Research Statement - Privacy Preserving Decentralized Systems

Gonçalo Pestana (gpestana@protonmail.com)

Dec 2018

Research goals

As a researcher and engineer, my long term goals are to study and build privacy preserving technologies that are secure and private by design. Given how pervasive Internet is in today's society, I believe that a free and ethical Internet is synonym with a free and ethical society. Thus, I am committed to help building an Internet that respects our fundamental rights to privacy and freedom.

Decentralized and P2P systems have demonstrated their potential [ref, ref] as the underlying paradigm for building online systems that respect it users' privacy. It has been shown [ref, ref snowden] the urgency to replace, complement or fix current online systems with respect to protecting users against private data and metadata leaks [ref], pervasive online surveillance [ref] and online censorship [ref]. While decentralized and P2P systems may be part of the solution, it has been discussed [1] and demonstrated [2, 3] that naive implementation of decentralized systems may harm privacy more than centralized systems. I am highly motivated to learn how current decentralized systems may harm user privacy and how to develop privacy enhancing mechanisms to address those problems. Moreover, I am interested in designing and developing privacy preserving and decentralized alternatives to current systems deployed at large scale, such as communication systems, authentication infrastructure, search engines and content discovery, and user data analytics infrastructure.

As P2P networks and decentralized systems (re)gain popularity among researchers and industry alike, I believe that it is important to design and implement decentralized systems not only preserve users' privacy, but also deliver on scalability, performance and usability. Failing to deliver on those properties will render decentralized systems unusable and unattractive for mass adoption or as a viable alternative to centralized systems.

The short term goals of my research are to study, design and help to implement privacy enhancing mechanisms and protocols for decentralized systems. The focus would be on studying and implementing privacy in the building blocks of P2P systems, while considering their security, scalability, performance and usability. From this starting point, I would expect the research to cover topics such as distributed systems, applied advanced cryptography (e.g. zero knowledge proofs, multi-party computation, threshold crypto, homomorphic encryption) and incentive design for agent collaboration in P2P networks.

Initial research directions

This section outlines ideas for initial directions:

Privacy preserving and censorship resistant search engines: Search engines have been (literally) the engine of the Internet and are the backbone of online services. Due to their importance, there has been interest in designing privacy preserving and censorship resistant web search engines (Lai et al. 2018) and to study current implementations regarding their security and privacy (Herrmann et al. 2014). There are, though, many challenges ahead to fulfill the vision of decentralized, secure, private and censorship resistant search engines: storage and communication constraints (Li et al. 2003), indexing and ranking mechanisms that suit the decentralized context, routing and scalability complexity (Herrmann et al. 2014),

and scalability are examples of the current open problems.

This research topic would consist of 1) investigating the challenges of indexing, discovering and distributing content in a decentralized networks; 2) study and design mechanisms for privacy preserving and censorship resistant search engines; 3) build primitives for indexing and querying data in a decentralized system at scale. I expect this topic to require research on anonymous computation techniques (e.g. homomorphic computation), privacy preserving content discovery and routing, anonymous communication systems and incentive design in P2P networks.

Practical privacy preserving DHTs: (Wang and Borisov 2012), (Mittal and Borisov 2009), studied protocols and mechanisms to achieve secure and anonymous low latency communication systems. Many other research studies focus (Shirazi et al. 2018) on the the same topic. However, real-world implementation of DHTs such as, for example, IPFS (Benet 2014) and Hyperswarm do not seem to adopt any of the privacy preserving mechanism studied in recent literature. My hypothesis are the following: 1) the penalty to paid for the added complexity and protocol performance is large; 2) current secure and private protocols assume centralized infrastructure (e.g PKI infrastructure); 3) there is no incentive and trust models to make sure peers collaborate to enhance the network privacy. I propose to study and measure privacy of real-world systems using DHTs. Based on the results and previous research, investigate novel privacy enhancing mechanisms, protocols and primitives that could improve privacy of DHTs, taking into consideration scalability, performance and security.

Systematization and privacy vulnerability research: (Troncoso et al. 2017) shows how lack of systematization on how to define and measure privacy in decentralized systems make it hard to design and implement such systems. Similarly to [tor, akamai, freenet, refs], I propose to investigate how data and metadata are leaked in decentralized and centralized systems. The main goals are 1) define a framework to study privacy in decentralized networks 2) to investigate and enumerate privacy vulnerabilities in current systems; 3) design and implement protocols and/or primitives to address the privacy vulnerabilities found.

From centralized to privacy preserving decentralized systems: Decentralized networks are becoming sound alternatives to centralized systems due to advances in P2P and incentive protocols, the constant increase of computation power and storage in edge devices and recent public cases disclosing how current systems are harming user privacy and security. Ledgers (Nakamoto, n.d.), contracts (Buterin 2013) file systems (Benet 2014), file sharing ("Bittorrent," n.d.), in-band PGP key distribution (Kulynych et al. 2018) are examples of decentralized systems which aim to replace analogous centralized approaches. However, as (Troncoso et al. 2017) shows, while decentralization can potentially improve privacy, integrity and availability, naive implementations may be counterproductive in regards to those properties. ClaimChain (Kulynych et al. 2018) is a sound example of how to replace centralized and brittle infrastructure with advanced cryptographic primitives, while adding interesting security and privacy properties which can be used in a decentralized context.

This research topic would focus on answering the question of "which centralized infrastructure can be replaced by privacy preserving decentralized systems and how?". PKI infrastructure [ref], search engines [ref], DNS infrastructure [ref] are examples of infrastructure that could improve in terms of privacy, availability and censorship resistance with a decentralized design.

Incentives in privacy preserving decentralized systems: Cryptocurrencies and smart contracts have - at least theoretically - shown the potential of game theory and incentive design for aligning interests of participants in P2P networks (Buragohain, Agrawal, and Suri, n.d.), (Park and Schaar 2010), (Ciccarelli and Cigno 2011), (Nakamoto, n.d.). Instead of focusing exclusively on incentives in decentralized networks as research topic, I believe it to be transversal to all the research directions mentioned previously.

Open research topics: I am open to relevant topics focusing on privacy preserving networks, privacy

enhancing technologies (PETs) and applied cryptography in the context of PETs.

References

Benet, Juan. 2014. "IPFS - Content Addressed, Versioned, P2p File System."

"Bittorrent." n.d. https://bittorrent.com.

Buragohain, C., D. Agrawal, and S. Suri. n.d. "A Game Theoretic Framework for Incentives in P2p Systems." *Proceedings Third International Conference on Peer-to-Peer Computing (P2P2003)*. IEEE Comput. Soc. https://doi.org/10.1109/ptp.2003.1231503.

Buterin, Vitalik. 2013. "Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform." https://github.com/ethereum/wiki/wiki/White-Paper.

Ciccarelli, Gianluca, and Renato Lo Cigno. 2011. "Collusion in Peer-to-Peer Systems." *Computer Networks* 55: 3517–32.

Herrmann, Michael, Ren Zhang, Kai-Chun Ning, Claudia Díaz, and Bart Preneel. 2014. "Censorship-Resistant and Privacy-Preserving Distributed Web Search." *14-Th IEEE International Conference on Peer-to-Peer Computing*, 1–10.

Kulynych, Bogdan, Wouter Lueks, Marios Isaakidis, George Danezis, and Carmela Troncoso. 2018. "ClaimChain." *Proceedings of the 2018 Workshop on Privacy in the Electronic Society - WPES'18*. ACM Press. https://doi.org/10.1145/3267323.3268947.

Lai, Ziliang, Chris Liu, Eric Lo, Ben Kao, and Siu-Ming Yiu. 2018. "Decentralized Search on Decentralized Web."

Li, Jinyang, Boon Thau Loo, Joseph M. Hellerstein, M. Frans Kaashoek, David R. Karger, and Robert Morris. 2003. "On the Feasibility of Peer-to-Peer Web Indexing and Search." In *Peer-to-Peer Systems Ii*, edited by M. Frans Kaashoek and Ion Stoica, 207–15. Berlin, Heidelberg: Springer Berlin Heidelberg.

Mittal, Prateek, and Nikita Borisov. 2009. "ShadowWalker: Peer-to-Peer Anonymous Communication Using Redundant Structured Topologies." In *Proceedings of the 16th Acm Conference on Computer and Communications Security*, 161–72. CCS '09. New York, NY, USA: ACM. https://doi.org/10.1145/1653662.1653683.

Nakamoto, Satoshi. n.d. "Bitcoin: A Peer-to-Peer Electronic Cash System."

Park, Jaeok, and Mihaela van der Schaar. 2010. "A Game Theoretic Analysis of Incentives in Content Production and Sharing over Peer-to-Peer Networks." *IEEE Journal of Selected Topics in Signal Processing* 4: 704–17.

Shirazi, Fatemeh, Milivoj Simeonovski, Muhammad Rizwan Asghar, Michael Backes, and Claudia Díaz. 2018. "A Survey on Routing in Anonymous Communication Protocols." *ACM Comput. Surv.* 51: 51:1–51:39.

Troncoso, Carmela, George Danezis, Marios Isaakidis, and Harry Halpin. 2017. "Systematizing Decentralization and Privacy: Lessons from 15 Years of Research and Deployments." *CoRR* abs/1704.08065. http://arxiv.org/abs/1704.08065.

Wang, Qiyan, and Nikita Borisov. 2012. "Octopus: A Secure and Anonymous DHT Lookup." *CoRR* abs/1203.2668. http://arxiv.org/abs/1203.2668.