Blockchain Access Privacy: Challenges and Directions

Article *in* IEEE Security and Privacy Magazine · July 2018

DOI: 10.1109/MSP.2018.3111245

CITATIONS
65

READS
956

3 authors, including:

Amir Herzberg
University of Connecticut
266 PUBLICATIONS **6,808** CITATIONS

SEE PROFILE

Blockchain Access Privacy: Challenges and Directions

Ryan Henry, Amir Herzberg, and Aniket Kate

Abstract—Privacy, facilitated by a confluence of cryptography and decentralization, is one of the primary motivations for the adoption of cryptocurrencies like Bitcoin. Alas, Bitcoin's privacy promise has proven illusory and, despite growing interest in privacy-centric blockchains, most blockchain users remain susceptible to privacy attacks that exploit network-layer information and access patterns which leak as users interact with blockchains. Understanding if and how blockchain-based applications can provide strong privacy guarantees is a matter of increasing urgency. Many researchers advocate using anonymous communications networks, e.g., Tor, to ensure access privacy. We challenge this approach, showing the need for mechanisms through which non-anonymous users can (i) publish transactions that cannot be linked to their network addresses or to their other transactions, and (ii) fetch details of specific transactions without revealing which transactions they seek. We hope this article inspires blockchain researchers to think 'beyond Tor' and tackle these important access privacy problems head-on.

Keywords—Privacy; Bitcoin; Anonymity; Tor; Private Information Retrieval

1 Introduction and Motivation

A *blockchain* is a distributed, append-only log of time-stamped records that is cryptographically protected from tampering and revision. In the eight years since blockchains were first proposed, their use as publicly accessible and verifiable ledgers for online financial transactions has become widespread. This rapid adoption has largely been spurred by the success of *Bitcoin*[1], a digital currency that—owing to its decentralized and pseudonymous nature, support for complex financial instruments (enabled by a powerful, built-in scripting language), and capacity to facilitate fast and inexpensive transactions across the globe—has proven to be a highly disruptive force in the finance and e-commerce sectors.

As Bitcoin and alternatives like *Ethereum*[2] and *Ripple*[3] continue to mature and grow in market value, it is becoming increasingly likely that blockchains as a means to facilitate financial transactions are here to stay. Yet blockchains represent far more than a mere monetary innovation; researchers and industry members alike are only just beginning to understand the true potential of blockchain-based distributed ledgers, with their strong integrity and availability guarantees and their ability to leverage community consensus to eschew centralized trusted curation. Indeed, beyond the sorts of payment transactions for which blockchains are already widely deployed, potential applications for blockchains abound in areas as diverse as electronic voting, certificate authorities, the Internet of Things, and smart systems. Moreover, the past few years were marked by announcements from numerous companies—ranging from startups like *R3*[4] to established technology firms like IBM, and financial institutions like Visa—about forthcoming products based on innovative blockchain designs that are specially tailored to meet organizational and business logic needs. The target applications for these products range from payment settlement through supply-chain management and beyond.

Just how private are today's blockchains? The ephemeral nature of users' pseudonymous identities in Bitcoin played a key role in its early success. However, eight years of intense scrutiny by privacy researchers has brought to bear an arsenal of powerful heuristics using which attackers can effectively link disparate Bitcoin transactions to a common user and, in many cases, to that user's real-world identity. Ultimately, instead of providing the bastion of privacy for financial transactions that its early adopters envisioned, Bitcoin and its altcoin brethren are in many ways *less* private than traditional banking, where government regulations mandate basic privacy protections. In an attempt to address this situation, the cryptography and privacy research communities have proposed and implemented several protocols aiming to improve blockchain privacy. These protocols all try to decouple users' pseudonymous identities from the specific transactions they make, thereby frustrating attempts to link transacting parties *based on data that appears in the blockchain*. However, none of the proposed protocols attempts to hide the identities of users from network-level adversaries *as the users publish or retrieve data from the blockchain*. Instead, the proposed protocols 'outsource' this crucial step, relying on an external anonymous communications network such as *Tor*[5]. However, running complex protocols over general-purpose, low-latency anonymity networks such as Tor is fraught with risks, and can expose users to subtle-yet-devastating deanonymization attacks,

[1] https://www.bitcoin.org/
[2] https://www.ethereum.org/
[3] https://ripple.com/
[4] https://www.r3.com/
[5] https://www.torproject.org/

thereby undermining the privacy guarantees of the entire blockchain system. We can do better!

Focus.pdf

Figure 1. Topology of a typical blockchain system. The two bold arrows (highlighted in **green**) illustrate sensitive information flows that must be protected in order to prevent attackers from leveraging network-level information to compromise the privacy of blockchain users.

2 Cryptography to the Rescue?

Most blockchains are, at their core, massively distributed and publicly accessible databases; therefore, beyond ensuring that the *data* they store do not, in and of themselves, betray user privacy, any research program that seeks to fully address blockchain privacy must additionally consider (at the very least) privacy for two fundamental types of transactions: *reading data from* and *writing data to* a blockchain.

In the context of cryptocurrencies like Bitcoin, the database represented by the blockchain is a publicly accessible and verifiable *ledger* of financial transactions. Specifically, whenever a transaction occurs, the originating party publicly *announces* the transaction to a handful of selected entities, who then spread the details of that transaction throughout the network via a gossip protocol. The transaction is ultimately aggregated with several other (unrelated) transactions into a discrete *block*, which then gets irreversibly appended to a *chain* comprising all earlier blocks. The chain of blocks can—indeed, to obtain strong integrity and availability, *must*—be replicated and shared in its entirety among many nodes in a network, thereby providing each node with a global, eventually consistent view of every transaction that has ever taken place. New transactions are reflected in all replicas of the blockchain within some predefined expected time, which can range from a few seconds (e.g., in Ripple) to a few minutes (e.g., in Bitcoin).

Each transaction is associated with a pair of *pseudonyms* (often called *wallets*), respectively identifying the sender and receiver of some digital assets. Users can generate new pseudonymous wallets with which to receive digital assets arbitrarily and at will; indeed, it is considered a best practice for Bitcoin users to generate a fresh, ephemeral wallet whenever they wish to conduct a new transaction. The primary motivation for generating such ephemeral wallets is to protect user privacy by making it difficult for an attacker to link together the various transactions involving a given user by simply examining the sender and receiver pseudonyms appearing in transactions recorded in the ledger. However, as Bitcoin and related altcoins grow ever-more prevalent, there is a growing concern that the "privacy" offered by this approach is illusory at best. Indeed, as mentioned previously, the past eight years of research into blockchain privacy has given rise to a veritable treasure trove of effective heuristics using which attackers can link Bitcoin transactions back to a common user, despite the widespread use of ephemeral wallets [1]–[3].

Figure 1 depicts a traditional blockchain architecture. (We use the qualifier "traditional" here to differentiate the blockchain architectures we consider from those involving *payment channels* and other layer-2 applications, which introduce a host of new privacy concerns that go beyond the scope of this article.) For the purposes of this article, we focus on the two arrows that are **bolded** and highlighted in **green**; specifically, we focus on the need for innovative mechanisms that allow users to

(i) *announce and publish transactions anonymously*, a task for which we envision a tailor-made anonymity mechanism that is integrated directly into the blockchain architecture; and to

(ii) *fetch transactions privately*, a task for which we envision using special private information retrieval (PIR) protocols designed and optimized to support efficient and expressive queries for transactions stored in a blockchain.

We note that a handful of second-generation altcoins—including *Zcash*[6] and *Monero*[7]—natively employ cryptographic techniques to prevent the *contents* of transaction on the blockchain from leaking private information about transacting parties. Likewise, the research literature contains several proposals (a selection of which we summarize in the next subsection) that aim to provide similar transaction privacy atop the deployed Bitcoin, Ripple, and Ethereum blockchains. While such approaches are indeed effective at protecting blockchain users against a subset of the deanonymization heuristics that plague mainstream deployed blockchains, we emphasize that the existing approaches, so far, focus on preventing the *data* stored in a blockchain from leaking private information—they do nothing significant to mitigate against inferences that leverage *network-level information* (e.g., IP addresses) or *access patterns* (e.g., specific blocks or portions thereof) revealed when users interact with the blockchain data. As such, the existing proposals all fall far short of solving the blockchain privacy problem in its entirety.

2.1 Existing protocols for transaction privacy

As the insufficiency of ephemeral pseudonyms became apparent to the Bitcoin community, a proposal called *CoinJoin* emerged as a potential solution. In CoinJoin, users route their transactions through a centralized *mixing service* (sometimes called a *tumbler*), which serves to obscure the relationships between the senders and receivers of those transactions before they are posted to the ledger. However, such centralized mixing services introduce a single point of trust and failure; indeed, the mixing service always knows the link between the sender and receiver of

[6] https://z.cash/

[7] https://getmonero.org/

each transaction and, perhaps more troublingly, there is nothing to stop the mixing service from stealing assets that users try to route through it. A series of progressively more sophisticated protocols have been proposed to address CoinJoin's limitations.

The first improvement was *Mixcoin*, which attempts to mitigate the risk of theft by holding the mixing service "accountable" if it steals a user's assets (though theft is still technically possible and the mixing service still learns who is transacting with whom). Building on a series of incremental improvements to this basic idea (including *BlindCoin* and *Blindly Signed Contracts*), a proposal called *TumbleBit* [4] finally addressed the accountability and anonymity weaknesses of Mixcoin in a manner fully compatible with Bitcoin; however, the TumbleBit approach requires upwards of 20 minutes (i.e., two Bitcoin block) per transaction on average and introduces additional transaction fees. The third author's own *CoinShuffle* and *CoinShuffle++* [5] take a different approach, having users perform a special multi-party computation among themselves so that no third-party mixing service is necessary.

The emerging privacy-centric cryptocurrencies, such as Zcash and Monero, employ cryptographic primitives such as zero-knowledge succinct non-interactive argument of knowledge (zk-SNARK), traceable ring signatures, confidential transactions and stealth addresses to offer significantly better privacy properties than those possible for Bitcoin transactions.

2.2 Inadequacy of existing proposals

The above transaction privacy protocols all aim to sever the link between senders and receivers as recorded in the transactions that get published to the blockchain. However, the approaches are all susceptible to attacks that reestablish links between transacting parties using network-level information and/or access patterns observed both as users announce their transactions and as they probe the blockchain to learn which of their transactions have posted to the ledger [2]. For example, an attacker who observes that a given user visits a website immediately before that website receives a donation via Zcash or Monero might surmise that the user made the donation; moreover, the attacker can all-but-confirm this suspicion if it later observes the same user checking whether the transaction in question has posted to the ledger.

To define away this elephant in the room, the developers of such privacy protocols typically assume that users communicate over an anonymous-communication protocol such as Tor; in fact, some privacy-centric altcoins—like Zcash, Anoncoin, and Torcoin—include native support for Tor and expect that all users interact with their blockchains *exclusively* through Tor. As an example, the Zcash website[8] clearly states (and we quote) that "a unique IP address can allow network observers to correlate your Zcash transactions with each other and with your other traffic" to which it adds that "advanced users may opt to connect through Tor to obfuscate their node's IP address, however, further exploration is needed on a vulnerability combining Bitcoin's Denial of Service mitigation (inherited into Zcash) and anonymous communication networks like Tor before we can recommend users who are not familiar with the attack to route their Zcash nodes through Tor."

This dependency on Tor for anonymity introduces some rarely-acknowledged-yet-undeniably-troubling weaknesses. One source of weakness stems from the fact that Tor is specifically designed to support *low-latency* communication, such as interactive web browsing and real-time instant messaging; indeed, it seems inherent (and real-world attacks seem to confirm) that such low-latency low-bandwidth anonymous communication systems can provide at most a relatively weak form of anonymity compared to high-latency approaches like Chaumian mix networks or high-bandwidth approaches like dining-cryptographers (DC) networks. Indeed, a recent paper by Das et al. [6] analyzed the so-called "anonymity trilemma" and concluded that, in the presence of a global passive (network-level) adversary, anonymous communications networks can hope to provide just two of three desirable properties: strong anonymity, low bandwidth overhead, and low latency overhead. Fortunately, because financial transactions are naturally able to tolerate moderate latency—indeed, so-called "permissionless blockchains", like the one used in Bitcoin, already impose latencies on the order of *several minutes* even without the use of an anonymous communications network—users need not settle for the relatively weak anonymity guarantees that low-latency systems like Tor can provide.

Further, Biryukov and Pustogarov [7] demonstrated how Bitcoin's "blacklisting" measures may ultimately leave users conducting Bitcoin transactions over Tor *more* vulnerable to active deanonymization attacks than those announcing their transaction non-anonymously. They describe man-in-the-middle attacks that exploit the Bitcoin network's built-in reputation-based DoS protection mechanism to force specific Bitcoin peers to ban Tor exit relays of the attacker's choice, thus forcing *all* Bitcoin traffic to exit the Tor network through a small set of attacker-controlled relays. Once in this privileged position, the attacker can launch several troubling privacy attacks, including deanonymization via traffic correlation (which is made easier because the attacker automatically controls one end of the communication), correlating multiple wallet addresses to a common user, and launching "double-spending" attacks by lying to thin clients about previous transactions involving a given wallet address.

Yet another problem arises from the fact that Tor is often blocked by IT departments within organizations or even subject to state-level censorship by authoritarian governments. This has direct negative consequences for the privacy of users connecting from such organizations or countries, even though the censorship is almost certainly intended to quell some other, unrelated usage of

[8] https://z.cash/support/security/privacy-security-recommendations.html

Tor. As a workaround for such censorship, Tor ships with support for some censorship-evasion techniques including *Tor bridges* and *pluggable transports*; however, the effectiveness of these mechanisms is far from perfect and censorship events continue to affect Tor users. In general, it seems unwise to advocate the wholesale use of censorship circumvention tools for activities that are typically not subject to censorship.

Moreover, a third-party anonymous communication network such as Tor may not be willing or able to support blockchain traffic on a large scale. A dual concern is some blockchain systems may be hesitant to use Tor since Tor has been used for nefarious purposes, ranging from ransomware and botnet command and control through to child pornography. As an anecdotal example of this, the third author has learned through communications with developers at Ripple that, despite being very keen on improving privacy for their clients, Ripple's developers are unwilling to leverage a Tor-based solution to do so.

Finally, due to their decentralized design, blockchain systems seem like prime candidates for fulfilling their own anonymity and privacy needs, avoiding the dependency on external services and providing performance and privacy/anonymity guarantees tailored to their own needs.

In short, we believe that effective blockchain privacy necessitates rethinking the one-size-fits-all approach of using external anonymous communications infrastructures to solve all problems requiring anonymity. Although anonymity does indeed love company, mixing two dissimilar types of traffic together does not necessarily improve anonymity for either type and, if not done very carefully and correctly, may in fact provide weaker anonymity than protecting each type of traffic with its own tailor-made solution.

3 Publishing Transactions Anonymously

By their very design, blockchain systems require extensive overlay networks through which participants announce transactions and agree on what transactions should ultimately appear on the blockchain. Thus, it seems natural to leverage the existing overlay structure to realize anonymous transaction publishing, rather than relying on an external service like Tor. We propose that blockchain privacy protocols should de-link users' network-level information from their transactions using mechanisms that piggybacks on the overlay network that is already in place for announcing transactions. The specifics of how such a mechanism might work vary, depending on the structure of the overlay network imposed by the *consensus protocol*—that is, depending on how participants decide which transactions qualify for inclusion in the blockchain.

Permissionless versus permissioned blockchains. Proposed and deployed blockchains fall into two distinct categories based on the mechanism they use to build a consensus around what data to immortalize in the blockchain: *permissionless blockchains* and *permissioned blockchains*.

The blockchains underlying Bitcoin and Ethereum constitute two prominent examples of *permissionless blockchains*. As their name implies, permissionless blockchains place no restrictions on who participates in the consensus process. Instead, unrestricted entities called *miners* collectively decide which blocks should be appended to the chain by providing an associated proof of work. In the case of Bitcoin, this proof of work takes the form of a "partial hash inversion", wherein the miners seek inputs that lead a cryptographic hash function to produce a digest whose numerical value does not exceed some global-parameter target.

Such a permissionless consensus guarantees that only valid blocks get appended to the blockchain (approximately) under the assumption that more than half of all mining resources in the network are controlled by honest—or, at least, non-colluding—entities.

The blockchains underlying Ripple and the Linux Foundation's *Hyperledger*[9] are two prominent examples of *permissioned blockchains*. In contrast to permissionless blockchains, permissioned blockchains do place restrictions on who participates in the consensus process. A group of highly available entities (with strong identities) collectively decide which blocks should be appended to the chain by leveraging a Byzantine fault-tolerant atomic broadcast protocol. This approach allows permissioned blockchains to reach consensus very rapidly, requiring as little as a few seconds for each transaction to be reflected in the ledger.

The contrasting security assumptions and efficiency guarantees of permissionless and permissioned blockchains make them well suited to different use cases and, indeed, the two varieties are prospering together: traditionally structured organizations/consortiums are increasingly adopting permissioned blockchains, while peer-to-peer solutions continue to leverage permissionless blockchains.

3.1 Publishing to permissionless blockchains

Permissionless blockchain systems (like Bitcoin and Ethereum) employ peer-to-peer (P2P) networks of relays to propagate transactions and blockchain updates throughout the network using a best-effort gossip protocol. Such P2P networks typically experience considerable churn, with relays joining, leaving, and rejoining the network at will; however, the average number of relays in the network at any given time can remain relatively high. For example, at the time of writing, the number of online relays in the Bitcoin network at any given time is about one-and-a-half times the number of Tor relays. (As of October 4, 2017, *Tor Metrics*[10] estimates about 6700 Tor relays versus the *Bitnode*[11]

[9] https://www.hyperledger.org/
[10] https://metrics.torproject.org/
[11] https://bitnodes.21.co/

estimate of about 9600 full Bitcoin nodes.) One might, therefore, consider employing the elaborate Bitcoin communication infrastructure toward improving the anonymity of users' announcements. Given the P2P nature of the network, we believe it may be possible to leverage the existing academic research on P2P anonymous communications networks. For instance, such a solution could be based upon Pisces [8], employing the social trust links to construct anonymous communication paths that are robust to compromise in the presence of route-capture attacks and Sybil nodes. However, given the dynamic and open nature of permissionless blockchains such as Bitcoin, establishing trust in relays will be a prominent challenge.

The *Kovri project*[12], an offshoot of the Monero and Bitcoin developers' recent interest in the Dandelion networking policies [9], clearly indicate the blockchain community's awareness of the problem; nevertheless, significantly efforts are necessary going forward. In general, it will be an interesting challenge to analyze and establish security, privacy, and viability of realizing a P2P anonymous communications system over permissionless blockchain systems.

3.2 Publishing to permissioned blockchains

Permissioned blockchain systems (like Ripple, *Corda*[13], and Hyperledger) employ a clique of highly available validator nodes for agreeing on transactions and blocks. These nodes employ traditional asynchronous Byzantine-tolerant consensus protocols to append a block of transactions to the blockchain. Here, validators select valid transactions to be agreed upon from those transactions forwarded by the users of the system. As typically transactions from several users are added to any given block, a simple approach to provide anonymity here will be to perform all the communication between users and validators over an anonymous communications network. However, we advocate improving efficiency and reducing the overhead by combining the consensus process for agreeing on transactions with the process of mixing users' announcements.

This problem can be modeled as an asynchronous multi-party computation (AMPC) problem, and can be solved using the generic AMPC techniques; however, we propose development of tailored solutions to further improve the efficiency. A possible tailored approach for agreeing on a randomly permuted set of transactions can involve combining Newton's identity method for power sums (as employed by Ruffing et al. [5]) with asynchronous verifiable secret sharing and asynchronous Byzantine consensus. Nevertheless, a key challenge will be to make these solutions scale well (possibly sublinearly) with the number of mixed transactions.

4 Fetching Transactions Privately

Blockchains differ from traditional databases in their use of cryptography as a means to eschew both centralization and trusted curators, all the while ensuring strong resistance to "tampering" (i.e., history rewriting). Yet this remarkable combination of attributes is guaranteed only for users that hold a complete local replica of the blockchain. With a blockchain currently over 100 GB and growing, this local-storage requirement is quickly becoming infeasible for casual Bitcoin users; as a result, many such users now employ so-called *thin clients*, which bypass the need to hold a local copy of the blockchain by forwarding blockchain queries to semi-trusted intermediaries.

Specifically, thin clients run in what is called *Simplified Payment Verification (SPV) mode*—so named after the section of the original Bitcoin whitepaper [10] that details it—wherein the initial syncing process connects to an arbitrary *full node* and downloads only the block headers (each of which includes a Merkle root committing to the actual block). The thin client then verifies that the given headers indeed form a blockchain (with sufficient difficulty value), after which they can request the details of transactions matching certain patterns (e.g., payments to or from particular addresses) from any full node. The full nodes reply to such requests with a copy of any relevant transactions together with Merkle branches linking those transactions to their associated block headers. This process exploits the Merkle tree structure to allow proofs of inclusion in a block without needing to provide the thin client with the full contents of the block.

The SPV approach has the distinct advantage that the cost of initial syncing scales linearly with the length of the blockchain (about 80 bytes per header, or 4.2 MB per year) and is independent of the size of the actual blocks. However, a naive implementation of SPV exposes thin clients to potentially devastating attacks on privacy. As a thin client will typically request details about precisely those transactions that correspond to keys it owns, it may end up revealing to the full node a complete list of its public addresses. In particular, *Bitcoin users that rely upon such thin clients are subject to deanonymization*. This is a serious risk; there have been numerous reports of high-rolling Bitcoin users being identified and targeted by miscreants to steal their digital fortunes.[14]

A tempting response is to route thin-client queries through an anonymity network like Tor; however, this leaves clients susceptible to low-cost deanonymization and double-spending attacks [7]. Indeed, the root problem for thin clients is not a lack of anonymity for the *querier* but, rather, a lack of privacy for the *queries*—anonymity, quite simply, solves the wrong problem. Instead, we observe that the problem of realizing private

[12] https://getkovri.org
[13] https://www.corda.net/
[14] https://bitcointalk.org/index.php?topic=16457.0

blockchain queries is imminently solvable using a well-known cryptographic primitive called *private information retrieval (PIR)*.

PIR is a cryptographic primitive that solves the seemingly impossible problem of letting clients query a remote database, while not exposing the clients' query terms or the responses they generate to the database operator. PIR has received considerable attention from the cryptography, privacy, and theoretical computer science research communities. Alas, despite a series of significant advances over the past two decades, existing PIR techniques are notoriously inefficient and, consequently, to date not one of the numerous PIR-based applications proposed in the research literature has been deployed at-scale to protect the privacy of users "in the wild".

As a result, transitioning the idea of using PIR to fetch blockchain transactions privately into practice still necessitates some basic research and rather substantial engineering and implementation efforts. Fortunately, some recent advances in PIR research yield the promise of PIR protocols that are sufficiently practical to deploy on databases of size commensurate with Bitcoin's blockchain.

4.1 Private blockchain queries from PIR

The key goals here are to create protocols that enable thin clients to (i) determine if particular transactions are reflected in the blockchain (and, if so, how many blocks have been appended since, a rough proxy for the computational effort that would be required to "undo" that transaction), and (ii) find out the balances associated with a set of public keys, reflecting all transactions that have occurred so far involving those keys.

This will involve defining appropriate data structures that lend themselves to being queried via PIR, as well as efficient mechanisms for keeping those data structures up to date as the blockchain grows. Although one could conceptually employ any PIR protocol for this purpose, thinking towards mass adoption among the millions of present and potential Bitcoin users, we suggest very strict requirements on acceptable communication and computation overhead. In effect, the target will be communication costs that are reasonable for a smart phone communicating over a mobile data connection, and computation costs low enough for a modestly equipped server to process tens or hundreds of queries every second. Such strict requirements preclude most existing PIR protocols; however, the recent introductions of (i) *distributed point functions* [11], (ii) Intel's *software guard extensions* (*SGX*) architecture[15], and (iii) the first author's *indexes of queries* [12] provide three very elegant—and, we believe, highly practical—ways to realize the kinds of PIR-based private blockchain queries we envision. Each approach brings its own performance characteristics and its own security assumptions, ranging from non-collusion, through computational assumptions, to trusted hardware. The research objective here will be to devise appropriate data structures to facilitate PIR-based queries over blockchain data, and then to implement and evaluate the suitability of the various approaches.

Moreover, by leveraging the anonymous communications framework we advocated in Section 3, it may be possible to realize lower-cost relaxations of information-theoretic PIR that satisfies a differentially private notion for private queries [13].

5 Concluding Remarks

General-purposes anonymous communications systems like Tor are not a panacea for communication privacy issues. Indeed, not all applications are anonymized equally well by low-latency anonymity networks, and not all privacy problems are adequately addressed by making users anonymous. In this article, we highlighted two prominent communication privacy issues that afflict current blockchain solutions: the problems of announcing blockchain transaction anonymously and fetching blockchain transactions privately. We proposed research directions that shift from the current norm of just saying 'do it over Tor' and instead seek to tackle these important problems head-on. In particular, for the problem of announcing blockchain transaction anonymously, we suggested to leverage blockchain consensus infrastructures instead of the external, general-purposes networks like Tor, while for the problem of fetching transaction privately, we offered directions towards making private information retrieval (PIR) schemes suitable and efficient for blockchain transactions.

While we only considered ways to address privacy challenges arising from network-level and access pattern leakage on traditional blockchains, new blockchain extensions—such as the *lightning network*[16], which has been recently proposed as a way to greatly improve the scalability of permissionless blockchains—introduce new subtle privacy challenges that will also require novel solutions. Although some solutions are already emerging towards improving privacy in these path-based transactions [14, 15], it is an interesting open challenge to devise scalable mechanisms for performing (multi-hop) payment-channel transactions *privately* against a network-level adversary.

Acknowledgements. This material is based upon work supported by the National Science Foundation under Grant Numbers 1718595 and 1719196, and by United States-Israel Binational Science Foundation (BSF) under Grant Number 2016718.

References

[1] Meiklejohn, M. Pomarole, G. Jordan, K. Levchenko, D. McCoy, G. M.Voelker, and S. Savage, "A fistful of Bitcoins: Characterizing payments among men with no names," *Communications of the ACM*, vol. 59,no. 4, pp. 86–93, 2016. Available: https://doi.org/10.1145/2896384

[15] https://software.intel.com/sgx

[16] https://lightning.network/

[2] Koshy, D. Koshy, and P. D. McDaniel, “An analysis of anonymity in Bitcoin using P2P network traffic,” in *Proceedings of FC 2014*, ser. LNCS, vol. 8437, Christ Church, Barbados, March 2014, pp. 469–485.Available: https://doi.org/10.1007/978-3-662-45472-5_30

[3] E. Androulaki, G. Karame, M. Roeschlin, T. Scherer, and S. Capkun, “Evaluating user privacy in Bitcoin,” in *Proceedings of FC 2013*, ser. LNCS, vol. 7859, Okinawa, Japan, April 2013, pp. 34–51. Available: https://doi.org/10.1007/978-3-642-39884-1_4

[4] E. Heilman, L. Alshenibr, F. Baldimtsi, A. Scafuro, and S. Goldberg, “TumbleBit: An untrusted Bitcoin-compatible anonymous payment hub,” in *Proceedings of NDSS 2017*, San Diego, CA, USA, February–March 2017. Available: https://doi.org/10.14722/ndss.2017.23086

[5] T. Ruffing, P. Moreno-Sanchez, and A. Kate, “P2P mixing and unlinkable Bitcoin transactions,” in *Proceedings of NDSS 2017*, San Diego, CA, USA, February–March 2017. Available: https://doi.org/10.14722/ndss.2017.23415

[6] D. Das, S. Meiser, E. Mohammadi, and A. Kate, “Anonymity trilemma:Strong anonymity, low bandwidth overhead, low latency—choose two,” in *Proceedings of IEEE S&P 2018* (to appear), San Francisco, CA, USA, May 2018. Available: https://eprint.iacr.org/2017/954

[7] A. Biryukov and I. Pustogarov, “Bitcoin over Tor isn’t a good idea,” in *Proceedings of IEEE S&P 2015*, San Jose, CA, USA, May 2015, pp.122–134. Available: https://doi.org/10.1109/sp.2015.15

[8] P. Mittal, M. K. Wright, and N. Borisov, “Pisces: Anonymous communication using social networks,” in *Proceedings of NDSS 2013*, San Diego, CA, USA, February 2013. Available: https://arxiv.org/abs/1208.6326

[9] S. Bojja Venkatakrishnan, G. Fanti, and P. Viswanath, “Dandelion: Redesigning the bitcoin network for anonymity,” In *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 1, no. 1, pp. 22:1–22:34, June 2017. Available: http://doi.acm.org/10.1145/3084459

[10] S. Nakamoto, “Bitcoin: A peer-to-peer electronic cash system,” https://bitcoin.org, Tech. Rep., November 2008. Available: https://bitcoin.org/bitcoin.pdf

[11] N. Gilboa and Y. Ishai, “Distributed point functions and their applications,” in *Proceedings of EUROCRYPT 2014*, ser. LNCS, vol. 8441, Copenhagen, Denmark, May 2014, pp. 640–658. Available: https://doi.org/10.1007/978-3-642-55220-5_35

[12] S. M. Hafiz and R. Henry, “Querying for queries: Indexes of queries for efficient and expressive IT-PIR,” in Proceedings of CCS 2017, Dallas, TX, USA, October–November 2017, pp. 1361–1373. Available: https://doi.org/10.1145/3133956.3134008

[13] R. R. Toledo, G. Danezis, and I. Goldberg, “Lower-cost ε-private information retrieval,” in *Proceedings on Privacy Enhancing Technologies (PoPETs)*, vol. 2016(4), Darmstadt, Germany, October 2016, pp. 184–201. Available: https://doi.org/10.1515/popets-2016-0035

[14] G. Malavolta, P. Moreno-Sanchez, A. Kate, M. Maffei, and S. Ravi, “Concurrency and privacy with payment-channel networks,” in *Proceedings of CCS 2017*, October–November 2017. Available: https://doi.org/10.1145/3133956.3134096

[15] M. Green and I. Miers, “Bolt: Anonymous payment channels for decentralized currencies,” in Proceedings of CCS 2017, October–November 2017, pp. 473–489. Available: http://doi.acm.org/10.1145/3133956.3134093

Dr. Ryan Henry is an Assistant Professor in the School of Informatics, Computing, and Engingeering at Indiana University Bloomington. His focuses on the systems challenges of applied cryptography, with an emphasis on using cryptography to build secure systems that protect the privacy of their users. Contact him at mailto:henry@indiana.edu.

Dr. Amir Herzberg is a Comcast professor for security innovation at the Dept. of Computer Science and Engineering, University of Connecticut. Previously he was with Bar Ilan University and with IBM Research. His research areas include network and web security, applied cryptography, privacy and anonymity, usable security and security for cyber-physical systems. Contact him at mailto:amir.herzberg@uconn.edu.

Dr. Aniket Kate is an Assistant Professor in the computer science department at Purdue university, USA. He designs, implements, and analyzes privacy and transparency enhancing technologies. Contact him at aniket@purdue.edu.