Analysis of Bitcoin Transaction Flows to Reveal Usage and Geographic Patterns

Shiva P. Bissessar

Supervisor: Dr. Nicolas T. Courtois

University College London

M.Sc. Information Security

Thesis

2013

Abstract

The crypto-currency Bitcoin, has achieved in a short space of time what no other previously proposed digital monetary system has been able to accomplish in terms of confidence, acceptance and usage. It decentralized nature allows it to facilitate trade and serve as a currency without central management. It is this very characteristic which places Bitcoin at the centre of the cross hairs of attention from various groups including law enforcement, tax evasion authorities, and investigative finance authorities.

This work serves to increase the body of research around methods to uncover Bitcoin usage patterns. In particular, a relative profile of the top geographic areas utilizing Bitcoin is presented as well as trends on dormancy in Bitcoin circulation. Additionally, relevant usage trends are discussed.

Keywords:

Bitcoin, law enforcement, investigative finance, tax evasion, computer forensics, geographic analysis of crime

To Adrika

Table Of Contents

Table Of Figures

Acknowledgements

I wish to thank Dr. Courtois for introducing me to this topic with such infectious passion and enthusiasm as well as for the guidance provided towards the production of this deliverable.

I also wish to thank the active community of forum users who post relevant responses and examples to other users' queries. Much information on the correct use of Python and Beautiful Soup was gained this way.

1 Introduction

Crypto-currency is a term coined to describe the new wave of proposed digital decentralized schemes to replace tradition currencies which employ cryptographic principles within their protocol to emulate requisite properties of said currencies without central management. The pioneer of this movement is Bitcoin which has achieved in a short space of time what no other previously proposed digital monetary system has been able to accomplish in terms of confidence, acceptance and usage.

1.1 Motivation & Goal

While at its core these currencies are products of computer science and cryptography, the implications of widespread adoption of any of these schemes delve into several disciplines including Economics, Finance, Legal, Regulatory, and even Politics. Indeed, Bitcoin is attracting attention from law enforcement, tax evasion and investigative finance authorities around the world at present.

My personal interest in this subject area steams from this multifaceted property, in that it requires knowledge beyond pure bits and bytes and allows the researcher to utilize and develop further, competencies outside the realm of computer science. As a M.Sc. Information Security candidate, I believe an appreciation for these other areas is essential and see in depth comprehension of Bitcoin as expanding my understanding of the surrounding aforementioned disciplines.

It is my hope that I am able to make a contribution to the established methods of researching Bitcoin and utilize said methods to make my own unique findings. I hope that my work in deciphering geographic and usage patterns from Bitcoin transaction data can be influential from an Information Security perspective in the undertakings of the previously mentioned groups who have a current heightened interest in Bitcoin and crypto-currencies.

1.2 Organization

Chapter 2 brings the reader up to date with essential economic, financial, legal and technical information on Bitcoin as relevant to the area of usage and geographic patterns. Chapter 3 gives details of the various methods, tools and processes utilized to collect and analyse Bitcoin transaction data. Chapter 4 presents the major findings from the analysis and offers initial thought on these findings when considered in isolation. Chapter 5 discusses the major findings presented in Chapter 4 in the wider context of other findings and other works presented in Chapter 2. Chapter 6 summarizes the major outcomes of this work, provides a self-evaluation and lays the foundation for possible future works. Chapter 7 contains a listing of referenced works and Chapter 8 serves as appendices for any important but non-essential information and storage media containing electronic copies of referenced material, data files and tables from which finding were made and any created program code.

2 Literature Review

2.1 Is Bitcoin Money?

Bitcoin is a decentralized virtual currency proposed in October 2008 by an individual or group of individuals using the pseudonym[1] Satoshi Nakamoto [38] which had its first transaction in January 2009. Since this time Bitcoin has been steadily increasing in popularity and has enjoyed a somewhat volatile ride in terms of its US dollar exchange rate, moving from around $20 in February 2013, to a high in excess of $250 in April 2013, to a current rate of around $118 in late August 2013 [30]. The meteoric rise in price earlier in this year has been circumstantially attributed to speculative interest in the currency by citizens of troubled EU economies as an alternative to their traditionally safe bank accounts. This perception was precipitated by the EU backed Cypriot proposal to seize percentages of ordinary citizens' bank savings as a condition for EU bailout of Cyprus [16]. If this is true it may indicate that Bitcoin was deemed to be a potential "store of value" which is a necessary characteristic of any form of money, as discussed further in §2.1.2. The attraction towards Bitcoin would have been the fact that unlike a regulated monetary system where a central bank would have been able to instruct commercial banks to seize percentages of the account holder's individual accounts, the Bitcoin system is comparatively largely unregulated with a unique absence of central authority to dictate matters pertaining to monetary supply.

2.1.1 Historical Context

Money evolved over time from objects of intrinsic value e.g. salt or cattle to more portable means of trading value such as commodity backed items. Gold certificates, for example, were used to represent some underlying value in gold which was redeemable upon demand by the bearer. Finally, fiat money became the latest physical incarnation of money; however these were no longer intrinsically representative of value of a commodity such as

[1] Gavin Andresen, current Chief Scientist at the Bitcoin Foundation, who assumed the role of Bitcoin spokesperson when Nakamoto went into seclusion, believes Nakamoto to be a single anonymous individual[57]. However, Security Researcher, Dan Kaminsky believes Satoshi to be a group of a group of 3-8 "quants" (finance and programming experts who design financial products) [46].

gold but was instead backed by the Government of a country as legal tender. Fiat money is therefore no longer pegged to some limited resource and allows for central authorities to print fiat money at will to expand the monetary supply based on the burden of future debt. The value in fiat money therefore comes from the trust that people have in it as a currency and their Government's ability to maintain value within the unit of exchange.

2.1.2 Essential Characteristics Of Money

This leads to the characteristic functions of money which have persisted over time though its various embodiments to today, that is, money must serve as; (i) a store of value, (ii) medium of exchange, and (iii) a unit of account. If one was to follow the principle of fiat money to its logical conclusion, in essence, one only needs to have a belief in a system for it to function as money, be it via trust in Government or otherwise. The need for intrinsic value is therefore eliminated which is the crux of the economic school of thought known as Chartalism and its modern day counterpart known as Modern Monetary Theory (MMT) [14]. In some regards, Bitcoin seems to follow this model as it is not backed by any Government; however the computational complexity used to create the currency is seen by others as the inherent value within the systems, as illustrated in §2.5.2.

2.1.3 Adoption And Acceptance

The April 2013 spike and equally drastic correction in the price of Bitcoin is easily decipherable in Figure 1 and unquestionably driven by demand, with the rational of alternate store of value put forward in §2.1. Bitcoin has also been establishing itself as a means for facilitating transactions with an increasing rate of businesses adopting Bitcoin as a form of payment, thereby satisfying the second characteristic of money as a medium of exchange. Indeed, the utility of Bitcoin as a medium of exchange can be seen to be greater than its usefulness as a store of value, with one security expert labelling it as analogous to cash being "teleported" between peers without the need for intermediaries [34]. One of the more noteworthy and legitimate organizations to adopt Bitcoin as a form of payment has been the web publishing service WordPress which announced in November 2012 that it

would begin accepting BTCs as a form of Payment [53][60]. Interestingly, WordPress ascribes their adoption of Bitcoin in accordance with their mission in the desire of "making publishing democratic", citing the fact that PayPal[2] and other similar network payment systems deliberately block or have restrictions on several countries thereby limiting the organization's ability to transact with individual bloggers from such countries [41]. In making such a decision to augment their payment channels via Bitcoin and thereby subvert inherent restrictions by traditional payment systems against certain nations, WordPress has effectively joined the ranks of monetary and financial reformist who embrace the decentralized nature of the Bitcoin system and seek to utilize said system to sidestep the burden of central oversight.

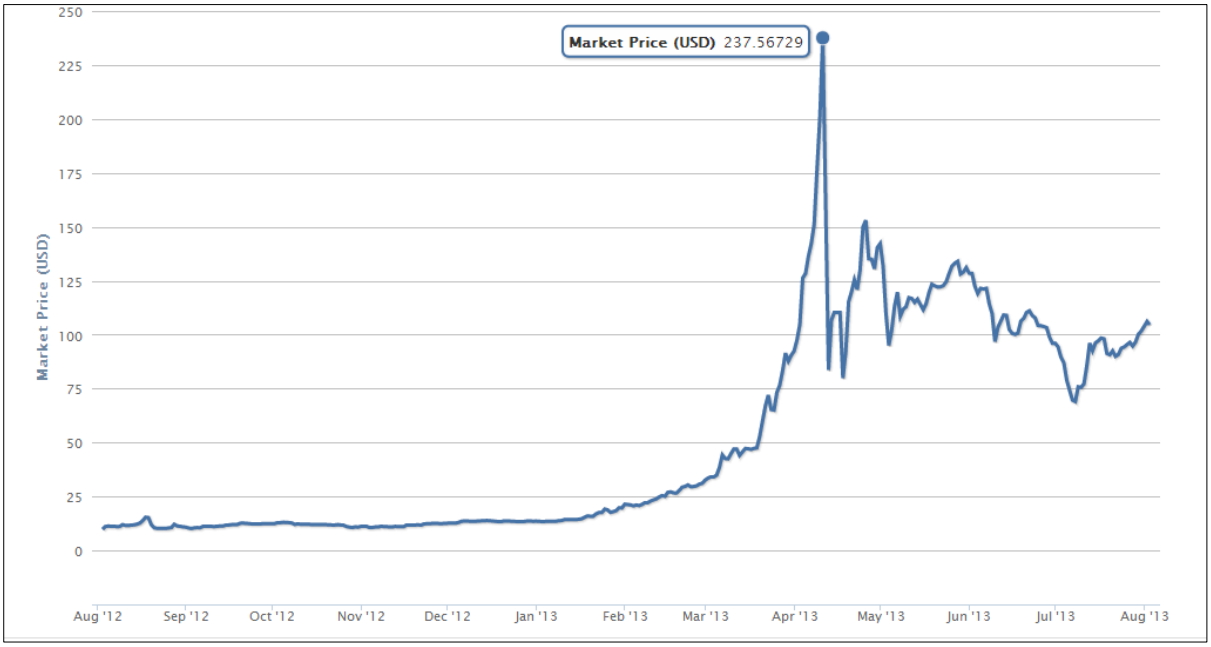


Figure 1: Bitcoin Market Price (USD) [13]

2.2 Virtual Currencies

Bitcoin, the first of the cryptographic based virtual currencies, is enjoying unprecedented popularity and acceptance by various groups with vastly different interests due to inherent

[2] PayPal itself is not to be considered as a virtual currency as it funded by credit transfers from a bank account making the funds within a PayPal account akin to electronic money as per definition by European Central Bank.

properties of Bitcoin itself. Indeed, these properties have also placed Bitcoin on the radar of the Federal Bureau of Investigations (FBI) who in April 2012 released an Intelligence Assessment paper citing that the decentralized peer-peer nature of Bitcoin in addition to the way the currency is created and distributed made it "distinctively susceptible to illicit money transfers, and manipulation through the use of malware and botnets"[3] [23]. However, aside from those interested in the possibility of facilitating payments for illicit trade; fiscal and monetary reformists, cryptography based currency enthusiast and even capitalist interest in the form of fund manager and individual financial investors are counted within the groups of individuals who would like to see Bitcoin succeed as a viable alternate currency. In October 2012, the European Central Bank (ECB) released a paper on virtual currencies where it noted three type classifications of virtual currencies with Bitcoin being cited as an example of their Type 3 classification. Within this class, "virtual money" has a bidirectional exchange with "real economy money" as well as has the property of being able to purchase both real life and virtual good and service [21]. Virtual currencies can also be considered as competitors with traditional currencies such as the Euro or US dollar.

2.2.1 Digital Divide

The aforementioned ECB paper establishes that there are two forms of "digital currency" namely "electronic money" and virtual money with the distinction between the two being based on the previously mentioned third characteristic of money i.e. the unit of account. In the case of electronic money, a legal link with "real economy money" is maintained and the unit of account is preserved in whatever currency the real economy money was initially in. The value converted from real economy money to electronic money is guaranteed to be equivalent. However, in the case of virtual currencies, there is an exchange which takes place between units of account whereby real economy money is used to purchase virtual currency and vice versa. Hence in the period between exchanging from one form to another and then back again, the value stored and subsequently retrieved is subject to the prevailing

[3] The report notes that WebMoney is a more established virtual currency for cyber criminals to continue utilizing in the near term.

exchange rate for the virtual currency. This means the retrieved value may be higher or lower than the initial purchase value.

2.2.2 For What It's Worth

In a mature fiat currency such the US dollar the wild fluctuations seen in the price of Bitcoin in 2013 would have been undesirable as the US dollar is seen as a unit of account for many transactions. For example, people are paid monetary compensation for their hours of work in accordance with prevailing rates of value for time with an understanding of the stability of the unit of account of the currency. While Bitcoin does possess the unit of account characteristic [21] and merchandise can be bought and sold using Bitcoin, the amount of Bitcoin charged for a good or service is normally pegged to the actual real economy currency equivalent for that good or service. Hence, if a pizza costs £5, you pay the Bitcoin equivalent of £5 for the pizza regardless of the exchange rate at that particular point in time.

2.3 Decentralization

2.3.1 Fixed Supply

In a traditional paper currency model the central bank assumes management of the monetary supply and the physical nature of the currency prevents double spending of the same unit of currency. With Bitcoin, the total number of units of electronic coins (BTCs) is fixed at 21,000,000 BTCs which gives it a non-inflationary characteristic distinct from traditional currencies and makes it more akin with the previous gold standard [21]. An integral aspect of central bank operations is the responsibility to manage the amount of money circulating within the system via monetary policy measures designed to reduce or increase the amount of liquidity within the system. With the absence of any notion of a central authority as within the Bitcoin ecosystem, the monetary supply can only increase towards a fixed ceiling at an almost predictable rate and it is left up to the free market principle to regulate demand and supply of the currency already in circulation. However

others have countered that a deflationary currency is not desirable as there would be shortages of the currency where demand would outstrip supply and lead to hording of currency rather than spending which brings commerce to a halt [12].

In 2011, at the Chaos Communication Congress conference, security researchers Hamacher and Katzenbeisser, delivered a presentation where they predicted that the total number of BTC in circulation must fall to zero at some point in the future due to BTC becoming lost[4] over time [10]. Examples of such loss occurrences include loss due to circumstances such as hard drive crashes or death of individuals without passing on the password to Bitcoin wallet. They argue that due to the fixed supply of Bitcoin, all BTCs will eventually vanish, the rate at which this happens simply being a function of probability. They suggest that to cater for such an eventuality the supply may have to be increased, however they caution that such an event would require some form of central oversight and management as well as require changes to the protocol to allow for the creation of new BTCs beyond the current fixed ceiling of 21,000,000 BTC.

2.3.2 Essential Intermediaries And Regulations

In the absence of the central management, currency exchanges play a vital role in the Bitcoin ecosystem facilitating exchange of real economy money to Bitcoin and vice versa. The largest of these exchanges is Mt. Gox which began operation in July 2010 and handles approximately 54% of all Bitcoin exchange volume although their share of the exchange market has dropped from as high as 80% in 2011 [29] [62]. The points at which sovereign currencies are traded for BTCs or where mining takes place are obvious control points for any regulations to be levied upon the Bitcoin ecosystem [47]. In March 2013, the US based Financial Crimes Enforcement Network (FinCEN) defined virtual currencies as compared to "real" currency and attempted to clarify the role of the various parties who participate within various types of virtual currency systems [63]. Their classifications sought to distinguish between "users", "administrators" and "exchangers" in such systems and revealed three types of virtual currency systems including the "De-Centralized Virtual Currency" type, which is the most logical classification for Bitcoin within their guidance

[4] Effectively permanently removing them from circulation within the system

paper. Decentralized Virtual Currency systems are defined as being absent of administrators and in the context of existing entities which they regulate as Money Sender Business (MSB) they provide the following definitions [63]:

> *"A user of virtual currency is not an MSB under FinCEN's regulations and therefore is not subject to MSB registration, reporting, and recordkeeping regulations. However, an administrator or exchanger is an MSB under FinCEN's regulations, specifically, a money transmitter, unless a limitation to or exemption from the definition applies to the person. An administrator or exchanger is not a provider or seller of prepaid access, or a dealer in foreign exchange, under FinCEN's regulations."*
>
> *FinCEN*

They further distinguish between persons who would engage in the production and usage of De-Centralized Virtual Currency for personal use, as falling under the classification of "user", as opposed to persons who would engage in the exchange or transmission of BTCs, who they classify as money transmitters. Money transmitters are considered to be a subset of MSB. In the context of Bitcoin this means that Mt. Gox and similar exchanges would be classified as money transmitters and would be subject to appropriate MSB regulations while those using Bitcoin for personal consumer transactions should be classified as regular users not falling under the remit of their regulations. Their classifications also seem to encompass person(s) who produce BTCs for the purpose of reselling for real currency as money transmitters. Some feel that the definitions are too vague in terms of addressing how small businesses view themselves.

2.3.3 Following The Rules

These FinCEN guidelines impact the exchange entities within the Bitcoin system which affects overall bidirectional flow between real economy currency and Bitcoin and ultimately overall adoption of Bitcoin as a viable alternative currency. The FBI Intelligence assessment and FinCEN guidelines indicate that virtual currencies are receiving a greater amount of attention in particular with respect to how they are used for illicit activities. In May 2013,

The US Department of Treasury and FinCEN announced the shutdown of the organization known as Liberty Reserve, in what was described as the "largest international money laundering prosecution in history" citing proceeds of laundering activities totalling $6,000,000 US [58][50]. It is alleged that Liberty Reserve was specifically established to facilitate the illegal transfer of money between parties engaged in various illicit trades via the use of a digital currency known as the LR. It is important to recognize that the intentional use of the LR within the closed Liberty Reserve system for illegal purposes sets it apart from Bitcoin, however this incident has pushed Bitcoin into the spotlight as a possible replacement virtual currency of choice for those seeking to participate in illegal activities while others, as previously mentioned in §2.2, believe that WebMoney may be increasingly used to facilitate these type transactions [64]. Indeed Mt. Gox has attempted to distinguish its operations from illegal activities by issuing a statement after the takedown of Liberty Reserve which indicated that they would now require identification verification for user account activities [37]. This could be seen as part of their attempt to achieve operational compliance, the need for which was highlighted earlier in May 2103 when FinCEN seized the US accounts of one of their subsidiaries due to the unlicensed activities which would normally be associated with a money transmitter. Mt. Gox has since filled the appropriate paper work with FinCEN to register its subsidiary as a money transmitter.

2.4 Yes It Is!

In August 2013, a US judge presiding over a case brought against Trendon T. Shavers, accused of running a Ponzi scheme under the guise of Bitcoin Savings and Trust ("BTCST"), ruled that, contrary to the defendant's submission that the court had no jurisdiction over the matter as Bitcoin was not a form of currency, Bitcoin was in fact a form of money [35]:

> *"It is clear that Bitcoin can be used as money. It can be used to purchase goods or services, and as Shavers stated, used to pay for individual living expenses. The only limitation of Bitcoin is that it is limited to those places that accept it as currency. However, it can also be exchanged for conventional currencies, such as*

the U.S. dollar, Euro, Yen, and Yuan. Therefore, Bitcoin is a currency or form of money, and investors wishing to invest in BTCST provided an investment of money."

Amos L Mazzant, US Magistrate Judge

This US based legal precedent declaring Bitcoin to be money will undoubtedly have implications to the entire legal landscape evolving around digital currencies, Bitcoin itself and its essential intermediaries. Subsequent recent international headlines from August 2013 illustrate this evolving landscape as; Thailand companies trading in Bitcoin ceased operations while the Bank of Thailand reviews Bitcoin's classification under existing laws [55]; the US State of New York subpoenaed companies trading in Bitcoin as they attempt to understand Bitcoin [22] and most recently the global precedent of Germany recognizing Bitcoin as "Private Money" [26]. In the context of such developments, the need for understanding Bitcoin's geographic trading patterns has assumed greater significance.

2.5 Technical Analysis

2.5.1 Mining

The monetary supply of Bitcoin was previously described as being fixed at a ceiling of 21,000,000 BTC which will be achieved by the year 2140. BTCs are produced or "mined" in accordance with participants' ("miners") ability to solve specific cryptographic hash puzzles upon which a reward of a quantity of freshly mined BTCs is realized. The reward issued for solving one of these puzzles is designed to decrease by 50% approximately every 4 years and currently stands at 25 BTC having gone through the first halving process in November 2012. This means that the cumulative supply of BTC mined will increase at a decreasing geometric rate up to the celling limit and date previously mentioned. A valid solution to the puzzle occurs when the cryptographic hash of an input and some previous puzzle information, yields a string with a certain number of preceding zeros. Hence, every time a puzzle is solved the result contains information from the previous puzzle. The puzzles and

their solutions are presented to the Bitcoin network as “blocks” and as they are related to each other, they form a “block-chain” which is ultimately a vehicle maintaining a public record or ledger of all transactions conducted on the Bitcoin network from since its inception.

2.5.2 Proof of Work

The mining of Bitcoin requires significant computing power to find a solution, however once a solution is found, proving that the solution is valid is a much simpler task. Additionally once a solution is found and the block is accepted by the network, incorporated within that block would be accompanying payment transactions between other random parties on the network which the miner has verified as authentic. The block also contains a hash of the previous block, which forms the linkage within the chain of blocks as depicted in Figure 2. Through such a construction the protocol achieves a proof-of-work characteristic in that each block within the chain builds upon the computational work executed in the past. A proof-of-work is the concept of expending some limited resource to demonstrate authenticity to participate within a scheme. Akin to a stamp on an envelope, the principle has application in the virtual world to prevent denial of service attacks by ensuring participants expend a moderate amount of computational effort to create a token which is then used to allow the sending of an email as detailed in Adam Back’s paper [4]. While others have critiqued the practicality of using such an approach in email systems [33], the Bitcoin system successfully implements the concept to yield the intrinsic value stored within Bitcoin; see §2.1.2. As the chain is representative of all the computation effort which has been expended in the past, the creation of a block therefore serves as a timestamp within the chain as to the legitimacy of the encapsulated verified transactions within the block chain up to that point in time. This concept is also fundamental in establishing the scarcity of the Bitcoin and prevention of double spending [7].

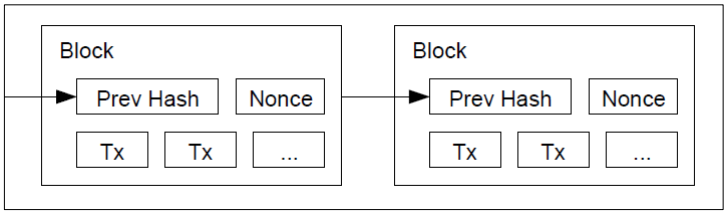


Figure 2: Bitcoin Block [38]

As miners work on solving the block they also verify transactions which constantly update within the block; hence once the solution is found and accepted the new block also legitimizes the transactions which they have verified and for this service they receive transaction fees. The size of this fee is decided by the payer and is the difference between the amount of BTCs from the input side of the transaction and the amount of BTCs which is signed over to other addresses in the output side of the transaction. Transaction fees provide incentive for miners to continue to participate in the Bitcoin network even after its supply limit is achieved and no more BTC can be mined. As shown in Figure 3, a person in control of addresses 1, 2 and 3 use the quantity of BTC stored within them to make a payment to the person in control of address 4. The 'change' from the transaction is specified by the payer and is sent back to him on an automatically generated new address (address 5). The difference between the sum of inputs and the specified outputs is the transaction fee which is assigned to verifier of the transaction (miner) upon solution of a block containing the transaction.

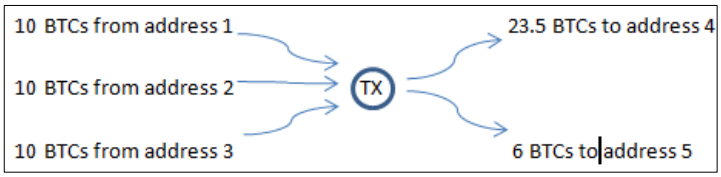


Figure 3: Inputs and Outputs of Transaction

2.5.3 Stability

To solve a block, miners produce a double SHA-256 hash of the previous block address along with a random nonce, which must yield an output less than a 256 bit number known as the "target". If the randomly produced hash from the miner is not less than the target, the client increments the nonce and generates another hash; this process repeats until the client produces a partial collision which satisfies the double hash less than target condition and submits this solution block to the network for inclusion in the block chain [32]. The random occurrence of producing a hash less than the target is made more difficult the lower the target; hence the "difficulty" is given by:

$$Difficulty = Maximum\ target / current\ target.$$

The desired rate at which blocks are solved is specified as 10 minutes, which if attained, means that 2016 block will be created in 2 weeks. Every 2016 blocks the target is changed to take into consideration the time take to produce this last set of blocks; if the blocks have been produced in less than 2 weeks the target is adjusted to make the blocks more difficult to solve and vice versa [20]. For any given difficulty level, the number of hashes to be performed within 10 minutes to solve a block, i.e. hash rate of the network, can be calculated and as of August 2013 the hash rate has cross the 300 Tera-hash marker [39].

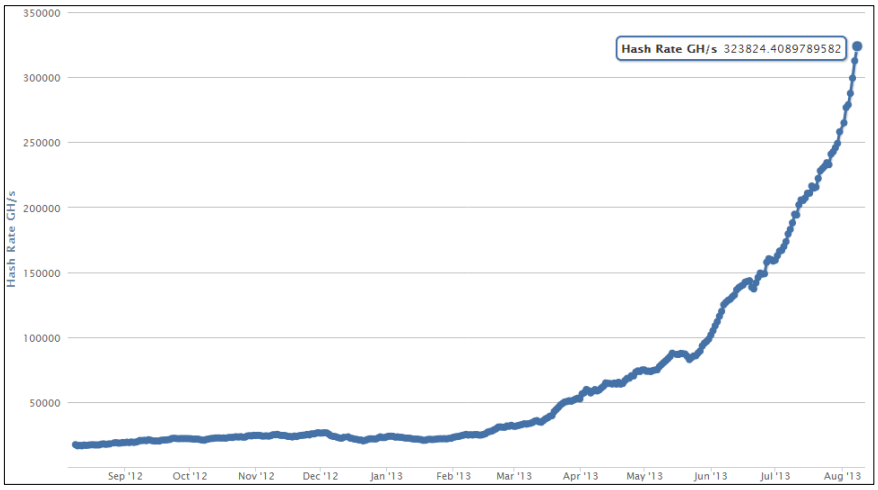


Figure 4: Bitcoin Network Hash Rate [33]

Figure 4 shows a distinct pattern of exponential growth of the hash rate which is attributable to the increase in computational power being dedicated to solving blocks where miners' computer systems have evolved from CPU based systems to GPU based systems to programmable FPGAs and finally to mining dedicated hardware ASICs [19]. Even before the arrival of ASICs, individual CPU were already becoming marginalized in the race for solving blocks due to the sheer computing power elsewhere active on the network. This led to the evolution of mining pools where miners pool computational resources to collectively mine blocks. In the Bitcoin 2013 Conference Security Panel discussion, concerns were raised of a small number of ASIC miners assuming significant network hashing power, followed by the suggestion of the possibility that in the near future the protocol may evolve to cater for multiple proof-of work functions designed to ensure that CPU, GPU and FPGA miners are not completely eradicated by the computational superiority of ASICs [48].

Although Bitcoin has previously been described as decentralized, there are aspects of the system which does require collective consensus of the nodes such as version revisions and acceptance of a solved block as part of the block chain. Nakamoto espoused a philosophy of

the "one-CPU-one-vote proof of works"[5] and the majority of nodes being "honest" thereby reducing the possibility of nodes colluding to form a 51% majority thereby giving them the ability to falsify the block chain [38]. He also stated that if such an eventuality was to occur, such a collective should find it easier to put their computing power towards mining rather than falsifying the block chain. The development of mining pools, where certain groups hold significant percentage of Bitcoin hashing power [48]; the reality of malicious intent as represented by Mt.Gox DDoS attacks April 2013 [36] and increasing malware attacks[6] strengthens the case for the possibility of such an attack whereby the attacker may target certain pools of significant hashing power to remove them from the equation and thereby increase their own percentage of total hashing power [32]. While some have stated that the 51% attack is a real concern others have countered that mining pools currently restrict themselves in terms of their percentage of hashing power to maintain confidence in the system which is essential to its viability [52][59].

2.5.4 Transactions

To spend some BTC, as illustrated in Figure 5, the payer (Owner 1) uses his private key to digitally sign the hash of a previous transaction (from where he received the BTCs) along with the public address of the payee (Owner 2). Owner 2 can verify that Owner 1's private key was used to transfer the BTCs over to him, hence he knows that Owner 1 was in possession of the BTCs in the first place. However, Owner 2 must rely on the verification process carried out by miners to ensure that Owner 1 did not previously transfer those particular BTCs to anyone else, i.e. Owner 1 is not attempting to double spend the BTCs. The miners achieve this by checking

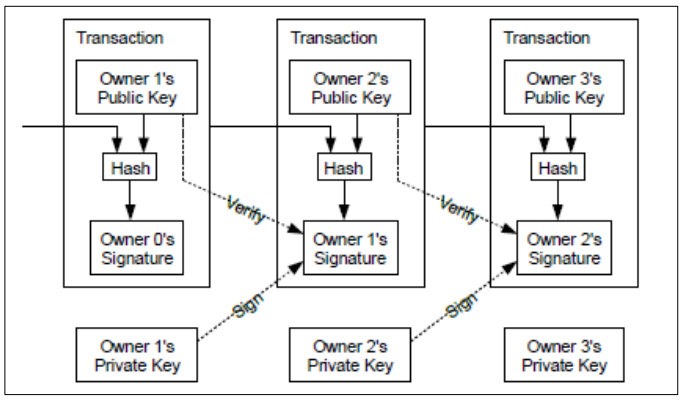


Figure 5: Spending Bitcoin [38]

[5] Though some have argued that the "one CPU one vote" system is no longer relevant to Bitcoin due to the massive computational superiority of ASIC miners [19]

[6] Malware in the Bitcoin space even extends to game developers being implicated in rogue mining software being installed to uninformed client systems [61]

the publicly broadcast block chain to ensure that the BTCs were indeed never previously spent and then verify the transaction which is then included the next block when it is solved. There is therefore an inherent lag in time between a payer signing over his BTCs to a payee and the payee being reasonably assured that the BTCs were never spent before, as the payee awaits at least 10 minutes to confirm if the transaction is included in the next block created.

As there may be more than one possible block solution submitted to the network a fork in block chain may form where one set of peers are working on one block and others are working on a different block[7]. The accepted block becomes the one which has more blocks subsequently chained to it. At the point when there is consensus that the longest path has been found, the shorter path of the fork is dropped altogether and all peers resume working on the block at the end of the longest path. This has implications on the acceptance of payment transactions whereby transaction are only considered to be confirmed after inclusion in 6 blocks but reduces the possibility of double spend transaction.

Each participant on the network will have a store of addresses within their Bitcoin client wallet and each address within the wallet will have a particular amount of BTCs associated with it. These would have been received from; a payee as payment within a transaction or themselves as change from a transactions or it may have come as a "coinbase" transaction which is the result of mining a block.

2.6 Security & Cryptography

A paper published at Financial Cryptography 2012, described Bitcoin in the context of succeeding where other "e-chash" systems have failed despite possessing "…no fancy cryptography…" [5]. However, notable security researcher Dan Kaminski stated he was left "excited" by his failed attempts to hack the "well designed" Bitcoin protocol and made an exuberant assessment of the source code as "fantastic" [32][31][46]. Both these assessment are essentially correct in that, from a security perspective, Bitcoin utilizes well

[7] Miners essentially vote on which block is the legitimate one as they chose to work on a particular one of the contentious blocks

founded cryptographic principles to facilitate the execution of simple and effective way to transfer payments between parties.

2.6.1 Signatures

The digital signing algorithm utilized by Bitcoin is based on Elliptic Curve Cryptography (ECC) using the secp256k1 curve [17]. The security of ECC is based on the difficulty of the Discrete Logarithmic Problem (DLP) or more specifically the Elliptic Curve Discrete Logarithmic Problem (ECDLP). ECDLP is considered to be much more difficult to solve than the integer factorization problem, upon which RSA relies, and the DLP in finite fields, upon which the Diffie-Hellman (DH) key exchange and ElGamal encryption scheme relies. Hence, ECC based encryption schemes, such as Elliptic Curve Digital Signing Algorithm (ECDSA), requires much smaller keys to provide the same level of security than RSA or DH and ElGamal [15]. The reduced key size requirement means that 256 bit ECC schemes are 64 times more computationally efficient than DH [51]. Bitcoin employs ECDSA with a 256 bit key and the increased computational efficiency allows for the use of Bitcoin wallets on device with less computationally power than the traditional PC such as mobile devices.

While ECDSA has been widely used and analysed over a long period of time and is therefore considered to be secure by the Cryptographic community, the actual implementation is where chinks in the armour can be found as proven in 2010, when SONY's implementation of ECDSA on the PS3 gaming console was proven to be insecure and private keys were subsequently derived due to poor pseudorandom number generation (PRNG) for a particular variable [18]. On 11th August 2013, Mike Hearn, a Bitcoin developer posted that Google Android OS implementation of Java SecureRandom class contained vulnerabilities which rendered Android based Bitcoin wallets app using this Android OS class to create key pairs vulnerable to poor PRNG [1][3]. Keys created using such a process would have been vulnerable to having the private key computed based on the knowledge of the Bitcoin public address, thereby placing transactions conducted with these addresses at risk. The bug was quickly fixed but not before at least 56 BTCs was confirmed as being stolen as a result of this vulnerability [2]. While attacks on the implementation of Bitcoin's cryptography are

inevitable it is assumed future ones will be treated in the same manner, i.e. patch the software, remove the bug etc.

2.6.2 Hashing

SHA-256 is used in the Proof-of-Work function in Bitcoin as described in §2.5.2 while RIPEMD160 is used in the construction of Bictoin addresses as illustrated in §8.1. The choice of SHA-256 as the hash function has been criticized as leading to the current situation today where thousands of individual CPU miners have been disenfranchised from the Bitcoin mining process as also previously described in §2.5.2.

2.7 Related Work

Despite the fact of all transactions being publicly broadcast to all participants on the network, privacy of the participants involved in transactions is provided through the use of Bitcoin addresses rather than using the participant real name. The Bitcoin address is constructed from a participant's ECDSA public key which is then hashed using SHA-256 and hashed again using RIPEMD-160 along with some further hashing and conversion steps as detailed in §8.1. This system is akin to being anonymous once there is no uniquely identifying piece of external information linking a user's identity to a Bitcoin address. Despite this quality of transactions being untethered to a person's real identity, researchers have utilized methods to identify "entities" as owners of particular Bitcoin addresses and hence if any of these addresses becomes identifiable with a real life identity (individual or group), said identity could now be linked to a Bitccoin entity and its other addresses and its transactions [44][45][40]. To increase anonymity on the network, participants are encouraged to create and use a new Bitcoin address with each transaction.

2.7.1 Anonymity

In 2011, Reid and Harrigan presented one of the first academic studies into the perceived anonymity of the Bitcoin network with the intention clarifying the fact that Bitcoin was not designed to be anonymous [44]. They examined the Bitcoin block-chain from inception to 12th July 2011 and inferred the presence of entities which controlled multiple Bitcoin addresses. These inferences were based on the transaction graphs they constructed which showed patterns of incoming and outgoing flows of BTCs from various addresses[8]. The flows were such that the individual addresses could be seen to be controlled by single entities. They also showed that associations between these entities and their corresponding real life identities could be made if any of their individual Bitcoin addresses were linked to real life identifying data.

2.7.2 Dormant Bitcoin

In 2012, Ron and Shamir analysed the public record of all Bitcoin transactions up to 13th May 2012 and produced transactional graphs of the data of with the intention of uncovering statistical properties and usage patterns [45]. Within this data set, they established that 609,270 Bitcoin address have received BTCs but have never participated in any outgoing transactions. The sum of BTCs accumulated within these addresses was deemed to be 7,019,100 BTCs or 78% of the total BTC supply within the data set. While stating that they do not claim these BTCs to be "out of circulation", they did note 59.7% of the then total supply could be classified as "old" or "dormant" coin, i.e. coins received by an address without a corresponding outflow up to 3 months before the 13th May 2012 cut-off date. When compensating for the possibility of legitimate "lost" coins in the system and thereby eliminating 1,657,480 BTCs from consideration, the resulting percentages for sum of Bitcoins accumulated within the identified addresses and the dormant coins remained high. They point out:

[8] For instance, at a very simplistic level and not taking into consideration the activity of mixing services; all input addresses in a transaction can be considered to belong to a single entity

"This is strong evidence that the majority of bitcoins are not circulating in the system…"

Ron and Shamir

In 2013, Ober et al[9] published findings based on their analysis of Bitcoin transactions using data up to 6th January 2013 (block 215,399) with a primary focus on investigating "global properties" of the Bitcoin transaction [40]. Similar to the techniques used by Reid and Harrigan, they used methods to identify entities via a "merging event", where previously un-associated public addresses are seen to become involved as inputs to a transaction. They cite such events as being attacks on unlinkability within Bitcoin, as the analysis of the public data reveals patterns between addresses which in turn could imply that a certain entity is in control of those addresses. They also look at dormant coins which have been out of circulation for 1 and 3 months identifying the absence of outgoing transactions from addresses, as a further decrease in the anonymity set of the network. Much like the previous study, their analysis showed a significant number of dormant coins in the system, i.e. approximately 6,300,000 dormant BTCs as of January 2013.

2.7.3 Dormancy, Supply and Geography

These studies stimulate the discussion of dormant coins in the system and the question of what percentage of these coins can be considered as lost? It is widely accepted that prior to the establishment of Mt. Gox, the user base for Bitcoin may have been less appreciative of Bitcoin having real value and hence would have been more lax with their respective supplies [45]. This could have led to Bitcoin wallets being deleted by accident or even intentionally leading to the possibility of a significant number of lost BTCs. On the other hand there is no way to currently know if in fact the majority of dormant coins are truly lost or simply being stockpiled by users akin to a savings account. The concept of lost coins remains intriguing none the less and in the context of the theory of the supply of Bitcoin evaporating over time as described in §2.3.1 the prospect of recovering lost coins and putting them back into circulation as a possible garbage collection and recycling mechanism could be seen as benefiting the community. One possible solution would require that

[9] Continuing the work cited in § 2.3.1, Hamacher and Katzenbeisser co-author this paper

someone deduce the private key associated with public BTC addresses associated with BTCs which have not been on the output side of a transaction for some extended period of time, say 5 years. It would also require the ability to computationally execute such an attack which at this point in time is infeasible.

Another question that arises, especially in the context of international headlines highlighted in §2.4 is that of understanding the geographic distribution of activity on the Bitcoin network. These areas form the focus of this research, which is to; utilize and adapt methods outlined in the related works as well as establish alternative methods to substantiate and/or uncover further relevant and interesting patterns of usage, dormancy and geography.

3 Methodology

3.1 Data collection

It was determined in advance that the analysis required would involve data collection from live websites and/or parsing of several files HTML based file offline, hence Phython was selected as the programming language for any necessary programming due its characteristic ability for parsing data and ease of use. A couple additional libraries were installed; Beautiful Soup [6] to facilitate the parsing of the HTML pages and GeoIP Database [25] to obtain an IP address/country association. Several forums, such as http://stackoverflow.com, were consulted to get further explanations and examples of the proper usage of Python and imported module commands.

Like the research mentioned in §2.7, the record of Bitcoin transactions needed to be examined in order to perform quantitative research and uncover any outstanding patterns. All Bitcoin blocks are made publicly available by Bitcoin Block Explorer [11] and this was downloaded by Dr.Courtois using the application wget [27] which resulted in 7.33GB of data. To sift and sort through this offline data store proved somewhat difficult to work with as blocks and transactions are linked via long strings of either Bitcoin addresses, block headers or their respective cryptographic hashes. Additionally, the raw blocks do not even contain the block number which is generally used to identify blocks. An example of the structure on one of these raw blocks is presented in §8.2. To avoid these problems, it was decided that the data should be collected online from [11] where individual blocks can be accessed directly via their unique block number and transactions from one Bitcoin address to another can be followed sequentially via hyperlinks.

3.1.1 Random Traversal Through The Blocks

The ease with which this website facilitated the navigation to any given block and its composite transactions resulted in the idea to make iterated random traversals within the block chain to trace BTCs all the way from their initial mining, i.e. coinbase transaction, to the point where they remain dormant associated with a Bitcoin address i.e. attain status of

"Not yet redeemed" in the output of a transaction. Such traversals, if performed with sufficient randomness and in sufficient quantity, would therefore be representative of the total set of all Bitcoin transactions. 4 hops of a traversal from block 197053 is shown in Figure 6.as an example

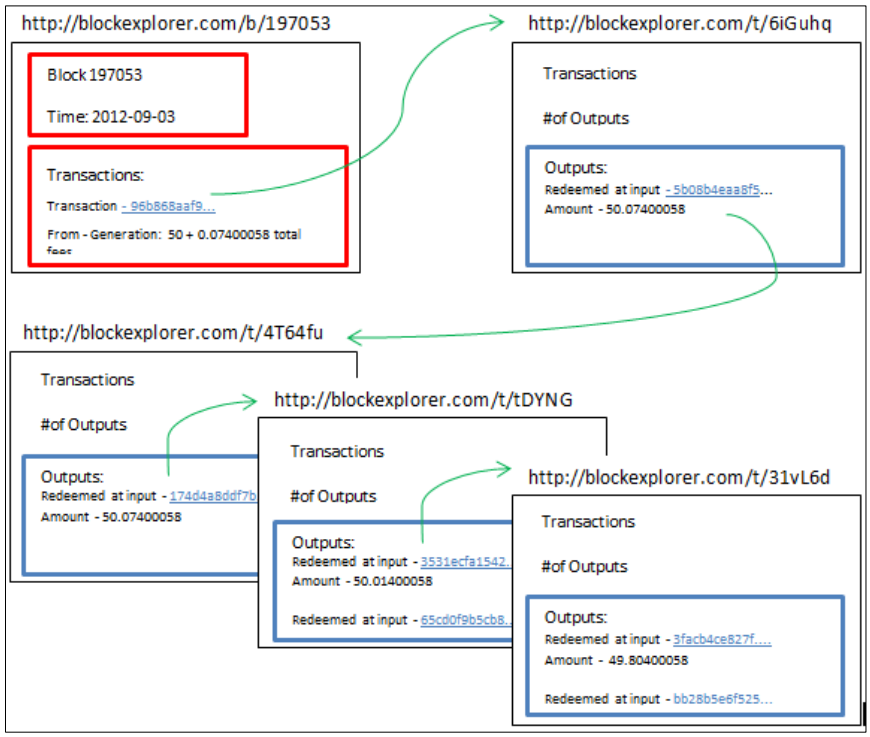


Figure 6: Random traversal of block 197053

To achieve a random traversal through the life of a BTC, the distribution of BTCs in the blocks needed to taken into consideration. Prior to block number 210,000, the coinbase value was 50BTC while every block with number greater than or equal to 210,000 has a coinbase transaction of value 25BTC. Hence, the random selection of blocks needed to take into consideration the proportion of higher value blocks over lower value blocks. This was accomplished by appropriate logic to check the proportion of blocks (which changed every 10 minutes) and direct the random choice of the block within the high value range or low value range accordingly. The piece of code which accomplishes this is presented in Figure 7.

```
#weighted probability of randomly selecting 50 BTC block
        probability50Block = 210000/(currentBlockNum)
        randomThrow = random.random()
        if randomThrow <= probability50Block:
           randomBlock = random.randint(1,209999)
        else:
           randomBlock = random.randint(209999, currentBlockNum)
        print('Randomly chosen  block: ', randomBlock)
```

Figure 7: Random weighted selection of Block

Additionally, once a block is chosen, due to the nature of the transactions in having several outputs, another weighted choice needed to be made as to which output branch to follow. In this case, the amount involved in the respective outputs was used as the weight in selecting which output transaction to follow. The traversal was considered to be concluded when an output transaction was chosen which contained the value "Not redeemed as yet" under the output header "Redeemed at input "rather than containing another transaction hash to follow.

Upon each iteration of the traversal, details at the block level were collected including; terminating block number, time at which BTC became dormant, amount of dormant BTCs and the number of iteration in the traversal. This data was collected for 500 bocks which contained a total of 69,049 transactions. This data can be found in worksheet "500_23rd"of [9] on the Appendix CD media.

3.1.2 Random Traversal Via IP Address

Dr. Courtois also provided offline data from another website, http://blockchain.info, which provided IP address information associated with transactions. The IP address captured in this data is that of the first node to relay the Bitcoin transaction to the website's network of nodes[10]. The dataset represented a concatenated random crawl though all transactions associated with a particular IP address where 83 IP addresses were examined. This dataset represented 500MB of data or 256,293 transactions. From this data the following information was collected for each IP address; country, sum of BTCs transacted and number

[10] The website lists the number of nodes currently connected on its home page, which has been noted several times to be in the range of 800 to 900 nodes, and as low as the 500 range at other times.

of transactions. This data can be found in worksheet “All tx” of [8] on the Appendix CD media. The CD also contains the source code for the Python program created for this task, “ip_data”.

3.1.3 Additional Transactional Data

Upon examination of initial results from the two previous processes, a modified “random traversal through the blocks” was contemplated. This lead to modification of the program to collect individual transactional data inclusive of IP address data from http://blockchain.info and cross reference said IP addresses with the GeoIP database to retrieve country information for each transaction. This greatly increased the time of data collection as well as resulted in a greater number of errors (“…connection was forcibly closed by the remote host”, etc.) which resulted in having to restart the traversals, thus providing a limiting factor to the data collection phase of this work. The addition data collected was useful in augmenting the IP address data collected in §3.1.2 to get a better appreciation of what is revealed by this IP address information. This resulted in the collection of detailed data of 57,573 transactions within another set of 500 randomly selected blocks. This data can be found in worksheet “500_27th ”of [9] on the Appendix CD media. The CD also contains the source code for the Python program created for this task, “transaction2”.

3.2 Analysis

1	Originating Block		went to sleep	Transaction End Point Amount	Number of Hops		Dormancy Begins	Dormant Days
84	231865		2013-04-17T20:17:43	3000	9		17/04/2013	93
85	153507		2011-11-16T05:38:46	40000	39		16/11/2011	463
86	231683		2013-04-16T18:32:38	3000	96		16/04/2013	94
87	163155		2012-01-21T07:20:28	188	11		21/01/2012	415
88	251849		2013-08-12T23:41:22	100	3359		12/08/2013	10
89	247050		2013-07-17T15:43:19	31487.7979	21		17/07/2013	28
90	252175		2013-08-14T20:45:52	9.15049113	269		14/08/2013	8
91	234084		2013-05-01T18:37:36	2650	258		01/05/2013	83
92	225975		2013-03-15T10:13:42	49.69409271	283		15/03/2013	116
93	19081		2009-07-08T22:26:53	50	0		08/07/2009	1078
94	230514		2013-04-09T21:03:36	69471.0822	33		09/04/2013	99

Figure 8: Looking for patterns in the data

The programs produced outputs in .csv file format which were then imported into Microsoft Excel as .xslx files. Within Excel visual inspection of the data along with the conditional formatting features were used to get an understanding of the data and get a first indication of what may be relevant patterns which could be further brought out using data aggregation and graphical analysis techniques; Figure 8. The same data was examined in several different layouts to find the best way aggregate and find such patterns. Excel allowed various statistics to be derived from the data such average number of traversal hops, average value of dormant BTCs and average dormancy age. Particularly interesting transaction flows were further investigated and illustrated using Graphviz which uses the .dot format for creating directed graphs [28][24]. A sample of code for Graphviz showing a portion of the digraph presented in §4.2 is shown in Figure 9.

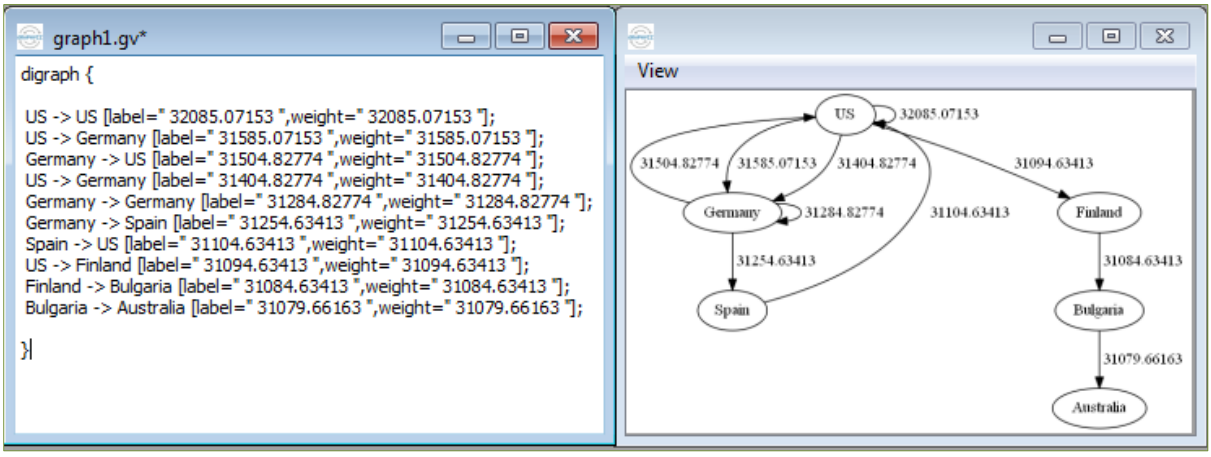


Figure 9: Graphviz sample code and output

It is acknowledged that attempting to detect and decipher patterns in Bitcoin transactions based on the IP address information obtained from http://blockchain.info may be somewhat speculative based on the fact that users who may be very intent on anonymizing their transactions and web presence have various applications available to them to accomplish same such as Tor [54]. It is also noted that any inference of country activity based on IP address information is not necessarily indicative of wide spread usage of Bitcoin within said country. Results ascribing usage to a country will therefore be limited to those where many instances of IP addresses belonging to a country are found. Having said this, it is believed that the results utilizing the IP data will still be illustrative of relative activity between countries.

In analysing the results the dormancy period was considered to be the time from which an address was deemed to have a store of BTCs which were "Not yet redeemed" to a date of 27th August 2013.

3.3 Limitations

The data collection phase had to be reduced due to time constraints. This was as a result of the length of time it took to establish an implementable methodology to collect the required data. The random weighted method of selecting blocks and the path to follow during a traversal, however, makes the reduced set of 1000 transactions truly representative of the full transaction data set.

Once the data was collected, the analysis phase took longer than planned as well. This was due to the time taken to find an appropriate ways to organize the data to reveal relevant and interesting geographic and usage patterns. These delays forced the original scope to become restricted such that an intended area of focus on the future supply of Bitcoin had to be eliminated.

4 Results

4.1 IP Address/Country Data Analysis

The initial analysis of IP address information from the 256,293 transactions related to 83 IP addresses derived by the method outlined in §3.1.2 was promising in that it revealed a pattern of geographic usage with 9 main countries as depicted in Figure 10 (8 other countries with transaction value less than 0.2% were omitted from this plot). The data suggests that USA and Germany lead the other countries both in terms of number and value of Bitcoin transactions. The UK , Russia and France all display a similar degree of high value usage of Bitcoin while Canada, although contributing 6.8% of the transactions analysed, only contributed around 3% of the total value of transactions analysed. Russia's usage of Bitcoin seems to be the exact opposite of Canada's, displaying qualities of a low quantity of high value transactions.

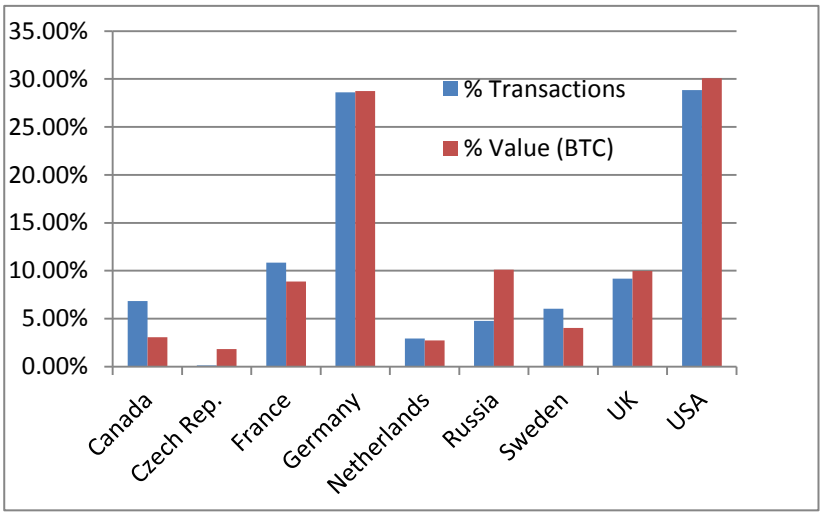


Figure 10: Geographic usage pattern via IP traversal, § 3.1.2

The additional IP address information gained from the method in §3.1.3 where 57,573 transactions were analysed, yielded a higher number of distinct countries than the previous method, 83 in total. The initial analysis showed some similarity with the IP traversal analysis with USA and Germany again being the major identified countries in terms of number and value of transactions. However, the presence of "Unknown" entities or IP addresses corresponding to 0.0.0.0 which could not be linked to any country was uncovered within 10,571 transactions. At 21%, the Unknown classification had a significant contribution to the percentage of total transactions it participated in; that figure making it on par with Germany and 7% lower than the highest contributor United States at 28%. But even more

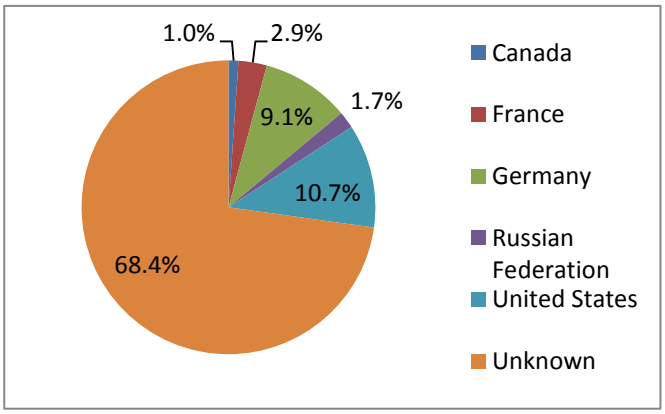


Figure 11: Percentage Transaction Value, § 3.1.3

significantly was the result that 68% of the value within the total transaction sum of was attributable to Unknown as shown in Figure 11. This clearly illustrates that the majority of high value transactions was conducted in such a way which maintained IP address/country anonymity.

Figure 12 further differentiates the two major countries uncovered, the United States and Germany, in terms of number of transactions and the value of the transactions.

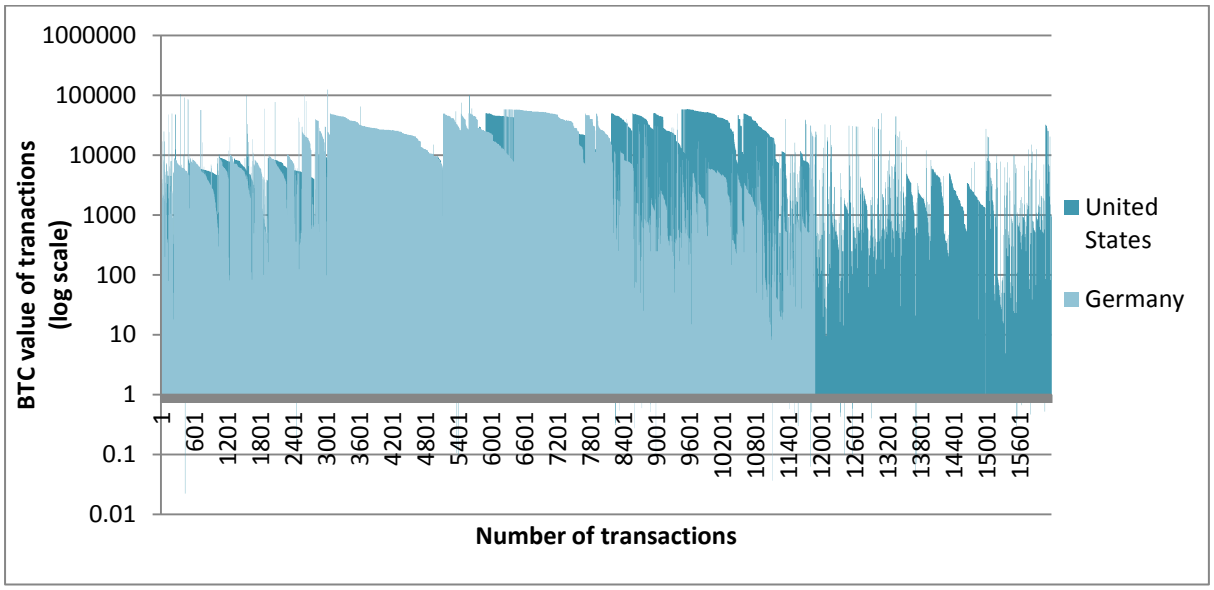


Figure 12: US and Germany comparison

4.2 Complexity Of Individual Transactions

At times, closer inspection of some of the transactions was required to highlight noteworthy characteristics. One such transaction is illustrated in Figure 13.

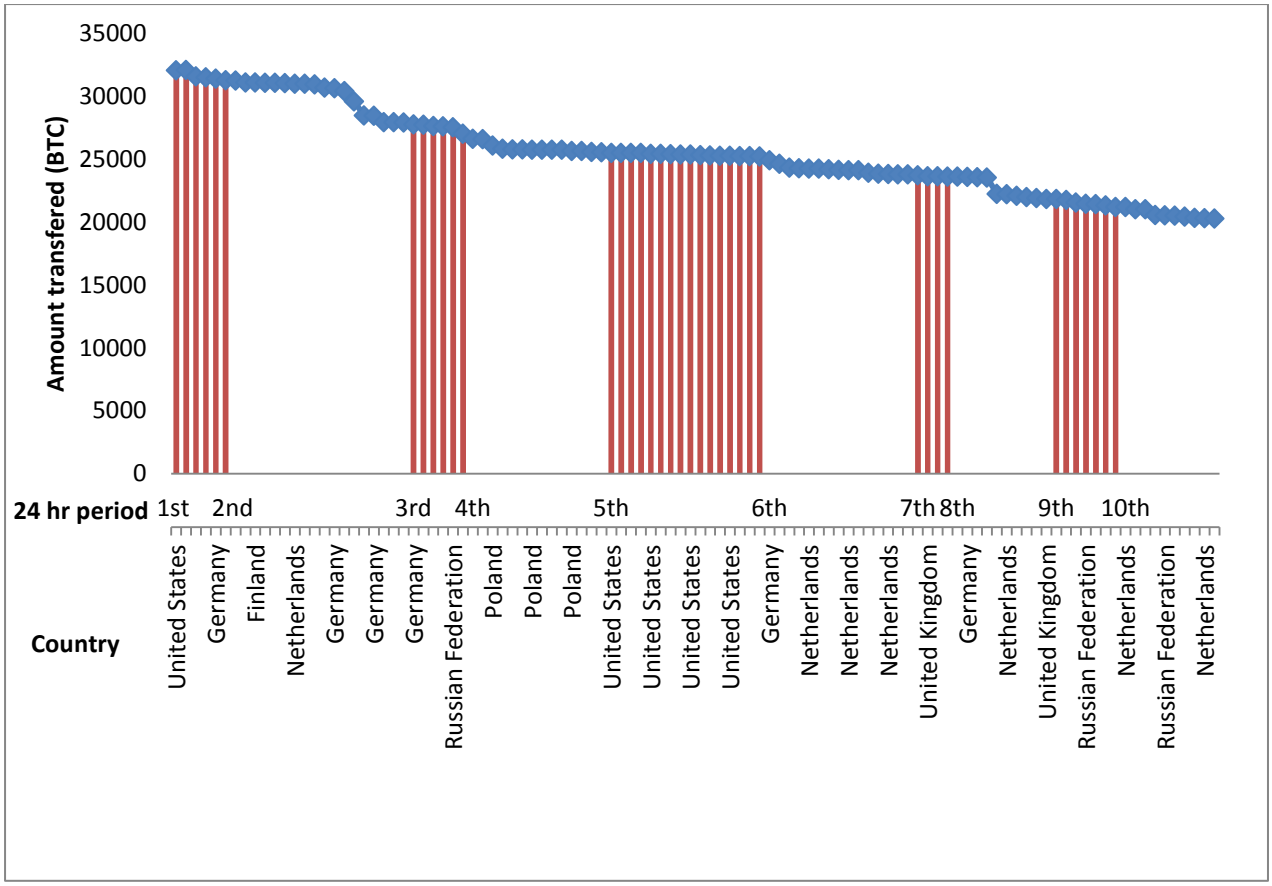


Figure 13: Following a transfer of 32,064 BTC ≈ 3.78M USD

On 12th August 2013 at 12:30:16, 32,064BTC was transferred to an address an over the course of 106 subsequent transactions over period corresponding to 10, 24hr cycles (illustrated above by red/whitespace alternating bands), this initial figure was reduced by 36% in transfers which seemingly spanned international borders. The transactions captured within the traversals are represented in a digraph in Figure 14, which shows the nodes and edges involved in the 106 transactions. The US node is the beginning of the flow while the Netherlands node is the end. The diagram became substantially more complex once the value of the BTCs being transferred was added; hence the values are not shown in this illustration. It is noted that this type of movement of funds is very typical and apparent when examining Bitcoin transactions[11]

[11] This particular chain of transactions actually continues beyond the data available within the trace and can be followed further on the Block Explorer website [56]

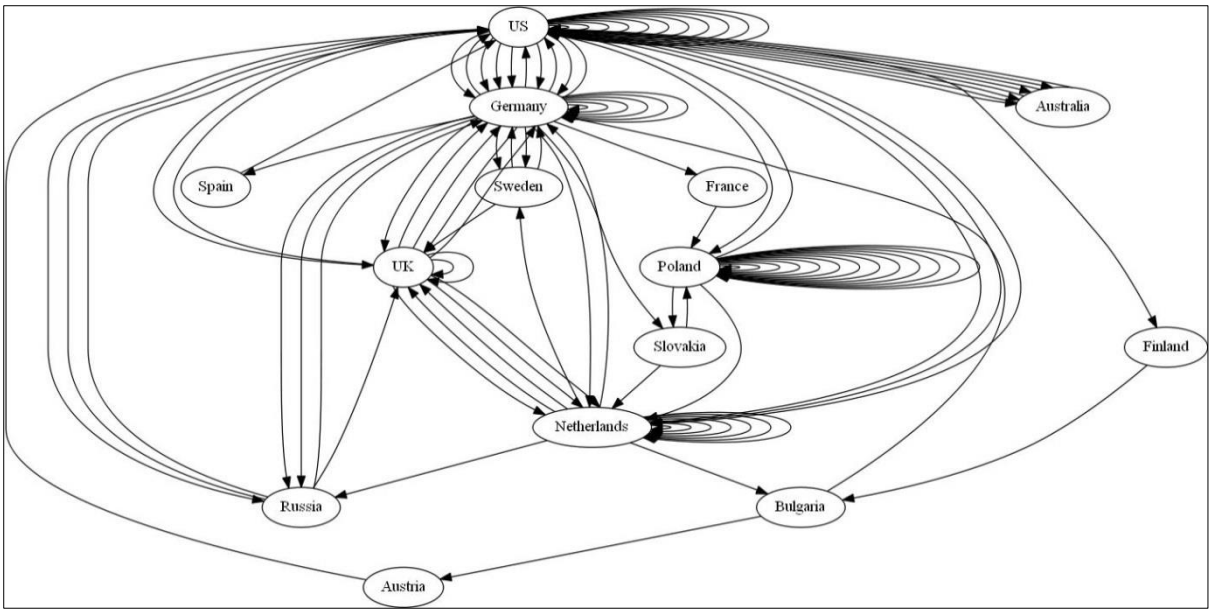


Figure 14: Digraph of transfer of 32,064 BTC ≈ 3.78M USD

4.3 Analysis Of Aggregated Transaction Data

The analysis of the 1000 blocks uncovered a high value of dormant BTCs stored within addresses which have not participated in an outflow transaction for an extended period of time greater than 90 days. Figure 15 illustrates this by showing the number of days for which a quantity of at an address was found to be dormant vs. the number of transactional hops it took to reach that dormant state. The size of the bubble is representative of the value of the dormant BTCs found at the end of the traversal. For illustrative purposes the range for dormant days has been restricted to less than or equal to 900 days and, as mentioned above, greater than 90 days. Similarly, transaction hops greater than 900 are not shown. To achieve the plot the data was ordered chronologically. While the plot does not show data from several outliers it is interesting to note that there is a significant store of value with a relatively high number of dormant days.

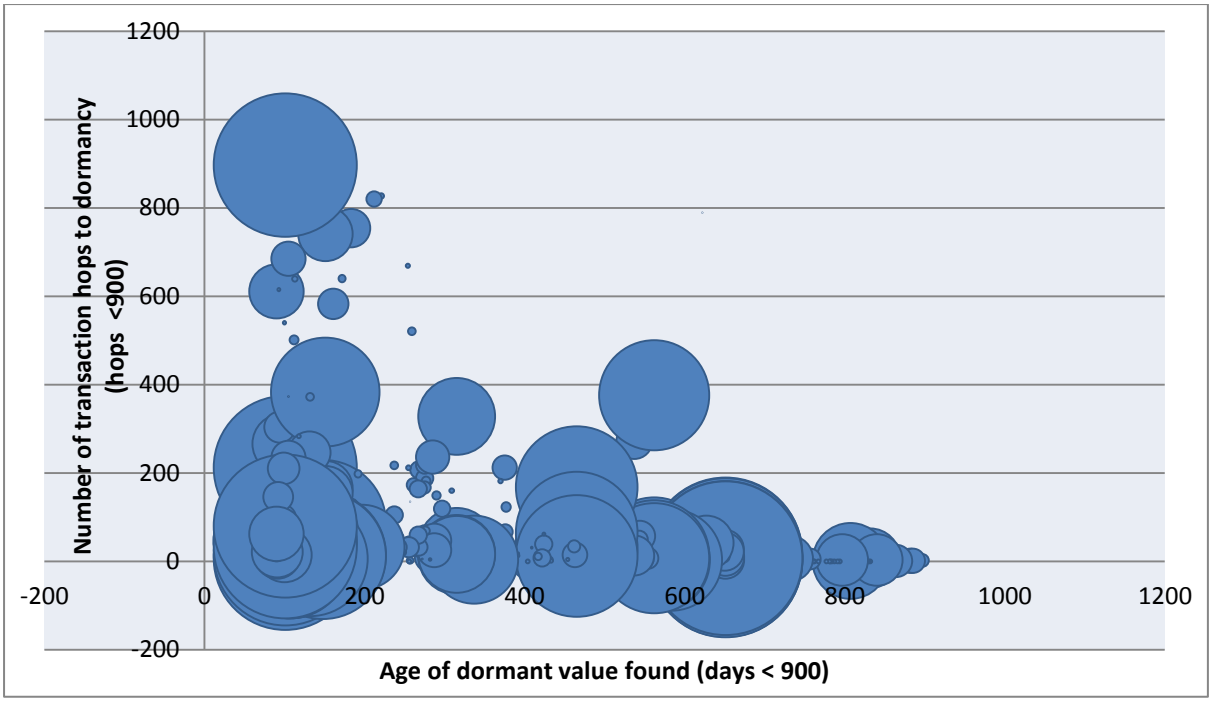


Figure 15: Dormant age vs. Number of transaction hops

4.3.1 Sliding Window Averages

The analysis also yielded interesting trends with respect to sliding window averages for; (i) the number of transaction hops taken to achieve a dormant state, (ii) the age of the dormant value found (days), and (iii) the actual value of the BTCs held at the dormant state. This analysis required that the results from the traversals be sorted in sequential order based on the original coinbase block's block number and only considered data from traversals which resulted in a dormant amount of BTCs being found which had an age equal to or greater than 90 days. The sliding window was set for the requisite criteria over the results of the previous 4 traversals.

Figure 16 shows an increase, at a decreasing rate, in the sliding window average of the number of hops required to achieve a dormant state over time. This implies that throughout the chronological order of the blockchain, there has been an increasing trend in the number of transfers performed before a set of BTCs become dormant. This could suggest increasing attempts to ensure anonymity of transactions by making many transfers and intermediary hops in the hope of masking the true source and destination of the funds involved.

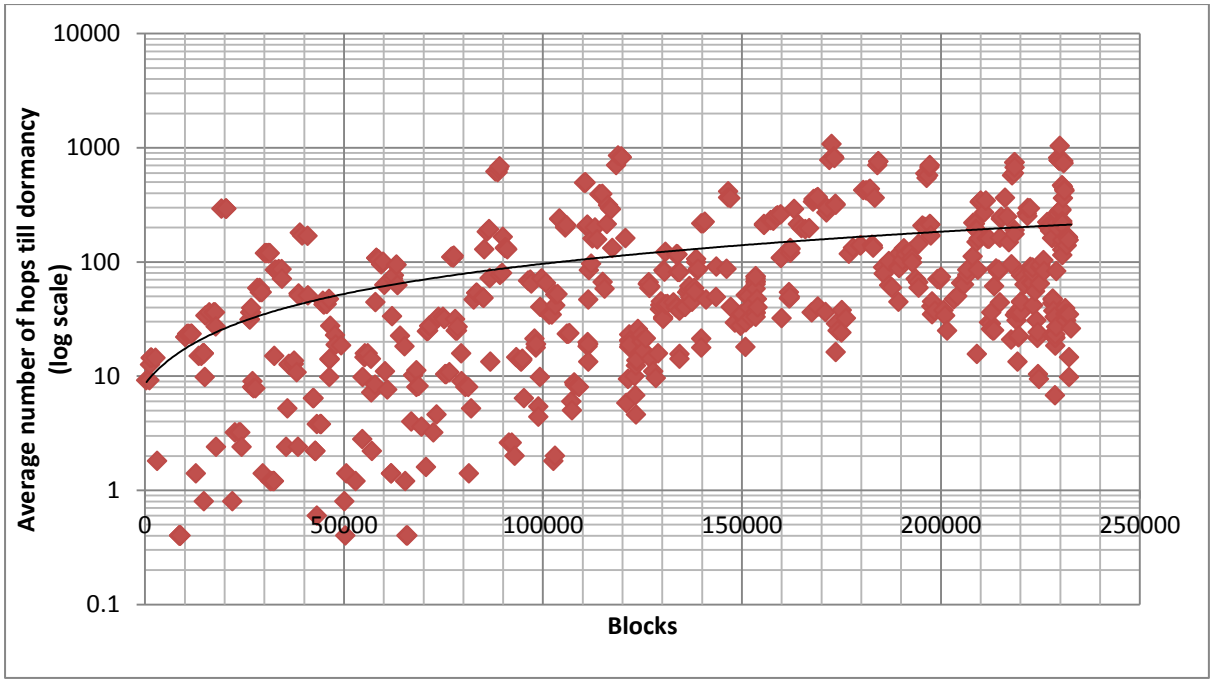


Figure 16: Average of hops required to attain dormancy

Figure 17 shows the average age of the dormant values found with a trend line highlighting that dormancy age decreases with time (at an increasing rate). This may be the result of a high degree of loss or saving in the early days of Bitcoin. Indeed 158 'zero hop' traversals for blocks below 210,000 were found; meaning their 50BTC coinbase value was never redeemed lending credence to the argument of higher loss or savings in the early days.

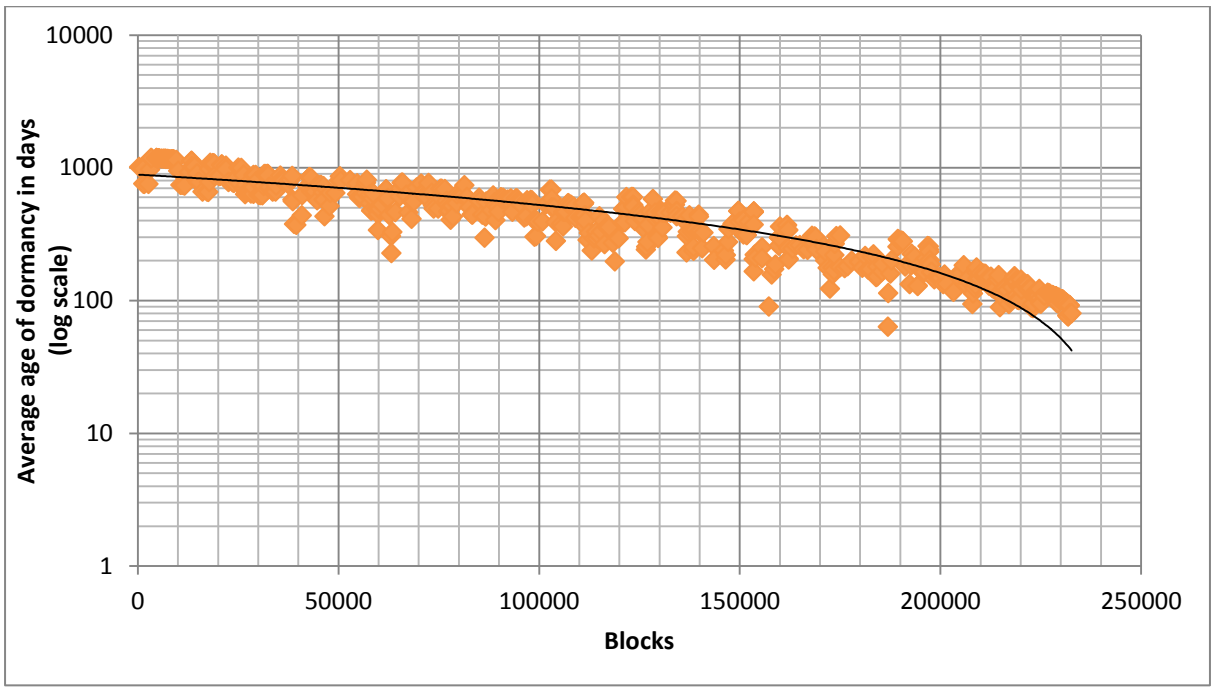


Figure 17: Average of age of dormant value (minimum age = 90 days)

Figure 18 shows the sliding window average of the dormancy age of the non-redeemed quantity of BTCs as revealed at the end of each individual traversal within a single Bitcoin address. Like the previous two figures, the logarithmic scale is again used and allows the wide spread of values to be better deciphered within their respective ranges rather than displaying the points all clustered together on linear scale. A very small but distinct and gradual increase over time is noticeable in the average quantity of BTCs found dormant. This could be indicative of larger quantities of BTCs being stored in single addresses which opposes the viewpoint that over time users have become more Bitcoin savvy and have learnt how to "hide their stash" in multiple small quantities.

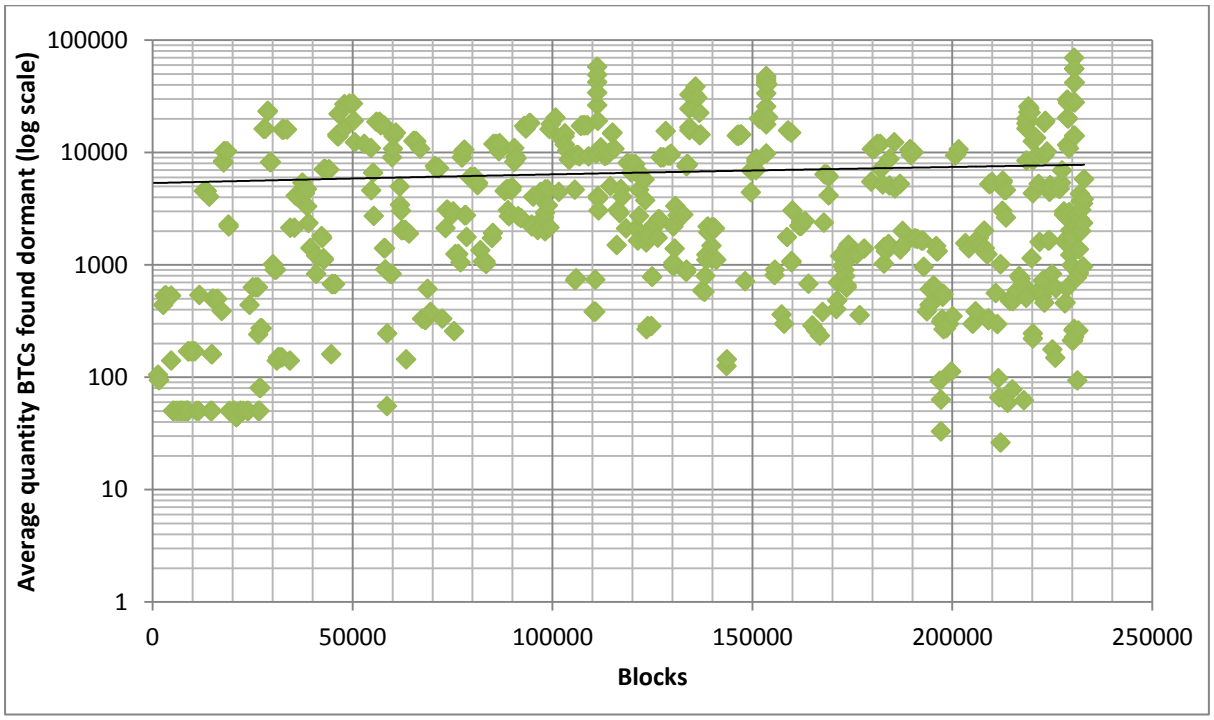


Figure 18: Average of quantity BTC found dormant at a single address

4.3.2 Seeking Correlations

The analysis also sought to decipher any correlation between pairs of results by examining; (i) dormant age vs. number of hops; Figure 19, (ii) number of hops vs. dormant amount (BTC); Figure 20, and (iii) dormant age vs. dormant amount (BTC); Figure 21. Again, only results coming from traversal with a dormant age greater than or equal to 90 days were considered. The corresponding trend line and R^2 values for each plot is also shown. It can be seen that dormant age and number of hops, with an R^2 value of 0.0645, are the most closely correlated pair of results of the three comparisons, while, dormant amount and number of hops shows a very weak correlation with an R^2 value of 0.0002

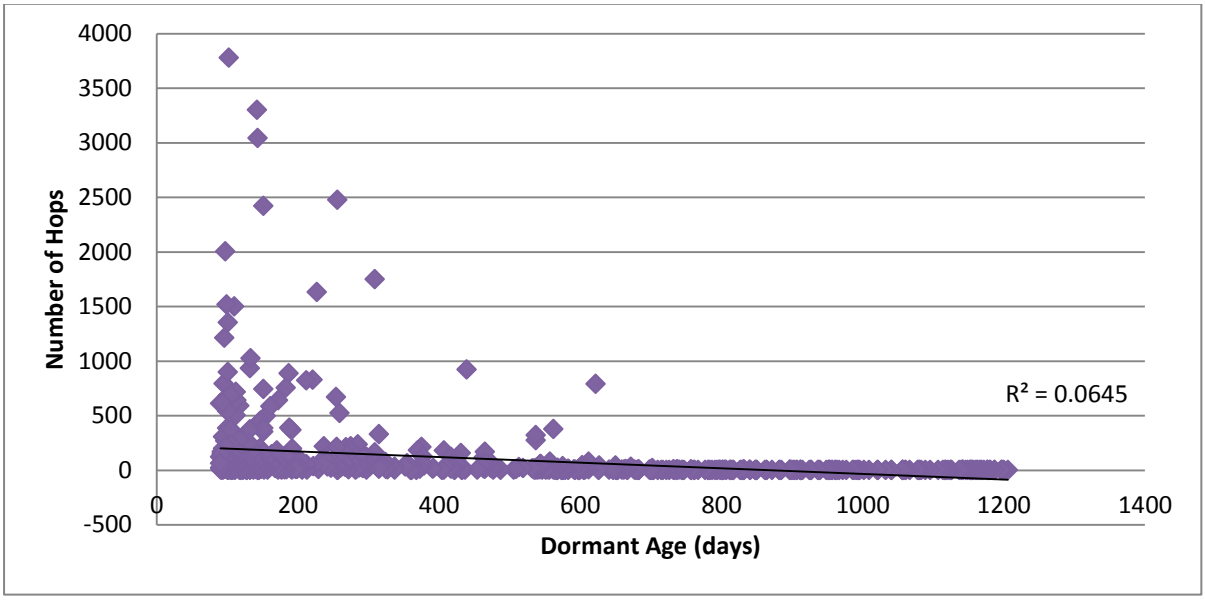


Figure 19: Dormant Age vs. Number of Hops

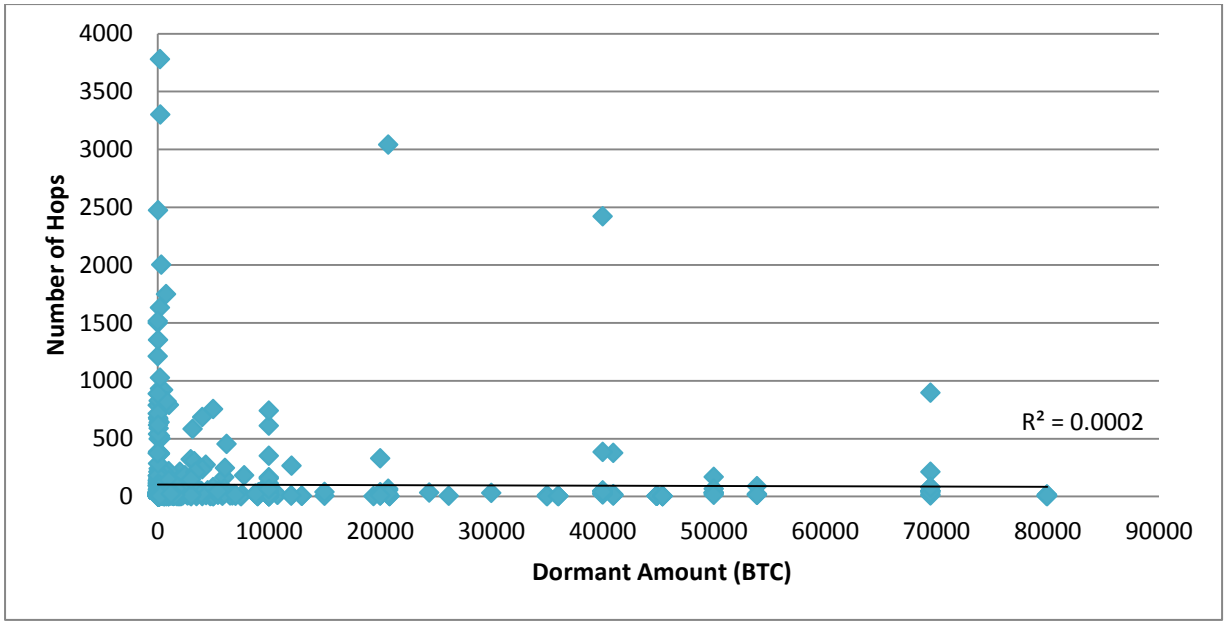


Figure 20: Dormant Amount vs. Number of Hops

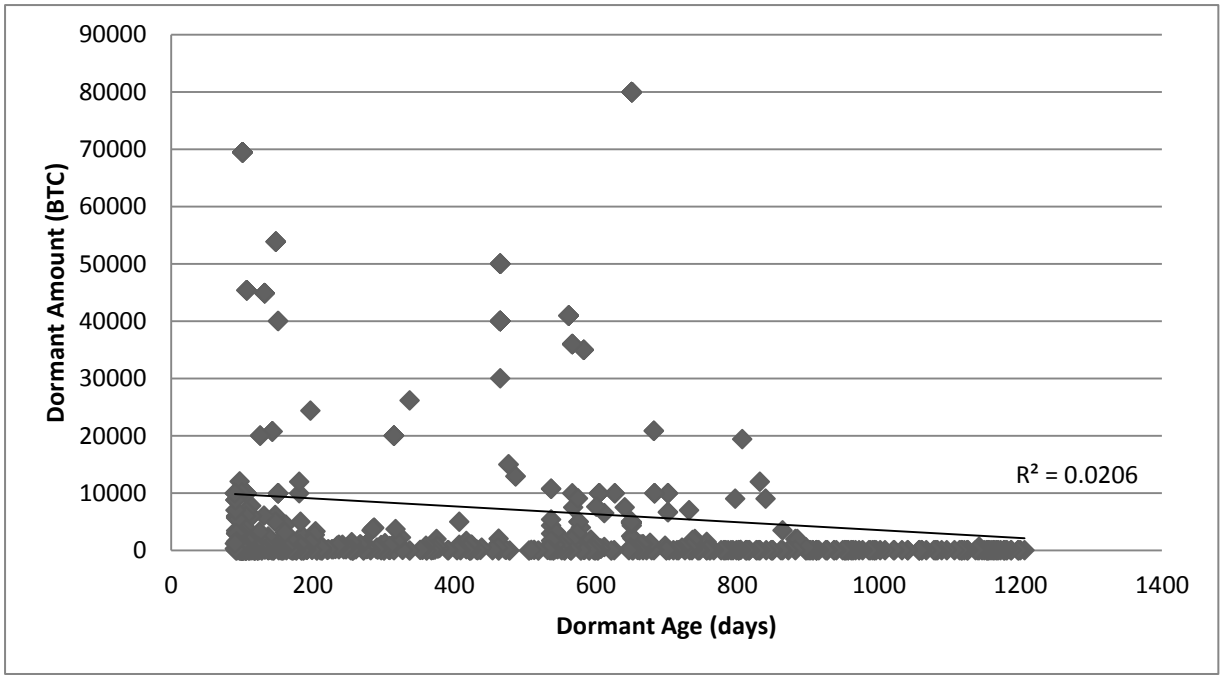


Figure 21: Dormant Age vs. Dormant Amount

5 Discussion

5.1 Geographic Patterns

The data shows a clear distinction between several countries as mentioned in §4.1 in terms of overall volume and value of transactions. The fact that there are a high number of high value transactions which cannot be linked to any particular country as depicted in Figure 11, does indicate attempts to mask the identity of the parties responsible for these transactions. While the caution presented in §0 does hold for these results, the countries identified by this data as the main global players i.e. the United States and Germany are two of the countries identified earlier in §2.4 as being at the forefront of international headlines for recognizing and attempting to regulate Bitcoin. Given the high degree of transaction complexity over a limited duration of time and the apparent great distances traversed as illustrated in §4.2 which supports the findings of [44][45] with respect to anonymity, it is important to accept these results with the proverbial ”grain of salt”.

With this caution in mind, it should be recognized that in a certain percentage of transactions, zero or limited attempt will be made to obfuscate the identities of the parties involved and this data will be present in results found. Hence, while the results may not indicate in absolute terms the actual value and volume of transactions attributed to the countries highlighted; it does imply the relative proportionality of usage between countries. To this end the methods outlined in §3 should be considered sound and useful in the context of uncovering global patterns of usage.

5.2 Usage Patterns

5.2.1 Higher Avg. Number Of Hops Before Achieving Dormancy In Recent Times

Figure 15 is very informative with respect to the age of a quantity of BTCs found to be dormant and the number of traversals required to find those BTCs at a particular Bitcoin address. It is essentially a subset of the data used to generate Figure 19 except that it also illustrates the value of each point on the plot. One can easily decipher there is a wealth of BTCs stored in a dormant state as highlighted within the research of [45][40] and similar to

the approach of these works, a minimum dormancy age of 90 days is set and only traversals meeting this criteria are depicted. The R^2 value of 0.0 645 for Figure 19 indicates that there is a correlation between the dormancy age and the number of hops with the trend line indicating that there are greater quantities hops being performed in recent times when the dormancy age is less. These findings are reinforced by Figure 16 within which a clustering of transaction, within the later blocks in the blockchain is visible. This shows the transactions have an increasing tendency to have a higher hop count before reaching a dormant state. As noted §4.3.1 this may indicate a higher degree of user attempting to maintain anonymity by engaging more transfers. This corresponds with the aspects of anonymity in Bitcoin research within [44][45] and is also supported by the illustrated example of a complex transaction as presented in §4.2. This finding also makes logical sense in that, users are expected to become more knowledgeable about Bitcoin given the amount of exposure Bitcoin has gotten in more recent time with more news coverage and higher rates of adoption and acceptance; §2.1.3.

5.2.2 The Avg. Quantity Of Dormant BTC Has Been Increasing Over Time

Figure 18 shows a very slight increase in the average quantity of Bitcoin found dormant at a particular address as blocks numbers increase over time. This is also shown in Figure 21 where the dormant quantity of Bitcoin is shown to be decreasing as the dormant age increases. The correlation between dormant quantity and dormant age is not as strong as previously mentioned correlation between dormancy age and number of hops, with the R^2 value in this instance being 0.0206 and indeed the trendline in Figure 16 indicates a very small change over time. This could be indicative of the phenomenon mentioned by Dan Kaminsky in [49] where he stated there is high degree of hoarding of Bitcoin occurring at present. The rationale for such hoarding would have been the price increase and fluctuations earlier in 2013 as noted in §2.1. Another reason for such a finding could be introduction of new players in the Bitcoin space such as investors who holding large quantities of Bitcoin in individual accounts.

5.2.3 Average Dormant Age Decreases Over Time

Figure 17 shows the tendency for the age of the dormant value found to decrease as the block numbers increase. To some extent this may be an expected result as the further back in block numbers you go, the higher becomes the probability of encountering a transaction occurring which resulted a dormant value of BTCs in an address for a period greater than 90 days. However, the data also shows a higher degree of coinbase transactions with a status of not redeemed yet, which lends support to the popular view that in the early days of Bitcoin there was a high degree of loss and storage. Other researchers [45] have cited this pattern of usage and have adjusted their data set to compensate for possible high loss affecting dormancy results by omitting data from transactions performed before Mt. Gox was established, as described in § 2.7.2.

5.2.4 Number of hops decrease with value of dormant quantity

With an R^2 value of 0.0002, Figure 20 shows a very weak correlation between the quantity of Bictoins found at a specific address and the number of hops it took to achieve dormancy. The trend line indicates that the number of hops decreases with the increasing quantity of dormant Bitcoin found. This result seems counterintuitive especially in light of the type of patterns of hops detailed §4.2 and given the low correlation between the variables in question, not much significance will be attributed to this finding.

6 Conclusion

The methods and techniques developed and employed in this work did reveal relevant usage and geographic patterns in Bitcoin transactions. In particular, a relative profile of the top geographic areas utilizing Bitcoin has been presented as well as trends on dormancy in Bitcoin circulation. To the best of my knowledge I do not believe that any one has previously attempted to analyse the geographic patterns of Bitcoin usage by countries.

It is considered that this work has relevance to various groups and disciplines including law enforcement, investigative finance, tax evasion, computer forensics and geographic analysis of crime. It is hope that this work can be utilized by these parties and/or others who are working to formulate a proper legal and regulatory framework within which Bitcoin's future viability can be assured.

7 Bibliography

[1] Android Key Rotation: *http://permalink.gmane.org/gmane.comp.bitcoin.devel/2714*. Accessed: 2013-08-13.

[2] Android random number flaw implicated in Bitcoin thefts: *http://nakedsecurity.sophos.com/2013/08/12/android-random-number-flaw-implicated-in-bitcoin-thefts/*. Accessed: 2013-08-13.

[3] Android Security Vulnerability: *http://bitcoin.org/en/alert/2013-08-11-android*. Accessed: 2013-08-13.

[4] Back, A. and others 2002. *Hashcash-a denial of service counter-measure*.

[5] Barber, S. et al. 2012. Bitter to Better—How to Make Bitcoin a Better Currency. *Financial Cryptography and Data Security*. Springer. 399–414.

[6] Beautiful Soup: *http://www.crummy.com/software/BeautifulSoup/*.

[7] Becker, J. et al. 2012. Can we afford integrity by proof-of-work? Scenarios inspired by the Bitcoin currency. *Workshop on the Economics of Information Security WEIS* (2012).

[8] Bissessar, S. all ip data 190813.xlsx.

[9] Bissessar, S. figures and tables.xlsx.

[10] Bitcoin - An Analysis [28C3]: 2011. *http://www.youtube.com/watch?v=-FaQNPCqG58&feature=youtube_gdata_player*. Accessed: 2013-07-10.

[11] Bitcoin Block Explorer: *http://blockexplorer.com/*.

[12] Bitcoin is ludicrous, but it tells us something important about the nature of money: *http://www.washingtonpost.com/blogs/wonkblog/wp/2013/04/12/bitcoin-is-ludicrous-but-it-tells-us-something-important-about-the-nature-of-money/*. Accessed: 2013-08-26.

[13] Bitcoin Market Price: *https://blockchain.info/charts/market-price*. Accessed: 2013-08-14.

[14] Bitcoin Obliterates "The State Theory Of Money:" *http://www.forbes.com/sites/jonmatonis/2013/04/03/bitcoin-obliterates-the-state-theory-of-money/*. Accessed: 2013-07-26.

[15] Buchmann, J. et al. 2006. Perspectives for cryptographic long-term security. *Commun. ACM*. 49, 9 (Sep. 2006), 50–55.

[16] Bustillos, M. 2013. The Bitcoin Boom. *The New Yorker Blogs*.

[17] Certicom Research SEC 2: Recommended Elliptic Curve Domain Parameters.

[18] Console Hacking 2010 - PS3 Epic Fail [27C3]: *http://www.youtube.com/watch?v=PR9tFXz4Quc&feature=youtu.be&t=35m25s*. Accessed: 2013-07-15.

[19] Courtois, N. The Unreasonable Fundamental Incertitudes Behind Bitcoin.

[20] Difficulty: *https://en.bitcoin.it/wiki/Difficulty#What_network_hash_rate_results_in_a_given_difficulty.3F*. Accessed: 2013-07-04.

[21] European Central Bank Virtual Currency Schemes.

[22] Every Important Person In Bitcoin Just Got Subpoenaed By New York's Financial Regulator: *http://www.forbes.com/sites/kashmirhill/2013/08/12/every-important-person-in-bitcoin-just-got-subpoenaed-by-new-yorks-financial-regulator/*. Accessed: 2013-08-22.

[23] FBI Bitcoin: Intelligence Assessment.

[24] Gansner, E. et al. Drawing graphs with dot.

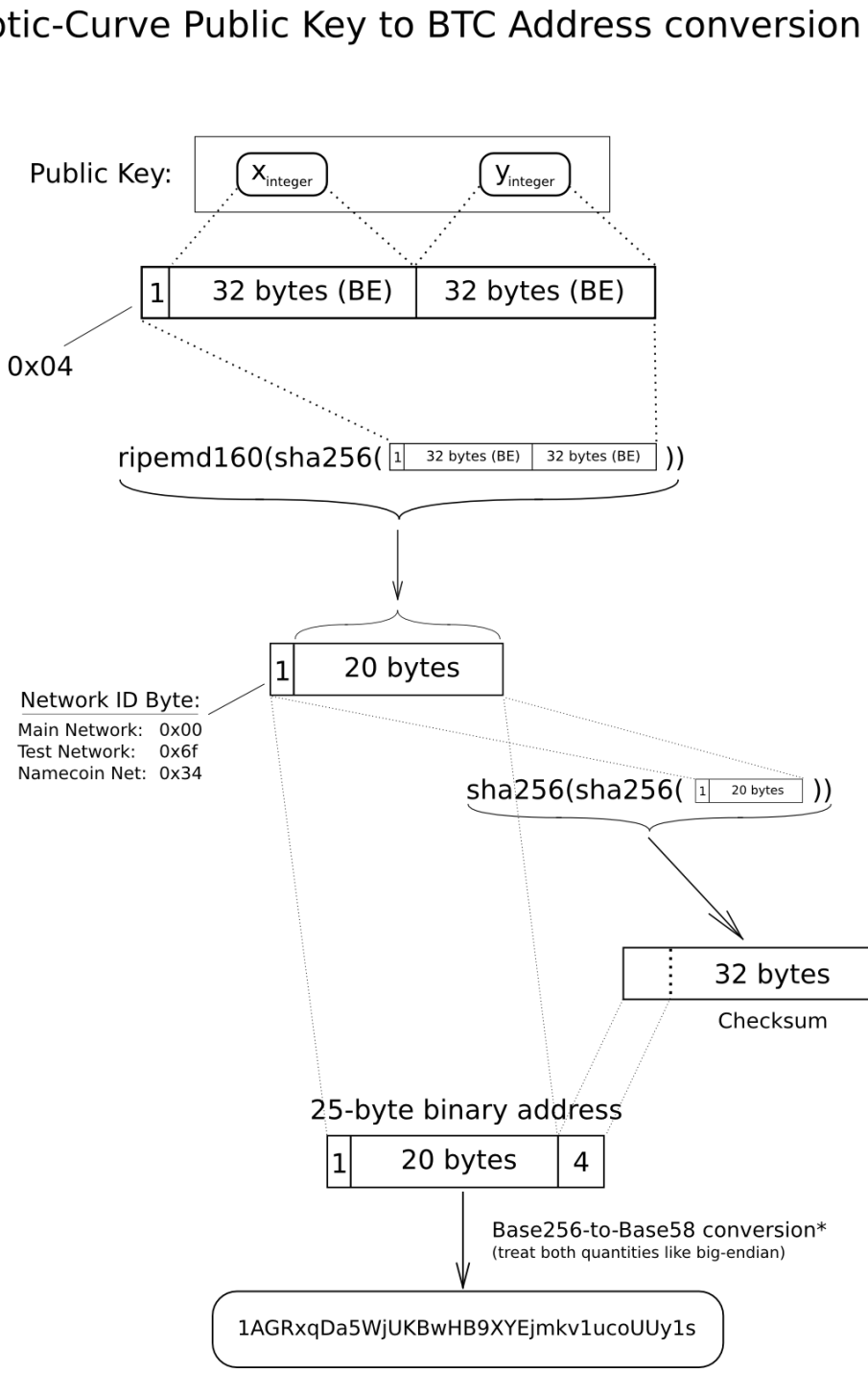
[25] GeoIP databases and web services: *http://www.maxmind.com/en/geolocation_landing*.

[26] Germany Recognizes Bitcoin As “Private Money”, Sales Tax Coming Soon | TechCrunch: *http://techcrunch.com/2013/08/19/germany-recognizes-bitcoin-as-private-money-sales-tax-coming-soon/*. Accessed: 2013-08-22.
[27] GNU Wget: *http://www.gnu.org/software/wget/*.
[28] Graphviz - Graph Visualization Software: *http://www.graphviz.org/*.
[29] Has the market lost confidence in Mt. Gox? - CoinDesk: *http://www.coindesk.com/has-the-market-lost-confidence-in-mt-gox/*. Accessed: 2013-08-26.
[30] How does Bitcoin work?: *http://www.economist.com/blogs/economist-explains/2013/04/economist-explains-how-does-bitcoin-work*. Accessed: 2013-08-03.
[31] I Tried Hacking Bitcoin And I Failed: *http://www.businessinsider.com/dan-kaminsky-highlights-flaws-bitcoin-2013-4*. Accessed: 2013-08-26.
[32] Kaminski, D. 2011. Some Thoughts On Bitcoin.
[33] Laurie, B. and Clayton, R. 2004. Proof-of-Work” proves not to work; version 0.2. *Workshop on Economics and Information, Security* (2004).
[34] Let’s Cut Through the Bitcoin Hype: A Hacker-Entrepreneur’s Take | Wired Opinion | Wired.com: *http://www.wired.com/opinion/2013/05/lets-cut-through-the-bitcoin-hype/*. Accessed: 2013-08-01.
[35] MEMORANDUM OPINION REGARDING THE COURT’S SUBJECT MATTER JURISDICTION: *http://ia800904.us.archive.org/35/items/gov.uscourts.txed.146063/gov.uscourts.txed.146063.23.0.pdf*. Accessed: 2013-08-18.
[36] Mt. Gox Statement Regarding Recent DDoS Attacks and Mitigation.
[37] Mt.Gox - Bitcoin Exchange: *https://www.mtgox.com/press_release_20130530.html*. Accessed: 2013-07-20.
[38] Nakamoto, S. 2008. Bitcoin: A peer-to-peer electronic cash system. *Consulted*. 1, (2008), 2012.
[39] Network Hash Rate: *https://blockchain.info/charts/hash-rate*. Accessed: 2013-08-08.
[40] Ober, M. et al. 2013. Structure and Anonymity of the Bitcoin Transaction Graph. *Future Internet*. 5, 2 (2013), 237–250.
[41] Pay Another Way: Bitcoin: *http://en.blog.wordpress.com/2012/11/15/pay-another-way-bitcoin/*. Accessed: 2013-08-26.
[42] PubKeyToAddr.png: *https://en.bitcoin.it/wiki/File:PubKeyToAddr.png*.
[43] rawblock 9000: *http://blockexplorer.com/rawblock/00000000415c7a0f3e54ed29a3165414e5952fce6848b697696fcc840011a5b5*.
[44] Reid, F. and Harrigan, M. An Analysis of Anonymity in the Bitcoin System.
[45] Ron, D. and Shamir, A. 2012. Quantitative Analysis of the Full Bitcoin Transaction Graph. *IACR Cryptology ePrint Archive*. 2012, (2012), 584.
[46] Security Panel - Bitcoin 2013 Confernece: *http://www.youtube.com/watch?v=si-2niFDgtI&feature=youtu.be&t=1m40s*. Accessed: 2013-07-26.
[47] Security Panel - Bitcoin 2013 Confernece: *http://www.youtube.com/watch?v=si-2niFDgtI&feature=youtu.be&t=14m2s*. Accessed: 2013-07-26.
[48] Security Panel - Bitcoin 2013 Confernece: *http://www.youtube.com/watch?v=si-2niFDgtI&feature=youtu.be&t=40m40s*. Accessed: 2013-07-26.
[49] Security Panel - Bitcoin 2013 Confernece: *http://www.youtube.com/watch?v=si-2niFDgtI&feature=youtu.be&t=0m50s*. Accessed: 2013-07-26.
[50] Surowiecki, J. 2013. Why Did Criminals Trust Liberty Reserve? *The New Yorker Blogs*.
[51] The Case for Elliptic Curve Cryptography: *http://www.nsa.gov/business/programs/elliptic_curve.shtml*. Accessed: 2013-07-10.

[52] The Challenges Ahead for Bitcoin: 2013. *http://www.cato-unbound.org/2013/07/29/jerry-brito/challenges-ahead-bitcoin*. Accessed: 2013-08-26.
[53] Top 10 Bitcoin Merchant Sites: *http://www.forbes.com/sites/jonmatonis/2013/05/24/top-10-bitcoin-merchant-sites/*. Accessed: 2013-08-13.
[54] Tor: *https://www.torproject.org/*.
[55] Trading suspended due to Bank of Thailand advisement: *https://bitcoin.co.th/trading-suspended-due-to-bank-of-thailand-advisement/*. Accessed: 2013-08-15.
[56] Transaction: *http://blockexplorer.com/tx/bd8cd3033ba2ad2b485f6bb5432f4819c83a2c2ac8ce5e402b567924011b2193*.
[1] *Triangulation 103: Gavin Andresen*. 2013. Available: http://www.youtube.com/watch?v=HYo7q5Lkk1w
[58] U.S. Attorney's Office - U.S. Department of Justice: *http://www.justice.gov/usao/nys/pressreleases/May13/LibertyReservePR.php*. Accessed: 2013-07-27.
[59] Waiting for the Shoes to Drop: 2013. *http://www.cato-unbound.org/2013/07/23/jim-harper/waiting-shoes-drop*. Accessed: 2013-08-26.
[60] What's Your Bitcoin Strategy? WordPress Now Accepts Bitcoin Across The Planet: *http://www.forbes.com/sites/jonmatonis/2012/11/16/whats-your-bitcoin-strategy-wordpress-now-accepts-bitcoin-across-the-planet/*. Accessed: 2013-08-10.
[61] 2013. Bitcoins illegally mined on network. *BBC*.
[62] Exchanges by volume. BitcoinCharts.
[63] FinCEN: Regulations to Persons Administering, Exchanging, or Using Virtual Currencies.
[64] 2013. Liberty Reserve "criminals" move on. *BBC*.

8 Appendices

8.1 Bitcoin Address Construction [42]



*In a standard base conversion, the 0x00 byte on the left would be irrelevant (like writing '052' instead of just '52'), but in the BTC network the left-most zero chars are carried through the conversion. So for every 0x00 byte on the left end of the binary address, we will attach one '1' character to the Base58 address. This is why main-network addresses all start with '1'

8.2 Raw Block # 9000 [43]


```
{

"hash":"a5f71a0b4ccda01fa6b491b6f4d320aa3320df81111d1c28dbbabfa47ec92d88",
  "ver":1,
  "vin_sz":1,
  "vout_sz":1,
  "lock_time":0,
  "size":135,
  "in":[
    {
      "prev_out":{

"hash":"0000000000000000000000000000000000000000000000000000000000000000",
        "n":4294967295
      },
      "coinbase":"04ffff001d026606"
    }
  ],
  "out":[
    {
      "value":"50.00000000",

"scriptPubKey":"045a6e07b7e579fd850014fd2bd27fbfbb0edc0e892217cd1bd5648b158
2fc57568c472fe5fb0125866ba3dc7ab14c7dd6fbeecbc01c82c422b5cc98403eaddb4b
OP_CHECKSIG"
    }
  ]
}
```