CRYPTIX
LABS

The Challenges of On-Chain Privacy

Privacy-preserving blockchain technology is becoming increasingly necessary for financial autonomy.

As cryptocurrencies scale and larger institutions, businesses, and state-actors become involved in developing these technologies for their own ends, the initial selling point of cryptocurrencies like Bitcoin – financial value based on cryptography which exists outside of the purview of governments and banks - becomes more and more apparent in its being a fantasy (spoiler alert: it always was). Privacy-preserving blockchains have, unfortunately, been far slower to develop, and have had even lower rates of uptake than public, transparent blockchains. There are numerous historical and practical reasons for this: early on in the blockchain hype, it was far more difficult to get funding an investment that would not make a quick return, due both to its seeming 'niche' appeal and the practicality that it took far longer to research and then implement complex cryptographically secure protocols than launching 'reddit on the blockchain'. Furthermore, most of the individuals involved in privacy technology were more interested in slowly creating good technology and building in a 'ground-up' manner, rather than rely on venture capital funding and promises that complex technology was just over the horizon.

This document serves as a general introduction to the idea of what privacy on a blockchain actually entails, a technical overview of the issues that privacy-preserving blockchains have to solve, as well as an overview of some of the technologies looking at resolving some or all of these issues. This is also the first in a series on privacy, each of which will focus in on deeper dives into a particular issue. Before diving into privacy-preserving blockchains however, it is useful to outline all of the potential ways in which the 'canonical' cryptocurrency – Bitcoin - can lead to information leakage, and ultimately traceable or de-anonymized transactions.

The Bitcoin blockchain is a public-by-design list of timestamped blocks of transactions. The details of each transaction – the sender's address, the recipient's address, the amount of Bitcoin transferred, and the timestamp at the time of transaction creation - are also public, in order to be historically verifiable. All addresses are themselves merely pseudonymous, with the public key of the address essentially acting as pseudonym. The communication protocol by which Bitcoin nodes broadcast transactions to other nodes involves sending this data in plaintext, and occurs over existing internet infrastructure, and will therefore be captured by whatever ISP is controlling said infrastructure. These aspects ultimately lead to a cryptocurrency which involves – at best – a single element of pseudonymity in the form of its addresses. Once an address can be linked to an individual or business (which, as discussed below, becomes trivial at scale), even this tiny element of privacy is removed, leaving the blockchain even more publicly transparent than the cashless society that Bitcoin was originally conceived of in order to undermine.

From this brief discussion of Bitcoin's transparency, we can therefore lay out some broad areas in which we can talk about what privacy actually means when we're discussing blockchains, specifically within the scope of protocol-specific privacy with regards to public blockchains.

These are:

1. Private transactions: both the sender and receiver's addresses should be hidden, to avoid address linking, and because the network that blockchain nodes broadcast transactions via cannot be presumed to be secure.

2. Private addresses: there should be no public list of addresses, or block explorer.

3. Private balances: if account addresses are public, the balances of these addresses should be private, as should all transactions these accounts have been involved in.

These three areas imply that a private blockchain would have many characteristics which are antithetical to the 'original' selling point of public blockchains: transparent and publicly verifiable addresses and transaction logs. However, as the examples below will show, this is ultimately necessary to retain the cypherpunk ethos underlying cryptocurrencies as they were originally conceived: decentralized finance, autonomous from existing financial institutions and not structured so as to have potential single points of failure or the need to rely on third-parties for transactional security.

There are numerous justifications for privacy-preserving blockchains. Some of the biggest ones are – outside of the broader points regarding human rights and desire for privacy – fundamentally of importance for the continued existence of privacy technology itself. The first has been touched on above: public blockchains like Bitcoin are merely pseudonymous. As they rely on public/private key cryptography, the need to broadcast your public key when receiving donations or payments links you to a particular address. Given that this information will most probably be passed via either a messaging service or publicly via something like Twitter, this link will exist (potentially publicly) in third-party servers outside of your control. Furthermore, the transaction involving your (now linked) address will exist publicly in the blockchain. As well as evidence that online merchants "often leak information to third parties that can be used to link customers to their public addresses" [5], well-funded private companies such as Chainanalysis [6] are utilizing data analysis techniques to further remove the small element of privacy that pseudonymity offers. Given the slew of scandals involving use of mass data analysis in recent years, this should be of great concern to the crypto community. Indeed, if what little trust in the cryptographic certainty offered by blockchains dissolves, this will greatly affect the ability for onboarding and expansion.

Furthermore, privacy-preservation is necessary for any given cryptocurrency to retain its fungibility. This is the 'interchangeability' of the currency, whereby one (e.g.) Bitcoin is equivalent in value to any other Bitcoin. However, many businesses or individuals may not want coins that have been 'tainted', either by being involved in unscrupulous transactions, or for more ideological reasons. Ultimately, this will lead to a bifurcation of the supply of the currency in question, and create even more volatile market situations, creating further road humps for wide-scale adoption.

Although these issues are pressing, there are multiple implementations of different blockchain protocols that are being developed in response to them. These include (but are not limited to) the 'original privacy coin', Monero, more corporate-oriented projects such as Z-Cash, and newcomers Grin and Beam, with each providing different answers to the issues raised above and will be discussed in depth in future documents. What the existence of this spectrum of privacy technology shows is that the quest for a computationally provable privacy solution is still very much in question, with a multitude of promising answers on offer. Ultimately however, it is also important to remember that these technologies are locked into a privacy arms race, with protocols patching in response to new academic research, and new and more intrusive methods of data analysis being uncovered every year.

[1] It is important to note that there is a lot of overlap between these areas, or privacy-preserving solutions within these areas.
[2] Whilst there is merit in discussing privacy in a broader sense, looking at techniques such as IP shielding or batching transactions, this discussion falls into the work already being undertaken by (e.g.) projects involving mixnets, and Tor. For the sake of brevity, this will not be discussed here.
[3] See this article for a more detailed example of this analysis: https://at-vc.com/intro-to-blockchain-what-are-privacy-coins-and-why-do-we-need-them/
[4] www.technologyreview.com/2017/08/23/149531/bitcoin-transactions-arent-as-anonymous-as-everyone-hoped/
[5] See 'When the Cookie Meets the Blockchain' (https://arxiv.org/abs/1708.04748) for details, and 'A New Attack Vector to Deanonymise Bitcoin Users'(https://medium.com/decentralize-today/a-new-attack-vector-to-deanonymize-bitcoin-users-9c6dc433d4b6) for a response.
[6] https://www.chainalysis.com/
[7] https://www.reddit.com/r/btc/comments/6ces98/reminder_segwit_breaks_fungibility_after_a_segwit/