Privacy Tech

Exploring the technology of privacy

A Privacy Engineer's Analysis of Bitcoin

R. Cronk, CIPM, CIPP/US
Privacy Tech | Mar 26, 2015

If you're not familiar with Bitcoin, then you should probably at least review this quick video intro before proceeding.

What is Bitcoin? (v1)

Simply put, Bitcoin is a peer-to-peer protocol that allows for the maintenance of a transaction ledger using a consensus algorithm designed to ensure ledger integrity. Too much of a mouthful? How about Bitcoin is an Internet-based currency with cash-like qualities? At least that is how it was introduced when it was launched into the world in 2009 by its pseudonymous creator.

As the creator's identity is shrouded in mystery, so too did he/she/they wish to hide the owners of Bitcoin from identification. However, while the real identity of Satoshi Nakamoto has remained a mystery, the same cannot be said for some of Bitcoin's more notorious users. With some gumshoe detective work and analysis of the Bitcoin blockchain (the very public transaction ledger that is built block by block), users can be de-anonymized. Some even more sophisticated attacks may exist, as well.

While not perfect, Bitcoin represents an improvement, from a privacy perspective, over the "disco-era" financial system currently in use.

For anyone who has had to comply with the PCI DSS standards to protect credit cards flowing through their networks, you have to realize that credit cards were designed in a different world, before rampant fraud became the norm. They were never designed for the Internet or commerce on the scale to which they are used today. I still remember as a kid, my grandmother pulling out a credit card in the 1970s only to have a store clerk thumb through a telephone book with expired and canceled credit cards listed in them. That book was published monthly.

We've come a long way.

With that in mind, let's examine some of the privacy risks found in traditional card-based systems and the subsequent controls applied in the Bitcoin system to reduce that risk.

Account Hijacking: An "Information Processing—Secondary Use" risk under the Solove Taxonomy, basically, this is what PCI and security controls on credit card numbers attempts to protect against, namely the risk that the account is taken over and used by a criminal to make fraudulent purchases. This is not only a risk to the individual account-holder but puts a significant burden on merchants who generally must bear the burden of fraud.

Bitcoin uses a **decentralized** architecture as a control. I mean this in terms of the useful information necessary to overtake an account. Bitcoin's public key cryptography means that ownership of the account is defined by who has control of the private key. The card-based system relies on a shared secret: the card number. Even with the advent of card security codes, that information has to be shared through the network to authenticate the owner. With Bitcoin, that information is never transmitted or stored except with the owner, reducing the likelihood of account hijacking.

Most of the major Bitcoin hacks have been a result of either re-centralization of Bitcoin accounts or sloppy operational security on the part of owners of large Bitcoin wallets.

Cross-Purchase Analysis: Many retailers link individual customer purchases through identifying information, such as a common card number used for distinct purchases. While retailers can use this information for better understanding of their constituency's purchasing patterns, it still opens up Aggregation Risk to the subject. Knowing that someone purchased a particular shampoo doesn't tell you much; knowing that someone purchased other products and a particular shampoo might suggest they are pregnant —something the customer may have reason to conceal.

While early implementations of Bitcoin software falls victim to this same privacy risk, newer software doesn't reuse addresses, so there is no single data element to turn to for cross-purchase analysis.

However, this technique is not perfect; it's more of **obfuscation** than anything else. A deeper analysis of the transaction ledger (the blockchain) could reveal a common identity between purchases. A consumer would have to turn to additional techniques, such as CoinJoin, which is not inherent in the Bitcoin protocol, to avoid this. Some alternative cryptocurrencies, such as ZeroCoin, build this into the native protocol.

Cross-Merchant Analysis: This can be achieved by aggregating data from retailers, an analysis on an individual's purchases across merchants. This also represents a **secondary use** risk as well as an **exclusion** risk because most consumers are unaware information is being shared with entities other than the retailers with whom they are doing business.

Given the public nature of the Bitcoin ledger, this analysis can similarly take place with Bitcoin. Unless an individual takes an extremely disciplined and less practical approach, a sophisticated analysis could reveal at least a portion of related transactions. In this case again, **obfuscation** is the principle control utilized by the Bitcoin protocol because multiple inputs and outputs can be combined in one transaction, making it difficult to discern which transactions are true from one party to another and which are intra-party transactions.

As transactions pile upon transactions, the probability of tracing funds becomes much more difficult.

Identification: I've told this story before, but several years ago my then-girlfriend received a Facebook friend request from our local pita shop guy. How did he know her name? Oh, let's see, it's printed on the credit card she handed over to him every time we went to the pita shop. Suffice to say we quit going there. The linking of the financial account information necessary to process to the credit card to a real world identity put my girlfriend at both **identification** risk and **secondary use** risk—and possibly **appropriation** risk.

Bitcoin, similar to cash, uses **data minimization** (specifically anonymization) to prevent these risks. There shouldn't be a need to hand over identifying information to complete the transaction. The goal of the retailer is to get paid for the goods or services delivered. The identity of the customer is superfluous to that objective.

As described above, Bitcoin uses data minimization, obfuscation and decentralization as the primary means of enhancing user privacy. While the future of Bitcoin remains to be seen, it does provide an interesting contrast from a privacy-risk perspective to the system of electronic payments we have in place now.

The author has a blog devoted to Bitcoin legal issues (rjceqs.com) and is co-organizer of the Atlanta Bitcoin Consumer Fair, an event coming up in April to showcase Bitcoin in the retail environment. You are invited to use the discount code "IAPP" for 50 percent off the standard ticket price.

Comments

Logged-in as: John Wunderlich

Share your thoughts