers are attempting to address the shortcomings of the antivirus engine by either employing multi core cloud processing or moving to other architectures perhaps better suited for the task. (Vasiliadis & Ioannidis 2010) presented a novel idea to achieve an exponential increase in performance capability by moving the antivirus engine's processing to the graphics card of the computer. In related work (da Silva dos Santos Silva 2009) proposed a distributed antivirus that will spread the processing of dynamic scanning across various cpu's spread all across the local area network of the organization. Both solutions are attempts to increase the available computing power to the antivirus while still maintaining a enjoyable user experience and while it might not solve all the problems, it could give dynamic analysis a greater chance of identifying malware.

3. Antivirus avoidance methodologies

As discussed previously almost all anti-virus engines rely on a type of signature in the existing executable to accurately identify a potential malicious threat. Due to this relationship between the fingerprint and identification several methods have become available to alter or even hide the offending fingerprint. (Brand 2011) mentions that well over 80 different antivirus avoidance techniques exist but no one antivirus tool exist to negate them all.

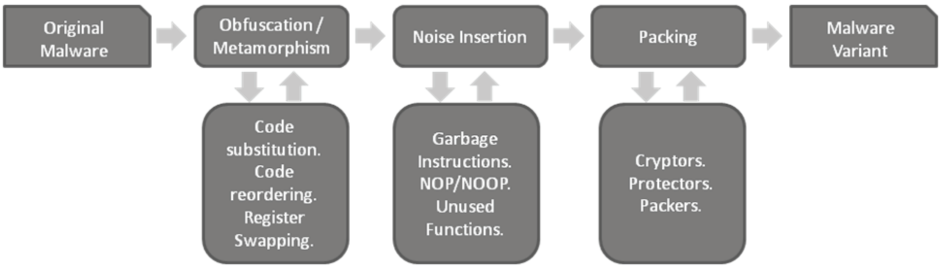

Original Malware
Obfuscation / Metamorphism
Noise Insertion
Packing
Malware Variant
Code substitution. Code reordering. Register Swapping.
Garbage Instructions. NOP/NOOP. Unused Functions.
Cryptors. Protectors. Packers.

Figure 2: Malware antivirus avoidance process

(Ollman 2009) states that Crypters, Protectors, Packers, Binders are some of the methods malware creators use to avoid detection. (Patel 2011) adds obfuscation, register swopping, junk insertion, subroutine transposition as extras wile (Tzermias et al. 2011) adds encryption, UTF-encoding and shellcode encoding. (Davis 2011) brings polymorphism and metamorphism to the collection that will allow code to generate dynamic signatures every time it is executed. The various techniques can be combined and re-combined with other techniques to produce a unique signature of the malware that no antivirus vendor has seen yet. A summary of some of the more prominent techniques are discussed below:

- **Crypters** - Are described as programs that add encryption to the code base of the malicious program to defeat the static analysis of the antivirus and prevent the detection of suspicious code markers by disassemblers such as IDA Pro. This is useful for getting the virus transmitted past network security boundaries but problematic on the target machine since the executable would still need to be decrypted before running.
- **Protectors** - Are programs that specifically alter the code in such a way to remove any information from the code that could be useful in debugging attempts. Should a debugging attempt be detected the protector could even have the ability to execute a different set of instructions to hide the true intentions of the malicious application or if a sandbox is detected, attempt to break out of the sandbox.
- **Packers** - According to the author are designed to reduce the size of the target executable to facilitate proliferation of the executable but it has the added benefit of reducing the surface area available to the antivirus vendors. With the addition of polymorphic encoding the packer not only reduces the size of the malicious executable but also constantly alters the signature of the executable. (Guo et al. 2008) explains that this particular technique further increases the size of the antivirus signature if the vendor is unable to unpack the malicious code decreasing overall system performance.
- **Binders** - Are described as tools that allow the embedding of malware content into files commonly searched for, so effectively creating one executable that contains multiple executables. An example of this is notepad.exe, office applications or even a pirated version of Windows that will perform the requested action but additionally install malware that was combined in the executable.
- **Noise insertion** – Refers to the ability to add various types of instructions to the existing code base that effectively does nothing to alter the application's behaviour but changes its code signature to something completely different. Examples of noise insertion in a application could be

a NOP insertion or a function call with no real meaning such as Sleep(0), “mov eax, eax” or “mul 0x1, ebx”.

When comparing the efficiency of the various common antivirus avoidance techniques (Hu 2011) states that packers, cryptors and protectors are some of the best avoidance techniques currently available. This efficiency is due to the fact that it prevents reverse engineering to such an extent that both static and dynamic analysis of the code becomes relatively useless. A summary of the effectiveness of the various types of avoidance techniques can be found in Table 2 that was compiled by (Ollman 2009).

Table 2: Effectiveness of antivirus avoidance techniques

	File Checksums	RegEx Signatures	File Heuristics	Behavioural Analysis	Debugger Analysis	Static File Analysis	Reverse Engineering
Code Metamorphism	✓✓✓	✓✓✓	✓✓				
Noise Insertion (code)	✓✓✓	✓✓✓	✓				
Compiler Settings	✓✓✓	✓✓	✓	✓			
Noise Insertion (Binary)	✓✓✓	✓					
Crypters	✓✓✓	✓✓✓	✓✓✓	✓✓✓	✓✓✓	✓✓✓	✓✓
Protectors	✓✓✓	✓✓✓	✓✓✓	✓✓✓	✓✓✓	✓✓✓	✓✓✓
Packers	✓✓✓	✓✓✓	✓✓✓	✓✓	✓	✓✓	✓
Binders	✓✓✓	✓✓	✓	✓			

4. Open source tools / frameworks to evaluate

Open source examples were chosen to allow full examination of the results and the experiment was divided into two categories. One part where only the malware executable is used and the other where the source of the malware is also available for processing. A further consideration was that the majority of malware is designed to establish a connection from an infected computer to a remote host and the software evaluated would need to perform a similar task. As such, to replicate the compiled executable scenario would require an open source product that establishes remote connections, and is normally flagged as malware by antivirus vendors. Netcat fits the description perfectly as it is both open source and constantly flagged by antivirus vendors as either malicious software or a hacking tool. For the second part of the experiment Metasploit was chosen since it not only has the ability to create remote connections but all the source is available and multiple encoders are also available. The meterpreter module is also constantly flagged as either malicious software, Swrort.A, EPACK.Gen2, Rozena.A or Swort-C malware variant and thus fits the specified criteria. The baseline meterpreter module is detected 62% of the time making it an ideal candidate to simulate malware antivirus avoidance.

Not all antivirus avoidance techniques are always applicable in all situations and it is worth mentioning the differences between the various types. If the source code of a piece of malware is available all the types of antivirus avoidance techniques are available. In the event that only the compiled binary is available, all is not lost as a range of the antivirus avoidance techniques are still applicable but it does somewhat limit the options. Further explanation follows in the tool discussion below.

- **Ultimate Packer for eXecutables** – (UPX 2011) has been available since 1996 is still one of the leading open source software packers. An updated version 3.08 has been released on 12 December 2011 and adds support for BSD systems.
- **PEScrambler** – According to (Nick 2011) is a tool that will obfuscate Win32 binaries by relocating portions of code and inserting anti-disassembly protection. Development has stagnated as of late and the project seems to be at a standstill
- **Social Engineering Toolkit (SET)** – Is an exceptional penetration testing framework that combines MSFEncode and MSFPayload with automation to such an extent that creating software executables, literally becomes menu driven. Additional parts of the Metasploit framework are also

incorporated into the SET framework such as Shellcodeexec.exe that is a lightweight executable with the ability to interpret shellcode and execute it in memory.

- **MSFVenom** – Compiled into the Metasploit framework the MSFVenom executable is the combination of MSFPayload and MSFEncode into one executable. The advantages are faster encoding times and a less complicated command line environment. Several encoders are present in the Metasploit framework, ranging from polymorphic encoders to simplistic XOR encoders.
- **Shaddam Source** – An open source protector that highlights the potential pitfalls available when making use of open source and freeware projects. The project does have some merit but hidden inside the code is various ways to not only encode and hide the selected executable but also to compromise the attackers computer.
- **Yoda** – A protector project that had its last update in 2006 and while that may sound old, it still has the ability to be problematic for certain anti-virus vendors. Full source is included in the download for anyone to investigate the intricacies of how the protector functions.

5. Experiment results when uploaded to VirusTotal

The results of the experimentation with open source software on the Netcat executable is depicted in Figure 3. The baseline measurement is when the Netcat executable is uploaded to Virustotal without any modifications and resulted in a 62% detection rate from the 43 participating virus vendors. Packing the executable with UPX made absolutely no difference and the detection rate remained at a constant 62%. Making use of the PEScrambler tool resulted in a 44% detection rate, an improvement of 18% or 8 less antivirus vendors that will flag Netcat as malicious software. Shaddam's protector managed a best of 41% detection rate but included a backdoor that wanted to infect the testing machine. Yoda's protector achieved a 55% detection rate and was outperformed by manual editing from a five year old tutorial by (Team 5150, 2006) on how to hide Netcat from antivirus software. The manual editing involved changing four INT 3 instructions to NOP's to remove the antivirus signature from the file and was surprisingly effective to this day.

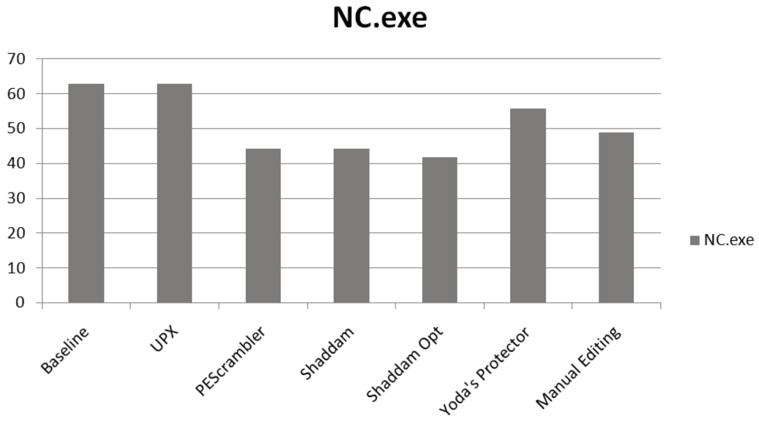


Figure 3: NC.exe VirusTotal detection rate

The Social Engineering Toolkit combined in the Metasploit Framework performed most admirably and achieved impressive results. The first option was to establish a meterpreter reverse connection by embedding it in a Win32 template executable and it was detected by 55% of the antivirus vendors participating on Virustotal. Secondly the SET framework option to encode a meterpreter module to Shellcode was employed and it achieved an extraordinarily low detection rate of only 2%. The problem with this type of antivirus avoidance is that it is more or less limited to being run from either a CD/DVD or USB drive and that there are two parts to the executable namely the Shellcode interpreter and the Shellcode. One of the options that defeated all antivirus detection was the 64 bit Windows Shell Reverse connection and it was as simple to create as starting the SET framework, selecting option 4 (Create a payload and listener) and then selecting option 6 (Windows Shell Reverse TCP X64).

For the MSFEncode example a normal meterpreter shell was encoded with the various encoding schemes available in the Metasploit framework. The results are interesting due to the fact that with no encoding a 69% detection rate was obtained that dropped to 62% with a polymorphic encoder encoded once. If the encoder was run ten times, the detection ratio climbed back to the same level as when no encoding was applied. The results were similar for the XOR encoder where the once off encoding led to a lower detection score than when the encoder was set to an iteration of ten times. As a side note the polymorphic encoder did not perform significantly better than the normal XOR encoder but a possible reason for this could be the small size of the executable.

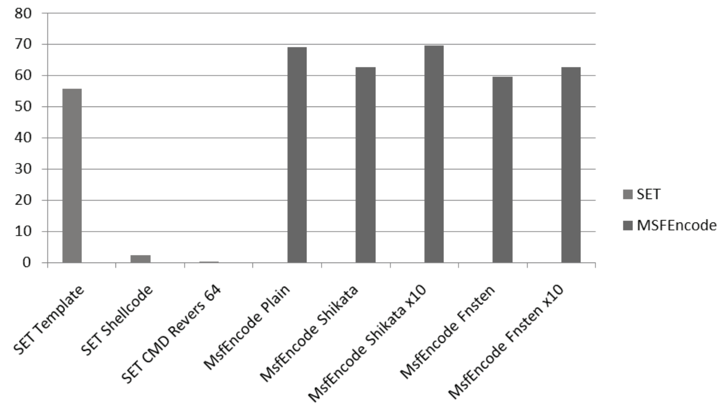


Figure 4: Social engineering toolkit detection rate

It should be noted that no manual editing was performed on the generated executables and if this was employed, a further drop in the detection rate could have been achieved. As an additional test a custom python module was written to open a socket on a target machine and connect to a remote machine. The lines of code required to perform this task in a Python module is less than 20 lines and similarly a custom receiver on the attacker machine is also coded in less than 20 lines. Once deployed on the target machine, no antivirus alerted the user to the newly created connection and Virustotal also gives the executable a clean bill of health with a 0% detection rate. With this very simple network communication enabled, an attacker can already perform various functions such as information gathering, Shellcode deployment and privilege escalation.

6. Conclusion

From the results obtained it was clear that antivirus software is effective against a multitude of threats but custom malware will in all likelihood not be detected. This could lead one to believe that malware created to target the whole internet indiscriminately will eventually be uploaded to a antivirus vendor and classified as malware to be detected from there on out. However malware that target specific organisations will have a much lower chance of being detected and thus a much higher chance of performing its malicious intentions for a longer period. Assessing the skill level of a malware creator is quite problematic, but from the examples investigated it is clear that an advanced degree is not required, especially if the tools such as the SET and Metasploit frameworks are considered. The software used are all opens source applications, and the code examples are freely available from the internet for anyone with a modicum of programming knowledge to follow. The conclusion is thus that with open source tools, it is quite possible to create malware that will completely avoid antivirus detection or at least avoid the antivirus solution deployed by the selected target organisation. If an attacker is determined enough to keep on trying it is all but a certainty that eventually antivirus avoidance will be achieved and organisations should keep this in mind when planning their security.

References

Brand, M. (2011) "*Analysis of avoidance techniques of malicious software*", [Online], Available at: http://www.ruxcon.org.au/2011-talks/analysis-avoidance-techniques-of-malicious-software/ (Accessed on 22/01/2012).

da Silva dos Santos Silva, C. (2009) "*RAVE - Final Thesis Report*", [Online], Available at: http://docs.di.fc.ul.pt/jspui/bitstream/10455/3289/1/RAVE_-_Final_Thesis_Report_-_Carlos_Silva.pdf (Accessed on 14/12/2011).

Daoud, E.A., Jebril, I.H. & Zaqaibeh, B. (2008) "*Computer Virus Strategies and Detection Methods*", International journal of open problems in computer science and mathematics, Vol 1, No. 2, pp.29-36.

Davis, M. (2011) "*Getting Dirty with GCC*", [Online], Available at: http://www.ruxcon.org.au/assets/Presentations/2011-2/ruxcon2011.pdf (Accessed on 06/01/2012).

Debrey, S. & Patel, J. (2010) Reverse Engineering Self-Modifying Code: Unpacker Extraction. In *17th Working Conference on Reverse Engineering*. Beverly, MA, 2010. IEEE, pp 131-140.

Guo, F., Ferrie, P. & Chiueh, T. (2008) A Study of the Packer Problem and Its Solutions. In *Proceedings of the 11th international symposium of recent Advances in Intrusion Detection*. Massachusetts, 2008. Springer-Verlag Berlin, pp 98-115.

Hu, X. (2011) *“Large-Scale malware Analysis, Detection, and Signature Generation”,* [Online], Available at: http://kabru.eecs.umich.edu/papers/thesis/thesis_0902_01PM.pdf (Accessed on 12/12/2011).

Kanich, C., Checkoway, S. & Mowery, K. (2011) Putting out a HIT: Crowdsourcing Malware Installs. In *Proceedings of the 5th USENIX Workshop on Offensive Technologies*. San Fransico, 2011, pp 6-13.

McAfee. (2011) “*McAfee Quaterly Threat Report*”, [Online], Available at: http://www.mcafee.com/us/resources/reports/rp-quarterly-threat-q3-2011.pdf (Accessed on 16/01/2012).

Nataraj, L., Yegneswaran, V., Porras, P. & Zhang, J. (2011) A comparitive assessment of malware classification using binary texture analysis and dynamic analysis. In *Proceedings of the 4th ACM workshop on Security and artificial intelligence*. Chicago, 2011. ACM New York, pp 21-30.

Nick, H. (2011) “*PEScrambler”,* [Online], Available at: http://code.google.com/p/pescrambler/source/checkout (Accessed on 28/12/2011).

Ollman, G. (2009) *“Damballa Serial Evasion Tactics*”, [Online], Available at: http://www.damballa.com/downloads/r_pubs/WP_SerialVariantEvasionTactics.pdf (Accessed on 11/12/2011).

Patel, M. (2011) “*Similarity Tests for Metamorphic Virus Detection*”, [Online], Available at: http://scholarworks.sjsu.edu/etd_projects/175 (Accessed on 18/01/2012).

Team 5150. (2006) *“Taking back netcat*”, [Online], Available at: http://team5150.com/~random/apps/netcat/Taking_Back_Netcat.pdf (Accessed on 11/12/2011).

Tzermias, T., Sykiotakis, G., Polychronakis, M. & Markatos, E.P. (2011) Combining static and dynamic analysis for the detection of malicious document. In *Proceedings of the European Workshop on System Security (EuroSec)*. Salzburg, Austria, 2011, pp 1-4.

UPX. (2011) *“Ultimate Packer for eXecutables*”*,* [Online], Available at: http://upx.sourceforge.net/ (Accessed on 16/01/2012).

Vasiliadis, G. & Ioannidis, S. (2010) GrAVity: A Massively Parallel Antivirus Engine. In *Proceedings of the 13th International Symposium*. Ottowa, Canada, 2010. RAID, pp 79-96.

Verizon. (2011) “*Verizon Data Breach Investigations*”. [Online] Available at: http://www.verizonbusiness.com/resources/reports/rp_data-breach-investigations-report-2011_en_xg.pdf (Accessed on 12/01/2012).

Overt Information Operations During Peacetime

Selma Tekir
Izmir Institute of Technology, Izmir, Turkey
selmatekir@iyte.edu.tr

Abstract: Information superiority is the most critical asset in war making. It directly addresses the perception of the opponent and in the long term the will of him to act. Sun Tzu's classical text states this fact by the concept of deception as the basis of all warfare. The success in warfare then is dependent on being aware of what's happening, accurately realizing the context. This is the intelligence function in broad terms and mostly open source intelligence as it provides the context. Competitive intelligence is based mainly on open sources and day by day the open source share in the intelligence product is increasing. Present diversified open sources & services represent a methodology shift in war. The two preceding ways have been overt physical acts against military targets in wartime and covert information operations conducted throughout peacetime against even nonmilitary targets respectively. The present methodology must be overt (open) information operations during peacetime. This coincides with a metaphor change as well. It proposes a transformation from a war metaphor into a game metaphor in which there are some playing rules. In fact, the existence of such rules helps in drawing the boundary of the field of competitive intelligence and thus making it a profession. Game metaphor is safer to adopt than war as it's easier to take responsibility in public disclosure scenarios in this case. By following this metaphor, you continue to stay in the boundary of legitimate competition. In other terms, you make a conscious preference in terms of war intensities by choosing to avoid the more intense war forms limited conflict, and actual warfare respectively. Finally, this preference is in accordance with the fundamental point of the Sun Tzu's entire argument: The vision of victory without fighting. To summarize, open source domination in the competitive intelligence lays the ground for the game metaphor that represents a transformation in warfare. The apparent outcome is overt information operations during peacetime. It emerges as the most important tool to fight against deception, thus success in information warfare in the contemporary world.

Keywords: information warfare, information operations, competitive intelligence, open sources, ethics

1. Introduction

Information superiority is the most critical asset in war making. It directly addresses the perception of the opponent and in the long term the will of him to act. The Army's Field Manual 3.0, Operations, (2001) as cited by (Thomas, 2003) describes the attainment of information superiority as capable of putting disparity in the enemy commander's mind between reality and his perception of reality. This definition identifies information superiority as a capability to cause a deviation from reality.

Deviation from reality brings about lack of plausability and persuasiveness, which results in loss in the public support and isolation. Thus, as Sun Tzu stated "deception is the basis of all warfare" (Zi and Mair, 2009).

The army doctrine (DoD, 2001) defines information superiority as "the operational advantage derived from the ability to collect, process and disseminate an uninterrupted flow of information while exploiting or denying an adversary's ability to do the same." This second definition highlights reciprocal intelligence processes.

Information superiority is outcome of a continuous, dependable intelligence process along with a counter-intelligence perspective. In the Notional Information Operations Model that is presented in the Information Operations Primer (Department of the Army, 2006), the friendly intelligence process is termed as "Information Management" while the adversarial one is described by Intelligence, Surveillance and Reconnaissance (Figure 1).

Both sides being a part of an information environment have different perceptions of a single reality, referred to as the situation. In order to achieve the goal of information superiority, information operations are conducted in the environment. Complete and correct assessment of the situation is called situational awareness and it's the high-level product of the information superiority reference model (Perry et al. 2004). The model is composed of three domains with associated functionalities (apparent in Figure 1):

- Physical domain-Ground truth (Entities, systems, intentions, plans, and physical activities.).
- Information domain-Collection and analysis capabilities.
- Cognitive domain-Decisionmaking, taking action.

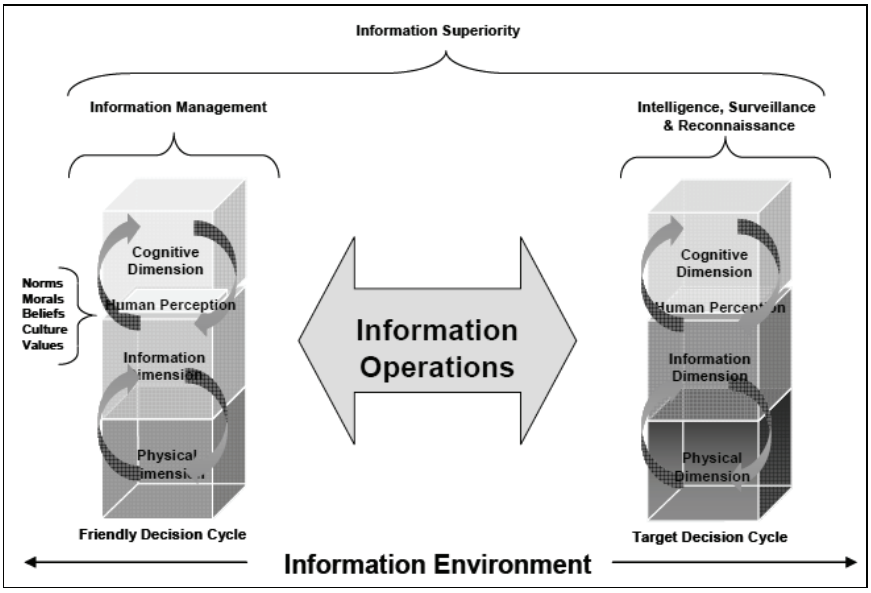


Figure 1: A notional information operations model (Department of the Army, 2006).

Situational awareness requires accurately realizing the context. Open Sources Information (OSI) can be used in this manner as it contributes the understanding of the problem, tells the current situation and the context. Evaluation of the OSI in the first place for the intelligence collection provides a good background and helps the effective and efficient tasking of the other, more difficult collection disciplines saving much of the resources.

Having recognized this potential, competitive intelligence is based mainly on open sources and day by day the open source share in the intelligence product is increasing. Present diversified open sources & services represent a methodology shift in war. The two preceding ways have been overt physical acts against military targets in wartime and covert information operations conducted throughout peacetime against even nonmilitary targets respectively (Waltz, 1998). The present methodology, owing to existing open sources, must be overt (open) information operations during peacetime. In order to clarify the implications of the proposed methodology, it's useful to consult the information warfare and information operations definitions.

The formal U.S. DoD information warfare definition is as follows (Waltz, 1998): Information warfare includes actions taken to preserve the integrity of one's own information system from exploitation, corruption, or disruption, while at the same time exploiting, corrupting, or destroying an adversary's information system and the process achieving an information advantage in the application of force.
The U.S. DoD defines information operations as the employment of the core capabilities of electronic warfare, computer network operations, psychological operations, military deception, and operations security, in concert with specified supporting and related capabilities, to affect or defend information and information systems, and to influence decision-making (Department of the Army, 2006).

The formal U.S. DoD definition of information warfare covers three different information operations namely defense, offense, and information dominance and no precedence assignment was introduced. It's also apparent in the definition of "Information operations", influencing decision-making was not recognized as the core functionality.

The issue of organizing these three information operations; identifying the boundaries, responsibility areas, resource tasking, and the communication and coordination process is helpful in defining an information warfare framework. A partial initialization of such a framework can be found in Libicki (2009).

Overt (open) information operations during peacetime methodology highlights the information dominance as the definitive component of the information operations as it puts an emphasis on information operations that support decision-making. Open source intelligence (OSINT) presents new opportunities in this area.

A second aspect of the proposed paradigm shift is the change in the treatment of war. Traditionally war can only be inferred to mean a “crisis or conflict” situation. This fact can be observed in the JP 1-02 definition of information warfare (DoD, 2001):

“Information operations conducted during times of crisis or conflict to achieve or promote specific objectives over a specific adversary or adversaries.”

Similarly, in the international law provided in the Geneva Convention and the additional protocols “armed conflict” characterizes a war incident (Haslam, 2000). However, here the aim is to avoid more intense war forms like limited conflict and actual warfare and to stay within the boundary of legitimate competition. This viewpoint can be considered in relation to the “before going to war” (jus ad bellum) criteria of Arquilla (1999) as it does not recommend progressing into the stage of waging war. Moreover, the idea reinforces the discipline of competitive intelligence.

2. Competitive intelligence

Competitive intelligence is the result of a trend to incorporate intelligence methodology and tools into private businesses. The trend was triggered by accelerated economic competition around the world that increased demand for better decision-making. Another factor is related to the high-cost of diversified and sophisticated tools through technology that the burden on the government budget has become not withstandable making the privatization a plausible choice.

The adoption of intelligence cycle by the business sector revealed the legal and ethical concerns. Particularly, the collection phase has come into question, OSINT has become the basis.

The ethics of information dominance (intelligence) is just a subset of the ethics of information operations. Within this work, the focus is on this particular subset. The broader area of “the ethics of cyberwarfare” is addressed in (Dipert, 2011).

There are apparently two metaphors in the competitive intelligence world, war and game respectively. War is inherited from the government intelligence and game is the new introduced one for the business sector. Adopting war metaphor, you continue to stay in the old world seeing the main collection effort as espionage. You admit to do whatever is required as long as the final aim is to defeat the opponent. This is not an appropriate vision and/or mission for a competitive intelligence professional.
The metaphor of a game sees competition in business as an exciting game, in which each competitor strives to achieve excellence, satisfy customers, and succeed as a result. The motive in this type of game is not to drive out the competition, but to work hard, play by the rules of the game, and do one’s best in order to succeed (Trevino and Weaver, 1997).

Today, the nature of intelligence has changed dramatically. The information operations are not evaluated within the frame of an armed conflict. They do not have to accompany an armed conflict. Globalization is dominant, global corporations are everywhere. Their interests are not restricted with national or even continental boundaries. Information warfare has been started to be evaluated within the context of organizational decision making.

This status was also supported by the three consecutive streams in Prescott (1999). The stated streams give the history of the intelligence field. The first is described as Sun Zi’s The Art of War, which articulates the philosophical framework for war making and intelligence. The second stream of intelligence puts national security concerns as a policy issue. The third stream places the business organization at center stage, which results in a systematic orientation towards business intelligence.

Warfare in the organizational context targets the perception, decision-making performance, operational effectiveness, and ultimately the will of the organizations.

In an adversarial competition or conflict, organization warfare seeks to understand and then change a target organization’s behavior to achieve the goals of the attacker; generally these goals include deterrence, dissuasion, deception, disruption, degradation, or total defeat of the targeted organization Particularly the adversary’s unity of command and purpose is targeted while preserving your own. In corporate competition, an organization may employ many of the principles and technical methods, herein, to ethically collect, analyze, and understand an adversary to employ ethical methods to influence the market and its competitor. The organization may also employ these methods to detect and deter an adversary's unethical efforts at manipulation (Kott, 2006).

The planner must consider the legality of organization influence operations and the potential collateral effects on civil populations to remain compliant with directly relevant international legal protocols (Geneva Conventions has a protection focus on civilian population, individual civilians or civilian objects).

In the organizational warfare, in targeting an organization for impact, there are two complementary approaches: Targeting the critical nodes of the organization, and focusing on the entire organization as a social entity. The former one puts an emphasis on individuals while the latter on the whole structure. People component deserves special attention as it is the one that transforms a business into a corporate asset.

As the competitive environment with globalization could be characterized by the game metaphor rather than the war metaphor (the traditional metaphor usually used for characterizing competition), it is increasingly important to include ethics in the corporation's strategy and potentially implement it in a way that achieves a competitive advantage for the company and adds value to the stakeholders (Azmi, 2006).

Contrary to the belief, game metaphor is not weaker than the war counterpart. Even using legitimate means one can acquire and process the required information. It's also advantageous to mutually agree on some set of rules as it also restricts the diversity of means available for the opposing parties. Thus, there is a decreased uncertainty in the competition.

Another concern is related to the characteristics of the digital environment. In the digital environment, the footprints cannot be concealed. Now or later it has high potential to be revealed. Being aware of this fact, staying in accordance with the rules of the game, one preserves his organizational reputation. Organizational reputation and its associated assets are of high importance.

Additionally, staying in the boundary of ethics helps in forming an organizational culture that will return benefits in the long term. This can be explained by corporate ethics rather than business ethics. On the one hand, business ethics has an external emphasis. Business ethics considers the gap between the corporation's ethical behaviour and the market place's perception of the corporation's ethics in its business operations. Corporate ethics, on the other hand, has an internal emphasis and this could be well managed toward a unique competitive advantage as anything related to people (corporate ethics through people) is very difficult to imitate and this raises the chances of achieving a sustainable competitive advantage (Azmi, 2006). By the use of corporate ethics, an investment on people is made and it will always have returns of vital value.

Game metaphor has an important mission as well: To draw the ethical boundary of the competitive intelligence discipline. CI practitioners need to have the profession recognized by external stakeholders as being legitimate and ethical. As the adoption of ethical standards is a hallmark of a profession, it paves the way for the competitive intelligence profession (Fleisher and Blenkhorn, 2000).

3. Conclusion

Competitive intelligence originated from the government intelligence. It borrowed the war metaphor from the traditional intelligence viewpoint. Game metaphor is in accordance with the conditions of the global economic environment, corporate cultures, civil emphasis, open sources & services, and thus more appropriate today. The above stated conditions started to put a pressure on all intelligence organizations that must also share them. Consequently, competitive intelligence has high potential to change its originating source by causing new reflections on the concept of war which can bring about redefinitions of information warfare and related concepts.

An example evidence is related to public disclosure. Game metaphor is safer to adopt than war as it's easier to take responsibility in public disclosure scenarios in this case. By following this metaphor, you continue to stay in the boundary of legitimate competition. In other terms, you make a conscious preference in terms of war intensities by choosing to avoid the more intense war forms limited conflict, and actual warfare respectively. Finally, this preference is in accordance with the fundamental point of the Sun Tzu's entire argument: The vision of victory without fighting (Zi and Mair, 2009).

To summarize, open source domination in the competitive intelligence lays the ground for the game metaphor that represents a transformation in warfare. The apparent outcome is overt information operations during peacetime. It emerges as the most important tool to fight against deception, thus success in information warfare in the contemporary world.

References

Arquilla, John. (1999) "Ethics and Information Warfare", [online], RAND Corporation, http://www.rand.org/content/dam/rand/pubs/monograph_reports/MR1016/MR1016.chap13.pdf (accessed on 23/03/12).

Azmi, Rania A. (2006) "Business Ethics as Competitive Advantage for Companies in the Globalization Era", [online], http://ssrn.com/abstract=1010073 (accessed on 31/01/12).

Department of the Army (2006) AY07, Information Operations Primer, Headquarters, Department of the Army, U.S. Army Printing Agency, Washington, D.C.

Dipert, R. R. (2010). "The Ethics of Cyberwarfare." **Journal of Military Ethics** Vol 9 Is 4, pp 384-410.

DoD (2001) Joint Publication 1-02: Department of Defense Dictionary of Military and Associated Terms, [online] http://www.dtic.mil/doctrine/new_pubs/jp1_02.pdf (accessed on 02/02/12).

Fleisher, Craig S., Blenkhorn, David L. (2000) **Managing Frontiers in Competitive Intelligence**, Quorum Books.

Haslam, Emily. (2000) "Information Warfare: Technological Changes and International Law", **Journal of Conflict and Security Law**, Vol 5 No. 2, pp 157-175.

Kott, Alexander. (2006) **Information Warfare and Organizational Decision-Making**, Artech House Publishers.

Libicki, M. (2009) "Cyberdeterrence and Cyberwar", [online], RAND Corporation, http://www.rand.org/pubs/monographs/2009/RAND_MG877.pdf (accessed on 22/03/12).

Perry, Walter L., David Signori and John E. Boon. (2004) "Exploring Information Superiority: A Methodology for Measuring the Quality of Information and Its Impact on Shared Awareness", [online], RAND Corporation, http://www.rand.org/pubs/monograph_reports/MR1467 (accessed on 31/01/12).

Prescott, J. (1999) The Evolution of Competitive Intelligence, Designing a Process for Action, APMP, Spring.

Thomas, T.L. (2003) "Is the IW Paradigm Outdated? A Discussion of U.S. IW Theory", **The Journal of Information Warfare**, Vol 2 Is 3, pp 109-116.

Treviño, Linda Hlebe, Weaver, Gary R. (1997) "Ethical issues in competitive intelligence practice: Consensus, conflicts, and challenges", **Competitive Intelligence Review**, Spring Vol 8 Is 1, pp 61-72.

Waltz, Edward L. (1998) **Information Warfare Principles and Operations**, Artech House, Inc.

Zi, Sun Mair, Victor H. (2009) **Art of War: Sun Zi's Military Methods**, Columbia University Press, New York.

Novel Tracking of Rogue Network Packets Using Danger Theory Approach

Solomon Uwagbole, William Buchanan and Lu Fan
Centre for Distributed Computing, Networks and Security, Edinburgh Napier University, Edinburgh, UK
s.uwagbole@napier.ac.uk
w.buchanan@napier.ac.uk
l.fan@napier.ac.uk

Abstract: Recently there has been heightened, continuous, and intrusive activity by remotely located rogue hacking groups, such as Anonymous and Lulzsec. These groups often aim to disrupt computer networks and gain access to private confidential data. A typical method used to steal confidential data is by SQL Injection (SI).This problem is likely to increase as Cloud Computing gains popularity, thereby moving organisations' network security boundaries, firewall, deeper into the internet cloud environment. There is thus a strong requirement for a real-time framework that detects and mitigates any intrusion activities as, and when, they occur. Conventional firewalls lock down ports and applications, but often does little against malicious packets stealthily concealed in legitimate network packets payload, thus a framework that solely depends on network packets payload analysis for malicious finger print, rather than traditional system calls and processes is required. This paper thus presents a novel framework that introduces the vaccination of Danger Theory's Dendritic Cell Algorithm (DCA) for the real-time detection and mitigation of network intrusions. The proposed framework draws an inspiration from the active and passive biological immune systems in which the human body has an efficient autonomous response to fight infections on encountering danger signals to indicate anomalies in cellular activities. This immunological principle is widely adopted in the computational field of study of Artificial Immune Systems (AISs). To achieve this novel bio-inspired computational framework of detection and response, there is research work in progress using.NET Framework implementation of DCA. There are two stages to this implementation which are creating detecting receptors input data to train DCA, and finally, using the trained DCA in real-time for detecting anomalous network packets payload. Take an example of database security exploit of SI that is discussed in this paper. Stage one involves creating detector precursor (receptors) by subjecting a database to be protected to a controlled SI scripts or code with the network packets payload of such exploits captured in real-time by using .NET custom built packets analyser. Stage two involves real-time monitoring of protected databases for anomaly (antigens) through the trained DCA by using r-contiguous rule to match receptors with antigens in the data pre-processing stage when immature Dendritic Cell (DC) is transformed to semi-mature or matured. The structure of SI packets is now constructed to easily isolate SI malicious packets from legitimate network packets payloads between known source and destination of confidential data request. The approach in brief is: protected data or assets are modelled as cells in tissues to be monitored; while rogue network packets triggers the computational modelled DCs to co-stimulate B and T Cells as to provide detection feedback to the protected cells. The outcome of this paper can be practically applied in: detecting an attempt to steal protected data and applications by a rogue remote intruder; and detection of man-in-the-middle attacks on applications that sit in cloud. The proposed bio-inspired approach to resolving SI computer systems security challenges is a research work in progress by this paper's author. The research proposes an easy adaptation of the system to any domain as the finger-print required for detection and training the system is now introduced by vaccination method.

Keywords: intrusion, detection, immunised network, danger theory

1. Introduction

The issue of computer security breaches was topical in 2011. Hacker groups such as Anonymous, Lulzsec, and others, have targeted a wide range of persons from private individuals to governments and corporate organisations with the intruders' main objectives being to steal confidential data and disrupt computer networks. Examples of victims include Sony, PayPal, FBI, World Bank, and so on. The most common method used is SQL Injection (SI) which is a technique where intruders append web application query strings with SQL command statements to circumvent both front-end and back-end validations. One thing that has constantly emerged from these numerous rogue hacking activities was the vast amount of data, containing both confidential personal and credit card details, which were stolen. In most cases no alarms were triggered to alert security experts or stakeholders that hacking activities were in progress. These intrusive activities by cyber criminals are often used for blackmailing; ransom demands; and political statements. Financial gains can be made by selling this illegally obtained information to allied organised criminals who then use it to commit crimes such as identity theft and fraudulent financial transactions on behalf of unsuspecting members of the public.

A hacker, or intruder, is an individual or group that breaches the security of computer systems by employing an array of techniques. These hacking attempts go beyond circumventing the conventional firewall allowed programs and ports, but instead piggy-back on legitimate packets payload for malicious activities including the theft of organisations sensitive data. A typical example is SI which is done by running an advanced Structured Query Language (SQL) script request to the database server to retrieve unauthorised data. Mere blocking of Microsoft SQL server communicating port 1433 on firewall is of no help. There seems to be little in the way of an immediate fix to these security breaches, as can be seen by repeated hacking activities resulting in the theft of vast amounts of data at Sony since April 2011 (Albanesius 2011). Robot Network (Botnet) is another example whereby a rogue user takes up the command and control of many computers on the internet at the detriment of unsuspecting legitimate users and owners of such systems. If there is a distributed network attack that could be probable to overwhelm the host or server as to result in Distributed Denial of Service (DDoS); a packet payload analysis could recognise these malicious packets and isolate them.

Therefore there is a need for an intrusion detection system that can monitor, in real-time, what is being transmitted from one source to its destination, and is able to learn and take proactive action. The spate of hacking activities in 2011 across different high profile organisations having considerable investments in research and development in security is evidence that there is a need for functional real-time network monitors that detect and abort intrusive activities. The widely used signature-based intrusion detections cannot handle the ever-evolving polymorphic network intrusion techniques employed by rogue hackers. In seeking a solution to remedy this problem; an analogy between Intrusion Detection Systems (IDS) and the biological human immune system has been explored in anomaly detections (Kephart 1994). This paper presents a novel implementation of Dendritic Cells Algorithm (DCA) primed with the features of SI scripts or code to determine the finger-print in computer network packets structure for robust and scalable detection.

2. Background

SQL injection (SI) is a database security exploit from front-end user interfaces whereby an intruder supplies the front-end web application with an advanced SQL statement or code appended to the normal URL query strings; thereby circumventing the back-end database validations to retrieve sensitive data. SI is an intrusion in which malicious code is inserted into strings that are transverse to an instance of SQL Server for parsing and execution (Microsoft n.d). There is no fool proof solution to SI as the vulnerabilities are attributed to: outdated system patches; bad coding; and out-of-sight of the issue practices by developers. Even if the problems are resolved in new enterprise applications implemented by developers adopting good coding practices; as the vulnerabilities have existed in the databases and applications to access them for years, it will be hard to totally eradicate the issues from legacy applications.

Figure 1 shows a normal user validated on the web page while in a compromised database the normal login page has been bypassed with an intruder adding to the Universal Resource Locator (URL) with a query string appended with an advanced SQL statement instead of the developers ADO.NET legitimate SQL statement for retrieving the data for validated users. Having circumvented the front-end web application from now onward the intruder capability to the database is endless. He can inject his own login credentials and virtually take over the database and harvest the confidential data content. The novel vaccinated Dendritic Cell Algorithm (vDCA) presented in this paper will allow retrofitting this approach to SI detection in both existing, and new database setups with possible SI vulnerabilities as shown in Figure4 below.

3. Related work

Researchers in computational fields have through a cross discipline of immunology and computer science developed algorithms by modelling computational abstracts from the immune system theories, processes and elements (Dasgupta, Yu and Nino 2011), and representing detection and recognition in geometrical shape space. These algorithms are constantly being innovated and have served as a reference point in applied AIS research to address computational issues in anomaly detections, computer security, optimisation and data mining, etc. These algorithms and models are as follows:

- Negative Selection Algorithm (NSA) (Forrest et al, 1994;Ji and Dasgupta 2007).
- Artificial Immune Network Algorithm (AIN) (Jerne 1974).

- Clonal Selection Algorithm (CLONALG) (De Castro and Von Zuben2000).
- Danger Theory (Aickelin and Cayzer2002) and DCA (Greensmith 2007).

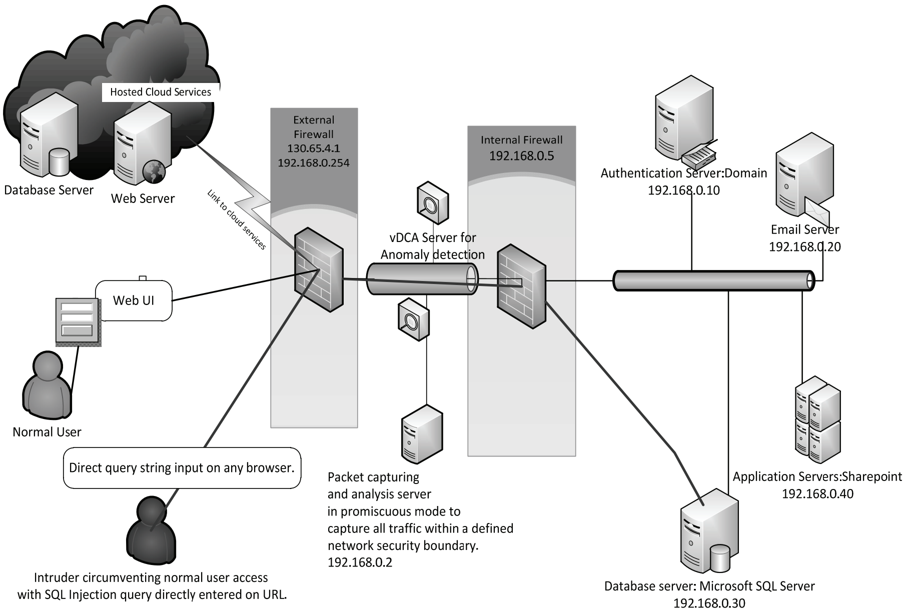

Hosted Cloud Services
Database Server
Web Server
Link to cloud services
External Firewall
130.65.4.1
192.168.0.254
Internal Firewall
192.168.0.5
Authentication Server:Domain
192.168.0.10
Email Server
192.168.0.20
vDCA Server for Anomaly detection
Web UI
Normal User
Direct query string input on any browser.
Packet capturing and analysis server in promiscuous mode to capture all traffic within a defined network security boundary.
192.168.0.2
Application Servers:Sharepoint
192.168.0.40
Intruder circumventing normal user access with SQL Injection query directly entered on URL.
Database server: Microsoft SQL Server
192.168.0.30

Figure 1: A Network architecture of the proposed implementation of vaccinated DCA

Research work in artificial immune system architecture referred to as (ARTIS) in which monitoring of network services, traffic and user behaviour are observed to detect any deviation from normal behaviour patterns(Hofmeyr and Forrest 1999, 2000; Hofmeyr, Somayaji and Forrest 1998). A further adaptation of ARTIS called LISYS examines the broadcasts source and destination of each TCP SYN packets to a detection node to check for anomalies. The latent time for detectors to confirm anomalies can be an issue here. Further work has been done employing LISYS with NSA in hybrid artificial immune system and Self Organising Map for network intrusion detection (Powers and He 2008)

These first generations of AISs above were adaptive immunity inspired which was modelled on the principles of the classical immunological concept of discrimination between Self/Non-Self (SNS) but subsequent second generations of AIS are links between innate and adaptive immunity. Burgess (1998) was critical of the implementation of NSA's SNS, and concluded it was too simplistic to explain the whole complex human immune system representation to solve computational issues; decided on an approach using Danger Theory (Matzinger 1994) which he considered appropriate to solve the computational complex abstraction from immune system. Following this model, Burgess (2000) implemented AIS on an autonomous and distributed feedback and healing mechanism, triggered when a small amount of damage could be detected at an initial attacking stage. The system he named CFengine was DT inspired based on statistical methods of detecting anomalies. It is now established that the innate immune system also controls the adaptive immune system (Schenten and Medzhitov 2011).

In the last decade new approaches to computer security anomalies detection has taken inspiration from Matzinger's danger theory on an immunological concept which is a new notion to immunological understanding; a shift in paradigm from the widely held SNS paradigm on immunology. Aickelin and Cayzer(2002) published their novel paper, inspired by Matzinger's danger theory, titled The Danger Theory and Its Application to Artificial Immune Systems called the DT. This paper drew reference from the human immune systems capability to respond to danger signals caused by necrotic cells (unnatural death of cells). There are many implementations of this DT by researchers in attempts to address issues relating to computer security but of which the DCA stands out in terms of functionality and results. The DCA (Greensmith 2007) is a bio-inspired innate immunity computational algorithm modelled on both the innate and adaptive principles. The DT concept in intrusion detection is modelled like the Dendritic Cell (DC) of the neuron seeking out danger signals when there is a sudden increase in computer network traffic. Algorithms inspired by DT are the DCA (Greensmith 2007) and Toll-like Receptor algorithm (Twycross 2007). The DT was extended for computer network anomaly detection in (Yeom 2007). Al-Hammadi, Aickelin and Greensmith(2010) explored Botnet detection using DCA. DCA has had a high success rate in intrusion detection but not in responding to an attack.

4. Danger theory (DT)

This paper restricts the AISs theory discussion to the DT as it is the approach used here in addressing SI. The discourse on immunity has led to two notable theories that stand out from the rest which are SNS and DT. DT proposed by Matzinger (Matzinger 1994) explains immunological responses triggered by the presence of a danger signal as a result of unnatural death of a cell; which is a biological processes called necrosis as against apoptotic which is the natural programmed death of cells.

The immune system as a robust validation machine was accounted for in Bretscher and Cohn (1970) in which it extended the two-signal model likening to danger model. The lymphocytes require two signals to become activated. These signals are antigen recognition (signal one) and co-stimulation (signal two) that is carried and presented by DC as Antigen Presenting Cell (APC). Co-stimulation is a signal that confirms that a signal deemed dangerous really is dangerous. Matzinger (1994) further applied the laws of lymphatic to the Danger Theory:

- Law 1: Two signals are needed to activate the lymphocyte. The survival of a lymphocyte depends on receiving the complete two signals or else it will not survive if it receives signal one without the co-stimulation of signal two. There has to be the existence of signal one for signal two to have meaning else in the absence of signal one, signal two will be ignored.
- Law 2: Signal one can originate from any cell. Thus, signal two comes from APC, however, the signal two for B cell activation emanate from T helper cells.
- Law 3: Activated (effectors) cells do not need signal two rather a trigger from distress cell, which revert to resting state after a short time. Immature cells are unable to accept signal two from any source as they are triggered by necrotic cell death.

Figure 2 illustrates DC migration from tissue to lymphatic node. The presence of a distressed cell or unnatural cell death (necrotic) activates the DC to engulf the debris which in the presence of MHC class is presented on the surface as APC which consequently co-stimulate Helper T cells. These, in turn, help the B cells to activate specific immune response. The computation models built on this biological phenomenon can be seen in DT with DCA as popular example.

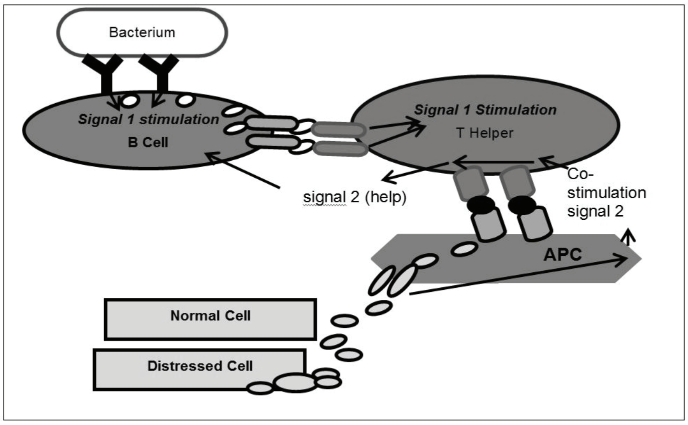


Figure 2: Danger model immune response activation diagram (Pereira 2011)

4.1 Danger model (DM)

The DT offers an alternative: the immune system reacts when danger is detected, not non-self *(*Danger Project n.d).This model included all body tissue cells and signalling to the immunological discussion. The in silico application of the Danger Theory was first presented in 2002 by Aickelin and Cayzer (2002). This paper pointed out some analogies to DM of artificial immune systems in the Danger Theory, which are highlighted in the following salient points:

- An APC is necessary to present an appropriate danger signal.
- The “Danger” signal may have nothing to do with danger.

- The appropriate danger signal can be positive or negative.
- Measures of proximity employed geometry shape space are used to mimic the danger zone.
- An immune response should not lead to a further loop of danger signals.

The bio-inspired algorithm presented in this model is the DCA widely explored in intrusion detections.

4.2 Dendritic cells algorithm (DCA)

The signals of DCA (Greensmith 2007) comprises of PAMP, danger and safe signals. PAMP indicates the presence of microbial and necrotic danger signal resulting from distressed cell death as to convert the immature dendritic cell to a reactive matured Dendritic Cell (DC) in the lymph node for analysis; while safe signals result in tolerant semi-mature DC. DCA is widely explored in intrusion detection or anomaly detection in computer network security and medical data classification. The DCA from its definition offers danger detection as against danger response. The novel approach discussed in this paper would offer, in addition to detection, a faster response with less computation. Table 1 below explains the immunological terminology used in relation to computational context in DCA with the suggested DCA with vaccination. While the DCA is composed of three main processes in computational abstraction; there is an additional process of vaccine integration to vDCA as:

- Controlled Inoculation of the Database asset with SI.
- Creation of DC.
- Exposure of the created DC to the tissue environment.
- Evaluate the data collected by the mature DCs.

4.2.1 Creation of DCs

In order to manipulate DCA there is a need for relevant abstract data modelled from DCs creation. These are: Maturation Stage; Antigen Storage; Cumulative Output Signals; and Migration Threshold as illustrated in Figure3 below.

- Maturation Stage starts with immature DCs which differentiate into either semi-mature or mature DCs based on the signalling type and location.
- Immature DCs are monitoring the tissue for signals which will decide whether they transform into semi mature or mature DCs. They do not partake in any danger evaluation.
- Semi-mature DCs are immature DCs that receive safe signal as to migrate to the lymph node. They have tolerance relationship with the antigens they carry.
- Mature DCs are immature DCs that receive danger signal as to migrate to the lymph node. They have tolerance relationship with the antigens they carry.
- Antigen Storage is the memory location in the tissue or monitoring environment.
- Cumulative Output Signals is a deciding factor in DCs migration to lymph node that must be either safe or danger signal.
- Migration Threshold is a time value of transition from sampling environment or monitoring environment to lymph node. It is the duration in the monitoring environment (tissue).

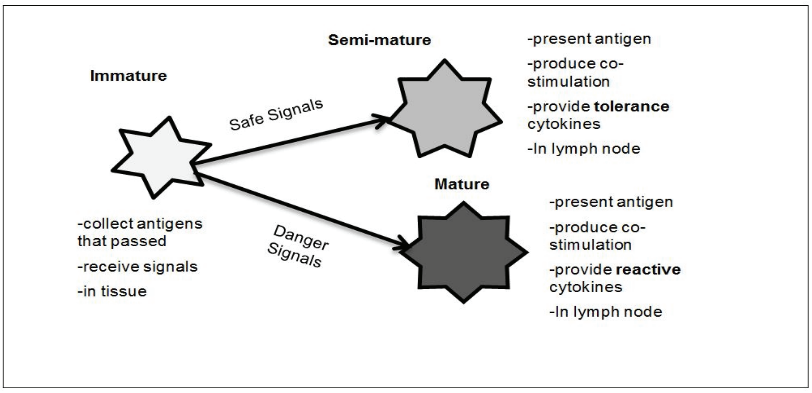


Figure 3: An abstract model of DCs maturation pathway

Table 1: Novel vDCA integrated with DCA immunological terminology and computational context.

Immunological Terminology	Computational Context
Vaccines	A controlled dose of SI scripts/code to produce a computer network packets pattern of antibodies making the system primed for easy detection of the vaccinated against feature.
Antibodies	A sterile recognised packets payload pattern feature of SI.
Tissue	A computational asset to monitor such as a database containing confidential data.
Antigen	A danger triggering entity or non-sterile anomalies in the computer network e.g. actual SI appended to query string.
Signal	Network traffic packets
Lymph node	Processing centre for computer network packet payloads

4.2.2 Exposure of the created DC to the tissue environment

The pseudo code in Algorithm 1 (Greensmith 2007) illustrates the computation steps in DC maturation stages. It evaluates the signal type to decide whether it is a danger signal or not. There are two types of signals which are input and output. Input phase (Line 1) is the triggering signals to immature DC to start maturation to semi-mature or mature depending on the signal type as shown in Table 2. Initialisation phase (Line 5) is the duration in the monitoring environment that is triggered on receiving input signals which is a value of the migration threshold. Tissue environment sampling phase (Line 7 - 12) is the collecting and storing of antigen from the environment. This is followed by updating the output (lines 2 and 4) according to Equation 1 by retrieving the signal input from the environment.

4.2.3 Evaluation of the data collected by DCs

The proposed system implementation will evaluate various DC cells based on the “mature context antigen value” (MCAV) which determines the anomaly coefficient for any given antigen type. MCAV is calculated by Equation 2 where MCAV value of a given antigen type is equals to the average of DCs cell that registered that antigen type divided by the mature context. The closer the value is 1, the more likely for it to be anomaly. Parameters used in the output Equation 1 is detailed in Table 3.

```
1.  Input : Signals of all types including controlled vaccines
2.  Output 1: Antigen context value of 0/1
3.  //closer to 1 is danger  antigen presence
4.  Output 2: Antibodies
5.  Initialize DC
6.  //The DC is in the tissue
7.  while CSM output signal < Migration Threshold do
8.  get antigen;
9.  store antigen;
10. get current values for input signals;
11. update cumulative output signals;
12. end while
13. //The DC enters the lymph node
14. if semi-mature output signal > mature output signal then
15. set cell state as semi-mature;
16. else
17. set cell state as mature;
18. end if
19. //The DC dies and communicates the information collected
20. kill cell
```

Algorithm 1: DCA

Table 2: vDCA input signals

Signal	Biological Property	Abstract Property	Computational Example
PAMP	Indicator of Microbes	Signature of likely anomaly	SI finger print
Danger Signals	Indicator of tissue damage	High level indicate “potential” anomaly	Rogue computer network traffic(unknown destination)
Safe Signals	Indicator of healthy tissue	High levels is an indicator of normally functioning system	Computer network packets size
Inflammation	Indicating general distress	A factor of all other input signals	User absent
vaccines	Controlled microbes	Specific signature	Specific SI finger print

$$Output_{i+1} = Output_i + (P_w * \sum_i P_i + D_w * \sum_i D_i + S_w * \sum_i S_i) * (1 + I)$$

Equation 1

Table 3: Parameters used in the output Equation 1

P	Input signal PAMP
D	Danger Signal
S	Safe Signal
w	Index for specific weight
i	Index for a given input signal

MCAV of antigen type = mature count / antigen count   Equation 2

5. Novel vaccinated DCA (vDCA) -detection

This paper introduces vaccinated DCA (vDCA) which offers to detect and respond to danger signals. In humans a shot of injection containing weakened or attenuated dose of bacteria or virus co-stimulate B and T cells into producing antibodies for the lifetime. Therefore inoculating babies when they are born has eradicated diseases such as measles. This paper explores inoculation discussion to DCA implementation as it will provide predictable computer network packets payload for a quick detection and elimination of SI or any domain of interest the approach is being applied to. Though the salient features of DCA in detecting anomalies is still a part of the system; the novelty approach now adds the ability to fine grain DCA to a particular problem of study without going back to the drawing board each time there is a new problem domain in which to apply DCA. Take an example of inoculating DCA implementation with a controlled known SQL statement and code used in SI exploits to emit pre-determined computer network packets pattern that is eliminated from computer network traffic as soon as they arrive as data stream across the computer network as shown Figure 5 below.

To draw an example of the practicality of the work in progress AIS approach to anomaly detection; a protected database is inoculated against known SQL Injections online or offline with all permutations of the network packets generated and captured using .NET SharpPcap (Morgan 2011) API as shown in Figure 4. These inoculated captured network packets are catalogued and stored in the repository to be used as receptors as illustrated in Figure 5 which are then matched against live antigens input data in a deployed environment .This measure of similarity leads to the r-contiguous symbol rule(Forest et al 1994) for the definition of the recognition region. This rule states that an antigen is recognised by a receptor if the length of the longest corresponding substring that the antigen has in common with the receptor is greater than or equal to the receptors. The value of receptors corresponds to the threshold value that determines the specificity when an antigen is matched against the receptor. Depending on this threshold value the packets may be classified towards safe signal or non-safe signal.

DCA has two versions in implementation (Greensmith 2007); a libtissue (Tywcross 2007) system implemented which is version 1.0, and C++ implementation which is the version 0.1. AIS researchers have validated the DCA (0.1) with different programming languages of their choice; C# is being used in this paper’s implementation. DCA is implemented in C# whereby the real-time input data is fed in by C# network packets capturing tool that uses SharpPcap (Morgan 2011) dynamic link library that provides cross platform interface from the .NET framework to low level network monitoring drivers of WinCap for Windows and LibCap for UNIX. The implementation is a server-server application that can exist as a service programs in existing servers but preferably dedicated servers for packet capturing and vDCA.

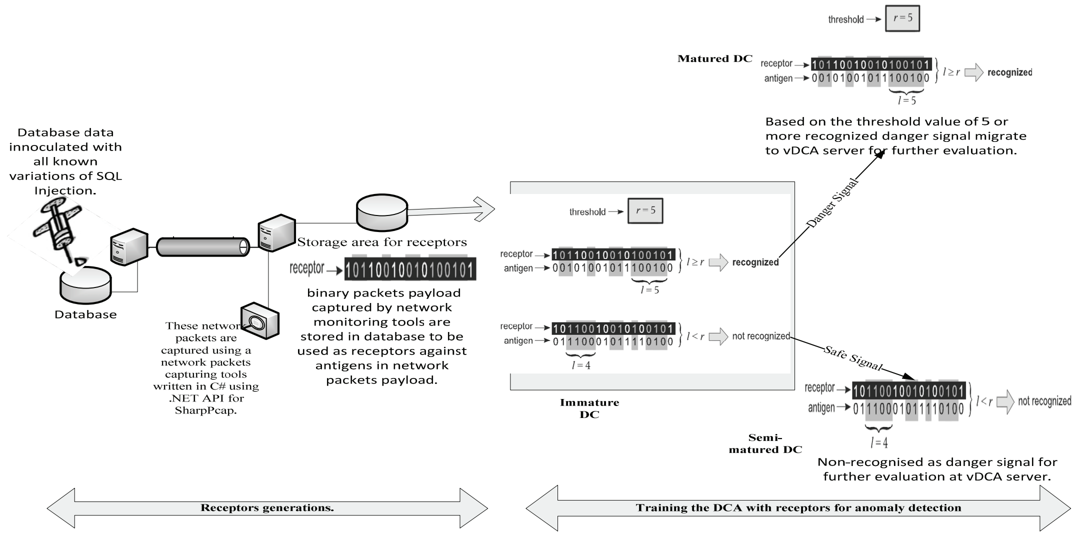

threshold → r = 5
Matured DC
receptor → 1011001001010100101
antigen → 0010100101011100100
l ≥ r recognized
l = 5
Based on the threshold value of 5 or more recognized danger signal migrate to vDCA server for further evaluation.
Database data innoculated with all known variations of SQL Injection.
Database
These network packets are captured using a network packets capturing tools written in C# using .NET API for SharpPcap.
Storage area for receptors
receptor → 1011001001010100101
binary packets payload captured by network monitoring tools are stored in database to be used as receptors against antigens in network packets payload.
threshold → r = 5
receptor → 1011001001010100101
antigen → 0010100101011100100
l ≥ r recognized
l = 5
receptor → 1011001001010100101
antigen → 0111000101011110100
l < r not recognized
l = 4
Immature DC
Danger Signal
Safe Signal
receptor → 1011001001010100101
antigen → 0111000101011110100
l < r not recognized
l = 4
Semi-matured DC
Non-recognised as danger signal for further evaluation at vDCA server.
Receptors generations.
Training the DCA with receptors for anomaly detection

Figure 4: Creating receptors in controlled pre-deployed environment

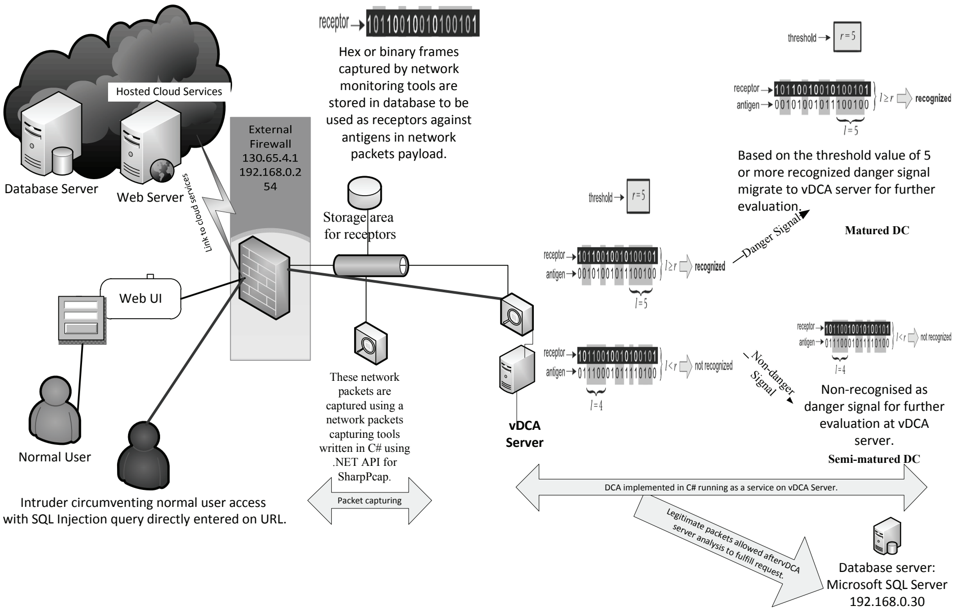

receptor → 1011001001010100101
Hex or binary frames captured by network monitoring tools are stored in database to be used as receptors against antigens in network packets payload.
Hosted Cloud Services
Database Server
Web Server
External Firewall 130.65.4.1 192.168.0.2 54
Link to cloud services
Storage area for receptors
Web UI
Normal User
Intruder circumventing normal user access with SQL Injection query directly entered on URL.
These network packets are captured using a network packets capturing tools written in C# using .NET API for SharpPcap.
Packet capturing
threshold → r=5
receptor → 1011001001010100101
antigen → 0010100101011100100
l ≥ r recognized
l=5
receptor → 1011001001010100101
antigen → 0111000101011110100
l < r not recognized
l=4
vDCA Server
threshold → r=5
receptor → 1011001001010100101
antigen → 0010100101011100100
l ≥ r recognized
l=5
Based on the threshold value of 5 or more recognized danger signal migrate to vDCA server for further evaluation.
Danger Signal
Matured DC
Non-danger Signal
receptor → 1011001001010100101
antigen → 0111000101011110100
l < r not recognized
l=4
Non-recognised as danger signal for further evaluation at vDCA server.
Semi-matured DC
DCA implemented in C# running as a service on vDCA Server.
Legitimate packets allowed aftervDCA server analysis to fulfill request.
Database server: Microsoft SQL Server 192.168.0.30

Figure 5: Receptors matching against antigens in deployed environment

6. vDCA -Response

The human immune system in detection does a lot of validation on suspected pathogens before deeming them confirmed pathogens. The next stage after the proposed detection approach has proved successful, is to adapt the system to drop anomalous packets payloads or malicious packets in real-time. The moment a malicious packet is detected a notification is sent to the operator of the system which triggers the system to carry out further validations with the gleaned data from the packet payloads. These anomalies detected are profiled which allow the system to drop a similar network packets whenever encountered.

7. Conclusion

Whilst the results are yet to be verified as the system implementation and testing of the vDCA is on-going research work by the author of this paper; the proposed vaccinated DCA will provide detection and response to SI computer network security threats. A vaccine provides lifetime immune protection to humans. Inoculation prevents humans from going through the pain of natural biological process of exposure to disease to produce antibodies as to build immunity. This, in computational context implies knowing rogue packets and avoid pre-processing; thereby optimising the performance of vDCA in detection and response. DCA detects anomalies; a compliment implementation of vDCA will detect and mitigate against SI.

References

Aickelin, U. and Cayzer, S. (2002)"The Danger Theory and Its Application to Artificial Immune Systems", *ICARIS,* University of Kent at Canterbury, U.K., September, pp. 9–11.

Albanesius, A. (2011) "Sony Hacked Again, Group Claims*", [*online].PCmag.com. Available from: http://www.pcmag.com/article2/0,2817,2386327,00.asp. [Accessed 15th August 2011].

Al-Hammadi, Y., Aickelin, U. and Greensmith, J. (2010) "Performance evaluation of DCA and SRC on a single bot detection", *Journal of Information Assurance and Security*, vol. 5, pp.11.

Bretscher, P. and Cohn, M. (1970)"A theory of self-nonself discrimination", *Science*, vol. 169, pp. 1042–1049.

Burgess, M. (1998) "Computer immunology",*In Proc. of the Systems Administration Conference (LISA-98)*, pp. 283-297.

Burgess, M (2000) "Evaluating cfegine's immunity model of site maintenance", In Proceeding of the 2nd SANE *System Administration Conference (USENIX/NLUUG*).

Dasgupta, D., Yu, S. and Nino, F. (2011) "Recent Advances in Artificial Immune Systems: Models and Applications", *Applied Soft Computing*, vol.11 no.2, pp.1574-1587.

Danger Project, [ONLINE] Available at: http://ima.ac.uk/danger. [Accessed 31 January 2012].

De Castro, L.N. & Von Zuben, F.J. (2000) "The Clonal Selection Algorithm with Engineering Applications", Citeseer, pp. 36-37.

Forrest, S., Perelson, A. S., Allen, L and Cherukuri, R. (1994)"Self-nonself discrimination in a computer", Proceedings *of the 1994 IEEE Symposium on Research in Security and Privacy*.

Greensmith, J. (2007) "The Dendritic Cell Algorithm", *PhD thesis,* School of Computer Science, The University of Nottingham.

Hofmeyr, S. A., Somayaji, A. and Forrest, S. (1998) "Intrusion detection using sequences of system calls", *J. Comput. Secur.* Vol.6, pp. 151–180.

Hofmeyr, S.A. and Forrest, S. (1999)"Immunity by Design: An Artificial Immune System", *Proceedings of the Genetic and Evolutionary Computation Conference*, vol.2, May, pp.1289-1296.

Hofmeyr, S.A. and Forrest, S. (2000) "Architecture for an artificial immune system. Evolutionary Computation", vol. 8, no. 4 pp.443-473.

Ji, Z. and Dasgupta, D. (2007) "Revisiting negative selection algorithms", *Evolutionary Computation*, vol.15, no. 2, pp.223-251.

Jerne, N.K. (1974) "Towards a network theory of the immune system", *Annales dimmunologie*, vol. 125C, no.1-2, pp.373-389.

Kephart, J.O. (1994) "A Biologically Inspired Immune System for Computers", Artificial *Life IV Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, pp.130-139.

Matzinger, P. (1994) "Tolerance, danger, and the extended family", Annual *Review of Immunology*, vol. 12, pp. 991–1045.

Morgan, C. (2011). *SharpPcap.*[Online].http://sourceforge.net/apps/mediawiki/sharppcap/index.php?Title=Main_Page. [Accessed 31 January 2012].

Pereira, G. (2011) "Artificial Immune System based on Danger Theory"*, INESC-ID.*

Powers, S. and He, J. (2008) "A hybrid artificial immune system and Self Organising Map for network intrusion detection", Information *Sciences*, vol.178, no. 15, pp.3024-3042.

SQL Injection, (2012) *SQL Injection*. [Online], Available at: http://technet.microsoft.com/en-us/library/ms161953.aspx. [Accessed 31 January 2012].

Schenten, D., Medzhitov, R. (2011) "The control of adaptive immune responses by the innate immune system",*Adv Immunol.* , vol.109, pp.87-124.

Twycross, J. (2007) "Integrated innate and adaptive artificial immune systems applied to process anomaly detection", *PhD thesis*, School of Computer Science, The University of Nottingham.

Yeom, K.W. (2007) "Immune-inspired algorithm for anomaly detection", *Stud. Comput. Intell. (SCI).* 57, 129–154.

Building an Ontology for Cyberterrorism

Namosha Veerasamy[1], Marthie Grobler[1, 2] and Basie Von Solms[2]
[1]Council for Scientific and Industrial Research, Pretoria, South Africa
[2]University of Johannesburg, South Africa
nveerasamy@csir.co.za
mgrobler1@csir.co.za
basievs@uj.ac.za

Abstract: Cyberterrorism and the use of the Internet for cyberterrorism is an emerging field. Often cyberterrorism activities overlap with traditional hacking and Information and Communication Technology (ICT) Infrastructure exploitation. As a result, the defining and differentiating characteristics of cyberterrorism can easily be misunderstood. The use of an ontology specifically developed for cyberterrorism, will provide a common framework to share conceptual models. By using an ontology, the internal and external environment of a field (in this case, cyberterrorism) can be captured together with the relationships between the environments. This paper proposes an ontology to identify whether a cyber event can be classified as a cyberterrorist attack or a support activity. The role of the cyberterrorism ontological model will be to provide a better structure and depiction of relationships, interactions and influencing factors by capturing the content and boundaries in the field of cyberterrorism. The ontology will be developed using a cyberterrorism framework covering influencing factors, together with a compiled network attack classification ontology. Classes will be drawn from research carried out on the use of ICT in the support of cyberterrorism. As defined in this research, a cyberterrorism attack consists of a high-level motivation that is religious, social or political. The individual/group can furthermore be classified as having a specific driving force depending of the level of extremism or revolutionary thinking. Thus, the ontology will take into consideration the motivating characteristics that play a significant role in contributing towards the definition of cyberterrorism. Overall, this paper promotes the understanding of the field of cyberterrorism and its relation to ICT manipulation, together with the use of the Internet to support terrorism in general. Ontologies enable a common view on a specific domain to generate knowledge that can be shared and reused. Ontologies can further be populated with specific dynamic instances of information and therefore can be used to generate real-world scenarios. In this paper, the proposed ontological model will form a knowledge base for the field of cyberterrorism and will provide instances that aim to convey realistic cyberterrorism situations and support examples.

Keywords: anti-forensics, Internet, terrorism, ICT, propaganda, social-networking

1. Introduction

The emergence of the cyberterrorism domain means that a new group of potential attackers on computer and telecommunication technologies may be added to the list of traditional criminals threatening Information and Communication Technology (ICT) Infrastructure (Janczewski, Colarik 2007). In addition, the use of the Internet as both enabler and support mechanism for cyberterrorism can potentially lead to misunderstanding of the field.

A further complication is the overlap between cyberterrorism activities and traditional hacking and ICT Infrastructure exploitation. As a result, the defining and differentiating attributes of cyberterrorism can be misunderstood. In many instances, there are defining characteristics that separate traditional criminal cyber attacks from cyberterrorism. So how does one specify the attributes of a cyberterrorist attack in order to identify the defining concepts? This paper proposes the use of ontologies to define whether a cyber event can be characterised as cyberterror, a support to terrorism or an unclassified other cyber event. This model is build by first looking at the background of ontologies and the motivation for building a cyberterrorism specific ontology. Thereafter, the classes of the ontological model are discussed: actors, effects, motivation, objectives, practices, targets and cyber events. The application of the model is practically demonstrated on examples. Finally, future work in terms of the cyberterrorism ontology is presented. The main contribution of this research is the supplementation of the existing knowledge base of both cyberterrorism and cyber attacks, by enabling the convenient classification of an attack facilitated by ICT through a cyberterrorism specific ontological model.

2. Background to ontologies

According to Gruber (1993), an ontology is defined as a formal and explicit representation of a shared conceptualization. Frantz and Franco (2005) argue that ontologies provide a shared and common understanding of a domain to be communicated among people and computers to facilitate knowledge sharing and reuse. In addition, Frantz and Franco explain that ontologies provide a formal explicit conceptualization (i.e. meta-information) that describes the semantics of information of the static domain capture of knowledge based systems. Moreover, Noy and McGuiness (2001) provide the following reasons for developing an ontology:

- To share a common understanding of the structure of information among people or software agents.
- To enable reuse of domain knowledge.
- To make domain assumptions explicit.
- To separate domain knowledge from the operational knowledge.
- To analyze domain knowledge.

Uschold and King (1995) proposed a four-step methodology for developing ontologies: identify the purpose of the ontology, build the ontology, evaluate and document. The step for building an ontology consists of three iterative sub-steps: ontology capture, ontology coding and integrating existing ontologies. Ontology capture is the identification of the key concepts and relationships in the domain of interest by producing precise unambiguous text definitions for such concepts and relationships. Ontology coding is the explicit representation of the conceptualisation in some formal language. Ontology integration refers to using other ontologies during the capture and coding process (Uschold, King 1995). The next section will explain these steps in more detail in terms of the cyberterrorism ontology.

3. Building a cyberterrorism ontology

Ontologies provide a common framework to share conceptual models. By using an ontology, the internal and external environment of a field can be captured in conjunction with the relationships between these environments. This paper proposes that an ontology can be used to identify and capture the content and boundaries in the field of cyberterrorism. The role of the ontology will be to provide a better structure and depiction of relationships, interactions and influencing factors, as suggested by Noy and McGuiness (2001) in Section 2.

The initial step of building an ontology is to determine the purpose thereof. The aim of the proposed ontology is to determine whether a cyber event can be classified as cyberterror or a support to terror. The next step is to build the ontology by capturing key concepts and relationships, coding the explicit representation of the conceptualisation in the ontology language (in this case, Protégé), and integrating it with other ontologies. Ontologies provide the ability to form a knowledge base for a specified field. This step enables the formal capturing of domain knowledge to promote sharing and exchange.

Protégé is ontology specific software that serves as a knowledge base editor and thus facilitates the capturing of an ontology. It was developed at the Stanford University for both the Windows and Linux environments. Protégé provides the ability to define classes, relationships and properties. It is openly available and can be downloaded from http://protégé.stanford.edu. Protégé also comes with visualisation packages such as GraphWiz that allows the asserted and inferred classification hierarchies to be visualised (Horridge et al. 2004). The visualisations help provide succinct images of the deductions drawn from the inserted data and specifications, see Figures 1 to 4.

The next step is the evaluation of the built ontology.The reasoning capabilities within Protégé are used to infer new information from the asserted ontology as part of the evaluation process. Protégé has a number of built-in reasoners or inference engines that can be used to make deductions and queries based on the input specifications (the asserted statements and definitions). In this research project, different reasoners drew the same conclusions and therefore did not influence the evaluation results. An ontology can contain information in an asserted form (stated as a fact) and thus it is valuable to operate on inferred relationships (derived as conclusion from given facts) rather than on the asserted relationships. This process minimizes the loss of information about what has been explicitly asserted by the users (Knublauch et al. 2005). An important consideration is therefore the background logic that is used for reasoning certain arguments.

The final step is the documentation of the ontology. According to Prieto-Diaz (2003), ontologies are built very much ad-hoc with the initial development of a controlled vocabulary for the subject area of interest. This is then organised into a taxonomy whereby key contents are identified and the concepts defined and related to create an ontology. Therefore, to initially develop the cyberterror ontology, a taxonomy was developed to identify the core concepts in the field of cyberterrorism. The process of building a taxonomy and ontology is very much intertwined. The various steps of building an ontology is iterative (Capture, Code and Integrate). The next section looks at the development of the cyberterrorism taxonomy and ontology using Protégé.

4. Classes in the cyberterrorism ontology

Previous research by van Heerden, Irwin and Burke (2012) was used as the basis for some of the underlying classes of the proposed cyberterrorism ontology. Van Heerden et al. proposed a Network Attack Ontology to classify computer-based attacks. For the cyberterrorism ontology, the core classes of Actor, Effect and Motivation were adopted and slight modifications were made to address specific requirements within the field of cyberterrorism.

The main classes in the proposed cyberterrorism ontology are the Actor, Cyber Event, Objective, Motivation, Practice, Effect and Target. For example, every Cyber Event would have an Actor entity, Objective, Motivation, Practice, Effect and Target. The goal of the ontology (based on the initial taxonomy whereby the main concepts are defined) was to determine whether a CyberEvent could be classified as a Cyberterror or a SupportTerror, based on its specified attributes in the other classes. Before explaining the functioning of the main class CyberEvent, a discussion on the development of each of the classes follows.

4.1 Actor

Van Heerden, Irwin and Burke (2012) formed the Actor Class with the following sub-classes:

- Commercial competitor
- Hacker
- *Script kiddie hacker*
- *Skilled hacker*
- Insider
- *Admin insider*
- *Normal insider*
- Organised criminal group
- Protest group

Figure 1 shows the actor classes and sub-classes that were carried over from the Network Attack ontology to the Cyberterror ontology. In ontologies, classes and sub-classes have an "is a" relationship. For example, every class in Protégé is defined as being a Thing and thereafter sub-classes are assigned to classes.

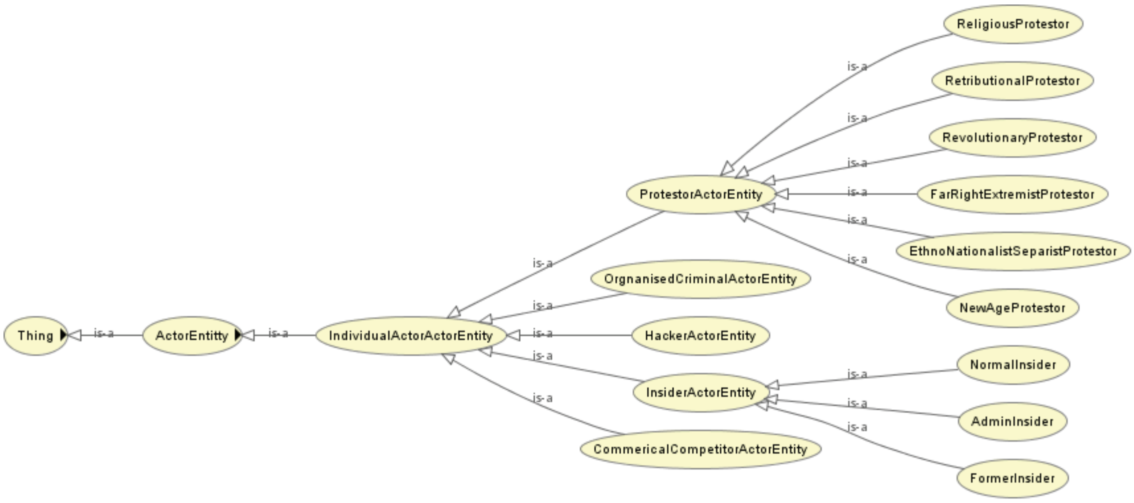


Figure 1: ActorEntity class

The original class Protest Group was extended to include examples of the type of groups that correspond to terrorist activities and included Religious, Ethno-nationalist separatist, Revolutionary, Far-right extremist, New Age and Retributional (Veerasamy 2009b). The original Actor class was also adapted to cater for individual and group activities by defining the core actor entity as an individual and a group entity as consisting of a number of individuals. In the next section, the possible effects are discussed.

4.2 Effects

Van Heerden, Irwin and Burke (2012) make use of the sub-classes Null, Minor, Major and Catastrophic in their Effects class of their Network Attack Ontology (see **Figure 2**).

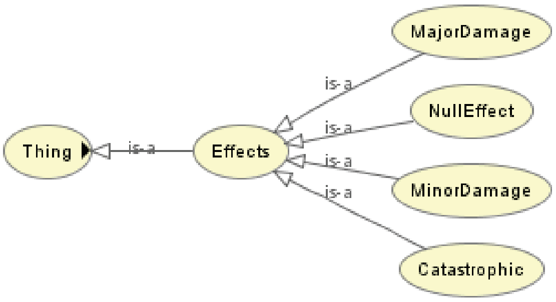


Figure 2: Effects class

Explanations taken from Mirkovic (2004) define "Null" as being no effect on the target, "Minor" to recoverable damage, "Major" to non-recoverable damage and “Catastrophic” refers to damage of such a nature that the target ceases to operate as an entity, for example declaration of bankruptcy. For this specific cyberterrorism ontology taxonomy, while the sub-classes were carried over, the definitions are adapted slightly:

- Null - no effect on target.
- Minor - recoverable damage to target (minimal financial implications and technical recovery required).
- Major - extensive financial or loss of reputation (more complex technical recovery required).
- Catastrophic - extensive damage such that the target failed to operate (massive damages-financial, technical and possibly life).

Now that the possible effects have been explored, the discussion moves on to an explanation of possible motivations.

4.3 Motivation

The Motivation class pertains to the high-level motivation or driving force of the actor. Often determining the motivation is subjective. However, a few high-level objectives have been identified from literature. Van Heerden, Irwin and Burke’s (2012) sub-classes for Motivation were Criminal, Ethical, Financial, Military and Recreational.

Denning (in (Gordon, Ford 2002)) talks about cyberterrorism being done to intimidate or coerce a government or its people to further political or social objectives. In addition, various terrorist groups are also strongly driven by religious beliefs, for example Al Qaeda prescribes to the principles of Islam. Therefore, while ordinary criminals or attackers may not have political, religious or social motivations, cyberterrorists do have these types of driving forces. Additional sub-classes that were added to the Cyberterrorism ontology included Political, Social and Religious. A summary of the motivation class is given in **Figure 3**. For example, an actor may have a more specific objection that stems from the high-level motivation. The different types of objectives are discussed next.

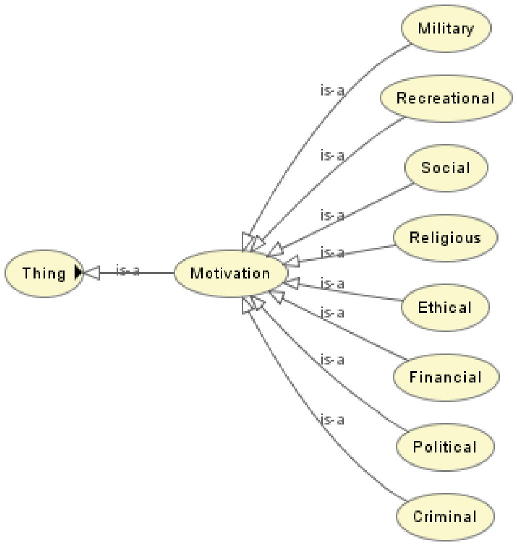


Figure 3: Motivation class

4.4 Objectives

The Objectives class refers to the low-level purposes of the attack and is sub-divided into Malicious Objectives (correspond to Cyberterror CyberEvents) and SupportTerror CyberEvents (correspond to a support activity). Based on previous research by Veerasamy (2009b) and Jenkins (2006), the Objectives class is divided as follows:

Malicious or attack objective

- Destroy
- Disrupt
- Force demands
- Interfere
- Intimidate
- Kill or maim
- Protest
- Publicity
- Steal
- Terrify

Support Objective

- Finance
- Intelligence
- Logistics
- Planning
- Propaganda
- Recruitment
- Social services (support for families of suicide attackers)
- Training

Objectives are not mutually exclusive. For example, a religious terrorist could be trying to interfere in operations, as well as finance the terrorist organisation. Therefore, a terrorist could have multiple objectives. The classification of a Cyberterror Event or SupportTerror Event will be also influenced by the effect, practice and motivation. The defining requirements for a Cyberterror CyberEvent and

SupportTerror CyberEvent are given in more detail in Section 4.7. The discussion now moves on to typical practices applicable to the cyberterrorist field.

4.5 Practices

Veerasamy (2009b) introduced some of the typical cyberterrorist practices as part of a framework covering influential factors in the field of cyberterrorism. These include the defacing of web sites, distribution of disinformation, spreading propaganda, denial of services using worms and viruses, disrupting of crucial services, corrupting of essential data, and stealing credit card information for funds. Furthermore, some of the uses of the Internet for cyberterrorism were classified as web literature, social-networking tools, anti-forensics and fundraising (Veerasamy, Grobler 2010). Based on the practices in literature, the Practices sub-class is structured as follows:

Anti-forensics

- Draft message folder
- Encryption
- IP-based cloaking
- Proxies and anonymisers
- Steganography

Data manipulation

- Denial of service
- Infections (worm, trojan or virus)

Fundraising

- Auctioneering
- Casinos
- Credit card theft
- Donations
- Drugs
- Phishing

Social networking

- Applications
- Blogs
- Forums
- Gaming
- Music
- Virtual personas
- Websites

Web defacement

- Cross-side scripting
- SQL injection

Web literature

- Biographies
- Encyclopaedias
- Essays
- Manuals
- Periodicals

- Poetry
- Statements
- Video

The list of practices is an indication of the range of practices that cyberterrorists typically utilise. Due to the growth of technology and digital capabilities, this list is not exhaustive and therefore can be extended as new practices are identified. Cyberterrorists usually have specific targets in mind when an attack is launched or in support of an attack. A discussion of the cyberterrorist targets follows.

4.6 Target

This class refers to the target of the cyber event or the type of system that the event occurs on. In order to distinguish between a criminal activity, a small-scale incident and a high-impact cyberterrorist event, the Target class is divided as follows (see **Error! Reference source not found.**):

- Government or critical
- Individual
- Organisational

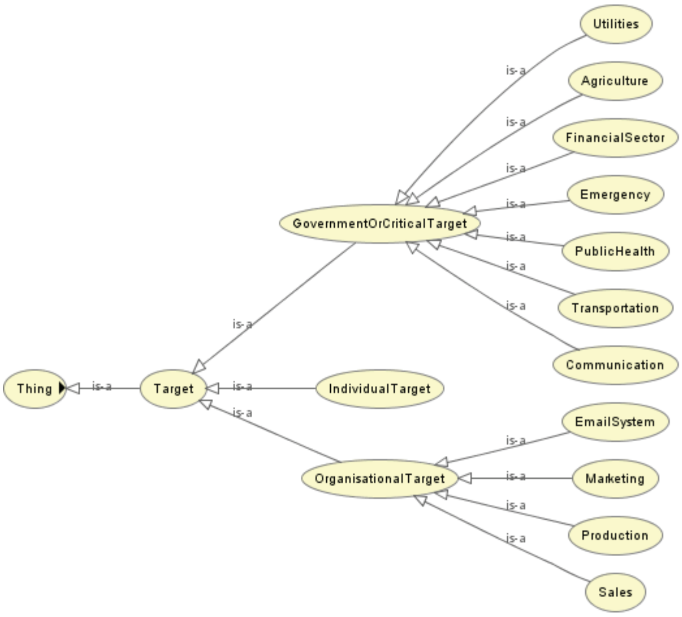


Figure 4: Target class

For example, terrorists would target high-impact infrastructure which would fall into governmental or organisational facilities. A virus affecting an unsuspecting single user's email address book would not be classified as cyberterrorist and therefore cyberterrorists targets would seek to have a detrimental effect on a critical target. The various industries and services that fall within the government or critical sector are Agriculture, Communication, Emergency, Finance, Public Health, Transportation and Utilities (Veerasamy 2009a). Furthermore, critical systems in an organisation that would be a prime target for a cyberterrorist attack include email systems, marketing, production and sales.

The various classes in the ontology have been discussed. The discussion now moves on to an explanation of the main class CyberEvent which links the previously discussed classes.

4.7 CyberEvent

The CyberEvent class forms a critical aspect as the aim of the overall ontology is to classify a cyber event as Cyberterror, a SupportTerror or as an unclassified OtherCyberEvent.

Cyberterrorism is defined as "A purposeful act, personally or politically motivated, that is intended to disrupt or destroy the stability of organizational or national interests, through the use of electronic devices which are directed at information systems, computer programs, or other electronic means of communications, transfer, and storage" (Desouza, Hensgen 2003). However, ICT infrastructure may not always be the target of an attack but can also serve a supporting role. For example, various techniques, networks and electronic devices may not always be used in a direct attack, it can still provide assistance in terms of communication, guidance, information gathering, preparation and financial backing (Veerasamy 2009a). Thus, it is imperative to differentiate between a cyber event being an actual cyberterrorist attack, and a cyber event that provides technology support to terrorism in general. In order to differentiate between a Cyberterror, SupportTerror and unclassified OtherCyberEvent, assertions were made in the ontology. The assertions relate to the motivation, objective, target, effect and practices.

With ontologies, the core assertions or defining attributes of classes need to be specified. The main class in the ontology is the CyberEvent class. Every CyberEvent could have more than one Actor, Objective, Practice and Target but only one Effect. The reasoning behind the assertion that a CyberEvent needs only one Effect, is that in order for a CyberEvent to be classified as Cyberterror it needs to have a major or catastrophic effect. For example, a cyberterrorist could be carrying out two activities simultaneously, fundraising and web defacement. The fundraising would be classified as a SupportTerror CyberEvent and the web defacement as Cyberterror CyberEvent. The fundraising may have a minor effect but overall the combined practices could have a major or catastrophic effect and the CyberEvent could be classified as Cyberterror CyberEvent.

Furthermore, in this ontology, the Cyberterror CyberEvent and SupportTerror CyberEvent attributes needed to be specified with detailed conditions in order for the cyber event to be correctly classified. To be classified as a CyberterrorEvent the following conditions were specified:

- The effect had to be major or catastrophic - terrorist attacks do not aim to have minor or null effect.
- The motivation had to be political religious or social - terrorists are primarily politically, religiously or socially motivated.
- The practice had to be data manipulation or web defacement - attacks usually constitute malicious behaviour like interfering with a web site of manipulating data using worms, Trojans or viruses.
- The target had to be an organisation, government or critical target - an attack on a minor individual computer or system would not cause terror.

To be classified as a SupportTerror CyberEvent, the event should have:

- A motivation that needed to be political religious or social - terrorists are primarily politically, religiously or socially motivated.
- A practice that had to be anti-forensics, fundraising or web literature - these are mainly practices that support terrorism support activities for recruitment, propaganda and planning.

Now that the various classes in the Cyberterror ontology have been introduced, this section is summarised briefly.

4.8 Overview of ontology

The classes Actor Entity, CyberEvent, Objectives, Motivation, Practice, Effect and Target were identified to be the main classes in the Cyberterror ontology and form the basis for the development of the ontology. The discussion now moves on to the application of the ontology through the classification of individual cyber events.

5. Ontology application: Classification of CyberEvent (individuals)

The description of the development of the ontology is given in the previous section. The main class is CyberEvent, which links all the other classes together. The CyberEvent class provides a means through which an event can be classified by the ontology as a Cyberterror, SupportTerror or OtherCyberEvent. Thereafter, individual examples can be instantiated with data and the reasoner tool in the ontology can be run in order to determine the cyber event's classification.

This section contains examples of individual-based incidents to show whether the reasoner can correctly classify the cyber event. Every cyber event has its own unique attribute specifications that would classify it as a Cyberterrorist CyberEvent, SupportTerror CyberEvent or an unknown OtherCyber Event. The attribute specifications of the individual cyber events are listed next.

5.1 Australian sewerage incident

Vitel Boden attacked the Australian Sewerage System in November 2001 (Lemos 2002). He was a former consultant on the project and after being refused a full-time position sought revenge. For this real-life example, the attribute specifications are as follows:

- Actor is a former insider.
- Motivation is social.
- Effect is major damage.
- Objective is a malicious objective of interfere.
- Practice is data manipulation (SCADA manipulation).
- Target is government or critical.

After the reasoner was run, it was inferred that the CyberEvent was of type Cyberterror.

5.2 Pakistan India example

The political conflict between Pakistan and India is represented in the ontology as follows:

- Actor is a protestor.
- Motivation is political.
- Effect is major damage.
- Objective is a malicious objective of destroy.
- Practice is web defacement.

After running the reasoner, it was inferred that this example could be classified as a Cyberterror CyberEvent.

5.3 Example religious

Similarly, another example was set up to test for the deduction of Support CyberEvents:

- Actor is a religious actor.
- Effect is minor damage.
- Objective is the support objective of finance.
- Motivations are financial and religious.
- Practice is casinos.

The inference engine in Protégé deduced that this example is a Support CyberEvent.

5.4 Overview of ontology

The ontology was set up to classify a cyber event as Cyberterror, SupportTerror or OtherCyberEvent depending on the details specified for its Actor, Motivation, Objective, Practice and Target attributes. Various examples can be instantiated in the ontology and therefore classified.

6. Future work and conclusion

This paper proposed the use of ontologies to clarify the field of cyberterrorism. It is relevant as it aims to classify a cyber event as a being cyberterrorism or a support to terrorism. The other attributes in the ontology also show the dynamic use of ICT by terrorist groups in manipulating systems to their advantage.

The attributes shown in the proposed ontology do not solely represent the only characteristics of a cyberterror attack but rather represent an abstraction of the most important considerations. By combining the ontology with other classification and development models, a better understanding of cyberterrorism can be gained. Later on, it is envisaged that the ontology proposed in this paper will open the dialogue for further discussion on the development of cyberterrorism by classifying attacks and identifying new practices, targets, motivations and objectives. In addition, the ontology captures important information about cyberterrorist motivations, objectives, practices, effects, targets and actors. The paper also provides practical insight into the communication and intimidation methods used by terrorists over cyberspace.

References

Desouza, K.C. & Hensgen, T. (2003), "Semiotic Emergent Framework to Address the Reality of Cyberterrorism", *Technological Forecasting and Social Change,* vol. 70, no. 4, pp. 385-396.

Frantz, A. 2005, "A semantic web application for the air tasking order", *10th International Command And Control Research And Technology Symposium Experimentation Track* .

Gordon, S. & Ford, R. (2002), "Cyberterrorism?", *Computers & Security,* vol. 21, no. 7, pp. 636-647.

Horridge, M., Knublauch, H., Rector, A., Stevens, R. & Wroe, C. 2004, "A Practical Guide To Building OWL Ontologies Using The Protégé-OWL Plugin and CO-ODE Tools Edition 1.0", *The University Of Manchester* .

Gruber, T.R. (1993), "A translation approach to portable ontology specifications", *Knowledge acquisition,* vol. 5, no. 2, pp. 199-220.

Janczewski, L. & Colarik, A.M. (2007), *Cyber warfare and cyber terrorism,* Information Science Reference.

Jenkins, B.M. (2006), "The New Age of Terrorism" in McGraw-Hill, New York, pp. 118–119.

Knublauch, H., Horridge, M., Musen, M., Rector, A., Stevens, R., Drummond, N., Lord, P., Noy, N.F., Seidenberg, J. & Wang, H.(2005), "The Protégé OWL Experience", *Proc. OWL: Experiences and Directions Workshop*.

Lemos, R. (2002), "What are the real risks of cyberterrorism", *ZDNet, August,* vol. 26.

Mirkovic, J. & Reiher, P. (2004), "A taxonomy of DDoS attack and DDoS defense mechanisms", *ACM SIGCOMM Computer Communication Review,* vol. 34, no. 2, pp. 39-53.

Noy, N.F. & McGuinness, D.L. (2001), *Ontology Development 101: A Guide to Creating Your First Ontology*, Stanford Knowledge Systems Laboratory Technical Report.

Prieto-Díaz, R. (2003), "A faceted approach to building ontologies", *Information Reuse and Integration, 2003. IRI 2003. IEEE International Conference on*IEEE, , pp. 458.

Uschold, M. & King, M. (1995), "Towards a methodology for building ontologies", *Workshop on Basic Ontological Issues in Knowledge Sharing*.

Van Heerden, R.P., Irwin, B. & Burke, I.D. (2012), "Classifying network attack scenarios using an Ontology", *Proceedings of the 7th International Conference on Information Warfare and Security.* Academic Conferences.

Veerasamy, N. (2009a), "A high-level conceptual framework of cyberterrorism", *Journal of Information Warfare,* vol. 8, no. 1, pp. 42-54.

Veerasamy, N. (2009b), "Towards a conceptual framework for cyberterrorism", *Proceedings of the 4th International Conference on Information Warfare and Security*, ed. L. Armistead, Academic Conferences International, 26-27 March, pp. 129.

Veerasamy, N. & Grobler, M. (2010), "Terrorist use of the Internet: Exploitation and Support Through ICT Infrastructure", *Proceedings of the 6th International Conference on Information Warfare and Security.* Academic Conferences, pp. 260.

Swarm UAV Attack: How to Protect Sensitive Data?

Robert Erra, Vincent Guyot, Loica Avanthey, Antoine Gademer and Laurent Beaudoin
SI&S and ATIS Lab, ESIEA, Paris, France
robert.erra@esiea.fr
vincent.guyot@esiea.fr
loica.avanthey@esiea.fr
antoine.gademer@esiea.fr
laurent.beaudoin@esiea.fr

Abstract: We consider the following scenario: a swarm of UAVs has a mission and a UAV from it has been captured: is it possible to secure (in a broad sense) the sensitive data and software in such a way as to avoid any information leak in this situation. In this article we study a possible solution using in a complementary way: an embedded secure token (a smartcard); Shamir's secret sharing algorithm associated to k-ary goodware.
Either of these tools can bring a very high level of security.

Keywords: military swarm, secret sharing scheme, secure token, k-ary goodware

1. Introduction

Nowadays, UAVs (*Unmanned Aerial Vehicles*) are increasingly used for military purposes and are spreading in civilian applications. The trends of these "next generation" vectors are the automation of the system, the miniaturization, the collaboration between several systems, and a general decrease in the production costs [GADEMER 2010, BEAUDOIN 2010].

The UAVs are a technological jewel of the armed forces. Therefore they are the subject of covetousness and become the target of various attacks. The reasons are numerous: spying (on the acquired data, on the mission's objectives, on the target, on the sender, on the backup location, etc.), falsifying (information, goals of the mission, etc.), disabling or falsifying the payload, destroying the vector (defense territory) or capturing it (for industrial information, for camouflage materials [MARTINEZ 2011], for reverse engineering hardware, etc.).

Yet today, UAVs are not fully automatic and are mainly operated by humans. That means the UAVs need to maintain data links with their owners. Those may be of two types: rising (orders) or descending (captured data). All these links are potential gateways for possible attacks on the integrity of the mission of the UAV, as can be seen with the hacking of the ROVER system [SHACHTMAN 2009, EGEA 2009] or, more worrisome, with an attack of the Parrot drone, bringing its deactivation and thus its fall [DELIGNE 2011]. The U.S. Department of Defense has also recognized that its Creech drone piloting base in Nebraska, was infected with a recalcitrant virus introduced through the USB stick used to load the flight plan of the unit! [SHACHTMAN 2011].

The other problem of the control link is that it can be accidentally lost or even confused by the enemy. In this case, the majority of U.S. UAVs are instructed to turn in a circle (or to follow predetermined paths) to get time to find the signal again or until they run out of fuel (which has probably causing the capture of an American reconnaissance UAV by Iran [USA UAV 2011]). In the case of a prolonged loss of signal, some models have even be instructed to return to the base in order to maximize the chances of recovering the signal ... and at the same time the unit. If not, an intervention by ground forces or air force is necessary to carry out the recovery or destruction of the device, which is extremely expensive.

To minimize the need to rely on data links between drivers, operators and vectors, the manufacturers are seeking to significantly increase the automation of UAVs, which means providing the vectors with a lot of valuable information, critical most of the time (flight plan, mission's objectives, starting coordinates, encryption keys, etc.). All these data require special protection, especially in case of UAV capture.

Let us imagine the following scenario: a UAV from a swarm has been captured on a mission. A lot of countermeasures against UAVs can be used [US 2003] and sensitive data have to be protected so: is it possible to secure (in a broad sense) the sensitive data and the software in such a way as to avoid any information leak in this situation?

In case of retrieval of the UAV two possibilities have to be considered:

- the UAV has shut down because of a crash: all the sensitive data are protected as the secure token is supposed to be tamper resistant; or have been destroyed ;
- the UAV has been captured unharmed and is still operational: in this case [US UAV 2011], all the sensitive data are not protected anymore and the reverse engineering will not be so difficult. We propose to use the Shamir's secret sharing scheme and the notion of k-ary goodware (which is a generalization of k-ary virus [FILIOL 2007, DESNOS 2009]) and we will show why in this case no sensitive information will be retrieved, even in the case where a small number of UAVs have been captured unharmed.

Different variants will be presented in this paper: 1) a swarm of UAVs for a unique target, 2) multiple swarms for a unique target 3) multiple swarms for multiple targets.

2. Secure token

Now, let us go back to our UAV swarm and let us suppose the following scenario: we have a swarm of N UAVs; the swarm has a mission. To protect the final goal of the mission, information M is encrypted using a symmetric algorithm E with a key K. So, each member of the swarm flies with the ciphered message C=E(K,M) , and it needs the secret key K to decipher the encrypted message.

The first idea is: we give to each member of the swarm the secret C and the key K. But there is a danger: if only one UAV is captured and not destroyed, the enemy can reverse the hardware/software to obtain the message M. There are a lot of ways to do it. We can use a smartcard also known as a secure token: the embedded secure token is in charge of the cipher key that is used to cipher the critical data (for example the flight plan, the backup destination, etc.).

The issue is the likely leak of information occurring when the UAV is captured by the enemy. In order to avoid such situation, the sensitive data, which need to be kept secret, have to be encrypted. If the UAV processes the encrypting key in RAM ; there is a risk of recovering the key from a captured UAV by forensics methods or by freezing the memory (techniques that enable to dump the content of the memory).

Using a smart card could prevent such a bad scenario. Actually, smart card is a temper resistant object, so it is not possible to analyze the data within. If the encrypting/decrypting key is inside the smart card and is only used through the smart card, the data is secure, captured UAV or not.

A smart card is not a simple storage card but rather a computing device able to perform computations in a very secure way.

Since the very beginning of this technology [MORENO 1974, 1975], smart cards have been designed to secure operations. Smart cards are tamper-resistant mobile tokens.

Both hardware and software measures [RE 2010] prevent from retrieving the content if appropriate credentials are not provided to the smart card. Sensors are able to detect physical illegal access by an intruder.

To provide security, smart cards are generally required to handle sensible cryptographic materials. To protect such materials, a smart card could be considered as a remote server delivering secret contents or services only to whom knows their existence and how to get them. A smart card is a black box responding to a given list of commands. But if some commands are secretly built-in within the smart card [GUYOT 2012], there is no way to uncover them without reading the programming code of the given application. This is a protection to prevent misusage from the smart card holder, but it could also be used to circumvent some tight data control procedures by hiding secret functions or data within smart cards looking genuine but being controlled by their holders.

3. From k-ary malware to k-ary goodware

We introduce here the concept of *k-ary goodware* which generalizes the notion of *k-ary malware* from [FILIOL 2007] which has given the following theoretical definition (and a Proof-Of-Concept):

- a k-ary malware is made of k different files, innocent looking ;
- Each of them can (inter)act independently or not and can either be executed in parallel or in sequence.
- Such a code is said to be sequential (serial mode) if the k constituent parts are acting strictly one after the other.
- It is said to be parallel if the k parts executes simultaneously (parallel mode).
- Not all the parts are necessarily executable. The cumulative action of each part defines the malware action.

A k-ary malware is a family of k files (some of them may not be executable) whose union constitutes a computer virus/mawlare and performs an offensive action that is equivalent to that of a true malware.

We will say now that a software is a goodware if it is not a malware and we propose the definition of *a k-ary goodware* that follows the definition of a k-ary malware, we just remove the *malicious* effects:

- a k-ary goodware is made of k different files;
- Each of them can (inter)act independently or not and can either be executed in parallel or in sequence.
- Such a code is said to be sequential (serial mode) if the k constituent parts are acting strictly one after the other.
- It is said to be parallel if the k parts executes simultaneously (parallel mode).
- Not all the parts are necessarily executable. The cumulative action of each part defines the goodware action.

We propose to use this notion to better protect the software and sensitive data that each UAV of a swarm will be given.

4. Secret sharing scheme

Desnos [DESNOS 2009] has presented a way to implement the notion of k-ary malware in a « secure way » using a cryptographic tool: he proposed to use secret sharing schemes (a PoC in Python has been presented to show the feasibility).

The first possible way to make a shared secret key is simple: we define N numbers $K_1, K_2, \ldots K_N$ such that

$$K = \sum_{i=1}^{N} K_i$$

So, each k-ary malware is given a “piece” K_i of the secret key K_S. But in this case if the enemy captures enough members then he has to try fewer keys for a brute force attack. A usual and very simple way to avoid this attack scenario [MOV 1996] is to take the modulo of everything:

$$K = \left(\sum_{i=1}^{N} K_i\right) \bmod p$$

with p a large enough prime number (and of course greater than K).

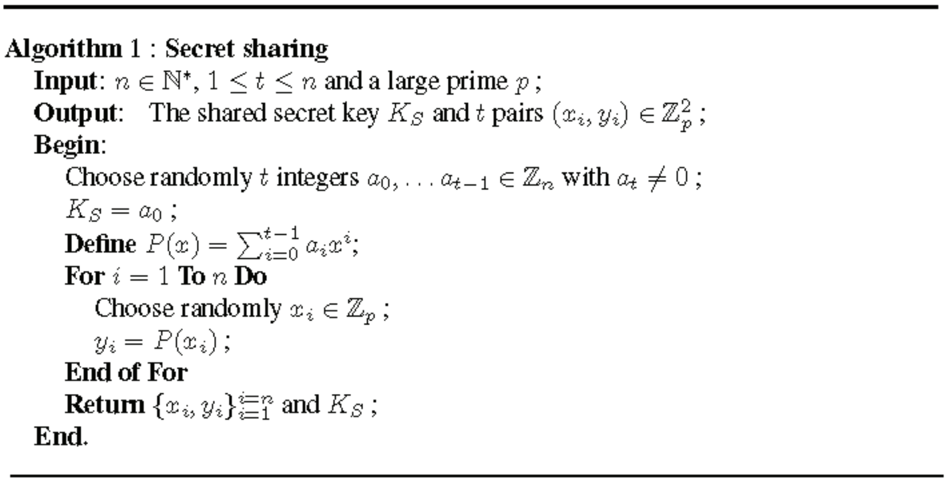
Now, if only one non destroyed UAV is captured then the enemy can hardly find the key. There is only one attack scenario; the enemy has to try all numbers between 1 and p-1.

But this countermeasure brings another problem: if (only) one UAV is intentionally or unintentionally destroyed or captured then the rest of the swarm is technically unable to compute the secret key K and so, the mission will fail.

To solve all these problems we propose to use a secret sharing scheme and more precisely we propose to use the Shamir's secret sharing scheme with threshold.

The computation of the shared secret key K_S and the *n* pairs $\{x_i, y_i\}_{i=1}^{i=N}$ is presented in the algorithm (1), these data satisfy the following property [MOV 1996]:

- each subset of t sharers can computes the shared secret key (we just have to ensure that none of them are equal, which is easy) ;
- It is computationally difficult to any subset of at most t−1 sharers to computes the shared secret key.



```
Algorithm 1 : Secret sharing
  Input: $n \in \mathbb{N}^*$, $1 \leq t \leq n$ and a large prime $p$ ;
  Output:  The shared secret key $K_S$ and $t$ pairs $(x_i, y_i) \in \mathbb{Z}_p^2$ ;
  Begin:
    Choose randomly $t$ integers $a_0, \ldots a_{t-1} \in \mathbb{Z}_n$ with $a_t \neq 0$ ;
    $K_S = a_0$ ;
    Define $P(x) = \sum_{i=0}^{t-1} a_i x^i$;
    For $i = 1$ To $n$ Do
      Choose randomly $x_i \in \mathbb{Z}_p$ ;
      $y_i = P(x_i)$ ;
    End of For
    Return $\{x_i, y_i\}_{i=1}^{i=n}$ and $K_S$ ;
  End.
```

The computation of the shared secret key for a subset of at least k players is easy: given t points, the Neville-Aitken algorithm allows to compute the unique interpolation polynomial P(x) of degree t−1 such that for all i □{1, …t}:

$$P(x_i) = y_i. \qquad (1)$$

The interpolation polynomial *P*(*x*), unique if all have the classical nuts property that none of them are equal is computed with the following recursion formula:

$$\begin{cases} P_{i,i}(x) &= y_i & 0 \leq i \leq t-1 \\ P_{i,j}(x) &= \dfrac{(x_j - x)P_{i,j-1}(x) - (x_i - x)P_{i+1,j}(x)}{x_j - x_i} & 0 \leq i, j \leq t-1 \end{cases} \qquad (2)$$

For the Shamir's secret sharing scheme, we just need to compute *P*(0) so, we can use the faster following "scalar" version of the Neville-Aitken algorithm, let us give a toy example with 3 pairs (x_0, y_0) , (x_1, y_1) and (x_2, y_2) all computation are done modulo p:

$$\begin{array}{lll} P_{0,0}(0) = y_0 & & \\ & P_{0,1}(0) & \\ P_{1,1}(0) = y_1 & & \\ & & P_{0,2}(0) \\ & P_{1,2}(0) & \\ P_{2,2}(0) = y_2 & & \end{array} \qquad (3)$$

So, we don't compute polynomials, which would be costly for processors with limited CPU on a UAV, we just compute with numbers.

5. Some scenario's and conclusion

So, if we envision the following possible scenario's:

- a unique swarm of UAVs for a unique target,
- multiple swarms for a unique target
- Multiple swarms for multiple targets.

How to use the tools we have briefly presented?

In each of UAVs of the swarm there is a software and sensitive data. The sensitive data can be ciphered, the secret key K_S has to be computed each time it is needed but a UAV has only a pair $\left(x_i, y_i\right)$. The pair $\left(x_i, y_i\right)$ and the sensitive data can be stored in a secure token.

But in a very paranoid mode we can also decide to cipher part of the software with the secret shared key, like the k-ary malware described in [DESNOS 2009]. So, a unique UAV will be unable to decipher the sensitive data but also the ciphered part of the software he has got.

We think that these tools, used together, give a high level of security because the enemy, if we use the Shamir's secret sharing scheme with a threshold value of $t \le N$, has to capture at least t UVAs to be able to read the sensitive data.

References

Beaudoin 2010 L. Beaudoin, A. Gademer, Towards Symmetrization Of Asymmetric Air Dominance: The Potential Key Role Playing By Home-Made Low Cost Unmanned Aerial Systems, Eciw, 2010

Deligne 2011 E. Deligne, Olivier Ferrand, The Ardrone Corruption, Hack.Lu, 2011

Desnos 2009 A. Desnos, Implementation Of K-Ary Viruses In Python, Hack.Lu, 2009

Egea 2009 Leonardo Nve Egea, Playing In A Satellite Environment 1.2, Iawacs, 2009

Filiol 2007 E. Fillol, Formalisation And Implementation Aspects Of K-Ary (Malicious) Codes, Journal In Computer Virology,3, 2, 75-86, 2007,Springer

Guyot 2012 V. Guyot, "Smart Card, The Stealth Leaker", Journal Of Computer Virology, Springer, To Be Published In 2012.

Gademer 2010 A. Gademer, Réalité Terrain Etendue: Une Nouvelle Approche Pour L'extraction De Paramètres De Surface Biophysiques Et Géophysiques A L'échelle Des Individus, Phd Thesis, 2010

Martinez 2011 L. Martinez, Does Iran Have U.S. Drone?, Abc News, 2011,

Shachtman 2009 N. Shachtman, Not Just Drones: Militants Can Snoop On Most U.S. Warplanes, Danger Room, 2009,

Shachtman 2011, N. Shachtman, Computer Virus Hits U.S. Drone Fleet, Danger Room, 2011

Moreno 1974 R. Moreno. Procede Et Dispositif De Commande Electronique (Patent 2.266.222). Inpi, France, March 1974.

Moreno 1975 R. Moreno. Data-Transfer System (Patent 4.007.355). Uspto, Usa, March 1975.

Mov 1996 Alfred J. Menezes, Paul C. Van Oorschot And Scott A. Vanstone, Handbook Of Applied Cryptography, Crc Press, 1996.

Re 2010 W. Rankl And W. Effing. Smart Card Handbook 4th Edition. Wiley, June 2010.

Us 2003 Countering The Tactical Uav Threat: Http://Www.Militaryphotos.Net/Forums/Archive/Index.Php/T-28018.Html

Us Uav 2011] Http://Www.Wired.Com/Dangerroom/2011/12/Iran-Did-Capture-A-Secret-U-S-Drone/

Us 2011] Http://Www.Iwatchnews.Org/2011/11/07/7323/Counterfeit-Chips-Continue-Plague-Pentagon-Weapons-Systems

PhD Research Papers

Proposal for a new Equation System Modelling of Block Ciphers and Application to AES 128

Michel Dubois and Eric Filiol
Laboratory of Operational Virology and Cryptology, Laval, France
michel.dubois@esiea-ouest.fr
eric.filiol@esiea.fr

Abstract: One of the major issues of cryptography is the cryptanalysis of cipher algorithms. Cryptanalysis is the study of methods for obtaining the meaning of encrypted information, without access to the secret information that is normally required. Some mechanisms for breaking codes include differential cryptanalysis, advanced statistics and brute-force. Recent works also attempt to use algebraic tools to reduce the cryptanalysis of a block cipher algorithm to the resolution of a system of quadratic equations describing the ciphering structure. As an example, Nicolas Courtois and Josef Pieprzyk have described the AES-128 algorithm as a system of 8000 quadratic equations with 1600 variables. Unfortunately, these approaches are, currently, deadlocks because of the lack of efficient algorithms to solve large systems of equations. In our study, we will also use algebraic tools but in a new way: by using Boolean functions and their properties. A Boolean function is a function from F_2^n to F_2 with n>1, characterized by its truth table. The arguments of Boolean functions are binary words of length n. Any Boolean function can be represented, uniquely, by its algebraic normal form which is an equation which only contains additions modulo 2 -- the XOR function -- and multiplications modulo 2 -- the AND function. Our aim is to describe a block cipher algorithm as a set of Boolean functions then calculate their algebraic normal forms by using the Möbius transforms. After, we use a specific representation for these equations to facilitate their analysis and particularly to try a combinatorial study. Through this approach we obtain a new kind of equations system. This equations system is more easily implementable and could open new ways to cryptanalysis. To test our approach we first apply this principle to the mini-AES cipher and in a second time to AES-128 algorithm.

Keywords: block cipher, Boolean function, cryptanalysis, AES

1. Introduction

The block cipher algorithms are a family of cipher algorithms which use symmetric key and work on fixed length blocks of data. As an example Rijndael is a symmetric block cipher that can process data blocks of 128 bits, using cipher keys with lengths of 128, 192, and 256 bits. One of the major issues of cryptography is the cryptanalysis of cipher algorithms. Cryptanalysis is the study of methods for obtaining the meaning of encrypted information, without access to the secret information that is normally required. Some mechanisms for breaking codes include differential cryptanalysis, advanced statistics and brute-force.

Recent works attempt to use algebraic tools to reduce the cryptanalysis of a block cipher algorithm to the resolution of a system of quadratic equations describing the ciphering structure. These approaches are infeasible because of the difficulty of solving large systems of equations -- for example, 8000 quadratic equations with 1600 variables for the AES-128 as described by Nicolas Courtois.

In our study, we will also use algebraic tools but in a new way by using Boolean functions and their properties. Our aim is to describe a block cipher algorithm as a set of Boolean functions then calculate their algebraic normal forms by using the Möbius transforms. To test our approach we apply this principle to the mini-AES cipher. Since November 26, 2001, the block cipher algorithm "Rijndael", in its 128 bits version, became the successor of DES under the name of Advanced Encryption Standard (AES). Its designers, Joan Daemen and Vincent Rijmen used algebraic tools to give to their algorithm an unequalled level of assurance against the standard statistical techniques of cryptanalysis.

Recent works suggest that what is supposed to be the AES strength could be is weakness. Indeed, according to these studies, to cryptanalyse the AES could be reduced to solving a system of quadratic equations describing the ciphering structure of the AES. These results are not implementable in real life and do not represent a true danger for the AES.

In our study, we will look at a reduced version of the AES: the Mini-AES. Our goal is to describe it under the form of systems of Boolean functions and to calculate their algebraic normal forms by using the Möbius transforms. The system of equations obtained is more easily implementable and once extended to the AES could open new ways to cryptanalysis of the AES.

2. Boolean function

2.1 Definition

A Boolean function is a function $F_2^n \to F_2$, with n>1, characterized by its truth table. The arguments of Boolean functions are binary words of length n. Thus, if we take n=2 we can define the Boolean function OR -- the logical OR -- characterized by its truth table as show on table 1.

Table 1: The truth table of the function OR

x_1	x_2	x_1 OR x_2
0	0	0
0	1	1
1	0	1
1	1	1

The support of a Boolean function supp(f) is the set of elements of x such that $f(x) \neq 0$, the weight of a Boolean function wt(f) is the cardinal of its support. Thus, the support of the OR function is supp(OR)={(0,1), (1,0), (1,1)} and its weight is wt(OR)=3.

2.2 Algebraic normal form

The algebraic normal form (ANF) of a Boolean function f in n variables is the only polynomial $Q_f : F_2[x_1,\cdots,x_n]/(x_1^2 - x_1,\cdots,x_n^2 - x_n)$ such as $\forall(x_1,\cdots,x_n) \in F_2^n$ we have:

$$f(x_1,\cdots,x_n) = Q_f(x_1,\cdots,x_n) = \sum_{(u_1,\cdots,u_n)\in F^n} a_u \prod_{i=1}^{n} x_i^{u_i}$$

We call degree of f, denoted deg(f), the degree of the polynomial Q_f which corresponds to the highest degree of monomials with nonzero coefficients from the ANF of f. Furthermore, the ANF of a Boolean function exists and is single.

In short, any Boolean function can be represented, uniquely, by its algebraic normal form under the form:

$$f(x_1,\cdots,x_n) = a_0 + a_1x_1 + a_2x_2 + \cdots + a_nx_n + a_{1,2}x_1x_2 + \cdots + a_{n-1,n}x_{n-1}x_n + \cdots + a_{1,2,\cdots,n}x_1x_2\cdots x_n$$

thus, the algebraic normal form of a Boolean function only contains additions modulo 2 -- the XOR function -- and multiplications modulo 2 -- the AND function.

By using the example of the function OR, we have:

$$f(x_1,x_2) = x_1ORx_2 = f(0,x_2) + f(x_1,0) + f(x_1,x_2) = x_1 + x_2 + x_1x_2$$

2.3 Möbius transform

It is with the help of the Möbius transform that we will calculate the ANF of a Boolean function.
The Möbius transform of the Boolean function f is defined by:

$$MT(f) : F_2^n \to F_2$$
$$u = \sum_{v \le u} f(v) \bmod 2$$

with $v \le u$ if and only if $\forall i, v_i = 1 \Rightarrow u_i = 1$.

From there, we can define the algebraic normal form of a Boolean function f in n variables by:

$$\sum_{u=(u_1,\cdots,u_n)\in F_2^n} MT(u) x_1^{U_1} \cdots x_n^{U_n}$$

3. Workings of the equations

The goal of our study is to propose a model of the mini-AES as a set of equations obtained by using the Boolean functions. This modelling is a first step; the final goal is to use this method on the AES.

3.1 Principle of elaboration

The principle adopted for the system of equations consists, from the truth table of a Boolean function, to calculate its ANF by using its Möbius transform.

3.2 Example with the MajAmong3 function

To understand more easily the mechanisms implemented, consider an example with the function MajAmong3. This function: $F_2^3 \to F_2$ is characterized by the truth table presented in the table 2.

Table 2: The truth table of the function MajAmong3

x_1	x_2	x_3	MajAmong3
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

By calculating the Möbius transform of the function we obtain the result of the figure 1.

x_1	x_2	x_3	MajAmong3	→	calculation of MT(f)				MT(f)
0	0	0	0	→	0	0	0	0	0
0	0	1	0	→	0	0	0	0	0
0	1	0	0	→	0	0	0	0	0
0	1	1	1	→	1	1	1	1	1
1	0	0	0	→	0	0	0	0	0
1	0	1	1	→	1	1	1	1	1
1	1	0	1	→	1	1	1	1	1
1	1	1	1	→	1	0	1	0	0

Figure 1: Calculation of the Möbius transform for MajAmong3

Once the Möbius transform obtained we take the elements of F_2^3 where $MT(MajAmong3) \neq 0$. In our case we have the triplets (0,1,1), (1,0,1), (1,1,0) from which we can deduce the equation:

$$MajAmmong3(x_1, x_2, x_3) = x_2x_3 \oplus x_1x_3 \oplus x_1x_2$$

4. Application to the mini-AES

4.1 The mini-AES

With the goal of helping students in cryptography and the cryptanalysts better to understand the internal mechanisms of the AES that Raphael Chung-Wei Phan presented, in 2002, his mini version of the AES. This version uses restrained parameters compared to the AES while preserving its internal structure. The mini-AES is a block ciphering algorithm based on the same mathematical primitives than its big brother the AES. The atomic elements with which the mini-AES works are elements of the finite field GF(2^4) called nibbles. As with AES, mini-AES uses a state array containing four nibbles, making it a 16 bits block ciphering algorithm.

The ciphering process of the mini-AES consists on two rounds involving the functions NibbleSub applying the S-BOX to the state array, ShiftRow performing a rotation of the cells of the state array and MixColumn multiplying each column of the state array by a constant matrix. The rounds architecture is presented on the figure 2.

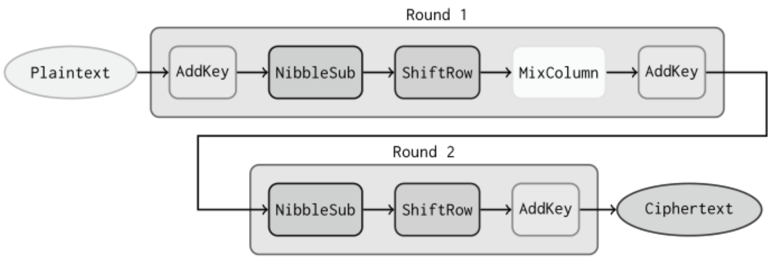


Figure 2: Architecture of the mini-AES rounds

4.2 Equations for the mini-AES

By taking the principle of generating equations showed in the paragraph 3, we will define the equations for the mini-AES. To simplify the process, we reduce the mini-AES to five Boolean functions: $F_2^{16} \to F_2^{16}$, one function for each round and three for the derivation key process.

The functions for the rounds X_1 and X_2 result from the conjunction of functions NibbleSub NS(), ShiftRow SR() and MixColumn MC() such as, by taking a 16 bits plain text block $b = (b_1, \cdots, b_{16})$, we have $X_1(b) = MC \circ SR \circ NS(b)$ and $X_2(b) = SR \circ NS(b)$. The three functions K_0, K_1 and K_2 describe the derivation key process such that, from a 16 bits key block $k = (k_1, \cdots, k_{16})$, we have the keys $K_{i_{i\in(0,1,2)}}(k) = (k_{i,1}, \cdots, k_{i,16})$ used in the rounds.

Ultimately, the mini-AES can be written as two equations R_1 and R_2 each one describing a round such that:

$$\begin{aligned}
b' &= R_1(b) = K_0(k) \oplus X_1(b) \oplus K_1(k) \\
&= (k_{1,1}, \cdots, k_{1,16}) \oplus (x_{1,1}, \cdots, x_{1,16}) \oplus (k_{2,1}, \cdots, k_{2,16}) \\
&= (b'_1, \cdots, b'_{16}) \\
b'' &= R_2(b') = X_2(b') \oplus K_2(k) \\
&= (x_{2,1}, \cdots, x_{2,16}) \oplus (k_{3,1}, \cdots, k_{3,16}) \\
&= (b''_1, \cdots, b''_{16})
\end{aligned}$$

With b, b' and b'' respectively the block of 16 bits in input, at the end of the first round and at the end of the second round and k the block of 16 bits of the key. We can then calculate the truth table of the Boolean functions K_0, K_1, K_2, X_1 and X_2 as in table 3 for X_1. Then by using the methodology chosen for the function MajAmong3, we obtain a set of 16 equations for each Boolean function. One equation for each bit of the block

Table 3: Some extracts of the truth table of the X_1() function

X_1(1026)	= 1000101101011001
X_1(1027)	= 0011110001011001
X_1(1028)	= 0101100101011001
X_1(1029)	= 1100110101011001
…	…
X_1(32000)	= 0100001000000111
X_1(32001)	= 0011111100000111
X_1(32002)	= 0010011100000111
X_1(32003)	= 1001000000000111
…	…
X_1(65000)	= 1111101100011000
X_1(65001)	= 1110001100011000
X_1(65002)	= 0101010000011000
X_1(65003)	= 0010100100011000

4.3 Formatting the equations

To facilitate the analysis of this set of equations and particularly to try later a combinatorial study we will use a specific representation for these equations.

The adopted principle consists of generating a file for each of the 16 equations of the R_1 and R_2 functions. Ultimately we will obtain 32 files. The content of each of these files consists of lines containing sequences of 0 and 1. Each line describes a monomial of the equation and the transition from one line to another means the application of the XOR operation.

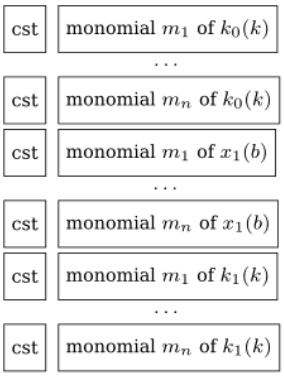

cst
monomial m_1 of $k_0(k)$
...
cst
monomial m_n of $k_0(k)$
cst
monomial m_1 of $x_1(b)$
...
cst
monomial m_n of $x_1(b)$
cst
monomial m_1 of $k_1(k)$
...
cst
monomial m_n of $k_1(k)$

Figure 3: File structure for the function $R_1(b)$

In order to facilitate understanding of the mechanism chosen we detail the implementation of the file corresponding to the bit b_1'$ of the function R_1 in the figure 4.

$$b_1' =$$
$$k_1 \oplus 1 \oplus x_{15}x_{16} \oplus x_{14} \oplus x_{14}x_{16} \oplus x_{13} \oplus x_{13}x_{15} \oplus x_{13}x_{15}x_{16} \oplus x_4 \oplus x_3x_4 \oplus x_2x_4 \oplus x_2x_3 \oplus x_2x_3x_4 \oplus$$
$$x_1x_3 \oplus x_1x_3x_4 \oplus x_1x_2 \oplus x_1x_2x_3 \oplus 1 \oplus k_{16} \oplus k_{14} \oplus k_{14}k_{15} \oplus k_{14}k_{15}k_{16} \oplus k_{13} \oplus k_{13}k_{14} \oplus k_{13}k_{14}k_{15} \oplus k_1$$

k_1	0	1000000000000000
1	1	0000000000000000
$x_{15}x_{16}$	0	0000000000000011
x_{14}	0	0000000000000100
$x_{14}x_{16}$	0	0000000000000101
x_{13}	0	0000000000001000
$x_{13}x_{15}$	0	0000000000001010
$x_{13}x_{15}x_{16}$	0	0000000000001011
x_4	0	0001000000000000
x_3x_4	0	0011000000000000
x_2x_4	0	0101000000000000
x_2x_3	0	0110000000000000
$x_2x_3x_4$	0	0111000000000000
x_1x_3	0	1010000000000000
$x_1x_3x_4$	0	1011000000000000
x_1x_2	0	1100000000000000
$x_1x_2x_3$	0	1110000000000000
1	1	0000000000000000
k_{16}	0	0000000000000001
k_{14}	0	0000000000000100
$k_{14}k_{15}$	0	0000000000000110
$k_{14}k_{15}k_{16}$	0	0000000000000111
k_{13}	0	0000000000001000
$k_{13}k_{14}$	0	0000000000001100
$k_{13}k_{14}k_{15}$	0	0000000000001110
k_1	0	1000000000000000

Figure 4: File corresponding to the bit b_1'

5. Application to the AES

The AES is an algorithm of symmetric block cipher. It encrypts and decrypts data blocks from a single key. Contrary to the DES, which is based on a Feistel network, the AES uses a network of substitution and permutation (SP-network). This includes substitution boxes, the S-Boxes, and permutation boxes, the P-Boxes. Each box takes a block of text and the key as input and provides a block of ciphered text as output. The information flow in a defined sequence of several P-Box and S-Box forms a round. This mechanism implements the principles of diffusion and confusion developed by Shannon. The objective of diffusion is to dissipate the statistical redundancy of the plain text in the ciphered text. Permutation operations ensure the diffusion. The objective of the confusion is to make difficult the relationship between the plain text, the key and the ciphered text. The confusion is obtained by substitutions, chosen with care.

Historically, the AES has two predecessors. The first is the cipher block algorithm Shark published in 1996 by Vincent Rijmen, Joan Daemen, Bart Preneel, Anton Bosselaers and Erik de Win. Shark uses blocks of 64 bits and a key of 128 bits. Like AES, it uses a SP-network with six rounds. The second algorithm called Square was published in 1997 by Joan Daemen and Vincent Rijmen. It uses an SP-network with eight rounds and works on blocks of 128 bits and also key of 128 bits.

5.1 The AES algorithm

The inputs and outputs of AES are 128 bits blocks and the key length can be 128, 192 or 256 bits. The basic unit of the algorithm is the byte. Blocks of data provided as input are transformed into arrays of four columns and four rows, each box containing a byte, whether 4*4*8=128 bits per arrays.At the beginning of operations of ciphering and deciphering of a block, the corresponding array of bytes is copied into the state array. The state array is a two-dimensional array of bytes containing n rows and m columns. For AES, n = m = 4. The operations of encryption and decryption are performed on this array and then the result is copied into an output array.

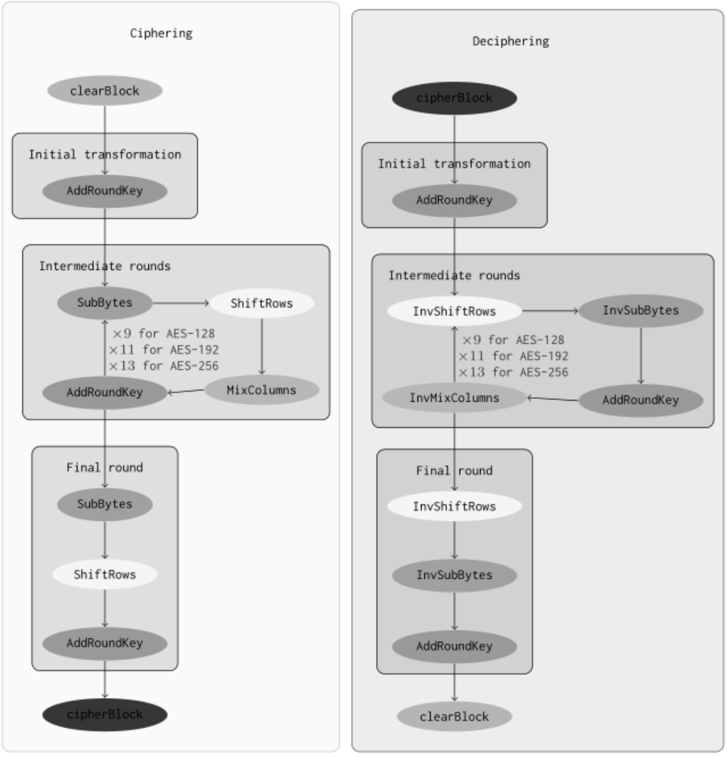


Figure 5: Ciphering and deciphering processes of AES

The ciphering operations are based on four predefined functions: AddRoundKey, SubBytes, ShiftRows and MixColumns. Each of these functions is executed on the state array. The ciphering cycle includes an initial transformation, some intermediate rounds and a final round. The initial transformation consists of applying the function AddRoundKey to the state array. The intermediate rounds execute, in the order, the functions SubBytes, ShiftRows, MixColumns and AddRoundKey on the state array. The final round differs from the intermediate rounds, by the removal of the function MixColumns in the transformations cycle. The number of rounds of AES depends of the key size. Thus, for a key of 128 bits, the number of rounds is 10, likewise, we have 12 rounds for a key of 192 bits and 14 rounds for a 256 bits key. The deciphering is realized by performing the inverse operations of the four encryption functions, in the inverse order. Thus, each function used in the ciphering operations disposes of its inverse function used for the deciphering: InvShiftRows, InvSubBytes and InvMixColumns. The AddRoundKey function stays unchanged. Like for ciphering, the deciphering process includes an initial transformation, some intermediate rounds and a final round. The initial transformation consists of applying the AddRoundKey function at the state array. The intermediate rounds execute, in order, the InvShiftRows, InvSubBytes, AddRoundKey and InvMixColumns functions on the state array. The final round differs of the intermediate rounds by the suppression of the InvMixColumns function in the transformations sequence.

5.2 Equations for the AES

Our aim is to apply to the AES the mechanisms described above for the mini-aes. However, with the mini-aes, we have Boolean functions from $F_2^{16} \to F_2^{16}$ and it is relatively easy to compute their truth tables. In the AES algorithm, the ciphering functions take 128 bits as input and 128 bits as output. So we should have Boolean functions from $F_2^{128} \to F_2^{128}$ and it is impossible to compute their Truth table. Indeed, such truth table contains 2^{128} inputs. We have to find solutions to describe each functions of the AES algorithm to obtain the same result as for the mini-aes algorithm. At the end, like for mini-aes, we obtain 128 files, each one describing the transformations of a bit block. The content of each of these files consists of lines containing sequences of 0 and 1. Each line describes a monomial of the equations and the transition from one line to another means the application of the XOR operation.

5.2.1 Solution for the SubByte function

The SubBytes transformation is a non-linear byte substitution that operates independently on each byte of the State array using a substitution table (S-box). This function is applied independently on each byte of the input block. So, we can reduce it as a Boolean function $F_2^8 \to F_2^8$ describing the S-Box of the AES. Thus we compute the truth table of the S-Box and then apply the Möbius transform on the result. By switching these results on the 16 bytes of the input block, we obtain 128 equations each one describing one bit of the input block. As an example the equation for low level bit is given on the figure below.

$$\begin{gathered}
1+x_{127}+x_{126}x_{127}+x_{125}+x_{125}x_{126}+x_{124}+x_{124}x_{126}+x_{124}x_{125}+x_{124}x_{125}x_{126}+ \\
x_{124}x_{125}x_{126}x_{127}+x_{123}+x_{123}x_{127}+x_{123}x_{126}+x_{123}x_{126}x_{127}+x_{123}x_{125}+x_{123}x_{125}x_{127}+ \\
x_{123}x_{125}x_{126}+x_{123}x_{124}x_{127}+x_{123}x_{124}x_{126}+x_{123}x_{124}x_{125}x_{126}x_{127}+x_{122}x_{127}+ \\
x_{122}x_{125}x_{127}+x_{122}x_{125}x_{126}x_{127}+x_{122}x_{124}x_{127}+x_{122}x_{124}x_{125}+x_{122}x_{124}x_{125}x_{127}+ \\
x_{122}x_{124}x_{125}x_{126}+x_{122}x_{123}x_{126}x_{127}+x_{122}x_{123}x_{125}x_{127}+x_{122}x_{123}x_{125}x_{126}x_{127}+ \\
x_{122}x_{123}x_{124}x_{125}x_{127}+x_{121}x_{127}+x_{121}x_{126}+x_{121}x_{126}x_{127}+x_{121}x_{125}+x_{121}x_{125}x_{127}+ \\
x_{121}x_{125}x_{126}+x_{121}x_{125}x_{126}x_{127}+x_{121}x_{124}x_{127}+x_{121}x_{124}x_{125}x_{126}+x_{121}x_{124}x_{125}x_{126}x_{127}+ \\
x_{121}x_{123}+x_{121}x_{123}x_{127}+x_{121}x_{123}x_{126}+x_{121}x_{123}x_{125}x_{126}+x_{121}x_{123}x_{124}x_{127}+ \\
x_{121}x_{123}x_{124}x_{126}+x_{121}x_{123}x_{124}x_{126}x_{127}+x_{121}x_{123}x_{124}x_{125}x_{127}+x_{121}x_{123}x_{124}x_{125}x_{126}x_{127}+ \\
x_{121}x_{122}+x_{121}x_{122}x_{126}+x_{121}x_{122}x_{125}+x_{121}x_{122}x_{125}x_{127}+x_{121}x_{122}x_{124}x_{127}+ \\
x_{121}x_{122}x_{124}x_{126}x_{127}+x_{121}x_{122}x_{124}x_{125}x_{126}+x_{121}x_{122}x_{123}+x_{121}x_{122}x_{123}x_{127}+ \\
x_{121}x_{122}x_{123}x_{126}+x_{121}x_{122}x_{123}x_{126}x_{127}+x_{121}x_{122}x_{123}x_{125}+x_{121}x_{122}x_{123}x_{125}x_{126}+ \\
x_{121}x_{122}x_{123}x_{124}x_{127}+x_{121}x_{122}x_{123}x_{124}x_{125}+x_{121}x_{122}x_{123}x_{124}x_{125}x_{127}+x_{120}x_{126}x_{127}+ \\
x_{120}x_{125}+x_{120}x_{125}x_{127}+x_{120}x_{125}x_{126}x_{127}+x_{120}x_{124}x_{126}+x_{120}x_{124}x_{125}+x_{120}x_{124}x_{125}x_{127}+ \\
x_{120}x_{124}x_{125}x_{126}+x_{120}x_{124}x_{125}x_{126}x_{127}+x_{120}x_{123}x_{127}+x_{120}x_{123}x_{126}x_{127}+ \\
x_{120}x_{123}x_{125}+x_{120}x_{123}x_{125}x_{127}+x_{120}x_{123}x_{125}x_{126}+x_{120}x_{123}x_{125}x_{126}x_{127}+ \\
x_{120}x_{123}x_{124}+x_{120}x_{123}x_{124}x_{125}x_{126}x_{127}+x_{120}x_{122}+x_{120}x_{122}x_{125}+x_{120}x_{122}x_{125}x_{127}+ \\
x_{120}x_{122}x_{125}x_{126}x_{127}+x_{120}x_{122}x_{124}+x_{120}x_{122}x_{124}x_{126}x_{127}+x_{120}x_{122}x_{124}x_{125}+ \\
x_{120}x_{122}x_{124}x_{125}x_{126}+x_{120}x_{122}x_{124}x_{125}x_{126}x_{127}+x_{120}x_{122}x_{123}x_{127}+x_{120}x_{122}x_{123}x_{126}+ \\
x_{120}x_{122}x_{123}x_{125}+x_{120}x_{122}x_{123}x_{125}x_{127}+x_{120}x_{122}x_{123}x_{125}x_{126}x_{127}+x_{120}x_{122}x_{123}x_{124}x_{127}+ \\
x_{120}x_{122}x_{123}x_{124}x_{126}+x_{120}x_{122}x_{123}x_{124}x_{125}+x_{120}x_{122}x_{123}x_{124}x_{125}x_{127}+x_{120}x_{121}+ \\
x_{120}x_{121}x_{125}+x_{120}x_{121}x_{125}x_{126}+x_{120}x_{121}x_{125}x_{126}x_{127}+x_{120}x_{121}x_{124}+x_{120}x_{121}x_{124}x_{126}+ \\
x_{120}x_{121}x_{124}x_{126}x_{127}+x_{120}x_{121}x_{124}x_{125}+x_{120}x_{121}x_{124}x_{125}x_{126}x_{127}+x_{120}x_{121}x_{123}x_{127}+ \\
x_{120}x_{121}x_{123}x_{125}x_{127}+x_{120}x_{121}x_{123}x_{125}x_{126}+x_{120}x_{121}x_{123}x_{124}x_{127}+x_{120}x_{121}x_{123}x_{124}x_{126}x_{127}+ \\
x_{120}x_{121}x_{123}x_{124}x_{125}x_{126}+x_{120}x_{121}x_{123}x_{124}x_{125}x_{126}x_{127}+x_{120}x_{121}x_{122}+x_{120}x_{121}x_{122}x_{126}+ \\
x_{120}x_{121}x_{122}x_{126}x_{127}+x_{120}x_{121}x_{122}x_{125}x_{126}+x_{120}x_{121}x_{122}x_{125}x_{126}x_{127}+x_{120}x_{121}x_{122}x_{124}+ \\
x_{120}x_{121}x_{122}x_{124}x_{127}+x_{120}x_{121}x_{122}x_{124}x_{125}x_{127}+x_{120}x_{121}x_{122}x_{123}x_{126}x_{127}+ \\
x_{120}x_{121}x_{122}x_{123}x_{125}+x_{120}x_{121}x_{122}x_{123}x_{125}x_{126}+x_{120}x_{121}x_{122}x_{123}x_{125}x_{126}x_{127}+ \\
x_{120}x_{121}x_{122}x_{123}x_{124}x_{126}+x_{120}x_{121}x_{122}x_{123}x_{124}x_{125}+x_{120}x_{121}x_{122}x_{123}x_{124}x_{125}x_{127}
\end{gathered}$$

Figure 6: The low level bit of the SubBytes function

5.2.2 Solution for the ShiftRows function

In the ShiftRows transformation, the bytes in the last three rows of the State array are cyclically shifted over different numbers of bytes (offsets). The first row, r = 0, is not shifted.For this function, we do not need to compute Boolean function. Indeed, the only operation of the ShiftRows transformation consists of switching byte through the State array. In our file this operation can be easily solved by a

XOR operation. As an example the second byte of the State array becomes the 6th after SiftRows function. Therefore to transform this byte we apply the following XOR to the second byte:

```
00000000000000000000000000000000000000001000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000100000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000010000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000001000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000100000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000010000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000001000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000100000000000000000000000000000000000000000000000000000000000000000000000000000000
```

5.2.3 Solution for the MixColumns function

The MixColumns transformation operates on the State array column-by-column, treating each column as a four-term polynomial. Each of these columns is multiplied by a square matrix. Thus we have for each column:

$$\begin{bmatrix} b'_i \\ b'_{i+1} \\ b'_{i+2} \\ b'_{i+3} \end{bmatrix} = \begin{bmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 01 \end{bmatrix} \bullet \begin{bmatrix} b_i \\ b_{i+1} \\ b_{i+2} \\ b_{i+3} \end{bmatrix}$$

So, for the first byte of the column we have this equation:

$$b'_i = 02 \bullet b_i \oplus 03 \bullet b_{i+1} \oplus 01 \bullet b_{i+2} \oplus 01 \bullet b_{i+3}$$

Since on GF_2^8, 01 is the identity for multiplication. This equation becomes:

$$b'_i = 02 \bullet b_i \oplus 03 \bullet b_{i+1} \oplus b_{i+2} \oplus b_{i+3}$$

We have the same simplification for all equations describing the column multiplication by this square matrix. At the end we only need to compute the truth table for multiplication by 02 and by 03 over GF_2^8.

5.3 Formatting the equations

To format the equations we use the same representation as for the mini-aes. Thus we have 128 files, one by bit of block. Each line describes a monomial and the shift from one line to another means a XOR operation. A file sample thus obtained is given in Annexe 1

6. Conclusion

We described the algorithm of the mini-AES and of the AES as Boolean functions and then we translated it into systems of equations using the Möbius transform of these functions. Now, the goal of our approach is to be able to perform a combinatorial analysis on the files of equations thus obtained. The use of a Boolean equation system of low degree is motivated by the fact that its solution is likely to be easier than the existing equation model. We are currently developing a new, combinatorial approach in Boolean equation system solving which seems to be promising.

7. Annexe 1

File of the low level bit block describing the first round of the AES

```
## addRoundKey
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001
## subBytes
1          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000100
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000110
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001000
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001010
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001100
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001110
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001111
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010000
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010001
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010010
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010011
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010100
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010101
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010110
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011001
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011010
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011111
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000100001
0          00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000100101
```

```
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000100111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000101001
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000101100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000101101
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000101110
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000110011
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000110101
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000110111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000111101
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001000001
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001000010
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001000011
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001000100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001000101
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001000110
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001000111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001001001
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001001110
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001001111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001010000
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001010001
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001010010
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001010110
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001011001
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001011010
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001011011
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001011101
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001011111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001100000
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001100010
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001100100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001100101
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001101001
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001101011
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001101110
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001110000
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001110001
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001110010
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001110011
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001110100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001110110
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001111001
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001111100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001111101
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010000011
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010000100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010000101
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010000111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010001010
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010001100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010001101
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010001110
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010001111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010010001
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010010011
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010010100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010010101
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010010110
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010010111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010011000
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010011111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010100000
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010100100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010100101
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010100111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010101000
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010101011
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010101100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010101110
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010101111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010110001
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010110010
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010110100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010110101
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010110111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010111001
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010111010
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010111100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010111101
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011000000
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011000100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011000110
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011000111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011001000
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011001010
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011001011
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011001100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011001111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011010001
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011010101
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011010110
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011011001
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011011011
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011011110
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011011111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011100000
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011100010
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011100011
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011100110
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011100111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011101000
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011101001
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011101101
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011110011
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011110100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011110110
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011110111
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011111010
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011111100
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000011111101
## shiftRows
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000100000000000000000000000000000000
## mixColumns
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000001000000000000000000000000
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010000000000000000000000000000000
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010000000000000000
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000100000000
0            00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000010000000
## end
```

References

Assmus, E.F. (1992) **On the Reed-Muller Codes**, Discrete Math

Courtois, N. and Pieprzyk, J. (2002) **Cryptanalysis of Block Ciphers with Overdefined Systems of Equations**, Lecture Notes in Computer Science

McCarty,P (1986) **Introduction to Arithmetical Functions**, Springer Verlag

FIPS (2001) **Advanced Encryption Standard**, http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf

Hamming, R. (1950) **Error Detecting and Error Correcting Codes**, The Bell System Technical Journal

Murphy, S. and Robshaw, M. (2002) **Essential Algebraic Structure Within the AES**, Crypto 2002 LCNS 2442

Phan, R. (2002) **Mini Advanced Encryption Standard (Mini-AES): A Testbed for Cryptanalysis Students**, Cryptologia

Rijmen, V. and Daemen, J. and Preneel, B. and Bosselaers, A. (1996) **The cipher SHARK**, {3rd International Workshop on Fast Software Encryption

Sakkour,B. (2007) **Etude et amélioration du décodage des codes de Reed-Muller d'ordre deux**, Ecole polytechnique

Shannon, C.E. (1948) **A Mathematical Theory of Communication**, Bell System Technical Journal

Shannon, C.E. (1949) **Communication theory of secrecy systems**, Bell System Technical Journal

Law of Armed Conflicts Applied to i-Warfare and Information Operations: How and Under What Legal Framework Should Surgical NATO and U.S. Military Drone Strikes be Conducted?

Berg Hyacinthe
Assas School of Law, CERSA/CNRS, Université Paris II, Sorbonne; France
cyberjuriste@yahoo.com

Abstract: When computers and computer systems are treated and utilized as weapons of war, it becomes more difficult to deny legal implications under domestic and international laws. In effect, many legal scholars have called for Rules of Engagement that would govern i-Warfare conducts, while taking into consideration the applicable notions of the laws of armed conflicts (LOAC). However, even the most powerful military in the world is still struggling with i-Warfare's legal ambiguities and technical challenges. As U.S. Army Gen. Martin Dempsey, chairman of the Joint Chiefs of Staff, acknowledged, during his Senate confirmation hearings, he is "not particularly well versed" on the subject of cyber warfare — a lacuna shared with his predecessor, Navy Adm. Mike Mullen. Nonetheless, they both agree on the need to combine diplomatic, military and economic prowess, in order to neutralize one of the most challenging threat to U.S. national security today: *asymmetric cyber warfare*. Several indicators continue to show that the U.S. strategic agendum aimed at "information dominance" would fail, if it does not include a competent level of knowledge of the laws that should — and will likely govern — the use of digital information technologies in warfare during this millennium onward. To acquire such knowledge, the most basic concepts found in such solid legal doctrines as Thomas Hobbes's *positive law* paradigm or/and John Locke's *moral obligation* principle may be helpful. Hence, this article highlights the danger associated with the introduction of new "high tech" weapons of war on the battlefield, in the absence of appropriate legal measures as required by laws of armed conflicts, military field manuals of modern States, and under Article 36 of 1977 Additional Protocol I to the Geneva Conventions of 1949. Furthermore, it discusses how and under what conditions military drone strikes may be conducted, to comply with domestic laws as well as international conventions and treaties. The author concludes with the following call to U.S. and NATO officials: protect your cyber warriors against obvious cybercrimes of war and "cyber-boosted" crimes against humanity by enacting appropriate and consensus-driven legislations.

Keywords: space sovereignty, information operations in space, space laws & treaties; ASAT & communications satellites; military drones; computers-as-weapons

1. Introduction and historical background

> *"If other states were to claim the broad-based authority that the United States does—to kill people anywhere, anytime — the result would be chaos".*
>
> *— Philip Alston, former UN special rapporteur on extrajudicial, summary, or arbitrary executions*

Following the Kosovo war, NATO officials faced criminal complaints for telecommunication posts destroyed in Prikili (Mandel *et al*. 1999). Operation Allied Force incorporated the first large-scale use of satellites as a direct method of weapon guidance. Space-based technologies currently activated in U.S. and NATO operated military drones contributed greatly to the Operation's military success in 1999.

Separately, when the Saudis appealed to America for imagery from U.S. surveillance satellites in Space, during the January 2009 border conflict with Yemen, the State Department warned that intervening in this border conflict, even if only by providing targeting information, could violate the Law of war. However, the French agreed to help, apparently under a different legal reasoning. Shortly thereafter, using this *precise satellite intelligence*, the Saudis were able to monitor the Houthis' hideouts, equipment dumps and training sites. Saudi warplanes then attacked with devastating effectiveness. Within a few weeks, the Houthis were requesting a truce, and by February 2009 this chapter of the border war was over. Indeed, to America's unpleasant surprise, a new French-Saudi intelligence channel was created. And, eight months later, U.S. authorities took a different stand against Baitullah Mehsud:

> *On August 5th [2009], officials at the Central Intelligence Agency (CIA), in Langley, Virginia, watched a live video feed relaying close-up footage of one of the most wanted terrorists in Pakistan. Baitullah Mehsud, the leader of the Taliban in Pakistan, could be*

> *seen reclining on the rooftop of his father-in-law's house, in Zanghara, a hamlet in South Waziristan. It was a hot summer night, and he was joined outside by his wife and his uncle, a medic; at one point, the remarkably crisp images showed that Mehsud, who suffered from diabetes and a kidney ailment, was receiving an intravenous drip. The video was being captured by the infrared camera of a Predator drone, a remotely controlled, unmanned plane that had been hovering, undetected, two miles or so above the house. The image remained just as stable when the CIA remotely launched two Hellfire missiles from the Predator. Authorities watched the fiery blast in real time. After the dust cloud dissipated, all that remained of Mehsud was a detached torso. Eleven [11] others died: his wife, his father-in-law, his mother-in-law, a lieutenant, and seven bodyguards (Meyer 2009).*

Today, the danger is too great to ignore the spirit of the law, as it applies to Space laws, the laws of armed conflicts, and the Geneva Convention (Wingfield 1999). Indeed, "peaceful purposes" may be superficially interpreted as to justify offensive military actions concocted in Space. However, the type of (non-aggressive) military activities admitted had been clearly described: "The use of military personnel for scientific research or for any other peaceful purposes shall not be prohibited. The use of any equipment or facility necessary for peaceful exploration of the moon and other celestial bodies shall also not be prohibited" (Outer Space Treaty 1967). In this context, "peaceful purposes" are linked to scientific research and exploration in the best interest of mankind, as opposed to Space-based offensive strikes in the name of international peace and security.

The military drones are components of a Space-based integrated weapon system. Yet, many credible questions loom over violation of the Law of war both in domestic and international legal regimes (DiCenso 1999). Of course, the proliferation and widespread use of military drone technologies will ultimately force U.S. authorities and their allies to regulate such weapons — the sooner the better. The author adopted juridical analogies and case studies as two relevant methodological approaches for this type of research.

2. U.S. and NATO's use of legally "untested" digital weapons of war

According to former CIA director Retired Gen. Michael Hayden, a former CIA Director under George W. Bush, the deadly military drone "strikes are arguably the most critical weapon in the U.S. arsenal against al Qaeda". However, critics of drone strikes, including a group of the American Civil Liberties Union (ACLU) lawyers, argue that these borderline legal acts war undermine the U.S. cause by (1) killing innocent civilians, (2) infringing on the sovereignty of other nations and (3) generating sympathy for Al Qaeda and other extremist organizations (American Civil Liberties Union 2010). Indeed, the United States is setting a negative precedent likely to be invoked by other nations acquiring similar technologies.

As established herein, U.S. and its NATO allies have been warned on a number of occasions of the danger associated with the introduction of new weapons and weapon technologies on the battlefield. The legality of their actions has been contested both formally and informally, with several legal scholars and jurists complaining through academic *fora* and arguing before U.S. federal Courts (Habid and Sayah 2012; American Civil Liberties Union 2010). Under international law, the aforementioned actions could yield into official diplomatic protests and/or *démarches* (Palwankar 1994), as the malaise persists.

When legalisms are used as means to circumvent the ban on direct use of Space in armed conflicts, whether as relay, instrument, weapon or gateway, Article IV of the Outer Space Treaty should apply (Outer Space Treaty 1967). Where the mere participation of military personnel in the installation of military satellites would not constitute a violation, the use of such satellites in connection with military drone strikes — already conducted under ambiguous legal conditions — further complicates the task of qualifying such conducts as legal.

Even where crimes are not committed, the conducts in question are concocted in such manner as to preclude war crimes investigators from discovering abuses and violations committed by remotely guided weapons, apparently very difficult, yet not impossible, to trace under existing laws. This situation is unacceptable. It is therefore irresponsible to create a situation where war crimes investigators are left with a robot (machine) as the *aggressor*, a signal (electronic transmission) as the *military order* and the equipment (source of the signal) as the *military commander*. This intentional process of dehumanization/depersonalization, as it relates to war crimes, cannot be permitted under any circumstance in modern societies. Where it may be argued that the signal is not deadly, the same cannot be said of the equipment that emits it. When chips or "tiny computers" are embedded with precision-guided munitions, computers become integral components of a deadly military arsenal. Indeed, components found in military drones today, are parts of a deadly Space-based offensive weapon system believed to be in violation of international laws, where mass casualty is involved.

3. Moral obligations: From Thomas Hobbes to John Locke and Col. Brian Ellis

Thomas Hobbes's *positive law* paradigm relates to a precept of general rule found out by reason, by which a man is forbidden to do that which is destructive of his life (Hobbes 1640). The development of very sophisticated Space-based weapons of mass destruction is a prime example of self-destructive endeavor, as emergent military powers continue to look Spaceward for *information dominance*.

Max Weber presented the political systems of modern Western societies as forms of "legal domination". Their legitimacy is based upon a belief in the legality of their exercise of political power. It is the rationality intrinsic to the form of law itself that secures the legitimacy of power exercised in legal forms (Weber 1964). However, Habermas counter-argued that legality can derive its legitimacy only from a procedural rationality with a moral impact (Habermas 1986).

John Locke departed from Hobbes in describing the state of nature as an early society in which free and equal men observe the natural law. He argued that people have rights, such as the right to life, liberty, and property that have a foundation independent of the laws of any particular society. Locke used the claim that men are naturally free and equal as part of the justification for understanding legitimate political government as the result of a social contract where people in the state of nature conditionally transfer some of their rights to the government in order to better insure the stable, comfortable enjoyment of their lives, liberty, and property. Since governments exist by the consent of the people in order to protect the rights of the people and promote the public good, governments that fail to do so can be resisted and replaced with new governments. Locke also defends the principle of majority rule and the separation of legislative and executive powers (Tuckness 1999).

This notion of separation of powers, anchored in Montesquieu's *Esprit des lois*, is present in the U.S. Constitution to ensure that appropriate legislations are not blocked by one branch, when dealing with such controversial issues as execution of American citizens without trial by American-operated military drones (Williams 2011). Indeed, very little effort has been deployed to come up with specific legislations covering military drone strikes launched by U.S. and NATO cyber warriors against "foreign" and national targets.

U.S. officials have been aware of this dangerous legal lacuna for some time now: Col. Bryan Ellis who recently served as Commander of the 35th Signal Brigade, 18th Airborne Corps, Fort Bragg, N.C., acknowledged that "Current U.S. criminal statutes apply to Information operations. Similarly, foreign criminal statutes will most likely apply to U.S. information activities" (Ellis 2001); Professor Michael N. Schmitt, Chairman of the International Law Department at the United States Naval War College, has made similar recommendations to U.S. and NATO officials (Schmitt 2004). In fact, Ellis went further, addressing U.S. officials in the following terms, "… misuse of information attacks could subject U.S. authorities to war crimes" (Ellis 2001).

In a natural state all people were equal and independent, and everyone [States and individuals, victims of drone strikes included] had a natural right to defend his "Life, health, Liberty, or Possessions" (Locke 1640). This became the basis for the phrase in the American Declaration of Independence: *Life, liberty, and the pursuit of happiness*. Hobbes and Locke's ideas would come to have profound influence on the Declaration of Independence and the Constitution of the United States. Quietly, emergent military powers start to build on the same principles to justify their own quest for happiness through advanced science and technology initiatives.

4. Objection to and protest against U.S. and NATO's military drone strikes

Unfortunately, several U.S. military and intelligence officials continue to suggest that laws intended to regulate the use of military drone operations both during armed conflicts and peacetime are not in America's best interest (Moore, Grimaila and Strouble 2009), thereby, ignoring thoroughly documented warnings and recommendations from Col. Brian Ellis and Professor Michael Schmitt, to avoid an exhaustive list.

The popular view pointing to "exploitable" opportunities in Pakistan, Afghanistan, Iraq, Yemen and elsewhere is understandable; yet, it does not justify the current state of anarchy in which military drones strikes are conducted. The *status quo* allows legalism and shortsightedness to reign over conventional wisdom and foresight for too long. The U.S., which suspended drone strikes in Pakistan following the November 26, 2011 NATO attack that left twenty four (24) Pakistan Army soldiers dead, later resumed the attacks over the objection of Pakistani officials and other concerned parties. Pakistani officials protested by directing U.S. military personnel to vacate the key Shamsi airbase that was used to launch the attacks.

Pakistani officials recently disputed claims of any agreement pertaining to the military drone strikes on their territory. Hence, the Pakistani Parliamentary Committee on National Security, a group of 18 members of Parliament responsible for reviewing relations with the United States, made the recommendation in a report to lawmakers in March 2012: "No overt or covert operations inside Pakistan shall be tolerated" (Habid and Sayah 2012). Hina Rabbani Khar, the Pakistani foreign minister, later added "No entity in Pakistan — in this current government, because I can only speak for this government — has ever given any tacit agreement to the authorization of drone strikes" (Habid and Sayah 2012). By saying "current government", the Pakistani foreign minister left the door open, when it comes to former Pakistani government or intelligence officials. Therefore, it is worth noting this prescient warning issued back in 2009:

> *Given that the CIA may have obtained some form of "covert" laissez-passer from political and/or military Pakistani officials to comply with international laws (e.g., dealing with the sovereignty issue, prior to initial strikes), legal analysts need to proceed with prudence. As a 99% clandestine agency, the CIA is not under any obligations to comment publicly on on-going sensitive operations, though the rule of law must always prevail according to statutory reporting requirements and congressional oversight outlined in the National Security Act of 1947 (Hyacinthe 2009, p.183).*

What about the evolution of drone and Space technologies throughout States overtly or covertly placed on the axis of evil? Will future laws only apply to their military scientists? Indeed, as emergent military powers acquire more sophisticated Space technologies, the legal ambiguities maintained over deadly military drone strikes are likely to play against the U.S. and its NATO allies. In fact, some of the Space-based military satellites are, *de lege lata*, strategic weapon components operating subtly on non-sovereign territories (Hyacinthe and Fleurantin 2008). As stated in the Outer Space treaty "The moon and other celestial bodies shall be used by all States Parties to the Treaty exclusively for peaceful purposes" (Outer Space Treaty 1967).

As David Ignatius reported, "A new arsenal of drones and satellite-guided weapons is changing the nature of warfare. America and its NATO allies possess these high-tech weapons, but smaller countries want them too" (Ignatius 2010). There has been a widespread, yet unfortunate, acceptance by the general public of a myth: military drones operate without heavy reliance on "eyes and ears" on the ground. In reality, contrary to the extraordinary autonomy projected through military propaganda and by profit-minded defense contractors, the efficiency of some of these drones comes with the ultimate sacrifice of human intelligence on the ground —meaning brave servicemen and servicewomen, infiltrating very dangerous organizations and forbidden enemy territories (Hyacinthe 2009, p.182).

Accordingly, Jane Harman, a former top member of the U.S. House Intelligence and Homeland Security committees, rightly cautioned about the need for strict guidelines in the use of drone strikes, which have increased under Obama's watch. "We could abuse this program," Harman said, "We've got to have a counter-narrative (to dissuade potential terrorists). We've got to live our values" (CNN Wire Staff 2011). On the NATO side, the Standardization Agreement (STANAG) 2449 sets out a minimum standard of necessary training for NATO forces and provides an outline of training programs useful to all military personnel, from noncommissioned to commissioned officers specifically. Training should also involve proper legal advice with regard to potential charges of "cyber war crimes" and "cybercrimes against humanity".

NATO recently placed a command to purchase several U.S. made military drones (Pawlak 2012). As described *infra*, America has faced several Court challenges over the legality of some of these weapons in international theaters, notably in Yemen as well as in Pakistan. Thirteen (13) NATO allies - Bulgaria, the Czech Republic, Estonia, Germany, Italy, Latvia, Lithuania, Luxembourg, Norway, Romania, Slovakia, Slovenia and the United States - will participate in the procurement of the unmanned drones and other equipment such as deployable ground stations that will support the Alliance Ground Surveillance (AGS) project. NATO officials, beyond their eager to sell/buy new weapons and transfer new weapon technologies, following a poor performance in Libya, should take a closer look at their obligation, when it comes to the "study, development, acquisition or adoption of a new weapon, means or method of warfare" under Article 36 of 1977 Additional Protocol I to the Geneva Conventions of 1949 (Daoust *et al*. 2002).

5. How / when should military drone strikes be conducted?

U.S. military drone strikes, delivering payloads of hellfire missiles, should be conducted only during wartime. Accordingly, a declaration of war is highly recommended. The U.S. military and its spy agencies are operating military drones during wartime in Afghanistan and Iraq as they have done during peacetime in Pakistan, Yemen and elsewhere. As a result, colossal damages including a heavy toll of human casualties have been recorded. Unfortunately, these operations are commonly effectuated under a legal regime that is, at best, ambiguous in important ways.

For instance, even where a secret pact between intelligence agencies would restrain Pakistani military retaliation against the U.S. and its NATO allies, from a foreign policy and legal perspective, fierce condemnation of these drone strikes must not be ignored. Under international law, it is through documented *diplomatic protests* that a competent Court will be able to hear complaints filed retroactively against belligerent States (McKenna 1962). Many Pakistani officials claim that the U.S. drone strikes violate their State sovereignty. Others claim indiscriminate killing of non-combatants including women and children. Though "targeted killings" have become a publicly accepted U.S. foreign policy, it does not guarantee compliance with international law. State Department lawyers can always argue in favor of unilateral military actions; but it will be difficult for them not to tarnish America's reputation as "peacemaker" in the process. The "good guy" image had been America's must effective weapon against its rivals.

As MacGibbon observed: "It is clear, however, that States are under no obligation to refrain from protesting until actual violations of their rights have been committed". The essence of such a protest is to give due notice to the respondent State that the protesting State regards the action as injurious to the interests of its nationals and contrary to international law. It, therefore, gives the Sate against whose action the protest is lodged an opportunity to withdraw or amend the objectionable action (MacGibbon 1953, p.299). U.S. Congress has established how the Commander-in-Chief may commit U.S. troop to the battlefield. In the best case scenario, a war begins only following authorization from Congress, according to the War Powers Act. And when military necessity obliges otherwise, Congress needs to be notified in a timely manner. Even covert Information Operations (IO) activities require (legal) presidential approval (Wingfield 1999).

6. Legal framework for military drone strikes

Under the United Nations Charter, the United States would normally be prohibited from using force inside Pakistan without obtaining Pakistan's consent. Accordingly, on a Complaint for Injunctive Relief dated March 16, 2010, before United States District Court, District of Columbia, ACLU lawyers argued in favor of strict guidelines that should govern such operations. The Complaint sought "a variety of records relating to the use of unmanned aerial vehicles to conduct targeted killings, including the legal basis for the strikes and any legal limits on who may be targeted; where targeted drone strikes can occur; civilian casualties; which agencies or other non-governmental entities may be involved in conducting targeted killings; how the results of individual drone strikes are assessed after the fact; who may operate and direct targeted killing strikes; and how those involved in operating the program are supervised, overseen or disciplined" (American Civil Liberties Union 2010).

In any civilized society, it would have been beyond shocking for wars to begin by a computer click and end in similar fashion, while unruly cyber warriors involved hide behind *nullum crimen*, *nulla pœna sine lege* (nothing is a crime without a preexisting law) and *nulla poena sin lege* (no punishment may be imposed without a preexisting law authorizing it). As argued *infra*, with regard to military drone strikes, the legal framework covering conduct, damage and responsibility must be established. The first step may involve a consensus-driven juridical notion of information warfare, supported by the appropriate taxonomy of various terms and concepts. U.S. Air Force scholars claim that "Beyond strict compliance with legalities, U.S. military activities in the information environment as in the physical domains are conducted as a matter of policy and societal values on a basis of respect for fundamental human rights. U.S. Forces, whether operating physically from bases or locations overseas or from within the boundaries of the US or elsewhere, are required by law and policy to act in accordance with US law and the Law of Armed Conflict" (Moore, Grimaila and Strouble 2009, p.71). However, Standing Rules of Engagement (ROE) issued by U.S. authorities (Lawyer 2000) are not laws. As such, they should not — even indirectly — be presented as cure to the current state of anarchy, against which legislations are much needed.

USC Title 10 (Armed Forces), USC Title 50 (War and National Defense), and USC Title 32 (National Guard) cover relations between the Congress and the Executive; but Title 50 was recently amended to ensure the Secretary of Defense can use Department of Defense components, including the National Security Agency (NSA), to accomplish intelligence missions without regards to statutory divisions of the respective intelligence responsibilities (Walker 2008; Moore, Grimaila and Strouble 2009, p.72). At issue are specific laws that should govern the use of digital information technologies, particularly military drones, on the battlefield. Short of a formal declaration of war, legalism will fail to justify military drone strikes during peacetime against a sovereign State. There is no consensus on a legal framework (Joyner and Lotrionte 2001) permitting such conducts. As a result, diplomatic protests (MacGibbon 1953) properly filed by wary and concerned States could further complicate NATO's recent position on this issue. The heated debate among EU members over the acquisition of several legally "untested" military drones is a clear warning sign.

7. Conclusion

In sum, as argued throughout this document, the military field manuals of most modern States, the U.S. Constitution, the Law of Armed Conflicts, as well as articles 2(4) and 51 of the Geneva Convention can pave the way to constructive juridical analogies and sound legal reasoning, in dealing with legality/illegality of military drone strikes by U.S. and NATO armed forces. Ultimately, wary jurists with expertise in this field will join forces with "savoir-faire" in favor of an adequate legal framework: thereby, establishing, in very clear and explicit terms, the laws and regulations that should govern the use of digital information and technologies as weapons of war, particularly in armed conflicts (Brown 2006). The author does not call for prohibition of military drones on the battlefield. He argues in favor of a legal framework, as required by law. It might have been a political victory to move ahead with the killing of an American citizen without a trial, in the case of Anwar al-Awlaki, despite an executive order banning assassinations. However, on the legal front, many challenges lay ahead for U.S. authorities, as potential enemy States look "drone-ward". Will Iran be able to use its drones under similar conditions to attack American targets anywhere? What about China and Russia?

The State Department's senior legal adviser, Harold Koh, plainly stated the Obama administration's view that it had authority to undertake drone attacks in countries where al-Qaida operatives were located (Williams 2011). However, the Administration's view does not replace the law. It is interesting to note that no legal reference had been given to justify such actions: a law is needed and Congress must act on it. Why? Emergent military powers are likely to use the same weapons, with a major difference: the target. It will become more difficult to dissuade them. Later, rogue States might be tempted to use the same weapons with impunity.

Lastly, the law must evolve with the people and institutions it is intended to protect. According to the French notion of "*sécurité juridique*", for example, there is a need for fail-safe legal measures intended to maintain, at the intersection of social interactions and technological innovation, new military conducts in compliance with the rule of law. As recent history suggests, the same military officials who are fighting against regulations today are likely to fight against retroactive applications tomorrow. Short-term U.S. military objectives set for Iraq, Afghanistan, Yemen, and Pakistan should not eclipse long-term strategic goals. Time has come to regulate military drone strikes on the battlefield. U.S. and NATO officials should to reckon with this new reality. When and if Pakistani and other concerned parties properly file a diplomatic protest over the deadly drone strikes, U.S. and NATO officials will have a new challenge to overcome under international. Therefore, the Secretary General of the United Nations could play an important role in addressing this serious matter to avoid a more complicated, disastrous situation.

References

American Civil Liberties Union (2010) "Complaint for injunctive relief", *Unites States District Court, District Court of Columbia*, [*online*], http://www.aclu.org/files/assets/Drones.1.Complaint.pdf, (last visited March 31, 2012)(case: 1:10-cv-00436-RMC).

Brown, D. (2006) “A Proposal for an International Convention to Regulate the Use of Information Systems in Armed Conflict”, *Harvard International Law Journal,* vol. 47, no.1, pp. 179-221.

Daoust, I., Coupland, R. and R. Ishoey, R. (2002) “New wars, new weapons? The obligation of States to assess the legality of means and methods of warfare”, *International Review of the Red Cross*, June 2002, vol. 84, p. 354.

CNN Wire Staff (2011) “Cheney: Obama should 'correct' criticism of Bush's anti-terror tactics”, [*online*], http://edition.cnn.com/2011/10/02/politics/cheney-obama-terror/index.html?iref=obinsite, (last visited March 31, 2012).

Ellis, B. (2001) “The International Legal Implications and Limitations of Information Warfare: What Are Our Options?”, [online], *U.S. Army War College*, http://www.iwar.org.uk/law/resources/iwlaw/Ellis_B_W_01.pdf (last visited Mar. 15, 2012).

Habermas, J. (1986) "Law and Morality", *The Tanner Lectures, Harvard University* [Oct. 1 & 2].

Habib, N. and Sayah, R. (2012) “Pakistan lawmakers to debate end to U.S. drone strikes”, *CNN*(03.26.2012), [online], (http://edition.cnn.com/2012/03/26/world/asia/pakistan-us-relations/index.html), (last visited March 30, 2012).

Hobbes, T. (1640) Elements of Law, Natural and Politic*, London.*

Hyacinthe, B. (2009) *Cyber Warriors at War*, Xlibris, pp. 238.

Hyacinthe, B. and Fleurantin, L. (2008) “Initial Supports to Regulate Information Warfare’s Potentially Lethal Technologies and Techniques”, *3rd International Conference on Information Warfare and Security, Peter Kiewit Institute of the University of Nebraska,* USA.

Ignatius, D. (2010) “Dazzling new weapons require new rules for war”, *Washington Post,* (Nov. 11, 2010), [online], http://www.washingtonpost.com/wp-dyn/content/ article/2010/11/10/AR2010111005500.html (last visited march 30, 2012).

Joyner, C. and Lotrionte, C. (2001) “Information Warfare as International Coercion: Elements of a Legal Framework”, 12 *Eur. J. Int'l L.* 825, pp.825-865.

Lawyer, A. (2000) “How to Keep Military Personnel from Going to Jail for Doing the Right Thing: Jurisdiction, ROE & the Rules of Deadly Force”, *The Army Lawyer*”, [online], http://www.loc.gov/rr/frd/Military_Law/pdf/11-2000.pdf (last visited march 15, 2012).

Locke, J. (1690) *Second Treatise of Government* (10th ed.), *Project Gutenberg*. [online], http://www.gutenberg.org/files/7370/7370-h/7370-h.htm, (last visited March 30, 2012).

MacGibbon, I. (1953) “Some Observations on the Part of Protest in International Law”, *British Year Book of International Law*, vol. 30, pp. 293-299.

Mandel, M. (1999) “Request that the Prosecutor investigate named individuals for violations of international humanitarian law and prepare indictments against them pursuant to articles 18.1 and 18.4 of the Tribunal Statute,” *ICT*, [online], http://jurist.law.pitt.edu/icty.htm, (last visited April 7, 2012).

McKenna, J. (1962) *Diplomatic Protest in Foreign Policy: Analysis and Case Studies*, Loyola University Press, pp. 222.

Meyer, J. (2009) “The Predator War”, *New Yorker,* [online], http://www.newyorker.com/reporting/2009/10/26/091026fa_fact_mayer, (last visited March 30, 2012).

Moore, T., Grimaila, M. and Strouble, D. (2009) “Laws and Regulations of USAF Military Operations in Cyberspace”, *4th International Conference on Information Warfare and Security*, Cape Town, South Africa, pp.68-76.

New York Times (2011) “Drones Transform How America Fights Its Wars”, [online], http://www.nytimes.com/slideshow/2011/06/20/world/20110620-DRONES-8.html (last visited March 31, 2012).

Outer Space Treaty (1967), *Geneva Convention,* United Nations.

Pawlak, J. (2012) “NATO to buy U.S.-made unmanned drone aircraft”, *Reuters (02/15/2012)*, [online], http://news.yahoo.com/nato-buy-u-made-unmanned-drone-aircraft-180711891.html (last visited April 3, 2012).

Palwankar, U. (1994) “Measures available to States for fulfilling their obligation to ensure respect for international humanitarian law”, *Review of the Red Cross*, No. 298, p. 9.

Schmitt, M. (2006) “Effects-Based Operations and the Law of Aerial Warfare”, *Washington University Global Studies Law Review,* vol. 5, p. 265.

Tuckness, A. (1999) “The Coherence of a Mind: John Locke and the Law of Nature”, *Journal of the History of Philosophy*, 37: 73–90.

Weber, M. (1964) *Wirtschaft und Gesellschaft* [English translation: economy and society], Cologne, ch. 3, pp. 2, 160ff.

Williams, P. (2011) “Can U.S. legally kill a citizen overseas without due process?”, *MSNBC* [online], http://openchannel.msnbc.msn.com/_news/2011/09/30/8063632-can-us-legally-kill-a-citizen-overseas-without-due-process (last visited march 28, 2012).

Wilson, T. (2001) “Threats to United States Space Capabilities”, *U.S. Space Commission,* [online], http://www.fas.org/spp/eprint/article05.html#8 (last visited Mar. 15, 2012).

Wingfield, T. (1999) "Legal Aspects of Offensive Information Operations in Space”, *Journal of Legal Studies* [USAFA], v. 9, pp 121-146.

Cyber Threat Management in Cognitive Networks

Anssi Kärkkäinen
Defence Command Finland, Helsinki, Finland
anssi.karkkainen@mil.fi

Abstract: Threats and attacks in cyberspace are growing in number and sophistication, originating different sources and encompassing intentional attacks as well as inadvertent causes. Despite ongoing development of security products, many organizations feel their infrastructure is inadequate for combating rapidly evolving threats. Next generation information and communication technology provides cognitive networking capabilities which also are challenging from security point of view. The cognitive network is defined as a network with a cognitive process that can understands current conditions, plan, decide, act on those conditions, and learn from the results of actions. This adaptive and self-acting behavior of the network requires new approaches to cyber threat management. Risk management and security mechanisms of the network must adapt to cyber threats and dynamically provide a coordinated response in real-time. The paper presents a framework for cyber threat management system for the cognitive networks. The framework consists of three layers. The first layer includes a single network node, the second one contains a cluster of nodes and the third one covers the entire network. The framework describes a risk assessment process, and includes also security policy aspects. Also, implementation challenges of the proposed framework are discussed in this study.

Keywords: cyber threat, cognitive network, threat management

1. Introduction

Next generation communication infrastructure is based on cognitive networking that provides better capabilities to use network resources efficiently. These smart networks include automated cognitive features that control and manage network elements independently. Traditional human based network management is required no longer. The network elements are able to adapt to the goals set by human end users or even services. Smart networks have lots of commercial drivers (e.g. use of heterogeneous infrastructure and technologies), but there might be even more interests in military environments in which networks are required to be dynamic, reliable and very adaptive.

Cognitive networks (CN) are a promising approach to improve network management and security. CN provides a smart communication platform which could observe its internal and external environment, plan, decide and adjust its parameters as a result of this process. The adjustment is done according to the desired goal, which could be set by users, applications or other services depending on situation. Because of this automated functioning, the hostile environment and dynamically changing goals, overall security of the network is challenging to obtain, although the cognitive layer is in theory able to take care of all the security requirements.

Cognitive networks create new challenges and threats from security perspective while human control over the network decreases. At the same time, threats and attacks in cyberspace are growing in number and sophistication, originating different sources and encompassing intentional attacks as well as inadvertent causes. Although the cognitive networks provide an ideal platform for cyber attacks (in sense of lacking human control) their adaptive nature increases survivability. In (Shore 2010) two strategies for survival are presented: survival by protection or survival by adaptation. Survival by protection involves the careful application of security mechanisms to restrict the effect of attacks and ensure the success of the essential components of the service. Survival by adaptation typically involves monitoring and changing the Quality of Service (QoS) available to applications, either increasing the QoS for essential services or reducing the QoS for non-essential services. The CN is a relevant example of this adaptive approach. Of course, one of the most difficult problems with CN is that of knowing when to adapt and what kind of adaptation to apply. Characteristics relevant to survivable system architectures include performance, security, reliability, availability, and modifiability.

Research on network security has been carried out a lot. This research also includes cognitive networks, but typically security research focuses on a specific part or element of a network system, and thus provides results from a very narrow point of view. Cyber threats are studied in many papers, but an overall threat management framework for CN is not presented. Thus, the purpose of this paper is to describe a framework that includes overall view of threat management. For cognitive networking the framework is layered to three layers; node, cluster and an entire network. The performance and implementation of the framework is not tested during this study, but the basic structure of the framework is presented, and functionality is described.

The structure of the paper is following. In Section 2 related work is presented. Section 3 discusses the security aspects of the cognitive networks and describes potential cyber threats related to the cognitive networks. Section 4 proposes a framework for threat management in cognitive networks. Section 5 discovers some implementation challenges and the final section concludes the paper.

2. Related work

Security threats of communication networks are widely discussed in various papers. Also, cognitive network research has focused more and more on security issues and challenges. Most of the studies concentrate on a certain piece of security in cyberspace, e. g. security threats and detection techniques (Fragkiadakis et al 2012) or control channel security (Safdar et al 2009).

Also threat management is studied in the other contexts than communication networks. For example, in (King et al 2005) a comprehensive threat management framework for a crop biosecurity national architecture is presented. The framework is relevant for biosecurity threats and has little to deal with cyber security. There is also research concerning critical infrastructure as power industry. In (Jiaxi et al 2006) vulnerability assessment methods for power industry are proposed. In (Atmaja et al 2011) a cyber security strategy for future distributed energy delivery systems is studied. These studies approach cyber threat at very high-level while the scope of this paper is to describe functionalities related to a cyber threat management framework.

Threat detection, cyber security threats and mission assurance are discussed in (Buford et al 2008), (Kawano et al 2005), (Morris et al 2011) and (Bodeau et al 2010). These studies still have quite narrow scope and they do not propose any overall framework for cyber threat management. Morris et al focus on the description of the system on cyber mission information needs, whereby collection, processing, management and mission model updates are based on cyber-related information from a variety of resources including commercial news, blogs, wikis, and social media sources. The result is a dynamic capability for cyber mission management that provides proactive, on demand cyber information to analysts, professionals, policy makers, and support personnel.

Bodeau et all present the cyber preparedness methodology that enables an organization to characterize the cyber threats, determine the level of preparedness necessary to ensure mission success, and facilitate strategic planning for cyber security to establish priorities for cyber security investment planning and management decisions.

Enterprise level information security management is discussed in several research papers (Choi et al 2008), but they approach threat management from business processes side. The purpose is to achieve the optimal level of security management in strategy planning.

3. Security aspects of cognitive networks

3.1 Cognitive networks

A basic idea behind cognitive networks is discussed in (Mahmoud 2007), (Thomas et al 2005) and (Thomas et al 2006). CN is a network can dynamically adapt its operational parameters in response to user and service needs or changing environmental conditions. The networks can learn from these adaptations and exploit knowledge to make future decisions. The applications of cognitive networks enable the vision of pervasive computing, seamless mobility, ad-hoc networks, and dynamic spectrum allocation, among others.

Cognitive networks are needed simply because they enable users to focus on things other than configuring and managing networks. Especially in a military tactical environment the soldiers should not concentrate on configuring the network devices, but their core business. Manual configuration also provides a higher risk for misconfiguration and may take too much time to redeploy systems in rapid military operations.

Unlike cognitive radios, CN does not restrict the scope in radio spectrum. A cognitive network tries to exactly perceive the current network situation and plan and decide to meet the end-to-end goals in an entire network aspect. A cognitive process in such networks could be viewed as the commonly known

OODA loop in which a network observes orients, decides and acts. Figure 1 shows the phases of the loop in context of cognitive networking.

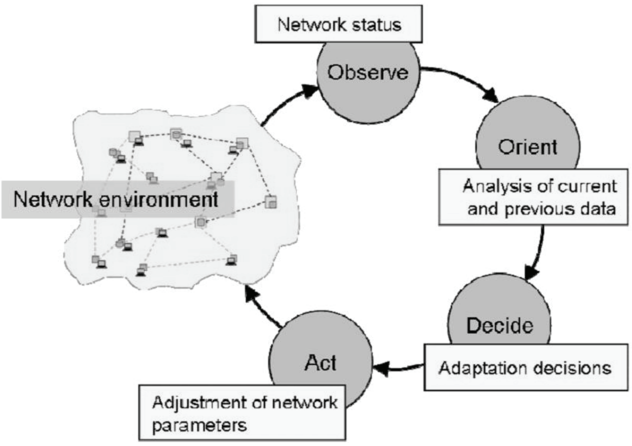


Figure 1: Cognitive process follows the OODA loop

The observation phase is critical because the effect of a cognitive network's decisions on the network performance depends on how much network state information is available. If a cognitive network has knowledge of the entire network's state, cognitive decisions should be more "correct" than those made in ignorance. For a large, complicated system such as mobile ad hoc networks, it is unlikely that the cognitive layer of the network would know the total system state. It could be very high costly to communicate status information beyond those network elements requiring it, meaning the cognitive network will have to work with less than a complete picture of the network resource status.

The orientation phase also plays an important role in the cognitive process. In this phase all observed information and previous knowledge are add together and analyzed. Filters and weighting are examples of methods used in the orientation phase. In the decision phase the best decision for the required end-to-end data flow capability is made. Finally, actions are taken in the acting phase. Action includes modifications of cognitive network elements. These elements associated with each data flow are allowed to act selfishly and independently (in the context of the entire network) to achieve local goals. The actions taken have straight effect to the observed.

3.2 Cyber threats in cognitive networks

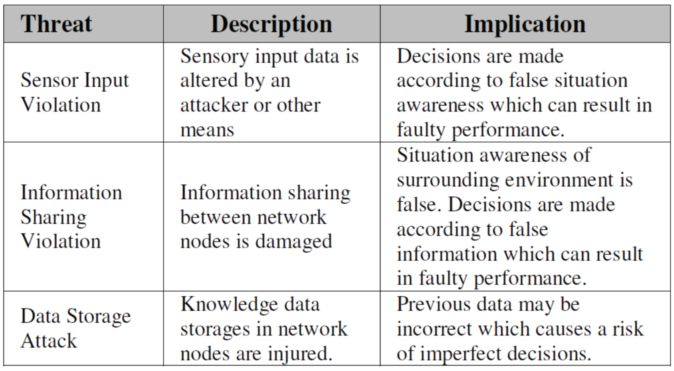
Cyber threats to communication networks are increasing all the time. In (Ventre 2011) aspects on cyber warfare are widely discussed. The book highlights less studied operational and planning features of a cyber attack. The attack is divided into three phases; intelligence, planning and conduct. When implementing CN the intelligence and conduct phases must be considered. CN security functions should include methods to prevent information and network intelligence, and to protect CN infrastructure against cyber attacks.

Security of CN is discussed in (Mody et all 2009), (Prasad 2008), (Burbank 2008), (Chaczko et al 2012) and (Clancy et al 2008). Cognitive networks face unique security problems not faced by conventional wireless or wired networks. In an ideal CN, security of the network is provided as a result of a cognitive process, which creates new threats. For instance, incomplete situation awareness may lead to the decision not to use any encryption although it is extremely required. Table 2 represents some major security threats related to cognitive networking.

In the CR network, locally-collected and exchanged information is used to construct a perceived environment that will influence both current and future behaviors, as well as the behavior of the nodes them. The training of an incorrectly perceived environment will cause the CR to adapt incorrectly, which affects short-term behavior. Unfortunately, the CR uses these experiences as a basis for new behaviors. Thus, if the malicious attack perpetrator is clever enough to disguise their actions from detection, they have the opportunity for long-term impact on behavior. Furthermore, the CR collaborates with its fellow radios to determine behavior. Consequently, this provides an opportunity to propagate a behavior through the network in much the same way that a malicious worm.

In wireless networks, the main concern is an attacker spoofing faulty sensor information, causing the radio to select an undesired configuration. The attacker can cause faulty statistics data to appear in the knowledge database of a network node by manipulating the receiving radio frequency (RF) signal. Since these radio signal statistics operate on raw RF energy, there is no cryptographic means of securing them.

Table 1: Major security threats in cognitive networks



Threat	Description	Implication
Sensor Input Violation	Sensory input data is altered by an attacker or other means	Decisions are made according to false situation awareness which can result in faulty performance.
Information Sharing Violation	Information sharing between network nodes is damaged	Situation awareness of surrounding environment is false. Decisions are made according to false information which can result in faulty performance.
Data Storage Attack	Knowledge data storages in network nodes are injured.	Previous data may be incorrect which causes a risk of imperfect decisions.

From military operations point of view, the automated decision-making process of the cognitive network causes a fundamental security threat. The process without a human operator may lead into the situation in which network behavior is out of control. The adversary can capture a node and modify it to function in a way the adversary desires.

4. Framework for cyber threat management

4.1 Overview

The overview of the proposed threat management framework for cognitive networks is illustrated in Figure 2. The framework consists of three layers which are a network, cluster and node layers. The layered structure is based on the fact that optimizing of the network is provided at three levels. Network system parameters including those related to security are optimized in a single node, in a cluster of nodes and finally concerning the entire network.

The threat management framework includes two main functionalities in each network node: the threat management process as a part of the cognitive process and the database element. The threat management process is based on the assessment process introduced in (Shore et al 2010), and it consists of threat identification, risk assessment and mitigation trade-off sub processes. The threat identification element receives information from several security sensors and databases, and then calculates and enumerates the threats and sets out intrusion/attack scenarios, and identifies the relevant vulnerabilities.

The risk assessment sub process quantifies the risk for each intrusion scenario through the use of event history databases, and policies and mitigation strategies. Quantifying the risk can be done using historical data or statistical sampling. Also, an expert opinion may be needed when relevant data is missing, but in the CN context this manual evaluation is not desired. The cyber event or incident may not always result in the same consequences. A number of consequences with differing probabilities (for instance, an attack on a network may result in a temporary outage of one workstation at one extreme, and a complete extended loss of the network at the other) may exists. In this cognitive network context, the expected damage from the event or incident is then the sum of the probabilities of each possible consequence.

At the final stage the mitigation trade-off sub process calculates the trade-off cost of mitigation against the risks. The process provides an adaptation map in which different responses to an incident are shown in a sense of costs. The costs of mitigation include such attributes as service availability, connectivity, security levels, etc.

The database element includes four main data storages. The policy database maintains the current security policies that need to be applied at each layer. The configuration database contains all configuration files to run the network node. Historical data is saved in the history database. Historical data is vital to create cyber threat scenarios during the learning process.

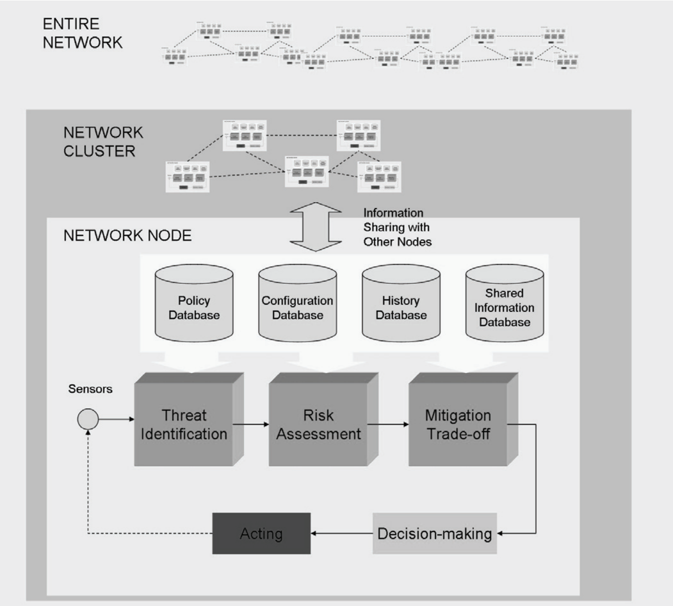


Figure 2: Overview of the framework

To keep the database element updated continuous information sharing between the nodes is required. Information sharing mechanisms should be built the way that the system contains no single point of failure. This is important in both commercial and military networks although the business drivers are different. There is also always trade-off between replicated control data and information payloads. In narrow band channels the CN should send control data as little as possible to guarantee Qos for services.

4.2 Threat identification

The basis of threat identification is a clear situational awareness of cognitive network's current state. The network should recognize all vulnerabilities in configurations and software. Also, threat libraries must be up-to-date, so that all known attach graphs are recognized. The operating environment also brings some differences concerning possible threats. In military environments there are always hostile parties trying to access and damage the networks and services. Each network node maintains a detailed list of all relevant threats including each possible intrusion/attack scenario and vulnerability which may be exploited during the current operation.

4.3 Risk evaluation

There are several methods to calculate the risk level, but the common understanding is that a risk consists of probability of a certain event and consequences caused by the event. In (Jiaxi 2006) an integrated risk assessment method is presented. The method is an integrative method to assess the cyber threat risk of any organization and thus it could be applied to cognitive networks. According to the method, the level of security risk is firstly classified into five categories.

Each category is assigned a value to indicate the relevant risk (see Table 2). PI is a performance index that is the reference value to identify the risk. The performance index may be calculated by different methods.

Table 2: The categories of cyber security risks

Risk Level	Very Low	Low	Normal	High	Very High
PI	<35	35-45	45-65	65-75	>75
LV	1	2	3	4	5

In next phase a cyber security risk matrix is created. The first row includes the percentage values of the cyber system risk belonging to each category. The second row consists of the probabilistic factors of incidents introduced by cyber events. The third row contains the influence factors of the incidents in cyberspace. An example of this cyber security risk matrix $\boldsymbol{M_s}$ is presented in (1).

$$M_s = \begin{bmatrix} 0 & 0.5 & 0.3 & 0.1 & 0.1 \\ 0.3 & 0.2 & 0.2 & 0.1 & 0.1 \\ 0.5 & 0.1 & 0.1 & 0.1 & 0 \end{bmatrix} \quad (1)$$

The integrated cyber vulnerability assessment can be calculated by applying the following formula:

$$I_{ir} = W_{cai} \times M_s \times LV^T \quad (2)$$

Where $\boldsymbol{I_{ir}}$ is the vulnerability index, ***LV*** is the security risk vector (value is taken from Table 2). $\boldsymbol{M_s}$ is the cyber security risk matrix and $\boldsymbol{W_{cai}}$ = [$\boldsymbol{w_r}$ $\boldsymbol{w_a}$ $\boldsymbol{w_l}$] is a vector, whose value indicates the weight of cyber security risk, damage risk and the damage influence.

The previous risk assessment method does not include any attributes related to cognitive processes. The adaptive behavior of the network should be perceived for example by adding an adaptation level factor to the equation (2). The structure of the factor is to be studied in further research.

4.4 Mitigation trade-off

Mitigation of an incident typically causes some trade-off between service availability and the risk level in the system. Networks and information systems are so complex combinations of hardware and software that it is even impossible to build a system without any vulnerability. Thus, it makes more sense to approach system security through risk and threat management. The mitigation of an incident may need service break outs or QoS level updates. The mitigation may also have high costs if new equipment or man power is needed. In some case, a threat is approved to appear if its probability is relatively small or consequences are estimated to be limited.

In this study the mitigation trade-off process is not discovered in more details. The trade-off direction could be included in network security policy. The security policy provides information about critical services, availability and confidentiality requirements and business driver of the communication network.

4.5 Security policy

Policy is a formal statement of operational requirements laid out in a formalized way. In (Shore et al 2010) traditional security policy is developed one step further by introducing the concept of policy owners and domains. The policy architecture consists of one organization wide policy supporting the organizational domain, which in addition to the set of organization wide common policy has subordinate policies for different sub unit domains, which themselves may further decompose policy into lower level policies and sub-domains. Also, external policy domains may exist, such as in the case of government regulations.

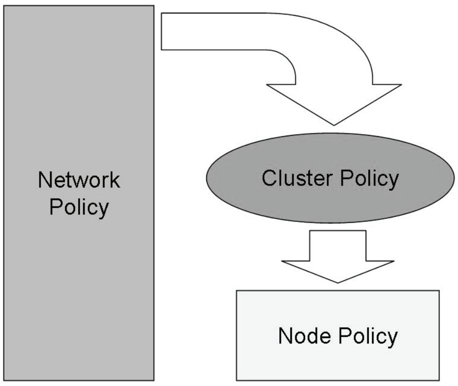


Figure 3: Policy hierarchy

In cognitive networking, reliability and survivability requirements create a demand for three policy domains: the Network, the Cluster, and the Node (see Figure 3). The Network policy sets critical network infrastructure obligations which then will drive the policy in the Cluster domain. The Cluster security policy sets requirements to each node in which the Node policy is created.

4.6 Information sharing

For cognitive networking information sharing between the network nodes is vital. Achieving better situational awareness requires continuously status information sharing. If information sharing is blocked the cognitive behavior of the network is lost and the nodes act as an independent cognitive element. In that case the end-to-end goals are not reached and optimizing the entire network is failed.

For threat management information sharing is even more important. Knowledge bases and threat libraries must be shared in real-time among all the nodes. This requires communication channels that are reliable and include capacity enough. Fixed communication systems enable more capacity than mobile (wireless) systems. Also, wireless channels provide a wider attack surface.

The information sharing channels set up new vulnerability points. An attacker may use these channels to jam the normal threat management process, thus the network could be even more vulnerable as threat scenarios are created incorrectly. The adversary could also tap to a communication channel and collect information about security threats and that way to find weakest spots in the network.

The information sharing channels require protection mechanisms as all the other communication channels. First, communication parties should be authenticated and authorized. Then, there should be a solution to protect data confidentiality. Typically, some cryptographic methods are used. Although these technologies already exist and could be implemented, these kinds of additional protocols and layers increase traffic between the nodes and spend the limited communication capacity.

Cryptography adds communication security (COMSEC) into data flows by providing information confidentiality, but especially in wireless systems the threats against transmission security (TRANSSEC) causes a high risk. Signals intelligence (SIGINT) may analyze and identify the intercepted frequencies, and electronic attack (EA) might be used against the communication channels. A typical example of EA is electronic jamming to prevent successful radio communications. To avoid TRANSEC threats totally, the communication link should be undetectable. Methods used to achieve TRANSEC may include frequency hopping and spread spectrum where the required pseudorandom sequence generation is controlled by a cryptography keys. Cognitive radios are promising platforms to provide low-probability of intercept and detection (LPI/LPD) waveforms as discussed and studied in (Petrin et al 2006).

5. Implementation challenges

Research on CN has already carried on years, but there are still few prototypes to demonstrate cognitive features. Thus, cyber threat management in the CN context is just taking initial steps. There are many implementation challenges to solve before the proposed threat management framework is in operational use. One of the major challenges with CN is decision-making process and learning functionality. It is possible to teach computers to act in a certain way in limited scenario, but how the system learns in a situation in which there is no data in prior.

One challenge is to build cross-layer functionality. Network systems consist of different layers (e.g. OSI), and the cognitive system should be able to optimize its parameters concerning all the layers. If a parameter is optimized according to information only from one layer, the system is never optimized for all the layers.

Implementation of threat management features also face huge challenges. For example, automated threat identification is not very simple process. The system should have a clear list of all threat types and possible vulnerabilities. At the same time, attackers are looking for new attack scenarios and graphs. It is challenging to implement cognitive threat management features that automatically recognize new threat types and vulnerabilities.

The network level threat management is also difficult to maintain solid if the CN is very mobile. Information sharing may be difficult when continuous connectivity is not guaranteed. In cyber space attacks may occur rapidly and threat scenarios should be updated in real-time.

6. Conclusion

Cognitive networking will provide smart functionality for future communication networks. Automated cognitive processes ensure better capacity allocation, less human configuring and management and more security features. But while the human control over the networks decreases, cyber threats are going to even more complicated. Security monitoring is maintained by cognitive devices, and the trust mechanisms in network elements are not based on trust between network operators.

As threat scenarios become more difficult to understand, requirement for a cyber threat management system is obvious. In this paper, a layered framework for cyber threat management was presented. The framework consists of three layers: node, cluster and network layers. At each layer there is a three-step threat management process. In the first phase possible threats are identified, and then risk assessment is generated. At the final phase, trade-off for incident mitigation is calculated. The process fits in the OODA loop of the cognitive process.

The threat management process for CN requires reliable information sharing between the network nodes. Unfortunately, this brings new threat scenarios as the information sharing channels could be used for hostile purposes. Another drawback of information sharing is that control messages generate loads of traffic which may decrease payload capacity.

As the proposed framework is still at very rough level, lots of future work is required. The future work includes several areas of the framework. The cognitive processes to create relevant threat situational awareness needs to be studied in more details. Also, risk assessment methods in a cognitive context require more research. In addition, collaborative threat identification among network nodes must be studied. This study is also missing some initial performance and functionality calculations to demonstrate and analyze the overall maturity of the proposed model.

References

Atmaja, T.D. and Fitriana, F. (2011) “*Cyber Security Strategy for Future Distributed Energy Delivery System*”, **International Conference on Electrical Engineering and Informatics** (ICEEI), pp 1 – 6.

Bodeau, D.J., Graubart, R. and Fabius-Greene, J. (2010) “*Improving Cyber Security and Mission Assurance via Cyber Preparedness (Cyber Prep) Levels*”, **IEEE Second International Conference on Social Computing** (SocialCom), pp 1147 – 1152.

Buford, J.F., Lewis, L. and Jakobson, G. (2008) “*Insider Threat Detection Using Situation-Aware MAS*”, **11th International Conference on Information Fusion**, pp 1 – 8.

Burbank, J. (2008) “*Security in cognitive radio networks: the required evolution in approaches to wireless network security*,” **3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications** (CrownCom), pp 1- 7.

Chaczko et al (2012) “*Security threats in cognitive radio applications*”, **14th International Conference on Intelligent Engineering Systems** (INES), pp 209 - 214.

Choi et al (2008) “A *Study on the Optimal Model for Information Security Management Level”*, **International Conference on Information Science and Security**, pp 238 – 244.

Clancy, T.C. and Goergen, N. (2008) “*Security in cognitive radio networks: threats and mitigation*”, **3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications** (CrownCom), pp 1 - 8.

Fragkiadakis, A., Tragos, E. and Askoxylakis, I. (2012) “*A Survey on Security Threats and Detection Techniques in Cognitive Radio Networks*”, **IEEE Communications Surveys & Tutorials**, Vol PP, Issue 99, pp 1 – 18.

Goldman, H., McQuaid, R. and Picciotto, J. (2011) “*Cyber Resilience for Mission Assurance*”, **IEEE International Conference on Technologies for Homeland Security** (HST), pp 236 – 241.

Jiaxi, Y., Anjia, M. and Zhizhong, G. (2006) “*Vulnerability Assessment of Cyber Security in Power Industry*”, **IEEE PES Power Systems Conference and Exposition**, pp 2200 – 2205.

King at al (2005) “*A Comprehensive Threat Management Framework for a Crop Biosecurity National Architecture*”, **Proceedings of IEEE International Geoscience and Remote Sensing Symposium**, Vol 3, pp 2105 – 2108.

Mahmoud, Q. (2007), “*Cognitive Networks: Towards Self-Aware Networks*”, **Wiley-Interscience**, 2007.

Mody et all (2009) “*Security in cognitive radio networks: An example using the commercial IEEE 802.22 standard*”, **IEEE Military Communications Conference** (MILCOM), pp 1 - 7.

Morris et all (2011) “*A perceptually-relevant model-based cyber threat prediction method for enterprise mission assurance*”, **IEEE First International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)**, pp 60 – 65.

Petrin et al (2006) ”*Cognitive Radio Testbed and LPI, LPD Waveforms*”, **IEEE Military Communications Conference 2006** (MILCOM 2006), pp 1 - 2.

Prasad, N. (2008) “*Secure cognitive networks*”, **European Conference on Wireless Technology**, pp 107 – 110.

Safdar, G.A. and O'Neill, M. (2009) "*Common Control Channel Security Framework for Cognitive Radio Networks*", **IEEE 69th Vehicular Technology Conference**, pp 1 – 5.

Shore, M. and Deng, X. (2010) "*Architecting Survivable Networks using SABSA*", **6th International Conference on Wireless Communications Networking and Mobile Computing** (WiCOM), pp 1 – 7.

Thomas, R., DaSilva, L. and MacKenzie, A. (2005) "*Cognitive networks*", **Proceedings of the First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks**, pp 352 – 360.

Thomas, R., Friend, D., DaSilva, L. and MacKenzie, A. (2006) "*Cognitive networks: adaptation and learning to achieve end-to-end performance objectives*", **IEEE Communications Magazine**, Vol 44, Issue 12, pp 51 - 57.

Ventre, D. (2011) "*Cyberwar and Information Warfare*", **John Wiley & Sons, Inc.**, 2011.

A Framework for the Detection and Prevention of SQL Injection Attacks

Emad Shafie and Antonio Cau
Software Technology Research Laboratory (STRL), Faculty of Technology, De Montfort University, Leicester, UK
eshafie@dmu.ac.uk
acau@dmu.ac.uk

Abstract: The use of Internet services and web applications has grown rapidly because of user demand. At the same time, the web application vulnerabilities have increased as a result of mistakes in the development where some developers gave the security aspect a lower priority than aspects like application usability. An SQL (structure query language) injection is a common vulnerability in web applications; it has been classified as the most dangerous type of vulnerability according to OWASP (Open Web Application Security Project) statistics (OWASP, 2010). An SQL injection vulnerability allows the hacker or illegal user to have access to the web application's database and therefore damage the data, or change the information held in the database. This paper will discuss a framework for the detection and prevention of common types of SQL injection attacks. The framework consists of three main components; the first component will check the user input for existing attacks, the second component will check for new types of attacks, and the last component will block unexpected responses from the database engine. Additionally, our framework will keep track of an ongoing attack by recording and investigating user behaviour. The framework is based on the Anatempura tool, a runtime verification tool for Interval Temporal Logic properties. Existing attacks and good/bad user behaviours can be specified using Interval Temporal Logic. Moreover, this paper will discuss a case study where various types of user behaviour are specified in Interval Temporal Logic and show how these can be detected.

Keywords: SQL injection, user input checker, runtime verification, database observer

1. Introduction

The permanent availability of web applications will increase the opportunity for everyone who is looking to exploit and damage these applications for illegal purposes. The common threat against the security of web application is the widespread occurrence of different types of web application vulnerability. SQL injection is a common vulnerability used to hack web application databases by executing a malicious SQL code injected by the attacker (Fu & Qian 2008,Clarke 2009).

An example of SQL injection attack is the following: a web application needs user input to perform its task. Accordingly, web applications usually provide a login page containing two text fields to allow the user to enter his/her user name and password. This user information will be sent to the web application database to check the user information. By submitting the user data, this data will be sent to the web application database using an SQL statement as follows:

Select * from UserTable where username= "entry_name" and userpassword =" entry_pass"
When this SQL statement is executed, the system will return the result of the query. If the user data is ok then the web application permits the user to access other pages at the website or the user data will be rejected and the login page reloads again. However, there is another scenario where the user enters the following at the user name field user name or '1'='1' - - then the SQL statement will be as follows:

SELECT * FROM UserTable WHERE username = "user name or '1'='1' - -"

At this stage the database engine considers any code after the keyword WHERE as a conditional statement, and because of "or 1=1 -- " the check condition is always equal to true. So, any code or condition after the double dash will be ignored. Consequently, the attacker will have unauthorized access to this web application.

The problem of this type of attack is that it cannot be handled or controlled by a firewall or other traditional communication security approaches as the attackers can gain access to the web application through the http protocol (Fu et al, 2007). Poor validation of the user input is the main reason of SQL injection attacks. Many approaches have tried to solve and block this type of vulnerability by using several different techniques ranging from static analysis approaches (Gould 2004, X 2007) to using dynamic approaches (Kosuga et al, 2007) or combining both of them in one technique (Halfond, 2005). This paper will discuss a novel technique to detect and prevent SQL injection attacks at the runtime using the Anatempura tool. Our framework does this by keeping track of an ongoing attack by recording and investigating user behaviour. The paper is divided in several sections as follows; Section 2 provides related work, Section 3 provides our approach in detail by describing all the framework components, Section 4 will discuss the Anatempura tool, Section 5 will illustrates our framework with the help of a case study, and in Section 6 we give a conclusion and discuss future work.

2. Related work

There are many approaches which exist that deal with SQL injection attacks, some of these approaches are as follows: Black Box testing approach, by gathering the information about all weak points in the website by using a web crawler to detect the vulnerable points that can be indictable (Huang et al, 2003). SQL Randomisation approach, by adding integer numbers randomly to each SQL keyword that is used in the query statement at that application, then during execution the application will rewrite the SQL statement using a SQL parser and random number to accept the SQL keyword according to that random instruction set. Thus, any SQL keyword without the number or out of the range will be rejected (Boyd & Keromytis, 2004).

Static Analysis approaches, by analysing web application to detect the vulnerable SQL query sentence in addition to validating the user input. Gould et al use JDBC checking tool to check statically for type correct queries in the SQL statement that are generated dynamically in Java. This technique is not effective enough to detect all types of SQL injection vulnerabilities because JDBC was not developed for this purpose or not for preventing attack, but it is usable. (Gould et al, 2004). X et al propose SAFELI tool which is a white box analysis tool that analyses an asp.net application. It depends on the analysis of the byte code of the application, and previous collected information about attack patterns. This collected information is used during the analysis as attack behaviour reference. In addition, SAFELI analyses the web application using the symbolic execution engine which clarifies all the application pages with its entry points (Fu et al, 2007).

Static and Dynamic Approaches, Halfond and Orso use static analysis procedure to build the SQL query model that determines the construct queries points which have direct access to the database. Successively, each of the construct queries points will be supported by runtime monitoring that investigates the queries before sending it to the database. This investigation checks those queries against any existing attack. However, the limitation is the monitoring step that depends on the result of the static analysis step, and that means if there is a fault in the first step then the fault will be in the other step (Halfond et al, 2005). Ruse and his colleagues suggest another solution against SQL injection attacks by developing an automatic model to capture any change of the sub-queries and its dependencies using CREST (test tool for C programs). The automatic model runs through three main steps starting from compiling SQL statement to be usable with C programs. The compiled sentences will be tested by using CREST which generates the test to check the injection possibility by detecting the injection causing requirements, and at the runtime the user input will be monitored to find any of the requirements that are detected in the previous step. The feature of this approach is there are no false positives like in other static analysis approaches, and it also looks at the semantic structure of the SQL query and not on the syntactic structure like other previous approaches. (Ruse et al, 2010). Lee et al, use the combination of static and dynamic technique by removing any of SQL attribute value of the SQL query at runtime and comparing it with a static SQL query (Lee, 2011).

Holzer and his colleagues have used a verification technique based on Computation Tree Predicate Logic (CTPL) which is extended from CTL Computational Tree Logic. This approach uses a model checker called Mocca that expects as input assembly source code. Mocca has been used to determine whether a security property expressed in CTPL holds or not (Holzer et al, 2007).

All the mentioned techniques are looking for the SQL injection attacks as one step or as a static attack but SQL attacks are dynamic or proceed in several steps which is the main assumption of our approach.

3. Detection and prevention framework (DPF)

3.1 Overview of DPF

DPF is used to monitor and block SQL injection attacks and it uses the Anatempura tool as a monitoring tool to block malicious users. DPF is initialized by specifying some of the existing attack patterns using ITL, and connecting the Anatempura tool to the web server to monitor the user input data. Broadly, DPF is divided into three main phases: the initial phase, the checking phase and the decision phase as show in Figure 1. The initial phase consists of several steps starting from the user input until the data arrives to the input checker which is the first step of the checking phase. In the checking phase, the system will analyse the data using the input checker which uses Anatempura to analyse the user entry against existing attack patterns, and thus to decide whether it is a bad or good input. The checking phase consists of three processes which are input checker, output checker and database observer and the result of the checking phase will be sent to the decision phase. The decision phase can be divided into two main parts which are the feedback and user behaviour components. The three phases will be explained in more detail below.

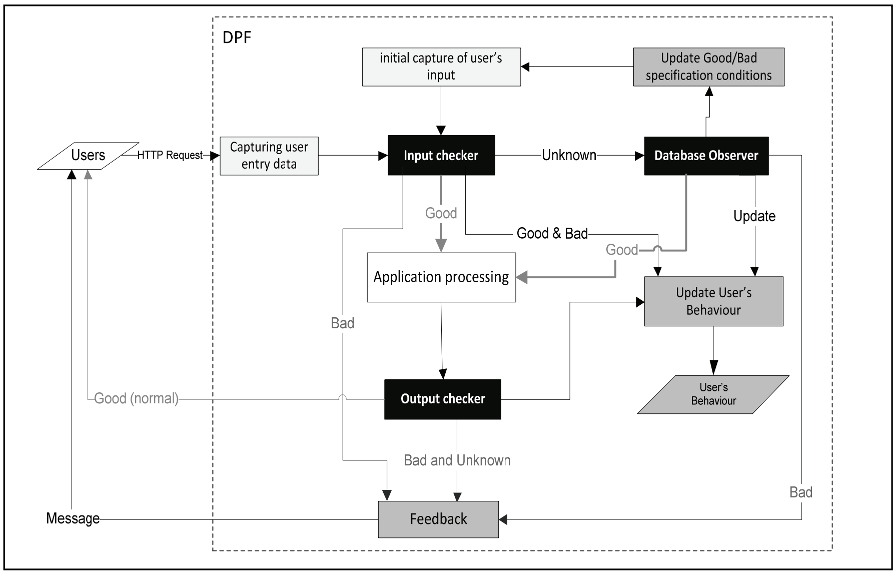


Figure 1: Detection and prevention framework

3.2 Description of each component

This section describes the components of our detection and prevention framework (DPF).

Initial capture of user input, is considered the basis of DPF to bootstrap the system by specifying the good and bad input at the character level using ITL. The good input should not include any symbol like single quotation and double quotation or star, or in other words the good input should not contain any of SQL keywords that can be used to attack the web application database. Furthermore, the bad input will be specified using ITL by specifying some of the existing attacks patterns (SQLlib, 2007). Note, the initial bad / good specifications that are specified using ITL will be used by the Anatempura tool.
Users, anyone who submits an http request to the web application.

Capturing Data will analyse the http request to extract the user entry data and convert it to Anatempura format (variable, value, timestamp). This component does not depend on Anatempura but on the web page application type. In other words, the data will be extracted using library calls offered by the programming language that is used to develop the web application.

Input checker will use the existing attack patterns that are prepared by the initial capture of the user input component. The checker will analyse the user entry against the existing attack patterns, and the outcome is as follows:

The entry data is good, so the data will be passed to the application server for normal processing, and the user behaviour will be updated accordingly. If the entry data type is a bad, then the data will be rejected and therefore DPF updates the user's behaviour and prepares to send a message to the user via the feedback component. The last possibility is that the entry data is unknown, then the database observer will detect the bad/good entry according to the observations of the observer. The following functions have been created to check the input:

Table 1: Checking functions

Input Seq.	Function Name	Function Aims
1	DecreaseSpaces()	To remove any extra spaces in the user input
2	LowerCase()	To transform all of the user input to lower case
3	Goodentry() and SearchTokens()	Check of the user input whether it includes SQL keywords
4	Badentry()	To check the input whether it includes any of the existing attack patterns
5	BehaviourDetection()	To model user behaviour

All of the mentioned functions in Table 1 are combined in one procedure CheckingModel.

Database Observer will check unknown entry cases which are not caught by the initial analysis of user entry data via the input checker. The main idea of the DB observer is to determine what will happen in the DB engine, and this will be done by monitoring each transaction that comes from the input checker as an unknown entry. Moreover, the checking of the transaction depends on the developer because the DB observer needs the developer to specify the expected result of each transaction that is run by the developed web application such as, the table name, running command type, number of the expected records, and the user type. The expected result will be compared with the runtime result. Furthermore, the comparison between the expected result and the runtime result is used to ensure that the database transaction was performed safely (if the runtime result is similar to what the developer expected) as shown in the Figure 2.

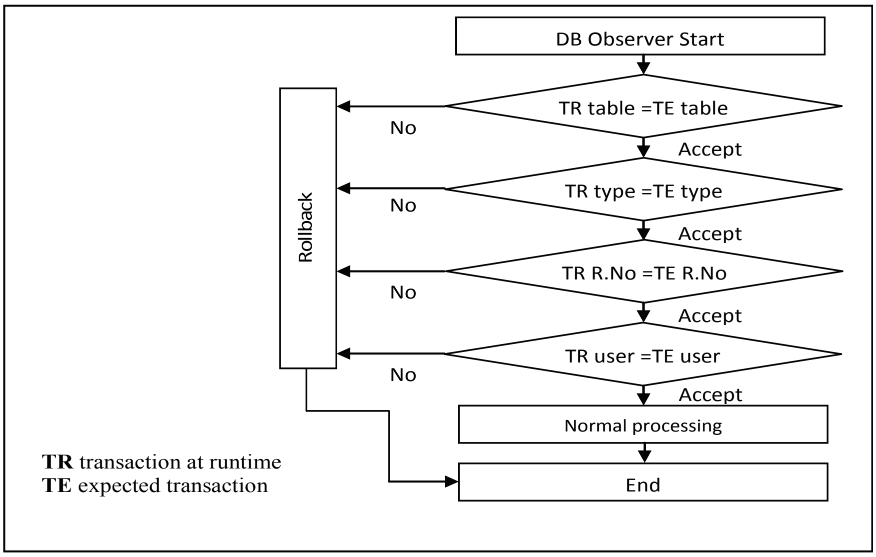


Figure 2: Database observer

Therefore, the database observer has to monitor four conditions that are specified by the developer or the programmer as follows:

- Transaction type of the real execution is the same as the expected one. For example, if the transaction type at the real execution is "Select" and the expected one is "Select" as well, then the database observer will accept this transaction at this step and continue. If they are different the database observer will do a rollback and respond to the input checker to reject the entry data and update the user behaviour.
- The table name of the real execution is the same as the expected one. For example, if the transaction table at the real execution is "users" and the expected table is also "users" the database observer will accept it and continue to the next step. If they are different then the database observer will do a rollback and this transaction will be rejected.

- The record number of the real execution is the same as the expected number. For example, the login page normally returns one record with the select statement, so if the real execution of the select statement returns the same number (one record), then the database observer will accept it, otherwise it will be rejected and the database observer will do a rollback.
- User type of the real execution is the same as the expected one. For example, if the user tries to change the password at the change password page the user type should be the same as the expected one. However, if the user tries to change another user's password then the database observer will catch this and this transaction will be rejected and the database engine will do a rollback.

Note the database observer can only deal with recoverable transactions so not DDL (Data Definition Languages) commands like create, drop, and alter table, because the DDL injected command cannot be recovered by rollback command, so these transactions should be rejected before accessing the database by the input checker.

Output checker will check whether the message sent to the user is safe or not. Moreover, the output checker will not analyse the response in the same way as the input checker, because it will block any message that contains details about database structure or type because these types of messages are not safe. Moreover, the output checker will block the unsafe messages using the library calls in the programming language which is employed to develop the web application. Feedback component prepares the message that will be sent to the user regarding the cases of bad entry data. Moreover, if the user enters the data using a bad method and the input is caught as unsafe then DPF will respond to this entry by using prepared messages relevant to the entry method.

User's behaviour will track each transaction in the system and detects the type of the transaction, i.e. whether it is a good or bad transaction. The tracking information will be used to model the user behaviour. Therefore, the user behaviour depends on the result of the input and output checker in addition to the result of the database observer. The DPF will use user information like IP address, user status (good, bad), attacking technique (primitive, advance), and time stamp of the transaction to build the user behaviour as shown in Figure 3.

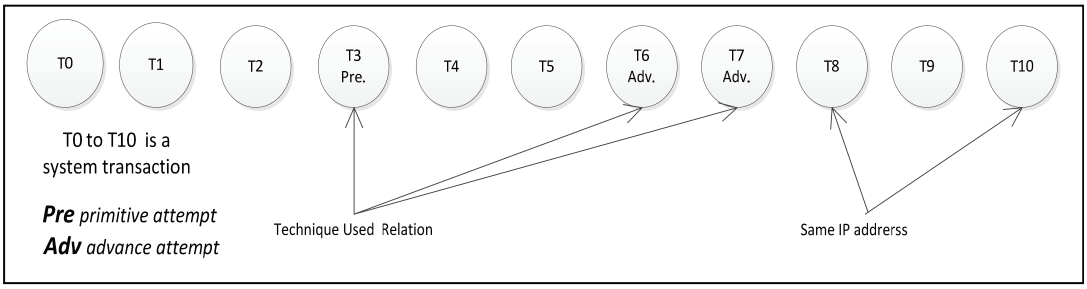


Figure 3: Transactions relation

Moreover, the DPF can determine user behaviour of three types i.e. normal, good and bad according to three criteria namely, the percentage of transaction, the sequence of the same transaction type, the transaction types as shown in Figure 4.

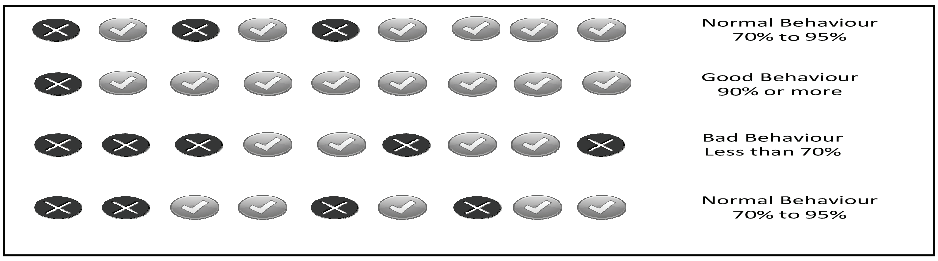

Normal Behaviour 70% to 95%
Good Behaviour 90% or more
Bad Behaviour Less than 70%
Normal Behaviour 70% to 95%

Figure 4: User Behaviour

Figure 4 shows user behaviour of various types: good behaviour if there are 95% or more of good transactions, the bad behaviour if less than 70% of good transactions and any other behaviour will be considered normal behaviour (between 70% to 95%). Additionally, user behaviour will also be classified as bad behaviour if there is a sequence of three or more bad transactions. Therefore, user behaviour can be used as a quick view of the transactions status and the proportion of hacking attempts that have been done so far. It can act as an early warning to the system.

DPF Updates

In DPF there are two types of updates which depend on the input data as follows:

Updating of User's Behaviour will update continuously the user behaviour within DPF according to the checking result of the checking processes of the input checker, database observer and output checker.

Updating of Existing attack patterns, this component is working in parallel with the database observer component. When the database observer finds any unsafe inputs, it will send this information to this component to update the existing attack patterns. Note, the updating of the attack patterns library will be done manually because the library that is used by the input checker is specified in ITL, thus the updating should use some translation into ITL to be usable with Anatempura.

4. Anatempura

Anatempura is a tool for the runtime verification of interval temporal logic formula. Interval Temporal Logic (ITL) is a flexible notation for both propositional and first-order reasoning about periods of time found in descriptions of hardware and software systems. Unlike most temporal logics, ITL can handle both sequential and parallel composition and offers powerful and extensible specification and proof techniques for reasoning about properties involving safety, liveness and projected time. Timing constraints are expressible and furthermore most imperative programming constructs can be viewed as formulas in a slightly modified version of ITL. Tempura provides an executable framework for developing and experimenting with suitable ITL specifications. In addition, ITL and its mature executable subset Tempura have been extensively used to specify the properties of real-time systems where the primitive circuits can directly be represented by a set of simple temporal formulae. In addition, various researchers have applied Tempura to hardware simulation and other areas where timing is important.

Anatempura, which is built upon C-Tempura, is a tool for the runtime verification of systems using Interval Temporal Logic (ITL) and its executable subset Tempura. The runtime verification technique uses assertion points to check whether a system satisfies timing, safety or security properties expressed in ITL. The assertion points are inserted in the source code of the system and will generate a sequence of information (system states), like values of variables and timestamps of value change, while the system is running. Since an ITL property corresponds to a set of sequences of states (intervals), runtime verification is just checking whether the sequence generated by the system is a member of the set of sequences corresponding to the property we want to check. The Tempura interpreter is used to do this membership test (Cau & Moszkowski 2011, Zhou et al 2005).

Runtime verification as used in AnaTempura does not suffer from the state explosion problem such as experienced by model checkers (Cimatti et al 2002, Holzmann 1997). A new state is computed on the fly using rewrite rules, so there is no need to compute the automaton with its states as is done by model checkers.

5. Case study

We discuss a case study to clarify how attacker's information can be used to model user behaviour. This case study describes an input scenario and assumes that the status of each input is already determined by Anatempura via the CheckingModel procedure as mentioned in Section 3. Table 2 shows a particular input scenario.

The input status column lists the status of each input. The following ITL formula determines whether two bad inputs are related by IP address:

$\exists$ IP. $\Diamond$ (Status(Input) = Bad $\wedge$ IP(Input) =ip) ; $\Diamond$ (Status(Input) = Bad $\wedge$ IP(Input) =ip)

This means if an IP in a certain state is equivalent to an IP in a previous then these inputs are related.

The following ITL formula determines whether two bad inputs are related by stored procedure

$\exists$ command. ◊ (Declare (Input) = command); ◊ (EXEC (Input) = command)

Table 2: Selective user’s inputs

Input Seq.	User IP	The input	status
	146.168.255.12	Normal	g
	146.168.255.13	Normal	g
	82.164.254.12	‘ or ‘1’=’1	b
	82.164.254.12	‘;drop table users;--	b
	212.164.254.14	Normal	g
	212.164.254.16	Normal	g
	212.164.254.14	any’; declare @NewStoreProcedure char(80) …….;	g
	67.164.254.14	normal	g
	146.164.2.46	‘ ;	b
	212.164.254.14	normal	g
	182.164.254.23	any’; EXEC (@NewStoreProcedure);	b
	212.164.254.14	Normal	g
	212.164.254.16	Normal	g

This means if an executing of command in a certain state exists and the declaration of the command is the same command in a previous state then these inputs are related.

In Table 2, input 3 and 4 are marked as bad, those attempts are one step attacks because they do not retrieve any information from the database and just try to inject the harmful code in the web application fields. However, those attempts have the same IP address which means there is a relation between them because both attempts have been done by the same user. The input 7 and 11 can be classified as related as well, because the attacker here tried to declare the stored procedure in the first attempt and in the second attempt he/she uses it. So there is a relation between these hacking attempts and this justifies the use of monitoring user behaviour.

6. Conclusion and future work

This paper has presented a new approach to detect and prevent SQL injection attacks. The paper highlighted some of the previous existing techniques used against these types of attacks. Our technique is classified as a runtime detection and prevention technique. It uses development of the Anatempura tool to specify existing attacks patterns and to model user behaviour. We illustrated our approach through a case study that describes a scenario where several inputs are related in any ongoing attack. Those relations have been defined in ITL to successfully model user behaviour. The paper shows our framework in detail by describing the detection processes and how those processes work with other components in the framework. Finally, our approach is still under development and we will enhance its functionality by running more experiments that model user behaviour and extra fine tuning to address the reporting of false negatives and false positives. Future work will consist of completing the implementation to evaluate it with real world SQL injection attacks scenarios.

References

Cau, A. & Moszkowski, B. (2011), ‘The Interval Temporal Logic Home Page’. http://www.cse.dmu.ac.uk/STRL/ITL/ Accessed [20-11-2011].

Clarke, J (2009), ‘*SQL Injection Attack and Defense’*. United States of America: Laura Colantoni. p63-76.

Cimatti, A., Clarke, E. M., Giunchiglia, E., Giunchiglia, F., Pistore, M., Roveri, M., Sebastiani, R. & Tacchella, A. (2002), ’ NuSMV 2: An OpenSource Tool for Symbolic Model Checking.’ *in* Ed Brinksma & Kim Guldstrand Larsen, ed, 'CAV' , Springer, pp. 359-364.

Fu, X. et al (2007), ‘A Static Analysis Framework for Detecting SQL Injection Vulnerabilities’, in Proceedings of 31st Annual International Computer Software and Applications Conference (COMPSAC 2007), pages 87 – 96.

Fu, X. & Qian, K. (2008), ‘SAFELI: SQL injection scanner using symbolic execution’. in Tevfik Bultan & Tao Xie, ed., 'TAV-WEB' , ACM, pp. 34-39.

Gould, C., Zhendong, Su. & Devanbu, P. (2004), ‘JDBC Checker: A Static Analysis Tool for SQL/JDBC Applications’, in 'ICSE', IEEE Computer Society, pp. 697-698.

Halfond, W. G. J. & Orso, A. (2005), ‘AMNESIA: analysis and monitoring for NEutralizing SQL-injection attacks’, in David F. Redmiles; Thomas Ellman & Andrea Zisman, ed., 'ASE', ACM, pp. 174-183.

Holzer, A., Kinder, J. and Helmut, V. (2007), ‘Using Verification Technology to Specify and Detect Malware’. Proc. The 11th International Conference on Computer Aided Systems Theory (EUROCAST), vol. 4739, pp. 497-507.

Holzmann, G. (1997), 'The model checker SPIN ', IEEE Transactions on Software Engineering, 23(5):279-294.

Kosuga, Y., Kono, K., Hanaoka, M., Hishiyama, M. & Takahama, Y. (2007), ‘Sania: Syntactic and Semantic Analysis for Automated Testing against SQL Injection’, in 'ACSAC' , IEEE Computer Society, pp. 107-117.

Lee, I., Jeong, S., Yeo, S. & Moon, J. (2012), 'A novel method for SQL injection attack detection based on removing SQL query attribute values', Mathematical and Computer Modelling 55 (1-2), 58-68.

OWASP Top 10 for 2010. https://www.owasp.org/index.php/Category: OWASP_Top_Ten_Project Accessed [09-09-2011].

Ruse, M., Sarkar, T. & Basu, S. (2010), ‘Analysis & Detection of SQL Injection Vulnerabilities via Automatic Test Case Generation of Programs’, in ‘SAINT’, IEEE Computer Society, pp. 31-37.

SQLlib-tool (2007), Open labs web application security. http://www.open-labs.org/sqlibf113b2.tar.gz Accessed [12-10-2011].

Zhou, S., Zedan, H. and Cau, A. (2005) Run-time analysis of time-critical systems. Journal of Systems Architecture, 51(5), pp. 331-345.

Work in Progress Papers

Information Systems Security Management (ISSM) Success Factor: Retrospection From the Scholars

Azah Anir Norman and Norizan Mohd Yasin
University of Malaya, Kuala Lumpur, Malaysia
azahnorman@gmail.com
norizan@um.edu.my

Abstract: Information System Security Management (ISSM) studies today have presented remarkable solutions in addressing security management (SM) problems. Many companies have designed SM procedures to protect their businesses from threats. Often, ISSM implemented by these businesses are based largely on common practices, current understanding and business requirements which seldom reach optimum levels. This presents risks as such practices often lead to resource wastage and security abuse. This paper attempts to review previous studies on ISSM implementation. This retrospection study aims to determine the most influential factors for successful ISSM implementation in a business. The study reviewed selected journal articles and conference papers in the field of information systems security. The three main classes of success factors in ISSM comprise technology characteristics, organizational structure and environmental influences. The success factors were collated from the ISSM success theoretical model which is based on selected IS theories. Fundamentally, technology, process and human elements that form the management mechanism were found to be vital for successful ISSM implementation. Retrospection of various scholars' practical-theoretical-experimental researches and views enables better understanding and the subsequent assimilation of success factors that influence successful ISSM implementation in a business context.

Keywords: information system security management, success factor, security management, information systems security

1. Introduction

The use of Information Systems (IS) is no longer bound to the company but has spread externally to customers and suppliers. Businesses today have developed and adopted many types of information systems to manage and fulfill customers' requests. Through information systems, business processes were improved thus increasing business productivity. Similarly, the explosive advent of Internet technology has seen a dramatic increase in the use of information systems.

The potential benefit of global marketing has become the most profitable advantage the Internet offers (Kaynak et al., 2005). However, despite being highly advantageous, free Internet access has invited various unethical activities over the past 20 years. Systems were breached and different types of losses were incurred during this time. As of December 2011, CyberSecurity Malaysia reported approximately 15218 cases involving Internet related threats as compared to 8090 reported cases in 2010 (MyCERT, 2012). A recent survey by (Kaspersky Lab, 2011) indicated that 46% of respondents identified cyber threats as potential business risks of the future. Almost half of the businesses in this survey agreed that cyber threats are the top three emerging risks.

The latest security survey by PWC, (2012) revealed that 33% respondents had not reported any security incidents, but there was an increase in reported incidents from 3% in 2010 to 8% in 2011. Alarmingly, the report also shows a decrease from 76% in 2010 to 72% in 2011, in the IT personnel's confidence regarding the readiness of their business to confront critical information threats. The decrease in confidence befitted the report, as only 40% of respondents are shown to be practicing effective SM while the remaining businesses are still trying to determine the best SM for their business.

The report also highlighted issues which are holding back SM in a business, emphasising the importance of SM success factors in guaranteeing effective SM. Adoption of security tools and techniques is worthless if businesses fail to understand the essence of ISSM success. This paper seeks to distinguish the success factors of ISSM through the scholars' lenses. Retrospection of previous studies leads to the identification of the possible success factors for successful ISSM implementation in a business.

2. Theoretical perspectives on information systems security management (ISSM) success

There are many SM frameworks including (Eloff and von Solms, 2000, Finne, 1998, Lech, 2000, Trcek, 2003, von Solms, 2005, Zuccato, 2007). To date, there have also been various reviews that identified the critical issues in SM (Wood, 1987), (Straub, 1990), (Loch et al., 1992), (Fitzgerald, 1995), (James, 1996), (Parker, 1997), (Straub and Welke, 1998), (Tryfonas et al., 2001), (Schlarman, 2002), (Baskerville and Siponen, 2002), (Kankanhalli et al., 2003), (Al-Salihy et al., 2003), (von Solms and von Solms, 2004), (Farn et al., 2004), (Keller et al., 2005), (Eloff and Eloff, 2005), (Chang and Ho, 2006), (Dzazali, 2006), (von Solms, 2006), (Torres et al., 2006), (Siponen and Oinas-Kukkonen, 2007), (Albrechtsen, 2007), (Hu et al., 2007), (Yeh and Chang, 2007), (Veiga and Eloff, 2007), (Anderson and Choobineh, 2008), (Siponen and Willison, 2008), (Dlamini et al., 2008), (D'Arcy and Hovav, 2008), (Fomin et al., 2008), (Barlette and Formin, 2008), (Werlinger et al., 2009), (Kraemer et al., 2009), (Ozkan and Karabacak, 2010), (Veiga and Eloff, 2010), (Tsohou et al., 2010), (Gillies, 2011), (Monfelt et al., 2011), (Yildirim et al., 2011), yet, many of these works have not discussed these issues in relation to SM success in a business context.

Thus this study aims to gather the perspectives of different scholars and simulate the collection of issues as the success factors of SM in business enterprises. While there is much research on the security standards and best practices for business, these standards do not include specific step-by-step information for implementing SM ((Siponen, 2006), ((Zuccato, 2007).

The concept of ISSM success in business is strongly related to the business structure. Scholars like (Ein-Dor and Segev, 1978) have stressed the importance of organizational factors in MIS success of an organization. Organization factors consist of organizational size, business type and management support. In later ISSM studies, organizational factors are again cited (Kankanhalli et al., 2003, Barlette and Formin, 2008, Werlinger et al., 2009) as the three factors deemed to be important in ISSM success and effectiveness. Since ISSM is an innovative process, innovation decision making is determined by technology, organization and environment (Tornatzky and Fleisher, 1990).

Once decision has been made, adoption of innovation is influenced by antecedents of Diffusion of Innovation (DOI). Successful adoption is influenced by the compatibility and complexity of technology and their relative advantage (Rogers, 1983), (Cooper and Zmud, 1990), (Agarwal and Prasad, 1998). Successful adoption also implies successful implementation of ISSM. In their evaluation of successful IS implementation, (DeLone and McLean, 1992, DeLone and McLean, 2003), highlighted the six IS success categories as system quality, information quality, users, user satisfaction, individual impact and organizational impact. In reference to the ISSM study context, only the applicable attributes under each theory are used for this study. The ISSM success theoretical model is conceptualized in figure 1 below.

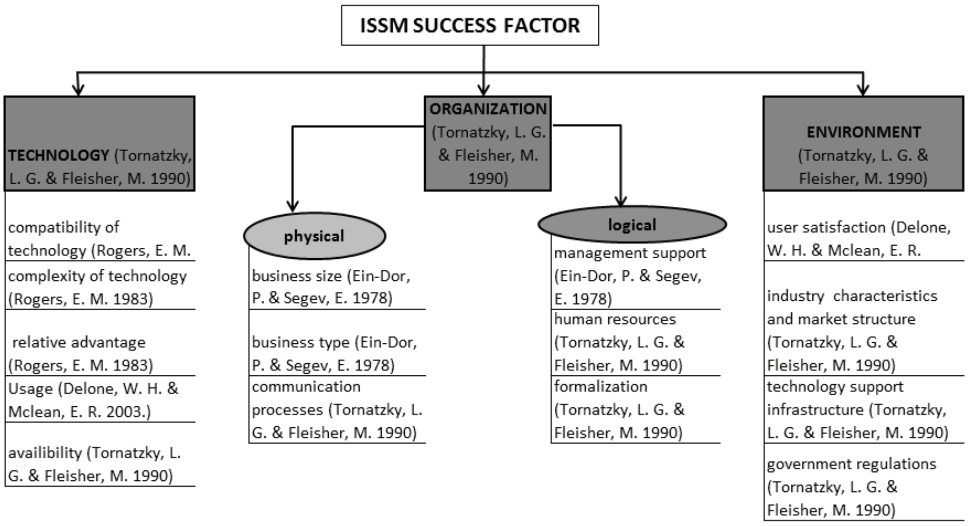

ISSM SUCCESS FACTOR
TECHNOLOGY (Tornatzky, L. G. & Fleisher, M. 1990)
ORGANIZATION (Tornatzky, L. G. & Fleisher, M. 1990)
ENVIRONMENT (Tornatzky, L. G. & Fleisher, M. 1990)
compatibility of technology (Rogers, E. M.
complexity of technology (Rogers, E. M. 1983)
relative advantage (Rogers, E. M. 1983)
Usage (Delone, W. H. & Mclean, E. R. 2003.)
availibility (Tornatzky, L. G. & Fleisher, M. 1990)
physical
business size (Ein-Dor, P. & Segev, E. 1978)
business type (Ein-Dor, P. & Segev, E. 1978)
communication processes (Tornatzky, L. G. & Fleisher, M. 1990)
logical
management support (Ein-Dor, P. & Segev, E. 1978)
human resources (Tornatzky, L. G. & Fleisher, M. 1990)
formalization (Tornatzky, L. G. & Fleisher, M. 1990)
user satisfaction (Delone, W. H. & Mclean, E. R.
industry characteristics and market structure (Tornatzky, L. G. & Fleisher, M. 1990)
technology support infrastructure (Tornatzky, L. G. & Fleisher, M. 1990)
government regulations (Tornatzky, L. G. & Fleisher, M. 1990)

Figure 1: ISSM success theoretical model

3. Findings

Analysis of previous scholarly articles concludes that ISSM success is attributed to three main segments, as follows:

- Technology characteristics: security infrastructure and tool/support mechanism
- Organization structure (physical): business size and business type
- Organization structure (logical): top management support, formalization and resources of the firm
- Environment influences: governance, enforcement and market structure

The three main segments, as depicted in figure 2, highlight the ISSM success factors in a business. The Venn diagram shows the overlapping middle area as the optimum ISSM state which encompasses technology, organization and environment to drive ISSM success in a business.

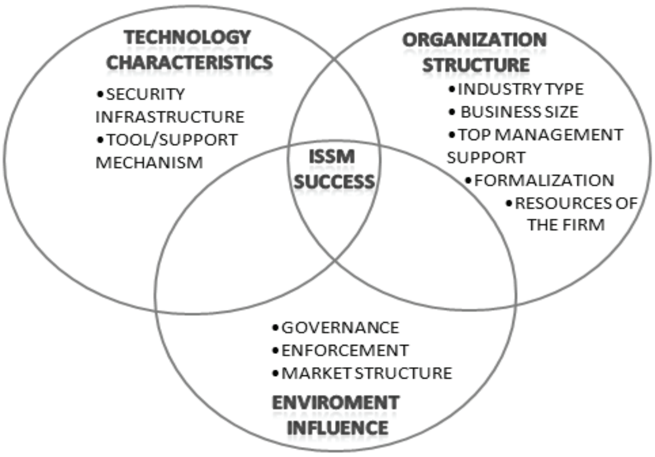


Figure 2: ISSM success factors

4. Summary

The efficacy of IS security in business may be determined by the identification of success factors. These factors ensure the balance between business processes and human control within an organization. Various researchers agree that the important elements determining success fall under three categories; technology characteristics, organization structure and environment influence. However, as studies on ISSM success is very much based on large and hierarchical organizations, there are a limited number of articles referring to small businesses from which examples may be sourced.

This research also sets forth the aim to further study the relationship between ISSM success and ISSM maturity in chosen business contexts. The research work adopted a sequential mixed-methodology to study the relationship between SM success factors with ISSM maturity. The quantitative investigation will determine issues in the focus research context which is the micro-SME, followed by in-depth interviews with selected business CEOs identified from previous investigations. This complementary method is adopted to ensure major issues are collated and discussed according to the analyzed theoretical perspectives.

Acknowledgements

The researchers would like to acknowledge University of Malaya, Malaysia for the Postgraduate Program Grant 2009/2010 in supporting this research.

Appendix 1

Table 1: ISSM success factors: Retrospection from scholars

	ISSM Success Factor	Organization - physical		Organization - logical				Technology		Enviroment		
				support	formalization	resources of the firm				governance		
No	success factor authors and year	industry type	organizational size	top management support	IS security objective /strategy/ policy and	information security awareness	staff competency and motivation	security infrastructure	tools/support mechanism	standards/ best practices	government enforcement	users involvement/ industry influences
1	YILDIRIM, E. Y., AKALP, G., AYTAC, S. & BAYRAM, N. (2011)			√	√	√	√			√		√
2	MONFELT, Y., PILEMALM, S., HALLBERG, J. & YNGSTRÖM, L. (2011)			√	√	√		√		√		
3	GILLIES, A. (2011)			√	√		√		√	√	√	√
4	TSOHOU, A., KOKOLAKIS, S., LAMBRINOUDAKIS, C. & GRITZALIS, S. (2010)			√	√	√	√	√	√	√		
5	VEIGA, A. D. & ELOFF, J. H. P. (2010)			√	√	√	√	√	√	√		
6	OZKAN, S. & KARABACAK, B. (2010)			√	√	√	√		√	√		
7	KRAEMER, S., P. CARAYON, ET AL. (2009).			√	√		√	√	√			
8	WERLINGER, R., K. HAWKEY, ET AL. (2009)		√	√	√	√	√	√	√			
9	BARLETTE, Y. AND V. V. FORMIN (2008)	√			√	√	√			√		√
10	FOMIN, V. V., H. J. D. VRIES, ET AL. (2008)	√		√	√	√	√	√	√	√	√	√
11	D'ARCY, J., & HOVAV, A. (2008).					√		√	√	√		
12	DLAMINI, M. T., ELOFF, J. H. P., & ELOFF, M. M. (2008).			√	√	√	√	√		√		
13	SIPONEN, M. & WILLISON, R. (2008)											
14	ANDERSON, E. E., & CHOOBINEH, J. (2008)			√	√			√	√	√		
15	VEIGA, A. D. AND J. H. P. ELOFF (2007)			√	√	√	√	√	√	√		
16	YEH, Q.-J. AND A. J.-T. CHANG (2007).	√	√				√	√	√			
17	HU, Q., HART, P. & COOKE, D. (2007)			√		√				√	√	
18	ALBRECHTSEN, E. (2007)			√		√	√	√				√
19	ZUCCATO, A. (2007)			√	√	√	√	√	√	√		√
20	MIKKO, T. S. AND O.-K. HARRI (2007).				√			√	√	√		
21	TORRES, J. M., SARRIEGI, J. M., SANTOS, J., & SERRANO, N. (2006).			√	√	√	√	√		√	√	√
22	VON SOLMS, B. (2006)			√	√	√		√		√		
23	DZAZALI, S. (2006)					√	√	√		√		
24	CHANG, S. E. AND C. B. HO (2006).	√	√	√	√		√		√			
25	ELOFF, J. H. P. & ELOFF, M. M. (2005)			√	√	√	√	√		√		
26	KELLER, S., A. POWELL, ET AL. (2005).				√	√	√	√	√	√		
27	FARN, K.-J., LIN, S.-K. & FUNG, A. R.-W. (2004)				√		√	√	√	√	√	
28	VON SOLMS, B. & VON SOLMS, R. (2004)				√	√		√	√			
29	AL-SALIHY, W., ANN, J. & SURES, R. (2003)			√	√	√	√	√		√		
30	KANKANHALLI, A., TEO, H.-H., TAN, B. C. Y., & WEI, K.-K. (2003).	√	√	√	√	√	√	√	√	√		
31	BASKERVILLE, R., & SIPONEN, M. (2002).				√				√	√		
32	SCHLARMAN, STEVEN. (2002)				√	√		√		√		
33	TRYFONAS, T., E. KIOUNTOUZIS AND A. POULYMENAKOU (2001)				√	√	√	√	√	√		
34	STRAUB, D. W. & WELKE, R. J. (1998)			√	√	√		√	√	√	√	
35	PARKER, D. B. (1997)	√		√		√		√		√		
36	JAMES, H. L. (1996)				√	√		√				√
37	FITZGERALD, K. J. (1995)			√	√	√	√	√		√		
38	LOCH, K. D., CARR, H. H. & WARKENTIN, M. E. (1992) .			√		√		√			√	
39	DETMAR W. STRAUB, J. (1990)			√	√	√	√	√				
40	WOOD, C. C. (1987)			√	√	√		√		√		

References

Agarwal, R. & Prasad, J. (1998) A Conceptual And Operational Definition Of Personal Innovativeness In The Domain Of Information Technology. **Information Systems Research**, 9, 204-215.

Al-Salihy, W., Ann, J. & Sures, R. (2003) Effectiveness Of Information Systems Security In IT Organization In Malaysia. The 9th Asia-Pacific Conference On Communications, APCC2003, 21-24 Sept 2003 Penang, Malaysia. IEEE, 716-720.

Albrechtsen, E. (2007) A Qualitative Study Of Users' View On Information Security. **Computers & Security**, 26, 276-289.

Anderson, E. E. & Choobineh, J. (2008) Enterprise Information Security Strategies. **Computers & Security**, 27, 22-29.

Barlette, Y. & Formin, V. V (2008) Exploring Suitability Of IS Security Management Standards For Smes. 41st Hawaii International Conference On Systems Sciences, 2008 Hawaii. IEEE.

Baskerville, R. & Siponen, M. (2002) An Information Security Meta-Policy For Emergent Organizations. **Journal of Logistics Information Management**, 15, 337 - 346.

Chang, S. E. & Ho, C. B. (2006) Organizational Factors To The Effectiveness Of Implementing Information Security Management. **Industrial Management & Data Systems**, 106, 345-361.

Cooper, R. B. & Zmud, R. W. (1990) Information Technology Implementation Research: A Technological Diffusion Approach **Management Science**, 36, 123-139.

D'arcy, J. & Hovav, A. (2008) Does One Size Fit All? Examining The Differential Effects Of IS Security Countermeasures. **Journal of Business Ethics**, Online First.

Delone, W. H. & Mclean, E. R. (1992) Information Systems Success: The Quest For Dependent Variable**. Information Systems Research**, 3, 60-95.

Delone, W. H. & Mclean, E. R. (2003) The Delone And Mclean Model Of Information Systems Success: A Ten-Year Update. **Journal of Management Information Systems**, 19, 9-30.

Dlamini, M. T., Eloff, J. H. P. & Eloff, M. M. (2008) Information Security: The Moving Target. **Computers & Security**, In Press, Accepted Manuscript.

Dzazali, S. (2006) Social Factors Influencing The Information Security Maturity Of Malaysia Public Service Organization: An Empirical Analysis. Association Of Information Systems- Australasian (ACIS 2006), 6-8 December 2006 2006 Adelaide, Australia. AIS Electronic Library (Aisel), 1-8.

Ein-Dor, P. & Segev, E. (1978). Organizational Context And The Success Of Management Information Systems. **Management Science**, 24 1064-1077.

Eloff, J. H. P. & Eloff, M. M. (2005) Information Security Architecture. **Computer Fraud & Security**, 2005, 10-16.

Eloff, M. M. & Von Solms, S. H. (2000) Information Security Management: A Hierarchical Framework For Various Approaches. **Computers & Security**, 19, 243-256.

Farn, K.-J., Lin, S.-K. & Fung, A. R.-W. (2004) A Study On Information Security Management System Evaluation--Assets, Threat And Vulnerability. **Computer Standards & Interfaces**, 26, 501-513.

Finne, T. (1998) A Conceptual Framework For Information Security Management **Computers & Security**, 17, 303-307.

Fitzgerald, K. J. (1995) Information Security Baselines. **Journal of Information Management & Computer Security**, 3, 8-12.

Fomin, V. V., Vries, H. J. D. & Barlette, Y.(2008) ISO/IEC 27001 Information Systems Security Management Standard: Exploring The Reasons For Low Adoption. EUROMOT 2008 Conference, 2008 Nice, France.: RSM Erasmus University.

Gillies, A. (2011) Improving The Quality Of Information Security Management Systems With ISO27000. **The TQM Journal**, 23, 367-376.

Hu, Q., Hart, P. & Cooke, D. (2007) The Role Of External And Internal Influences On Information Systems Security- A Neo-Institutional Perspective. **Journal of Strategic Information Systems**, 16, 153-172.

James, H. L. (1996) Managing Information Systems Security: A Soft Approach. Information Systems Conference Of New Zealand, 1996. Proceedings, 1996. 10-20.

Kankanhalli, A., Teo, H.-H., Tan, B. C. Y. & Wei, K.-K. (2003) An Integrative Study Of Information Systems Security Effectiveness. **International Journal of Information Management**, 23, 139-154.

Kaynak, E., Tatoglu, E. & Kula, V.(2005) An Analysis Of The Factors Affecting The Adoption Of Electronic Commerce By Smes: Evidence From An Emerging Market. **International Marketing Review**, 22, 623-640.

Keller, S., Powell, A., Horstmann, B., Predmore, C. & Crawford, M. (2005) Information Security Threats And Practices In Small Businesses. **Information Systems Management**, 22, 7-19.

Kraemer, S., Carayon, P. & Clem, J. (2009) Human And Organizational Factors In Computer And Information Security: Pathways To Vulnerabilities. **Computer & Security**, 28, 509-520.

Lab, K. 2011. Global IT Security Risk. Kaspersky Lab ZAO. [Accessed on 19 March 2012]

Lech, J. J. (2000) Information Security Framework For Health Information Systems. Healthcare Information Systems: Challenges Of The New Millennium. IGI Publishing.

Loch, K. D., Carr, H. H. & Warkentin, M. E. (1992) Threats To Information Systems: Today's Reality, Yesterday's Understanding. **MIS Quarterly**, 16, 173-186.

Monfelt, Y., Pilemalm, S., Hallberg, J. & Yngström, L. (2011) The 14-Layered Framework For Including Social And Organizational Aspects In Security Management. **Information Management & Computer Security**, 19, 124-133.

Mycert, M. C. E. R. T. (2012) Mycert Incidents Statistics [Online]. Malaysia: Mycert. [Accessed 27 January 2012 2012].

Ozkan, S. & Karabacak, B. (2010) Collaborative Risk Method for Information Security Management Practices: A Case Context within Turkey. **International Journal of Information Management**, 30, 567-572.

Parker, D. B. (1997) The Strategic Value Of Information Security In Business. **Computer & Security**, 16, 572-582.

PWC, P. C. (2012) The 2012 Global State of Information Security Survey. Key findings from the 2012 Global State of Information Security Survey. PriceWaterHouse Coopers. [accessed on 19 March 2012]

Rogers, E. M. (1983) Diffusion Of Innovation 3rd Edition, Macmillan Publishing Co., Inc.

Schlarman, S.(2002) The Case for a Security Information System **Information Security Journal: A Global Perspective**, 11, 44-50.

Siponen, M. (2006) Information Security Standards Focus On The Existence Of Process, Not Its Content. **Commun. ACM**, 49, 97-100.

Siponen, M. & Willison, R. (2008) Information Security Management Standards: Problems And Solutions **Information And Management**, 46, 267-270.

Siponen, M. T. & Oinas-Kukkonen, H. (2007) A Review Of Information Security Issues And Respective Research Contributions. **SIGMIS Database**, 38, 60-80.

Straub, D. W. (1990) Effective IS Security: An Empirical Study. **Information Systems Research** 1, 255-276.

Straub, D. W. & Welke, R. J. (1998) Coping With Systems Risk: Security Planning Models For Management Decision Making. **MIS Quarterly**, 22, 441-469.

Tornatzky, L. G. & Fleisher, M. 1990. The Process Of Technological Innovation, United States, Lexington Books.

Torres, J. M., Sarriegi, J. M., Santos, J. & Serrano, N. 2006. Managing Information Systems Security: Critical Success Factors And Indicators To Measure Effectiveness **Lecture Notes In Computer Science**, 4176/2006, 530-545.

Trcek, D. (2003) An Integral Framework For Information Systems Security Management. **Computer & Security**, 22, 337-360.

Tryfonas, T., Kiountouzis, E. & Poulymenakou, A. (2001) Embedding Security Practices In Contemporary Information Systems Development Approaches**. Information Management & Computer Security**, 9, 183-197.

Tsohou, A., Kokolakis, S., Lambrinoudakis, C. & Gritzalis, S. (2010) A Security Standards' Framework To Facilitate Best Practices' Awareness And Conformity**. Information Management & Computer Security** 18, 350-365.

Veiga, A. D. & Eloff, J. H. P. (2007) An Information Security Governance Framework. **Information Systems Management**, 24, 361-372.

Veiga, A. D. & Eloff, J. H. P. (2010) A Framework And Assessment Instrument For Information Security Culture **Computers & Security**, 29, 196-207.

Von Solms, B. 2005. Information Security Governance: COBIT Or ISO 17799 Or Both? **Computers & Security**, 24, 99-104.

Von Solms, B. (2006) Information Security - The Fourth Wave. **Computers & Security**, 25, 165-168.

Von Solms, B. & Von Solms, R. (2004) The 10 Deadly Sins Of Information Security Management. **Computers & Security**, 23, 371-376.

Werlinger, R., Hawkey, K. & Beznosov, K. (2009) An Integrated View Of Human, Organizational And Technological Challenges Of IT Security Management. **Information Management & Computer Security**, 17, 4-19.

Wood, C. C. (1987) Information Systems Security: Mangement Success Factors. **Computer & Security**, 6, 314-320.

Yeh, Q.-J. & Chang, A. J.-T. (2007) Threats and countermeasures for information system security: A cross-industry study. **Information & Management**, 44, 480-491.

Yildirim, E. Y., Akalp, G., Aytac, S. & Bayram, N. (2011) Factors Influencing Information Security Management in Small- and Medium Sized Enterprises: A Case Study from Turkey International **Journal of Information Management**, 31, 360-365.

Zuccato, A. (2007) Holistic security management framework applied in electronic commerce. **Computers & Security**, 26, 256-265.