Quantitative Analysis of the Full Bitcoin Transaction Graph

Dorit Ron and Adi Shamir

Department of Computer Science and Applied Mathematics,
The Weizmann Institute of Science, Israel
{dorit.ron,adi.shamir}@weizmann.ac.il

Abstract. The Bitcoin scheme is a rare example of a large scale global payment system in which all the transactions are publicly accessible (but in an anonymous way). We downloaded the full history of this scheme, and analyzed many statistical properties of its associated transaction graph. In this paper we answer for the first time a variety of interesting questions about the typical behavior of account owners, how they acquire and how they spend their Bitcoins, the balance of Bitcoins they keep in their accounts, and how they move Bitcoins between their various accounts in order to better protect their privacy. In addition, we isolated all the large transactions in the system, and discovered that almost all of them are closely related to a single large transaction that took place in November 2010, even though the associated users apparently tried to hide this fact with many strange looking long chains and fork-merge structures in the transaction graph.

Keywords: Bitcoin, digital coins, electronic cash, payment systems, transaction graphs, quantitative analysis

1 Introduction

Bitcoins are digital coins which are not issued by any government, bank, or organization, and rely on cryptographic protocols and a distributed network of users to mint, store, and transfer. The scheme was first suggested in 2008 by Satoshi Nakamoto [1], and became fully operational in January 2009. It had attracted a large number of users and a lot of media attention [2] [3] [4], but so far it was difficult to get precise answers to simple questions such as: How many different users are there in the system? How many Bitcoins are typically kept in each account, and how does this balance vary over time? Are most Bitcoins kept by a few large users? Do they keep their Bitcoins in "saving accounts" or do they spend them immediately? How many users had large balances at some point in time? What is the size distribution of Bitcoin transactions, and how many of them are micropayments?

In this paper we answer all these (and many additional) questions. We use the fact that all the transactions ever carried out in the Bitcoin system are available on the internet (in an anonymous way). On May 13^{th} 2012 we downloaded the full public record of this system, which consisted of about 180,000 HTML files. After parsing and processing these files, we built a graph of all the Bitcoin addresses and transactions up to that date. We then used the intrinsic properties of the scheme in order to identify many cases in which we can show that different addresses must belong to the same owner, and used this information to contract the transaction graph by merging such addresses, in order to get a more accurate picture of the full financial activity of all the owners. After obtaining this new graph, we analyzed many of its statistical properties. In this paper we describe the most interesting and informative distributions we found in a series of tables. In addition, we isolated all the large ($\geq 50,000$ Bitcoins) transactions which were ever recorded in the system, and analyzed how these amounts were accumulated and then spent. We discovered that almost all these large transactions were the descendants of a single large transaction involving 90,000 Bitcoins which took place on November 8^{th} 2010, and that the subgraph of these transactions contains many strange looking chains and fork-merge structures, in which a large balance is either transferred within a few hours through hundreds of temporary intermediate accounts, or split into many small amounts which are sent to different accounts only in order to be recombined shortly afterwards into essentially the same amount in a new account.

There was one previous reported attempt to download and analyze the full Bitcoin history which was described in [5]. They created the graph of transactions on July 12^{th} 2011, which was before the scheme really caught on. Thus, the total number of Bitcoins participating in all the transactions in our graph is about three times larger than in their graph. In addition, we expect the transactions in our more mature graph to better represent typical use of the system, whereas their graph represents primarily the experiments run by early adopters. However, the biggest difference between our papers is that they were primarily interested in privacy issues and showed that it is possible to identify various users in spite of the official anonymity of the scheme, whereas we are primarily interested in the statistical properties of the Bitcoin transaction graph.

Another analysis of the Bitcoin transaction graph was presented at the Chaos Computer Club Conference in Germany in December 2011 [6]. Again, they were primarily interested in how to defeat the anonymity of the network (based on the same idea of collapsing addresses belonging to a common owner), but also included some interesting comments about the economic principles behind the scheme, the effect of lost coins on its operation, weaknesses in its protocols, and the general topological properties of this transaction graph.

The paper is organized as follows. In Section 2 we describe the structure of the Bitcoin network. In Section 3 we summarize the main statistical distributions we extracted from the downloaded network, which describe many interesting and even surprising properties of the scheme. Finally, in Section 4 we present the graph of the largest transactions and analyze its strange structure.

2 The Bitcoin Network

Bitcoin is a decentralized electronic cash system using peer-to-peer networking, digital signatures and cryptographic proofs to enable payments between parties without relying on mutual trust. It was first described in a paper by Satoshi Nakamoto (widely presumed to be a pseudonym) in 2008. Payments are made in Bitcoins (BTC's), which are digital coins issued and transferred by the Bitcoin network. Nodes broadcast transactions to this network, which records them in publicly available web pages, called block chains, after validating them with a proof-of-work system.

Participants begin using Bitcoin by first acquiring a program called a Bitcoin wallet and one or more Bitcoin addresses. Bitcoin addresses are used for receiving Bitcoins, in the same way that e-mail addresses are used for receiving e-mails. Even though Bitcoin is considered to be an experimental payment system, it is already deployed on a large scale (in the sense that the current value of all the coins issued so far exceeds 100,000,000 USD) and attracts a lot of media attention. Its proponents claim that it is the first truly global currency which does not discriminate its users based on citizenship or location, it is always running with no holidays, it is easy to secure with very low usage fees, it has no chargebacks, etc. On the other hand, its detractors claim that it is widely misused to buy illegal items and to launder large sums of money, and that it is too easy to steal Bitcoins from wallets via cyber attacks.

Unlike fiat currency, which has been declared to be legal tender by a government despite the fact that it has no intrinsic value and is not backed by reserves, the Bitcoin system has no centralized issuing authority. The network is programmed to increase the money supply in a slowly increasing geometric series until the total number of bitcoins reaches an upper limit of 21 million BTC's. Bitcoins are awarded to Bitcoin "miners" for solving increasingly difficult proof-of-work problems which confirm transactions and prevent double-spending. The network currently requires over one million times more work for confirming a block and receiving an award (currently 50 BTC's) than when the first blocks were confirmed.

The exchange rate of Bitcoins has fluctuated widely over the years, from merely $0.01 to over $30 per BTC. Today (October 2012) it is worth a little over $12 per BTC. The entire activity in the Bitcoin network is publicly available through the internet and is recorded in the form of a block chain, starting at block 0 [7] (created back on the 3^{rd} of January 2009). Each block reports on as little as a single transaction to as much as over a thousand transactions, and provides hyperlinks to other blocks and to other activities of each address.

Many users adopt the Bitcoin payment system for political and philosophical reasons, as well as pragmatic ones, and some small businesses have started to accept Bitcoins for their physical or virtual merchandise. One underground website which accepts only Bitcoins, called Silk Road, enables users to buy any drug imaginable by using the Tor network to protect their anonymity [3]. In a 2011 letter to Attorney General Eric Holder and the Drug Enforcement Administration, senators Charles Schumer of New York and Joe Manchin of West Virginia called for an investigation into Silk Road and the Bitcoin network. Schumer described the use of Bitcoins at Silk Road as a form of money laundering (ML) [4]. Consequently, Amir Taaki of Intersango, a UK-based Bitcoin exchange, put out a statement calling for the regulation of Bitcoin exchanges by law enforcement. In fact, the most widely used Bitcoin exchange, Mt.Gox, warns new users who request a new account "Please be advised that accessing your account via the Tor network and/or public proxies may result in a temporary suspension of your account, and having to submit anti ML documents."

A transaction in Bitcoins is a generalization of a regular bank transaction in the sense that it allows multiple sending addresses and multiple receiving addresses in the same transaction. The senders and receivers of transactions are identified through their public keys from public/private key pairs, which we refer to as addresses. It specifies how many Bitcoins were taken from each sending address and how many Bitcoins were credited to each receiving address, without the details of who gave how much to whom. An address may receive Bitcoins which are either newly generated or have a specific sending address. Each owner can have an unbounded number of addresses owned by him. In fact, it is considered good practice for an owner to generate a new address, i.e., public-private key-pair, for every transaction. Owners are advised to take the following steps to better protect their identity: they do not have to reveal any identifying information in connection with their addresses; they can repeatedly send varying fractions of their BTC's to themselves using multiple (newly generated) addresses; and/or they can use a trusted third-party to mix their transactions with those of other owners. On the other hand, some owners volunteer to reveal their ownership of some particular addresses, e.g., when they advertise their merchandize, ask for donations, or act as Bitcoin exchanges (such as Mt.Gox, which owns a huge number of addresses).

A very important feature of the Bitcoin network is that a transaction involving multiple sending addresses can only be carried out by the *common owner* of all those addresses, as it is demanded by the Bitcoin system that "Whoever sent this transaction owns all of these addresses". This legal requirement is also technically ensured by the fact that each received amount must have a cryptographic digital signature that unlocks it from the prior transaction. Only the person possessing the appropriate address is able to create a satisfactory signature, and thus funds can only be spent by their owners. Under this assumption, it is possible to go over the entire list of transactions and merge the sets of addresses serving as senders of a single transaction into a single owner. This can cause a cascade of new mergings which we follow until the list of owners stops shrinking. This yields a lot of information about common ownership of addresses, but there is no guarantee that two addresses which do not get merged in this process in fact belong to different owners. By aggregating all the addresses and transactions which can be traced to the same owner, we can get a more informative picture of his total assets and financial activities. If we have any external information about the real ownership of any one of these merged addresses, we can get a fuller picture of the Bitcoin activity of that particular individual or organization. For example, since WikiLeaks publicly advertised one of its addresses when it asked for donations, we could determine that WikiLeaks owns at least 83 addresses, was involved in at least 1088 transactions and had an accumulated income in all these addresses of 2605.25 BTC's.

We acquired the complete state of the Bitcoin transaction system on May 13^{th} 2012, which contained all the transactions carried out in the system since its inception on January 3^{rd} 2009 until that date. This required downloading 180,001 separate but linked HTML files, starting from block number $180,000$ [8] and following the links backwards to the zeroth block initiating the system in January 2009. Each file was parsed in order to extract all the multisender/multireceiver transactions in it, and then the collection of transactions was encoded as a standard database on our local machine. We then ran a variant of a Union-Find graph algorithm [9] in order to merge all the addresses which are known to belong to the same owner, and to combine all the transactions which can thus be associated with him (but without eliminating the internal transfers, which become self loops in the graph). All the statistics described in the next section are derived from such a reduced transaction graph rather than from the original graph represented by the raw HTML files.

3 Statistics Calculated Over the Bitcoin Transaction Graph

At the time we downloaded the graph there were 3,730,218 different public keys, each associated with a different address: 3,120,948 of them were involved as senders in at least one transaction, while the additional 609,270 appear in the network only as receivers of BTC's. By running the Union-Find algorithm, we were able to associate the 3,120,948 addresses with 1,851,544 different owners. Since the other 609,270 addresses were never used as senders, they could not be merged with any other addresses by the Union-Find algorithm, and thus they all remained as an owner with a single address. By adding these singletons, we get a total of 2,460,814 (possibly) different owners, which implies that each one of them has on average about 1.5 addresses. However, there is a huge variance in this statistics, and in fact one owner owns 156,722 different addresses. By analyzing some of these addresses and following their transactions, it is easy to determine that this owner is Mt.Gox, which is the most popular Bitcoin Exchange site (responsible for almost 90% of all the exchange operations in the network). The full distribution of the number of addresses per owner is given in Table 1.

In our reduced transaction graph, each m-to-n transaction has a single sender (since the m sending addresses necessarily belong to the same owner) and at most n receivers. It can thus be decomposed into at most n different transactions from the single owner of the m senders to the owners of the n receivers. In case some of the receiving addresses are identified as sharing a common owner (using the owner-addresses map), their amounts are accumulated to create a single common transaction, and if some of the receivers are identified with the single sender, we create a single self loop with the combined amounts. The resulting

Table 1. The distribution of the number of addresses per owner

Larger or equal to	Smaller than	Number of owners
1	2	2,214,186
2	10	234,015
10	100	12,026
100	500	499
500	1,000	35
1,000	5,000	41
5,000	10,000	5
10,000	50,000	5
50,000	100,000	1
100,000		1

graph has 7,134,836 single sender and single receiver transactions, out of which 814,044 (about 11%) involve Deepbit (the largest Bitcoin mining pool), and 477,526 (about 7%) involve Mt.Gox. About 10% of the transactions are self loop. The transaction graph is not connected as it is composed of 133,742 different connected components, many of size one. For instance, there are as many as 43,710 components (about 33%) consisting of a single address which are used only for accepting (one or several batches of) freshly minted Bitcoins, and which have never participated in any incoming or outgoing transactions.

There are many types of statistics and graphs about the Bitcoin network which can be readily downloaded from the internet [10] [11]. However, these types of statistics tend to describe some global property of the network over time, such as the number of daily transactions, their total volume, the number of Bitcoins minted so far, and the exchange rate between Bitcoins and US dollars. We can go much further than that, since the entire transaction graph can be used to determine the financial history of each owner including all of his sending/receiving activities along with the daily balance of Bitcoins in his various addresses and how they vary over time. Having this graph at hand enables us to study various statistical properties of the network, which are not easy to determine by following a small number of online links in the Blockexplorer representation of the Bitcoin network. In the rest of this section, we describe some of our findings so far, but we expect to have a much deeper and richer analysis of the data in the near future.

Here is our first surprising discovery. The total number of BTC's in the system is linear in the number of blocks. Each block is associated with the generation of 50 new BTC's and thus there are 9,000,050 BTC's in our graph of owners (generated from the 180,001 blocks between block number zero and block number 180,000). However, if we sum up the amounts accumulated at the 609,270 addresses which only receive and never send BTC's, we see that their owners have actually put aside in some kind of "saving accounts" 7,019,100 BTC's, which are almost 78% of all existing BTC's. 59.7% of all the coins are "old coins" which were received more than three month before the cut off date (May 13^{th} 2012), and still had not triggered any outgoing transactions. This means that there are much fewer BTC's in circulation than previously presumed. Yet, the total number of Bitcoins participating in all the transactions since the establishment of the system (except for the actual minting operations) is 423,287,950 BTC's. This implies that each coin which is in circulation had to be moved a much larger than expected number of times.

Another interesting finding is that the total number of Bitcoins received by most owners is negligible. As can be seen from Table 2, 36% of all owners received fewer than one BTC (currently worth about 12 USD) each throughout their lifetime, 52% received fewer than 10 BTC's and 88% fewer than 100. At the other end of the distribution there are only four owners who received over 800,000 BTC's and 80 owners who received over 400,000.

Table 2. The distribution of the accumulated incoming BTC's per owner

Larger or equal to	Smaller than	Number of owners
0	1	893,763
1	10	389,302
10	100	881,273
100	1,000	255,826
1,000	10,000	36,713
10,000	50,000	3,593
50,000	100,000	181
100,000	200,000	55
200,000	400,000	30
400,000	800,000	76
800,000		4

Similarly, as can be seen in Table 3 the current (on May 13^{th} 2012) balance of almost 97% of all owners was less than 10 BTC's. This number decreases to 88% if instead of looking at one specific moment, we look at the *maximal balance ever seen* throughout an owner's lifetime. This statistics is summarized in Table 4. In addition, it can be seen that there are only 78 owners with current balance larger than 10,000 BTC's. This number grows to 3,812 when looking at the maximal balance ever seen.

Another measure that may indicate the level of activity of an owner is the number of transactions he has been involved with. Its distribution is presented in Table 5. It is remarkable that 97% of all owners had fewer than 10 transactions each, while 75 owners use the network very often and are affiliated with at least 5,000 transactions.

We have also calculated the distribution of the size of the transactions in Bitcoin as summarized in Table 6. Again, it is evident that many transactions are very small, and 28% are smaller than 0.1 BTC each. The Bitcoin scheme actually enables sending *micro* transactions, which are of the order of 10^{-8} BTC (this is the smallest fraction into which a BTC can be broken, and is called

Table 3. The distribution of the current (on May 13th 2012) balance of BTC's per owner

Larger or equal to	Smaller than	Number of owners
0	0.01	2,097,245
0.01	0.1	192,931
0.1	10	95,396
10	100	67,579
100	1,000	6,746
1,000	10,000	841
10,000	50,000	71
50,000	100,000	5
100,000	200,000	1
200,000	400,000	1
400,000		0

Table 4. The distribution of the maximal balance of BTC's ever seen per owner

Larger or equal to	Smaller than	Number of owners
0	0.1	547,763
0.1	10	668,247
10	100	945,083
100	1,000	259,142
1,000	10,000	36,769
10,000	50,000	3,513
50,000	100,000	163
100,000	200,000	40
200,000	400,000	26
400,000	500,000	68
500,000		2

Table 5. The distribution of the number of transactions per owner

Larger or equal to	Smaller than	Number of owners
1	2	557,783
2	4	1,615,899
4	10	222,433
10	100	55,875
100	1,000	8,464
1,000	5,000	287
5,000	10,000	35
10,000	100,000	32
100,000	500,000	7
500,000		1

a Satoshi). When we also consider midsize amounts, we see that 73% of the transactions involve fewer than 10 BTC's. On the other hand, large transactions are rare at Bitcoin: there are only 364 transactions larger than 50,000 BTC's. We have carefully inspected all these large transactions and describe our findings in the next section.

Table 6. The distribution of the size of the transactions in the Bitcoin scheme

Larger or equal to	Smaller than	Number of transactions
0	0.001	381,846
0.001	0.1	1,647,087
0.1	1	1,553,766
1	10	1,628,485
10	50	1,071,199
50	100	490,392
100	500	283,152
500	5,000	70,427
5,000	20,000	6,309
20,000	50,000	1,809
50,000		364

It is interesting to investigate the most active owners of Bitcoin, those who have either maximal incoming BTC's or maximal number of transactions. 19 such owners are shown in Table 7 sorted in descending order of the number of accumulated incoming BTC's shown in the third column. The leftmost column associates the owners with letters between A to S out of which three are identified: B is MT.Gox, G is Instawallet and L is Deepbit. Eight additional owners: F, H, J, M, N, O, P, and Q are pointed out in the graph of the largest transactions (Fig. 1) which is presented in the next section. The second column gives the number of addresses merged into each owner. The fourth column presents the number of transactions the owner is involved with.

Table 7 shows that Mt.Gox has the maximal number of addresses, but not the largest accumulated incoming BTC's nor the largest number of transactions. Owner A in the first row of Table 7 owns the next largest number of addresses, about 50% of those of Mt.Gox's, but received 31% more BTC's than Mt.Gox. Deepbit had sent 70% more transactions than Mt.Gox. It is interesting to realize that the number of addresses of 13 of these owners is a fifth or more of the number of transactions they have executed, which may indicate that each address indeed serves just for a few transactions. It is also clear that six out of the 19 owners in the table have each sent fewer than 30 transactions with a total volume of more than 400,000 BTC's. Since these owners were using large transactions, we were able to isolate them and to follow the flow of their transactions, see Section 4 below. On the other hand, owner A has never sent any large transactions and thus has not been included in our graph of the largest transactions.

Table 7. The list of most active owners in Bitcoin, which have either maximal incoming BTC's or maximal number of transactions. Some of the letters in the leftmost column: F, H, J, M, N, O, P and Q refer to the red letters in Fig. 1 pointing these owners out.

Owner ID	Number of Addresses	Accumulated Incoming BTC's	Number of Transactions
A	78,251	2,886,650	246,012
B (Mt.Gox)	156,722	2,206,170	477,526
C	13,289	941,013	77,525
D	12,520	867,996	48,347
E	191	692,864	1,353
F	12	660,000	23
G (Instawallet)	23,649	633,606	92,593
H	9	580,000	59
I	10,561	514,066	49,550
J	4	500,021	6
K	134	479,254	1,039
L (Deepbit)	2	452,929	814,044
M	9	442,000	10
N	128	432,161	137
O	10	432,286	14
P	1	432,078	3
Q	14	430,490	23
R	2,124	321,866	300,486
S	1,037	20,308	197,334

4 The Graph of the Largest Transactions in Bitcoin

We have identified and analyzed all the largest ($\geq 50,000$ BTC's) transactions in the Bitcoin system, (there were 364 such transactions as described in the last column of Table 6), and followed their flow. We started with the *earliest* such large transaction, the one of 90,000 BTC's made on November 8^{th} 2010. By tracing each of the other 363 large transactions in this category, we were able to show that 348 were actual successors of this initial transaction. The resulting directed graph is depicted in Fig. 1. This graph reveals several characteristic behaviors of the flow in the Bitcoin transaction graph: long consecutive chains of transactions, fork-merge patterns that may include self loops, setting aside BTC's and final distribution of large sums via a binary tree-like structure.

Long Chains. A common prominent practice of Bitcoin owners is to create chains of consecutive transactions as can be seen in Fig. 7: An initial amount of 50,000 BTC's is rapidly transferred from one address to another leaving out some small amounts. In this example 350 such transactions are carried out within the first two days during which the initial amount of 50,000 BTC's is reduced to 34,000 BTC's. In the next three weeks an additional 100 transactions follow and the amount is further reduced to merely 15,000 BTC's. A similar chain of length 120, with initial amount of 500,000 BTC's which decreases to 340,000 BTC's at the end of the chain, is shown in Fig. 1. Note that some of the transactions in this chain are carried out by Mt.Gox. Additional such chains can be found in Fig. 2, Fig. 3, Fig. 4 and Fig. 5, with lengths of 3, 15, 23, 26, 80 and 88 transactions.

Fork-Merge Patterns and Self Loops. Another frequent scenario in Bitcoin is transferring a large number of BTC's from one address to another via several intermediate addresses, each receiving part of the entire amount and then sending it, mostly in full, to the same destination whether directly or via other mediators. Examples can be seen in Fig. 6, Fig. 8 and Fig. 9. A harder to follow fork-merge pattern is presented in Fig. 5: An owner is sending 90,000 BTC's to himself three times in self loops. Each time he splits it into different amounts, 76+14, 72+18 and 69+21. He uses the same address for the small amounts and different addresses for the large amounts. Then he exchanges the entire 90,000 BTC's at Mt.Gox. Finally, the 90,000 BTC's are being transferred via a chain of 90 transactions using 90 different addresses (which may or may not belong to the same owner), where at each a 1,000 BTC's are sent back to the first owner, recombined into essentially the very first amount of 90,000 BTC's.

Keeping Bitcoins in "Saving Accounts". Another long chain of transactions from the beginning of March 2011 can be seen in Fig. 3. This chain is different from the above ones, since at 28 out of its 30 steps, it puts aside 5,000 BTC's in what seems to be "saving accounts". The accumulated sum of 140,000 BTC's has never been sent since. These Bitcoins are an example of our discovery that 78% of all Bitcoins are not circulating.

Binary Tree-Like Distributions. Often amounts of BTC's are distributed among many addresses by splitting it into two similar amounts at each step. This results in a binary tree-like structure as depicted in Fig. 10 and in Fig. 4.

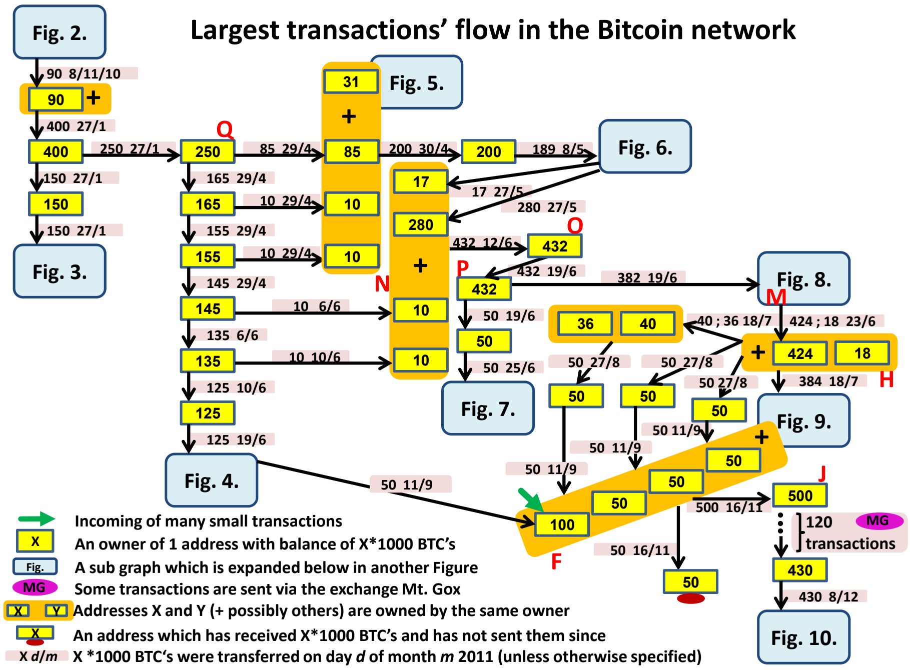

Fig. 1. The backbone of the graph of almost all largest transactions in the Bitcoin scheme (those which are larger than 50,000 BTC's). The red letters refer to some of the most active owners in Bitcoin as listed in Table 7.

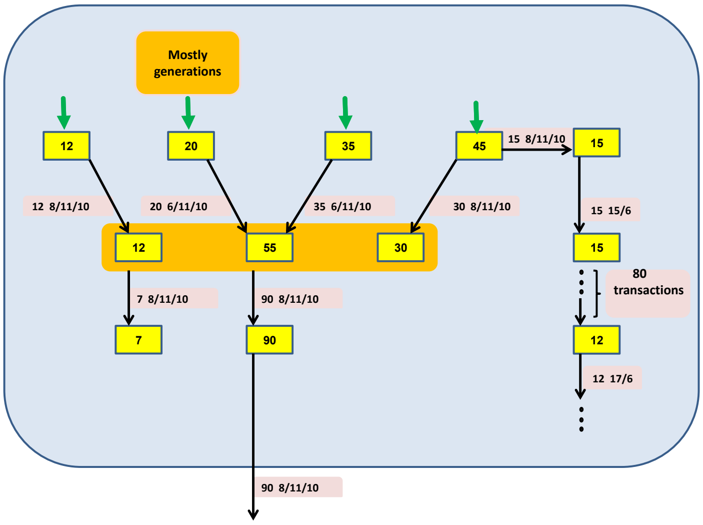

Fig. 2. A Sub graph of Fig. 1: A trace back of some flows of BTC's leading to the first large transaction of 90,000 BTC's on November 8^{th} 2010.

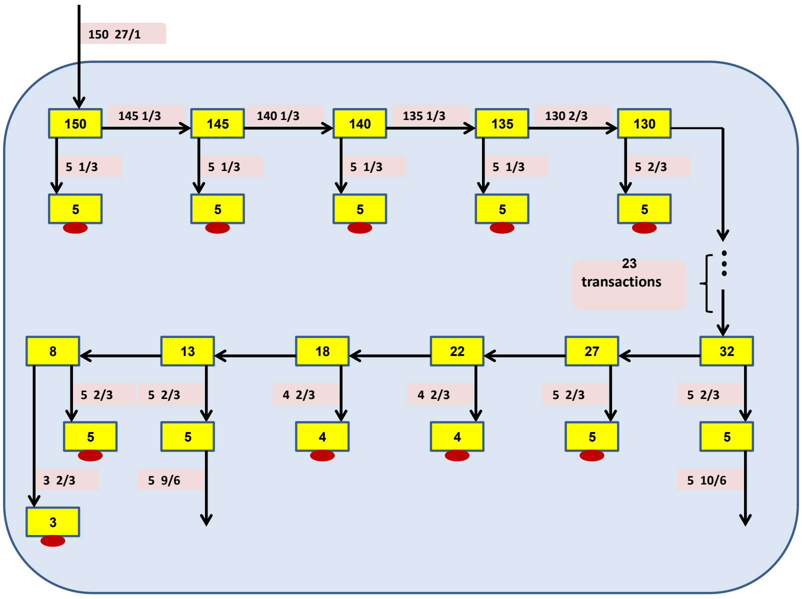

Fig. 3. A Sub graph of Fig. 1: A long chain of transactions where each address puts aside a small amount of BTC's. Those amounts sum up to 140,000 BTC's.

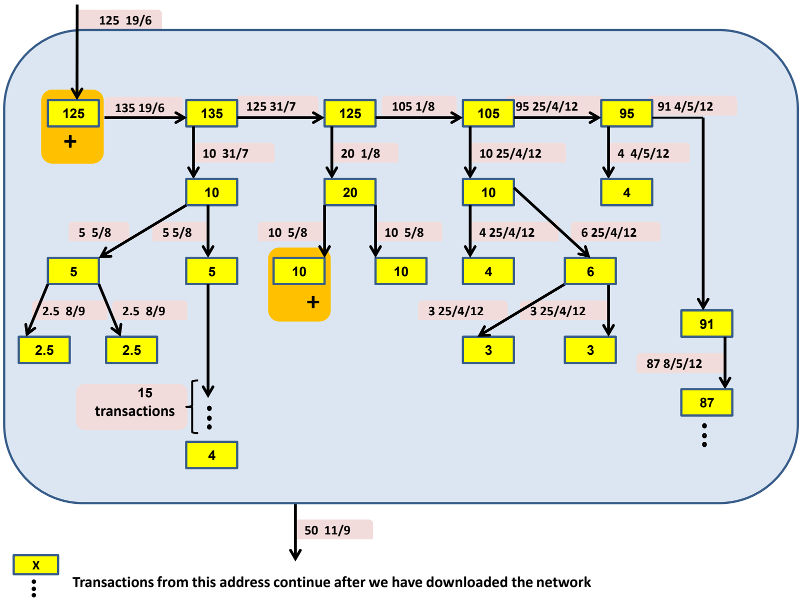

Fig. 4. A Sub graph of Fig. 1: A long chain of transactions where each address transfers most of its BTC's forward. The rest is distributed in a binary tree-like structure.

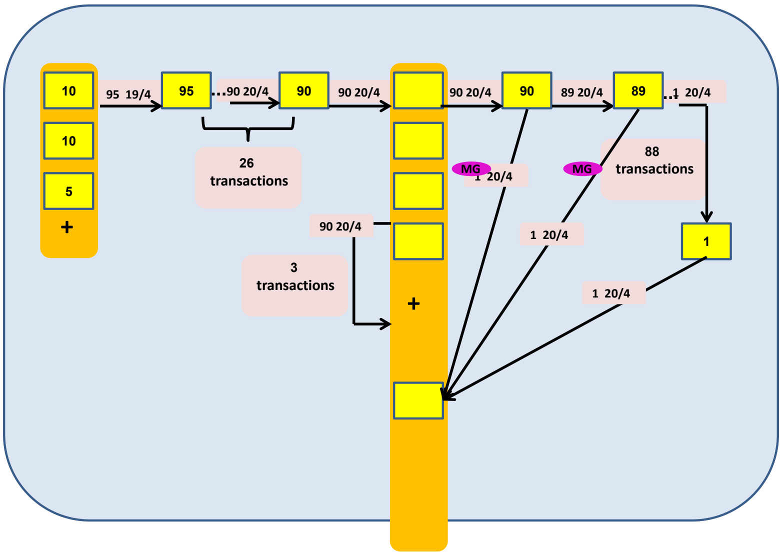

Fig. 5. A Sub graph of Fig. 1: An owner is sending 90,000 BTC's to himself in a self loop, then transfers it forward but gets it back via 90 transfers of 1,000 BTC's each, all carried out on the same day. 31,000 of it is then transferred forward.

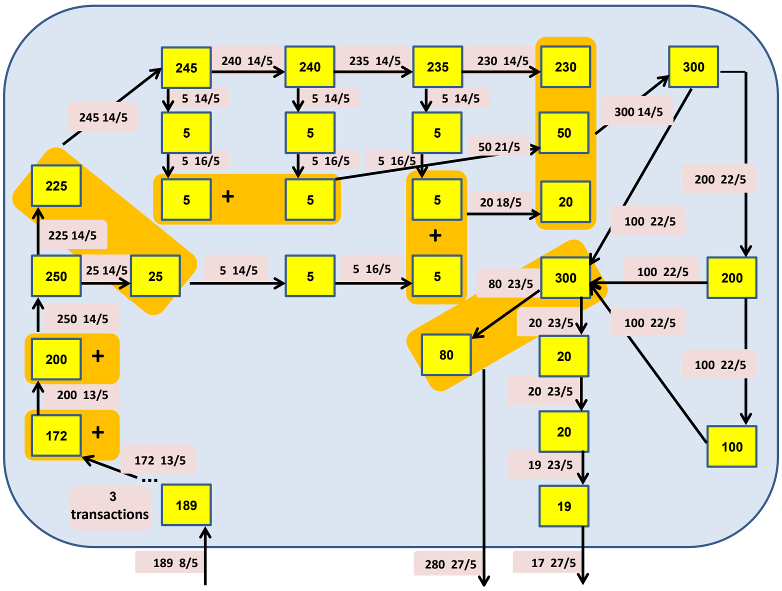

Fig. 6. A Sub graph of Fig. 1: Large amounts of BTC's are transferred from one address to another by sending parts of it to intermediate addresses, which are then being merged into the same destination.

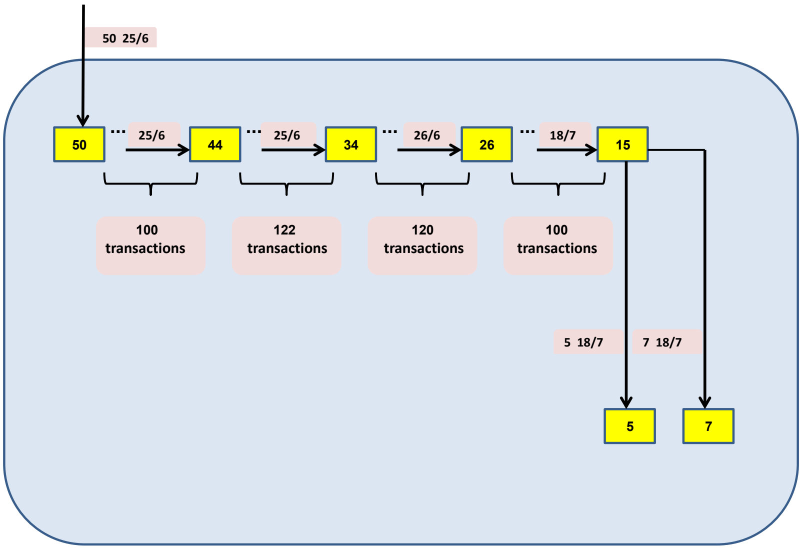

Fig. 7. A Sub graph of Fig. 1: Large amounts of BTC's are rapidly transferred in a very long chain of hundreds of transactions in a very short period of time.

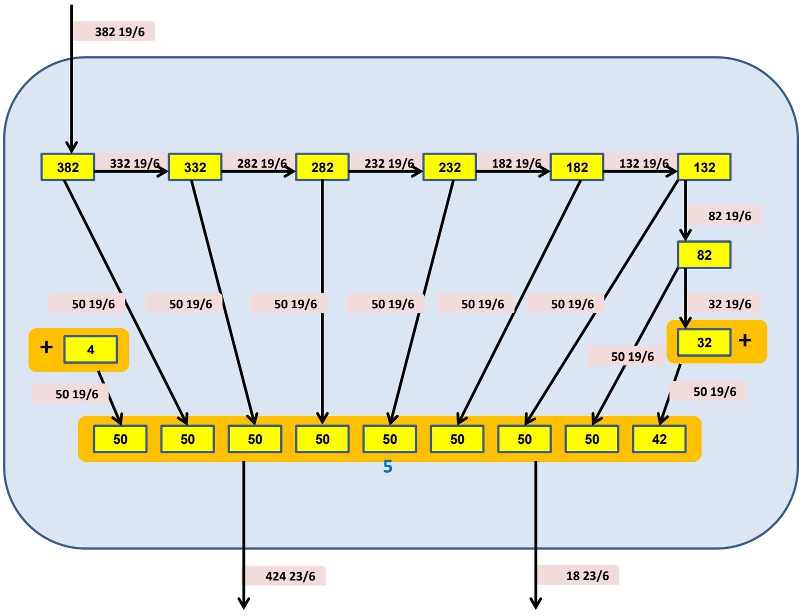

Fig. 8. A Sub graph of Fig. 1: A very large amount of BTC's is transferred by splitting it into equal amounts each directed to a different address belonging to the same owner, then most of the accumulated sums are transferred to a single receiver.

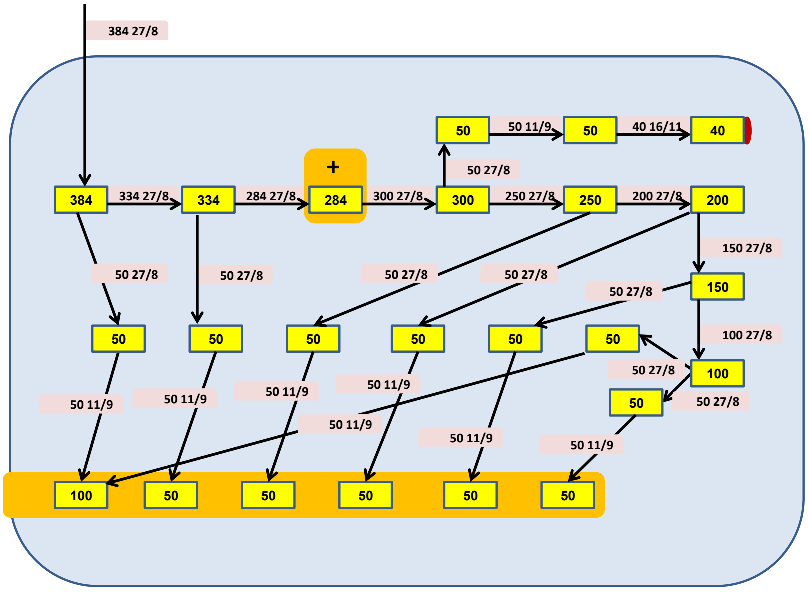

Fig. 9. A Sub graph of Fig. 1: A similar scenario as described in Fig. 8 but with more intermediate addresses.

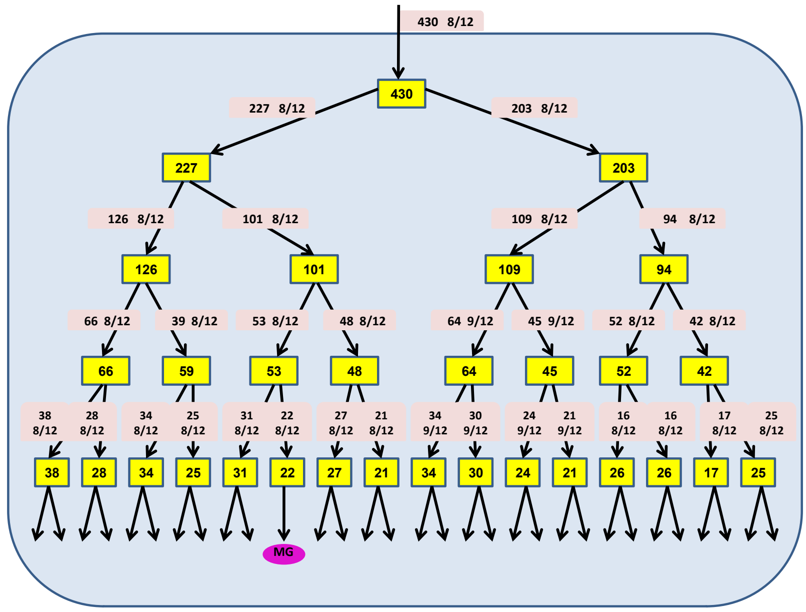

Fig. 10. A Sub graph of Fig. 1: The largest amount of transferred BTC's is finally distributed among many addresses via a binary tree-like structure.

5 Conclusions

The Bitcoin system is the best known and most widely used alternative payment scheme, but so far it was very difficult to get accurate information about how it is used in practice. In this paper we describe a large number of statistical properties of the Bitcoin transaction graph, which contains all the transactions which were carried out by all the users until May 13^{th} 2012. We discovered that most of the minted Bitcoins remain dormant in addresses which had never participated in any outgoing transactions. We found out that there is a huge number of tiny transactions which move only a small fraction of a single Bitcoin, but there are also hundreds of transactions which move more than 50,000 Bitcoins. We analyzed all these large transactions by following in detail the way these sums were accumulated and the way they were dispersed, and realized that almost all these large transactions were descendants of a single transaction which was carried out in November 2010. Finally, we noted that the subgraph which contains these large transactions along with their neighborhood has many strange looking structures which could be an attempt to conceal the existence and relationship between these transactions, but such an attempt can be foiled by following the money trail in a sufficiently persistent way.

Acknowledgments. This research was supported by the Citi Foundation. We would like to thank Ronen Basri, Uriel Feige, Michal Irani, Robert Krauthgamer, Boaz Nadler, Moni Naor and David Peleg from the Computer Science and Applied Mathematics Department of the Weizmann Institute of Science for many interesting and informative discussions. We would also like to thank Aharon Friedman for his major help in acquiring and processing the Bitcoin data base.

References

1. Nakamoto, S.: Bitcoin: A Peer-to-Peer Electronic Cash System, 2008.
2. Wallace, B.: The Rise and Fall of Bitcoin, Wired Magazine, 23 November 2011, http://www.wired.com/magazine/2011/11/mf_bitcoin/all/
3. NPR Staff: Silk Road: Not Your Father's Amazon.com, 12 June 2011, http://www.npr.org/2011/06/12/137138008/silk-road-not-your-fathers-amazon-com
4. Brett, W.: Senators seek crackdown on "Bitcoin" currency, Reuters, 8 Jun 2011, http://www.reuters.com/article/2011/06/08/us-financial-bitcoins-idUSTRE7573T320110608
5. Reid, F., Harrigan M.: An Analysis of Anonymity in the Bitcoin System, arXiv:1107.4524v2 [physics.soc-ph] 7 May 2012.
6. Hamacher, K., Katzenbeisser, S.: Bitcoin - An Analysis, 29 Dec 2011, http://www.youtube.com/watch?v=hlWyTqL1hFA
7. Bitcoin's block number 0, http://blockexplorer.com/b/0
8. Bitcoin's block number 180,000, http://blockexplorer.com/b/180000
9. Cormen, T.H., Leiserson, C.H., Rivest, R.L., Stein, C.: Introduction to Algorithms, Second Edition. MIT Press and McGrawHill, 2001. Chapter 21: Data structures for Disjoint Sets, pp. 498-524.

10. Forbes: Top 10 Bitcoin Statistics, `http://www.forbes.com/sites/jonmatonis/2012/07/31/top-10-bitcoin-statistics/`
11. Block chain: Bitcoin charts `http://blockchain.info/charts`