Privacy Preserving Distributed Hash Table [DRAFT]

Gonçalo Pestana, Hashmatter (goncalo@hashmatter.com)

Dec 2018

Abstract

Distributed Hash Tables (DHT) are an important building block for the decentralized web and P2P systems. DHTs are overlay networks that allow peers to store and lookup information stored in a decentralized network, where participant nodes are responsible for routing requests and storing data. As a result of its P2P nature, naive design and implementation of DHTs leak user behavior information to other participants. Questions such as "who is participating in the network?"; "who is requesting a particular set of data"; and "who is storing and providing a particular set of data?" can be answered by anyone with relatively low resources. The data leaked by DHTs can be used to target and censor individuals, communities and services. On the other hand, privacy preserving networks are designed and implemented in such ways that user privacy is protected against targeted and mass surveillance and censorship. We argue that 1) the comeback of decentralized networks is a crucial opportunity for building privacy preserving networks from the ground-up and; 2) that DHTs are an important primitive in the decentralization landscape and thus implementations of the DHTs should take user privacy in consideration. This work aims at building the foundations to research and implement privacy preserving DHTs. We review protocols and techniques used to design and build secure and privacy preserving DHTs and their current implementations. We also discuss future directions for design and implementation of privacy preserving DHTs, taking into considerations how it impacts properties such as decentralization, scalability and performance of the network. We conclude that …

1. Introduction

- decentralized systems and user privacy
  - decentralized systems pose new threats to anonymity (Greschbach, Kreitz, and Buchegger 2012)
- DHTs as building block to the decentralized web
- what are DHTs
- examples applications
- However, since nodes need to collaborate to route requests and store data, a lot of data is leaked.
- How to design and build privacy preserving DHTs?
- measure performance and privacy (threat model)

The remaining paper is structured as follows. First, we describe how a general purpose DHT works and outline the most popular protocols used nowadays and their respective applications. Secondly, we describe the threat model of DHTs from a security and privacy perspectives. Thirdly, we review protocols and techniques that aim at preserving complete or a subset of privacy properties in DHTs. Finally, we outline open questions and future research directions.

X. Distributed Hash Tables

- any peer can participate in the network
- constrained overlay vs non-constrained

X. Threat model

- high level properties of a privacy preserving DHT (use well known treat models e.g. used by cwitch)

- hide not only the content but also the context of the information.

1) hide initiator information;
2) hide source information;
3) safe against security attacks;
4) decentralized - not relying on centralized services/registries;

X.x Secure DHTs

X.x Privacy preserving DHTs

X. Privacy preserving protocols and techniques

X.x Octopus DHT

Octopus DHT (Wang and Borisov 2012)

X.x Freenet "Darknet" [restricted, friend-to-friend routing]

X.x Shadow Walker [internal onion]

X.x External mixnets and onion routing

X.x Advanced cryptography primitives [zk, MPC, homomorphic computation]

X. Decentralization, scalability and performance

X. Discussion and future work

Conclusion

References

Greschbach, Benjamin, Gunnar Kreitz, and Sonja Buchegger. 2012. "The Devil Is in the Metadata—New Privacy Challenges in Decentralised Online Social Networks," March.

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