Fall 2019 Project Proposal

Alexander J. Gaidis
Brown University

Abstract

I plan to replace *guarded circuit exchange* in Yodel with a verifiable private information retrieval (PIR) scheme to reduce server costs.

1 Summary

Voice over IP (VoIP) allows two parties to use thier voices to communicate via the internet. By encrypting the packets passed between one another, participants of a conversation can ensure their voice data achieves a desirable level of privacy; an adversary that intercepts packets will only see jumbled mess for the payload. However, in traditional VoIP systems, metadata is leaked which can be just as sensitive as the actual packet payload for activists, journalists, whistle-blowers and others. In the words of General and Former Director of the CIA Michael Hayden, "we [the U.S.] kill people based on metadata" [7]. Systems to hide metadata are not new [1] with the 2nd generation onion router being the most recent approach that gained wide-spread adoption and notoriety. However, following the Snowden revelations of 2013 [3] the problem of metadata hiding has renewed interest in academia.

Most recently, Yodel [2] was published in an attempt to protect metadata in VoIP. This communication system is round-based, using mixnet and cover traffic generation to mask metadata against a strong global adversary (each server has a 20% chance of being controlled which matches the threat profile of Tor). One of the techniques used to produce cover traffic is what the authors dub *guarded circuit exchange*. In this protocol, a user creates two circuits so as to have a backup in case communication on one fails. This way the user's traffic patterns always look the same. However, guarded circuit exchange presents a significant bandwidth cost to running a server due to each user constructing two circuits with millions of possible ongoing voice calls at a time.

This semester, I propose to replace guarded circuit exchange with a verifiable private information retrieval (PIR) protocol to reduce the number of circuits that need to be made and maintained in Yodel. In this solution, each user should only need to maintain one circuit for communication with another party.

My end goal is to work on a project that is a boon to privacy and security while still trying to have it be conference-submission worthy.

2 Investigation Overview

Investigating this problem space will revolve mostly around understanding PIR on a deeper level; metadata-private communication seems relatively straight-forward after various readings done for CSCI 2950-v [5] and CSCI 2390 [4]. The most helpful thing would be to meet with Seny Kamara for his vast cryptography knowledge to ensure my understanding of the possibilities and limitations of PIR are correct. To find out more information about Yodel beyond the paper I will be reaching out to David Lazar, the main author. For a deeper understanding of large-scale systems and networks I will reach out to Malte Schwarzkopf. And finally, for a sanity check and to make sure my designs aren't foolish, I will run things by systems security guru Vasileios Kemerlis.

I am pretty new to PIR, so as per the timeline below I will most likely spend the rest of October familiarizing myself with the space and learning the necessary math. Luckily, it seems that PIR is not new to the metadata-privacy-concerned world as papers like [6] provide a good basis for its use. However, this project does seem risky and that it could be much bigger than I anticipate if I struggle with the cryptography or end up redesigning more of the system than I originally intended.

3 Timeline

The following is how I plan on trying to squeeze all of this work into the next few weeks.

October 11 - 21: background reading (encrypted communication and PIR), meeting/communicating with relevant people, Yodel source-code reading

October 21 - 31: architecture description including PIR protocol

November 1 - 15: code solution

November 16 - 30: code solution and debug

December 1 - 3: write-up and presentation preparation

References

[1] David L. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. In *Communications of the ACM*, pages 84–90, 1981. https://www.freehaven.net/anonbib/cache/chaum-mix.pdf.

[2] Yossi Gilad David Lazar and Nickolai Zeldovich. Yodel: Strong metadata security for voice calls. In *ACM Symposium on Operating Systems Principles (SOSP)*, 2019. https://people.csail.mit.edu/nickolai/papers/lazar-yodel.pdf.

[3] The Courage Foundation. Revelations. https://edwardsnowden.com/revelations/.

[4] Matei Zaharia Jelle van den Hooff, David Lazar and Nickolai Zeldovich. Vuvuzela: Scalable private messaging resistant to traffic analysis. In *ACM Symposium on Operating Systems Principles (SOSP)*, 2015. https://oaklandsok.github.io/papers/unger2014.pdf.

[5] Joseph Bonneau Sascha Fahl Henning Perl Ian Goldberg Nik Unger, Sergej Dechand and Mathew Smith. Sok: Secure messaging. In *IEEE Symposium on Security and Privacy*, 2015. https://oaklandsok.github.io/papers/unger2014.pdf.

[6] Srinath Setty Sebastian Angel. Unobservable communication over fully untrusted infrastructure. In *USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, 2016. https://www.usenix.org/conference/osdi16/technical-sessions/presentation/angel.

[7] Johns Hopkins University. The johns hopkins foreign affairs symposium presents: The price of privacy: Re-evaluating the nsa, 2014. https://youtu.be/kV2HDM86XgI.