What may have happened at mt gox

by Cameron Winklevoss - Founder, Winklevoss Capital

It may be months or years before government authorities and industry experts across the globe get to the bottom of what happened to the ~750K BTC that purportedly went missing from Mt. Gox. By now, the basics of the exchange’s downfall are well publicized, but the actual substance and fate of the nearly $500 million dollars of bitcoin are still shrouded in mystery.

Six month graph from Winkdex

On February 24, 2014, a Mt. Gox “ Crisis Strategy Draft” first leaked on the Internet and subsequently was confirmed to be a genuine, but not final, draft of the company’s strategy. According to the Crisis Strategy Draft, the exchange lost ~750K BTC due to Transaction Malleabillity, an issue with the Bitcoin protocol that has been known (and dismissed as a low-priority issue) since 2011.

The explanation goes like this: a user(s) was able to manipulate the transaction ID of a given bitcoin withdrawal from the exchange, resulting in a “ mutant” version of the withdrawal being confirmed by the network with the original version being dropped. As a result, the exchange would register this withdrawal as failed (when in fact it was confirmed into the blockchain with a different transaction ID) and allow the user(s) to make another withdrawal. Done over-and-over by a user(s) undetected and without exchange safeguards, Transaction Malleability exploits could theoretically deplete the exchange’s BTC deposits. Assuming the exchange used some element of cold storage, the exchange would have to continue to unknowingly replenish their “ leaking” hot wallet to the point where the cold storage was completely emptied out.

Not surprisingly, many have challenged this account of events, especially considering the vast sum of money lost (~99% of all deposits) and the lack of clarity on how big of a “ problem” Transaction Malleability actually is/was. As I mentioned before, this issue had been known since 2011, and Bitstamp, for example, publicly stated on February 11th, 2014 that “ no funds have been lost and no funds are at risk” before managing to implement a “ simple solution” to resume BTC withdrawals four days later. While it is entirely possible that Mt. Gox was victim of Transaction Malleability exploits, it is almost certainly a false explanation for its failure.

So with that said, below are possible theories as to what may have happened (the level of detail below is still largely high-level so as not to get lost down the rabbit hole):

1) Transaction Malleability

2) Hack

a) Poorly designed and/or accessible Cold Storage

b) Technical and security mismanagement

c) Cloak and Dagger

3) Inaccessible Cold Storage

4) Some Combination of 1, 2, 3

5) Inside Job

1) Transaction Malleability theft can occur when an attacker node grabs a transaction coming from an exchange and instead of relaying it to their peers , the attacker subtly “ mutates” the transaction resulting in it having a different transaction ID (i. e. transaction fingerprint) before relaying it to peers. And while the attacker has altered the fingerprint, they have not altered its essential meaning or validity.

The nature of the mutation is akin to the following:

Pinnochio’s makes the best pizza.

If we replace the word best with the word good, we get:

Pinnochio’s makes good pizza.

The meaning of this new sentence is somewhat different than the first, but this new sentence is still true. In the world of Bitcoin, it only matters whether a given transaction is true or not true.

The attacker then proceeds to broadcast this new “ mutant” transaction to the network. If the attacker is connected to a large number of peers (i. e. “ super node”) then it can spread it’s “ mutant” to everyone else faster than the original. Once the “ mutant” becomes more prevalent, it will likely be mined into the next block on the blockchain. Once mined, it is now considered “ real” while the original transaction will remain unconfirmed, be forgotten and eventually rejected.

Nine different avenues of malleability have been identified, but the most recent flood of “ mutants” starting on February 9th, 2014 have been linked to one particular type known as PUSHDATA2. Ken Shirriff posted an analysis of the attack, graphed hour by hour. This occurred two days after Mt. Gox halted bitcoin withdrawals and Ken found no large traces of it occurring historically over the blockchain. If it was “ Malleability-related theft that [supposedly] went unnoticed [by Mt. Gox] for several years” it would have been another form, perhaps less conspicuous. And while there doesn’t (yet) appear to be a historical analysis of the volume of all mutants over the blockchain, it is somewhat beside the point.

It doesn’t really matter what malleability corner-cases Gox may have (repeatedly) fell for, the default position of a system when a transaction goes unconfirmed should be to figure out why. As the explanation goes, Mt. Gox’s customer service, accounting and transaction confirmation systems were so inadequate as to allow users to repeatedly use the Transaction Malleability exploit without an eyebrow being raised. Hypothetically, if the Mt. Gox systems were incredibly incompetent, a user could continue to use the Transaction Malleability exploit to withdraw the same bitcoins over and over again, so long as its “ mutant” transaction data was confirmed before the original transaction data. It just remains hard to believe a system could be designed with such an egregious flaw not encountered by the competition.

Furthermore, a Transaction Malleability exploit that emptied Mt. Gox’s wallets would leave a tremendous digital trail starting from the fact that the exploiting users would be known users of Mt. Gox with, presumably, some level of identity verification. Mt. Gox has not publicly called for a criminal investigation into any of its users alleged to have exploited Transaction Malleability.

2) Hack could have occurred several years ago which management has been trying to dig themselves out of over time. As the price of BTC increases, and potential additional hacks occur, what starts out with good intentions increasingly becomes a nightmarish ponzi scheme. Poor accounting and/or incompetence make the company increasingly less transparent w/ reactive and evasive PR. This leads to an increasing loss in confidence and when fiat is frozen ultimately leading to a suspension of withdrawals on June 20th, 2013 users become cagey and start running for the exits. A temporary reinstatement of measured withdrawals buys some time, but eventually the situation reaches the point of insolvency and the music stops.

Potential Hacks

a) Poorly Designed or Accessible Cold Storage

It’s possible that Mt. Gox’s cold storage wasn’t really “ cold”. Karpeles stated on IRC chat in one instance that he was going to access the cold storage via a “ firewall”, which would indicate that the supposedly cold stoarge was in fact “ hot”.

b) Technical and Security Mismanagement

The Mt. Gox development team consisted chiefly of Mark Karpeles who is not a crypto or security expert and was widely considered to be in way over his head. Most likely there were serious security flaws and vulnerabilities in the system.

c) Cloak and Dagger

Most startups don’t take great pains to lock down their employee’s computers or have 24/7 monitoring. Given the shortcomings mentioned above, consider the “ janitor attack” whereby a infiltrator posing as cleaning personnel is able to carry out a USB attack to place malware on a development machine and successfully corrupt random number generator MtGox used to generate internal bitcoin accounts. Instead of using 256bits of entropy, it now uses 50bits which to the naked eye would still look random, but no one is ever going to know (only the person looking for it can find it). The attacker then scans the blockchain for these weaker addresses and then makes off with the money. Yes it might seem extreme, but very possible given how comparatively soft a target Mt. Gox appears to have been. An engineer could for example simply apply for an interview. A bank or art gallery heist is arguable much harder to pull off. There is plenty of historical precedent for elaborate multi-million dollar heists in the past.

d) Many other possibilities, this is just to name a few.

3) Inaccessible Cold Storage is a very real possibility given that Mark Karpeles was Mt. Gox’s own Wizard of Oz, believed to be the only person who had total access and control of the tech and what was going on behind the scenes. Given that he had no version control software (i. e. ability for devs to overwrite each other’s code), up until recently no staging environment (risk of easily introducing bugs into system), etc. it is conceivable that he built a cold storage system that he can no longer access for whatever the reason. (e. g., due to lost or malformed private keys).

Karpeles stated in an interview after the “ Crisis Strategy Draft” leak that “ Well, technically speaking it’s not ‘ lost’ just yet, just temporarily unavailable”. It is possible that the bitcoins may not be lost forever, if access to the cold storage can be restored . If access cannot be restored, however, the BTC may have blasted into bitcoin heaven.

4) Some Combination of 1, 2, 3 may best explain the loss. That is to say a variant of a hack possibly coupled with Transaction Malleability could have forced Mt. Gox to dip into cold storage. When going to the deepest or coldest of their cold storage, they may have realized it was a) inaccessible or b) gone

5) Inside Job seems unlikely in this case because there is rarely a getaway car in Bitcoin. To pull off a major heist, you would first need a good wheelman, something you would not find in the Mt. Gox organization. And, despite the common misunderstanding that Bitcoin is “ anonymous” it is in fact “ pseudonymous”:

1) All movement of bitcoin are recorded on the blockchain;

2) You can trace the flow of funds (akin to genealogy) of all bitcoins forever;

3) While we don’t necessarily know who is on either side of a given transaction, we do know all transactions that occur and movements of bitcoin

Furthermore, bitcoin addresses can become tied to real world identities when the bitcoins are converted to fiat or spent, or when a bitcoin user accidentally or intentionally ties his identity to the digital address. Recent research also indicates that data analysis can use statistical processes to affiliate common users' bitcoin addresses and even tie IP addresses to bitcoin transactions; here, here and here.

Dan Kaminsky's analogy to a stolen piece of art is always helpful. You can steal a Picasso, but good luck trying to sell it and cash out at an auction house. These properties alone make this theory less than attractive.

Lastly, an additional theory (perhaps one of the more far out ones) some people entertain is that a government (US, Japan) wanted to seize the bitcoin as part of either an investigation of Silk Road or MtGox's bank misflings and he cooperated, hence " just temporarily unavailable".

Conclusion

So where does this leave us? An Inaccessible Cold Storage that managed to bleed bitcoins into the ether is no doubt the worst outcome possible and could potentially explain a Transaction Malleability cover story. The upside is that 1, 2, 5 are all to an extent traceable and by virtue of this in theory recoverable. All of this will take time, no doubt countless hours on the part of authorities, forensic accountants and thankless volunteers who scour the blockchain looking for leads. If/when a list of lost or stolen coins is compiled exchanges around the world may blacklist them (i. e. " coinvalidation"). Regardless, many customers sadly may never receive more than cents on the dollar despite knowing very well where their bitcoin are. A painful reminder of the growing pains of Bitcoin1. 0 but the final closure of a chapter and the start of what we have all been waiting for, Bitcoin2. 0.