Traditional bias toward journals in citation databases diminishes the perceived value of conference papers and their authors.

BY BJORN DE SUTTER AND AÄRON VAN DEN OORD

To Be or Not To Be Cited in Computer Science

THE SUSTAINABILITY AND nonconformism of conferences as premier publication venues in computer science is the subject of intense debate.[3,4,17,18] Evaluating scientists for promotion and budget allocation involves metrics like journal impact factors[14] and h-indexes[7] based on citation counts retrieved from Scopus and the Web of Science (WoS).

Many computer scientists view the historical focus of these databases on journals as a professional disadvantage, even though many conferences have been included in Scopus since 2004 and WoS since September 2008, including older ones entered later.

» **key insights**

■ The supposedly reliable Scopus and WoS include incomplete citation records for indexed CS journal and conference papers.

■ Despite the difficulty of automatically retrieving Google Scholar records, these records can still be used to correct the missing records in Scopus and WoS.

■ Corrected citation counts allow for much fairer evaluation of CS researchers and related conferences.

Inclusion of proceedings and journals in Scopus and WoS is often viewed as a stamp of approval and relevance. By contrast, databases like CiteSeer[x] and Google Scholar (GS) also cover books, technical reports, and other less-important manuscripts. Moreover, whereas Scopus and WoS are generally viewed as providing correct information, GS is known to include erroneous records.[11]

Higher citation counts than those in Scopus and WoS can be obtained by extending the coverage of a citation count by, say, including citations of non-indexed publications[11] and by combining databases.[10] Despite the need to manually cleanse GS records of erroneous and irrelevant records,[2,5,10,11] GS is useful for extending

coverage as well.[1,4,10,11,18]

Meho and Rogers[10] concluded in 2008 that choosing WoS or Scopus did not have a significant effect on the citation-based ranking of human-computer interaction researchers. However, in a case study involving library and information researchers, Meho and Yang[11] observed the opposite, finding that conclusions drawn for one scientific domain cannot be generalized to other domains.

Complementing coverage studies, this article explores the inaccuracy of citation records, along with their effect on the perceived impact of CS conferences and on author ranking. Figure 1 outlines the difference between coverage and accuracy for an author of a book B and a journal article J1, with their citations visualized at the top of the figure. B is cited by another article J2 and by a technical report TR. J1 is cited by a conference paper C. TR is a preliminary version of C, with the same title and authors. The list of references in C is shorter than TR, so TR cites B, but C does not cite B.

The middle segment of the figure reflects GS citation records, with GS mistakenly attributing the citations by TR to C. GS also covers publications of lesser importance (such as books). Due to its erroneous and for certain policies irrelevant records, the GS citation count in the example is not reliable.

WoS records are visualized in the bottom segment of the figure. The first observation is that WoS does not index less-important manuscripts (such as TR and B). However, WoS

Undercitation in some databases seems to be caused mostly by their use of inferior parsing technology.

sometimes does keep track of their citations by indexed papers (such as the citation of B by J2). With the manual-count method of Meho and Rogers,[10] these citations can still be counted, should citation-analysis policy demand it. Second, the citations of B by C and of J1 by J2 were never added to WoS. For policies that neglect citations by papers not indexed in WoS, the missing citation of B does not matter, but the missing citation of J1 always matters. Both the citing and cited papers have the WoS stamp of approval, so the citation should be counted. But when for some reason the database lacks a correct record of the citation, as in this example, it is not counted, and the author suffers professionally from undercitation. The study by Meho and Yang[11] on library and information researchers said 0.5%, 4.4%, and 12% of relevant citations were, at the time of the study missing from GS, Scopus, and WoS, respectively, due to database errors.

Here, we evaluate undercitation resulting from such an error. Complementing the studies mentioned earlier, we recently uncovered a significant undercitation bias in Scopus and WoS against covered CS conferences, demonstrating how it weakens the CS community's effort to win greater appreciation for conference papers. We also found how variations in undercitation of individual authors make the ACM Digital Library (DL), Scopus, and WoS unreliable information sources for citation-based metrics. We also present an automated method that combines the coverage of GS with the quality assurance of Scopus and WoS to detect undercitation resulting from missing citations.

We do not question Scopus or WoS coverage. The analyses we perform for any such database involve only publications indexed in that database. Hence all undercitation results presented here are independent of database coverage. Moreover, we do not take a position for or against citation-based metrics, though their usefulness has been questioned,[12] and many refinements have been proposed.[13,14] Our results demonstrate only that unless a corrective method is used, as we do here, to correct raw counts obtained from Scopus and WoS, their in-

Figure 1. Publications (vertices) and citations (edges) recorded by databases.
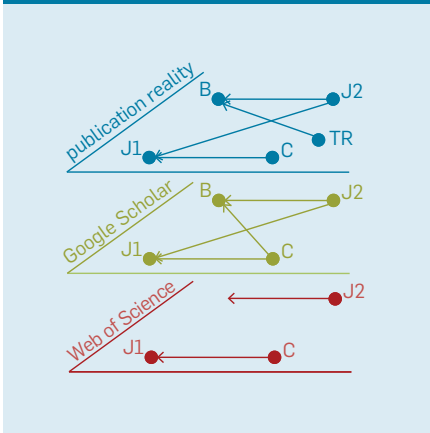
accuracy makes them unsuitable for CS research evaluation.

Relative Relevant Undercitation

To study the accuracy of database citation records, we measure the records' relative relevant undercitation (RRU); the RRU of a database query is the fraction of all (cited, citing) paper pairs for which both cited and citing papers are indexed in the database but for which the database has no record of the citing paper in the cited-by list of the cited paper. This fraction equals the underestimation of the citation count reported by the database within its own coverage; in Figure 1, the citation of J1 by J2 is missing in WoS, but the citation by C is present, resulting in an RRU of 50%.

To compute RRUs, we developed a Python tool for querying six online databases—the ACM DL, CiteSeer[x], DBLP, GS, Scopus, and WoS—by mimicking a researcher manually browsing a database by sending similar HTTP and parsing retrieved (HTML) data. Given a reference list of an author's papers, the tool first queries the databases by title; for papers not found by title, it tries searching by cited author. The search is limited to the papers in the reference list to prevent counting publications by other authors with the same name or initials.[10]

For each paper found in a database, the tool retrieves its cited-by list. In its extended mode of operation, it downloads the BibTeX or End-Note descriptions provided by the database for all entries in that list. In its fast mode the tool instead parses the HTML pages to identify the citing papers. As those pages display information in a less-uniform way than EndNote or BibTeX, the fast mode can produce less-accurate results. However, this mode is considerably faster for most databases than its extended-search mode, as fewer HTTP queries are needed. Most databases try to detect and block seemingly automated querying. To work around this filter, the tool is designed to sleep a random amount of time, say, 25 to 35 seconds between consecutive queries to GS. The result of this first search phase is a list of (cited, citing) paper pairs of citations, each recorded by at least one database. This list constitutes the tool's estimate of an author's publication genuine citation count.

In a second phase, the tool searches all databases for all papers occurring in the list. This search by title automates a search comparable to manual searches in other studies.[11] When both the cited paper and the citing paper of a citation are found in a database, the tool considers that citation relevant for that database. For each relevant citation, the tool searches the cited-by list of the cited paper. When the citing paper is in it, the tool labels the citation as found in the database. In such cases the citation is also included in automated citation counts provided by that database. When the citing paper is not found in the cited-by list, the tool labels the citation as relevant but missing. A database's RRU for a reference list of papers is the number of missing relevant citations divided by number of relevant citations.

As some citations may not be recorded in any searched database, our tool can underestimate RRUs. The risk of underestimation can be avoided with the manual-count method of Meho and Rogers,[10] though their experience suggests their labor-intensive method will not identify a significant number of additional relevant missing citations.

Due to erroneous database records, our tool can also unintentionally overestimate the number of relevant but missing citations in a database, thereby overestimating its RRU; we quantify this potential overestimation later.

Experiments

We used the tool in 2010 and 2011 to perform three complementary experiments: First, we set it to search all aforementioned databases for three authors, using its extended-search mode. Though we focus here on citation accuracy, the experiment also enabled us to compare a database's coverage on the basis of what the three authors would consider their own relevant output. Due to the tool's long running times—several weeks—we were able to study only three authors in the experiment. We next searched GS and WoS for 14 editors-in-chief of various CS transactions published by ACM and the IEEE Computer Society. Using the tool's fast mode, we thus limited searches to publication lists we obtained from DBLP. The experiment was less accurate and covered fewer databases than the first experiment but included many more authors and publications, enabling us to validate the trends we observed in the first experiment. Finally, we performed a similar experiment for GS and WoS for eight ACM and IEEE transactions published from 2000 to 2002 to study the influence of RRU on journal impact factors.

Experiment one. Three colleagues at Ghent University who began publishing around 1990 assembled a reference list of their own peer-reviewed conference and journal publications; Table 1 lists the number of publications in each database. In the Computer Systems Lab at Ghent University, we have permanent access to the ACM DL, and WoS, as well as to various free databases. This experiment was carried out from September 18 to October 13, 2010, when we also had temporary access to Scopus.

GS, Scopus, and WoS provide excellent coverage of journal papers. In line with the findings of others,[4–6,10,15]

Table 1. Number of journal/conference papers in authors' reference lists and in online publication databases.

Author (domain)	Reference	WoS	Scopus	ACM	Google	CiteSeer[x]
Koen De Bosschere (compilers, computer architecture)	66/143	57/89	50/54	40/51	57/112	7/38
Bart Dhoedt (distributed computing, networks)	53/187	51/113	44/100	23/47	43/143	3/17
Wilfried Philips (image and video processing)	64/285	56/130	58/106	13/17	58/214	6/26

GS covers more conferences than other academic databases. Unlike some other studies[10] we did not find more extended conference coverage in Scopus compared to WoS. This might have been due to increased coverage in WoS in the studies.

Table 2 lists the citation counts in each database; in line with previous studies, the Scopus and WoS citation counts were only a small fraction of those in GS. Our tool thus relied heavily on the unreliable GS to compute RRUs.

Table 3 contrasts the numbers of relevant citations to the found numbers and are partitioned into four categories—J2J, C2J, J2C, and C2C—to distinguish whether citing and cited publications are journal (J) or conference (C) papers. The last column on the right combines Scopus and WoS, where a citation is considered relevant/found as soon as it is relevant/found in at least one of the two and missing if found in neither. Table 4 lists the h-indexes computed by the tool for the databases; the found h-indexes were based on found citations and the corrected h-indexes on relevant citations. The corrected h-indexes correspond to the h-indexes

we would obtain if a database would fix all its missing citations. As the tool gives us a list of missing relevant citations, we requested corrections of citation records through the WoS correction-request form. Most of our requested corrections were applied within weeks. We used the tool to collect the numbers presented here before that correction. As the h-indexes are based on coverage, the corrected h-index in one database may be smaller than the counted h-index in another database.

The large RRUs in ACM, Scopus, and WoS indicate that missing relevant citations are an important cause of undercitation. We also observed a large variation in the RRUs of individual authors, affecting their ranking based on WoS and WoS+Scopus h-indexes. Unlike the ACM DL, which should be able to handle conferences with the same completeness as journals, Scopus and WoS reflect much more undercitation for conferences than for journals. So independent of their coverage, Scopus and WoS put conference-oriented authors at a disadvantage. Worse is their J2C undercitation. Large numbers of J2C citations is one of the strongest argu-

ments for convincing scholars and researchers from non-CS disciplines to value CS conference papers, though it is precisely those citations that are most underestimated. For example, Koen (listed in the tables) might try to convince a promotion committee that conferences should be valued like journals in his domain by pointing to his high #J2C/(#J2J+#J2C) ratio of 43% in WoS. However, this ratio is not nearly as convincing as the 60% he achieved with corrected WoS records.

We see three potential causes for many of the missing citations: First is overcitation in other databases or inclusion of nonexisting citations; the next experiment demonstrates these possibilities occur to a limited degree. The remaining causes are the incorrect parsing of correct references and the occurrence of incorrect and incomplete references in papers, or so-called miscitations. Some papers have been miscited in more than 165 different ways,[16] with more miscitations among non-English names[8] and in papers with more authors.[9] For example, our own work is often miscited because the "De," "van," and "den" are incorrectly treated as middle names or because they are capitalized incorrectly. The RRUs we found in WoS and Scopus are more than an order of magnitude higher than the 0.5% and 4.4% found by Meho and Yang.[11] But that difference is not surprising; of all the scientific disciplines, librarians and information scientists probably produce the most accurate citations. Note, however, that our experiments

Table 2. Total number of citations per author and citations in each database.

Author	Total	WoS	Scopus	ACM	Google	CiteSeer[x]
Koen	2280	384 (17%)	531 (23%)	217 (10%)	2156 (95%)	138 (6%)
Bart	1056	278 (26%)	313 (30%)	62 (6%)	842 (80%)	20 (2%)
Wilfried	1937	745 (38%)	975 (50%)	10 (1%)	1370 (71%)	46 (2%)

Table 3. Number of citations per author, category, and database and corresponding RRUs.

		WoS				Scopus				ACM				WoS + Scopus			
		J2J	C2J	J2C	C2C	J2J	C2J	J2C	C2C	J2J	C2J	J2C	C2C	J2J	C2J	J2C	C2C
Koen	relevant	209	205	312	278	201	249	190	226	73	110	142	147	246	288	352	339
	found	101	114	77	92	151	192	75	113	42	65	56	54	180	232	132	179
	RRU	52%	44%	75%	67%	25%	23%	61%	50%	42%	41%	61%	63%	27%	19%	63%	47%
Bart	relevant	194	141	127	104	156	122	112	107	24	34	25	26	211	156	149	132
	found	140	88	19	31	131	96	37	49	15	12	14	21	188	144	45	70
	RRU	28%	38%	85%	70%	16%	21%	67%	54%	38%	65%	44%	19%	11%	8%	70%	47%
Wilfried	relevant	411	240	211	101	507	323	209	141	6	4	10	9	531	385	275	204
	found	350	239	82	74	453	304	124	94	3	0	4	3	496	369	177	153
	RRU	15%	0%	61%	27%	11%	6%	41%	33%	50%	100%	60%	67%	7%	4%	36%	25%

consider only citations recorded by at least one database. So even if the occurrence of incorrect or incomplete references inflates a database's RRU, it apparently did not stop GS or other databases from recording those citations. Undercitation in some databases therefore seems to be caused mostly by their use of inferior parsing technology.

Experiment two. We ran a second experiment in April 2011 for seven editors-in-chief of ACM transactions and seven editors-in-chief of IEEE Computer Society transactions, aiming to validate the previously obtained RRUs on a larger sample set and assess our tool's accuracy in the presence of erroneous GS records. Due to the large amount of data, we ran the tool in its fast mode. Based on 14 reference publication lists obtained from DBLP, the tool collected 36,931 citations for 1,778 papers in GS and WoS, labeling 18,342 citations as relevant to WoS, of which 9,669 were in WoS, and the remaining 8,673 as missing.

Among the 8,673 missing citations, 1,678 cited conference papers were not cited according to WoS. For such papers, the possibility must be considered that conference data was entered into WoS long after the conferences took place and hence after the citing papers were entered. To estimate the likelihood of this potential cause of RRU, we performed an additional check, building on the assumption that if the late entering of conference data is a major cause of RRU, the result would likely be WoS reporting

all papers of the covered conference editions as having zero citations. For each of the 14 editors-in-chief, we selected one conference paper with no citations according to WoS and with the most citations according to GS. For these papers, which were published from 1985 to 2009 and covered 445 of the 1,678 suspect citations, we manually verified that at least one paper of the same conference edition was cited at least once according to WoS. The result was positive in terms of finding citations for all 14 conferences, indicating that late entering of conference data is not likely a major cause of RRU in WoS. Even if late entering of data was a significant cause of undercitation in WoS, similar late entering apparently did not prevent GS from being more complete.

To estimate how much erroneous GS records inflate the RRU in WoS, we manually checked 15 randomly selected citations per author the tool had labeled as missing from WoS. From these $14\times15=210$ supposedly missing citations, 19 had been labeled incorrectly as such; the corresponding 95%-confidence interval based on the normal approximation is $9.5\pm3.9\%$. To compensate for this

overestimation, we corrected the number of missing relevant citations as reported by the tool with 9.5% for computing the RRUs reported later in this article.

Most incorrect labels resulted from confusing multiple manuscripts with the same title and authors. Whereas the tool's fast mode was responsible for the confusion, the more extended-search mode as used in the first experiment would have prevented the error. However, in most cases of incorrect labeling, GS simply provided incorrect citation information. In the majority of such cases, GS provided a link to the citing document on CiteSeer[x]. We inspected the .pdf documents cached on CiteSeer[x], discovering they indeed cited the cited paper. We also discovered these .pdfs are not from the published conference or journal papers CiteSeer[x] claimed them to be but rather from technical reports and Ph.D. theses with the same title and authors but with longer reference lists due to lack of a page limit on the documents. While Google provides little public documentation on its information sources, the correlation between the errors in CiteSeer[x] and GS points in the direction of CiteSeer[x] as

Table 4. Found/corrected h-indexes for the various databases.

Author	WoS	Scopus	WoS + Scopus	ACM	Google	CiteSeer[x]
Koen	12/16	13/17	15/19	8/14	24/25	7/7
Bart	9/12	8/11	11/12	4/6	13/15	3/4
Wilfried	13/15	15/17	16/18	2/4	20/20	4/5

Figure 2. WoS RRUs for 14 editors-in-chief.
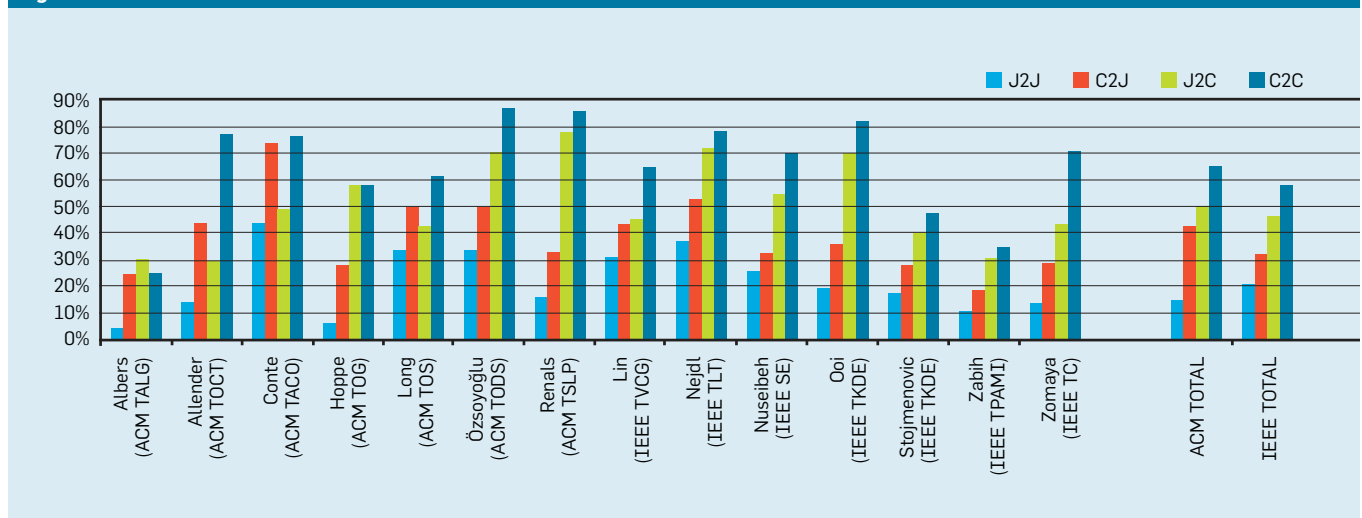
the culprit for a considerable fraction of overcitation in GS. Whatever the cause, however, this overcitation with 9.5% is much smaller than the undercitation we found for other databases.

The conclusions of our first experiment remain valid, as confirmed in Figure 2; also, for the editors-in-chief, there was significant varying RRU for all types of citations, and the RRU was particularly large for J2C and C2C citations. Figure 3 visualizes the h-indexes and h-cores of the 14 editors-in-chief. An h-core consists of an author's x papers each cited x or more times, with x being the author's h-index.[7] The bars in Figure 3a represent the h-indexes computed on found citations in WoS, and the bars in Figure 3b represent corrected h-indexes based on relevant citations, confirming h-indexes based on found citations suffer significantly from undercitation. Some authors suffer more than others, to the point their ranking is altered significantly; for example, Ooi was in next-to-last place according to uncorrected WoS citation counts but in third place after correction.

Figures 3a and 3b also show the contribution of conference papers to h-indexes. Each blue/orange box indicates a journal/conference paper in the h-core, ordered left to right from most cited to least cited. For example, Albers has an h-index of 11 in WoS; of the 11 papers in her h-core, the first, fifth, sixth, and 10th most-cited are journal papers. Based on WoS counts, 43% of all papers in the h-cores are conference papers, and based on the corrected counts, 60% are conference papers. Of the five most-cited papers per author, WoS reports 36% are conference papers, whereas the corrected citation counts report 56% are conference papers. These numbers are much higher than those presented by Bar-Ilan[1] as obtained from WoS in late 2008/early 2009. This higher count might result from increased coverage in WoS from 2009 to 2011, though we could not verify this conclusion. These results again confirm that using uncorrected WoS citation counts to estimate the importance of conferences for an author can lead to significant underestimation for that author. Moreover, such underestimation varies significantly from author to author; for Zomaya in Figure 2 and Figure 3, WoS attributes 2/12 h-core papers to conferences, which is close to the corrected number of 2/13. However, for Ooi in Figure 2 and Figure 3, WoS attributes 4/11 to conferences, which does not even approximate the corrected number 15/20. We conclude that GS should be used as a complementary source of information to obtain accurate citation counts, even when policy stipulates the citation analysis is limited to WoS coverage. Though this second experiment did not include the ACM DL or Scopus, this conclusion can be extended to those databases, of which the first experiment revealed compa-

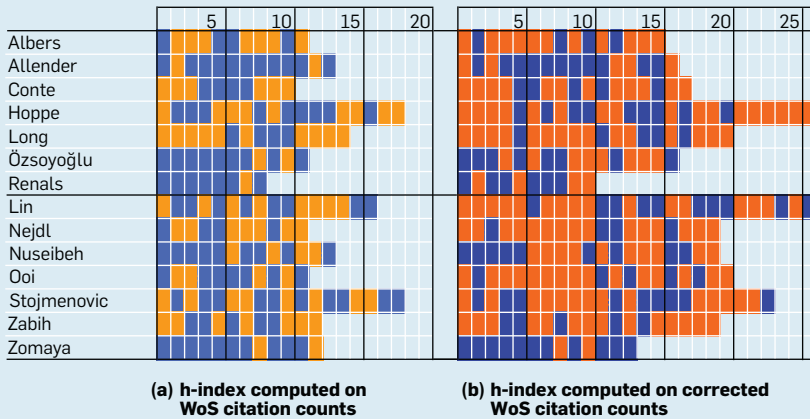**Figure 3. WoS h-indexes and h-cores for 14 editors-in-chief.**

(a) h-index computed on WoS citation counts

(b) h-index computed on corrected WoS citation counts

Figure 4. WoS RRUs for articles in eight transactions volumes, 2000–2002.
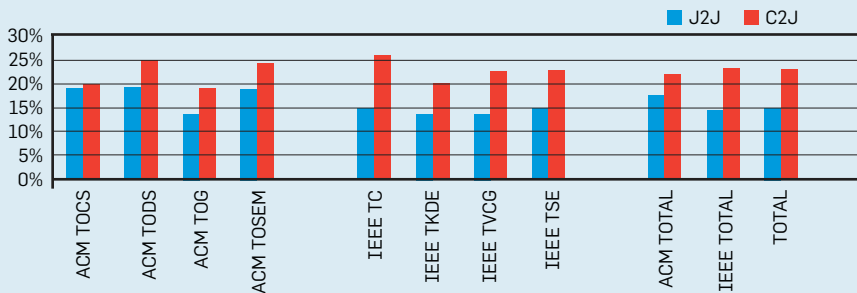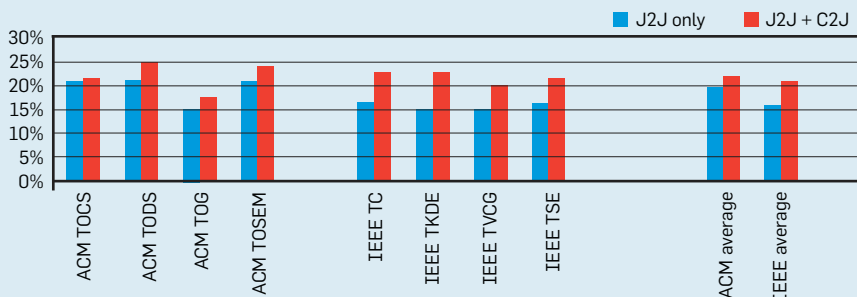

Figure 5. Underestimation of average journal impact factors, 2002–2003.

rable levels of undercitation.

Experiment three. We applied a similar search for the articles published in eight ACM and IEEE transactions from 2000 to 2002, selecting these years to allow collection of citations over a significant period of time and to include four ACM and four IEEE transactions from related, largely overlapping domains. We excluded editorials and republished proceedings to ensure a fair comparison. For the 135 ACM articles and 770 IEEE articles in the considered volumes, our tool's fast mode collected 42,658 citations in GS and WoS in April 2011. Before applying a correction with 9.5%, our tool labeled 19,215 citations as found and relevant to WoS, and 5,193 as relevant but missing.

Figure 4 outlines the resulting, corrected RRUs, which are comparable to those of the other experiments, but the underestimation of C2J citations is over 10% less than what we observed in the first two experiments. The RRUs also reflect considerably less variation, suggesting C2J citations of the selected ACM and IEEE transactions are recorded more accurately in WoS than are the citations of other journals in which the sampled authors were published.

For J2J citations our tool confirmed much less variation between the ACM and the IEEE transactions than we observed for individual authors in the first two experiments. However, based on a T-test, the difference in average J2J RRU between ACM and IEEE is statistically significant; IEEE transactions papers are undercited less than ACM transactions papers. To determine whether this difference might have resulted from differences in citation formats, policies, or culture between ACM and IEEE, we analyzed the sources of the citations. While this analysis indicates ACM and IEEE papers favor citing within their own organization, the numbers were inconclusive with respect to the cause of the different RRUs.

Finally, we studied the effect of undercitation on journal impact factors by computing their average underestimation for 2002 and 2003. These impact factors are based on citations of papers published from 2000 to 2002, the years for which our tool crawled the databases. We used two methods: one in which only J2J citations of full articles (excluding letters, editorials, and republished proceedings) were counted, and one in which J2J and C2J were counted. For each method, we computed the impact factors based on citations the tool found in WoS and on corrected WoS counts. Including the C2J counts resulted in impact factors between 2.37x and 4.35x higher, averaging 3.63x higher for ACM transactions and 3.39x for IEEE transactions. Figure 5 outlines the underestimation of the impact factors resulting from missing citations. In line with the previous results, the underestimation was significant (15%–21%) when only J2J citations are counted, and higher (17%–25%) when C2J citations are included. Due to their higher RRU in WoS, the tool reported the impact factors for ACM transactions are underestimated considerably more than those for IEEE transactions.

These results indicate it is in the interest of ACM and IEEE to include C2J citations to compute journal impact factors and ensure they are recorded more accurately.

Conclusion

ACM, Scopus, and WoS must develop better reference-parsing technology to fix the significant undercitation in their databases. Due to the variations in undercitation, the ACM DL, Scopus, and WoS, and even combinations thereof, are unreliable information sources for the most commonly used citation-based metrics in CS. Moreover, Scopus and WoS databases reflect a significant bias against covered conference proceedings, resulting in underestimation of their impact.

Supposedly unreliable, broad databases like GS can be used to identify and correct undercitation problems in Scopus and WoS without undermining the virtues of their selective inclusion of high-impact conferences and journals. However, due to the inherently slow access to databases like GS, even an automated tool like ours is slow, to the point of being not generally applicable.

Finally, we found a correlation between transactions publishers and their transactions' undercitation, to the disadvantage of ACM. **C**

References
1. Bar-Ilan, J. Web of Science with the Conference Proceedings Citation Indexes: The case of computer science. *Scientometrics 83*, 3 (June 2010), 809–824.
2. Bar-Ilan, J. Which h-index? A comparison of WoS, Scopus, and Google Scholar. *Scientometrics 74*, 2 (Feb. 2008), 257–271.
3. Birman, K. and Schneider, F. Program committee overload in systems. *Commun. ACM 52*, 5 (May 2009), 34–37.
4. Freyne, J., Coyle, L., Smyth, B., and Cunningham, P. Relative status of journal and conference publications in computer science. *Commun. ACM 53*, 11 (Nov. 2010), 124–132.
5. García-Pérez, M. Accuracy and completeness of publication and citation records in the Web of Science, PsycINFO, and Google Scholar: A case study for the computation of h-indices in psychology. *Journal of the American Society for Information Science and Technology 61*, 10 (Oct. 2010), 2070–2085.
6. Harzing, A.-W. and van der Wal, R. Google Scholar as a new source for citation analysis. *Ethics in Science and Environmental Politics 8*, 1 (June 2008), 61–73.
7. Hirsch, J. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America 102*, 46 (Nov. 2005), 16569–16572.
8. Kotiaho, J. Papers vanish in mis-citation black hole. *Nature 398*, 6722 (Mar. 1999), 19–19.
9. Kotiaho, J., Tomkings, J., and Simmons, L. Unfamiliar citations breed mistakes. *Nature 400*, 6742 (July 1999), 307–307.
10. Meho, L. and Rogers, Y. Citation counting, citation ranking, and h-index of human-computer interaction researchers: A comparison of Scopus and Web of Science. *Journal of the American Society for Information Science and Technology 59*, 11 (Sept. 2008), 13–28.
11. Meho, L. and Yang, K. A new era in citation and bibliometric analyses: Web of Science, Scopus, and Google Scholar. *Journal of the American Society for Information Science and Technology 58*, 13 (Nov. 2007), 2015–2125.
12. Meyer, B., Choppy, C., Staunstrup, J., and Van Leeuwen, J. Research evaluation for computer science. *Commun. ACM 52*, 4 (Apr. 2009), 31–34.
13. Moed, H. and Van Leeuwen, T. Improving the accuracy of Institute for Scientific Information's journal impact factors. *Journal of the American Society for Information Science 46*, 6 (July 1995), 461–467.
14. Moed, H., De Bruin, R., and Van Leeuwen, T. New bibliometric tools for the assessment of national research performance: Database description, overview of indicators, and first applications. *Scientometrics 33*, 3 (July 1995), 381–422.
15. Noruzi, A. Google Scholar: The new generation of citation indexes. *Libri 55*, 4 (Dec. 2005), 170–180.
16. Price, N. What's in a name (or a number or a date)? *Nature 395*, 6702 (Oct. 1998), 538–538.
17. Vardi, M. Conferences vs. journals in computing research. *Commun. ACM 52*, 5 (May 2009), 5–5.
18. Wainer, J., Goldenstein, S., and Billa, C. Invisible work in standard bibliometric evaluation of computer science. *Commun. ACM 54*, 5 (May 2011), 141–146.

Bjorn De Sutter (bjorn.desutter@elis.ugent.be) is a professor in the Computer Systems Lab of the Department of Electronics and Information Systems at Ghent University, Ghent, Belgium.

Aäron van den Oord (aaron.vandenoord@elis.ugent.be) is a Ph.D. student in the Computer Systems Lab of the Department of Electronics and Information Systems at Ghent University, Ghent, Belgium.