Senior Project Proposal

Text-Based Front End for Private Watermark Circumvention Protocol

Seth Lifland
February 4th, 2015
Advisor: Bryan Ford
Director of Undergraduate Studies in Computer Science: James Aspnes
Yale University

I. Background

In the world of modern computing, it is increasingly impossible for users of the internet to maintain anonymity in any of their online activities. The constant monitoring activities of Internet Service Providers, governments and their dedicated agencies such as the NSA, and individual websites and applications have created a world in which virtually no online activity, whether as innocuous as watching funny cat videos on YouTube or as sinister as conspiring to plot a terrorist attack, goes unrecorded.

Based on a talk entitled 'A Most Intrusive Security Surveillance System,' given at Yale in the fall of 2014 by William Binney, the former NSA Technical Director of the World Geopolitical and Military Analysis among other positions, it became clear to me that the surveillance programs of the NSA and other agencies are not only incredibly intrusive, but also have consistently been declining in effectiveness while increasing in intrusiveness.

According to Mr. Binney, the surveillance programs under his watch worked on the basis of monitoring and linking together networks of known conspirators and unknown parties, only allowing the personal identity and information of a suspect, particularly a suspect who was also a U.S. citizen, to be de-anonymized if the suspect was implicated in communications with a certain threshold number of known terrorist or criminal figures. Binney went on to explain that the current NSA model is to monitor and record metadata on virtually all communications in which any word from a primitive, context-free set of buzzwords appears, with no regard for anonymity and no attempt to create rational networks of conspirators. The result is that the NSA is drowning itself in noise with this new strategy, because the volume of data recorded and cached by this model is simply too massive to be adequately monitored by the NSA's human workforce.

At present, the NSA has rendered itself completely unable to predict, preempt, and avert acts of terror by drowning itself in garbage data, but it has made itself vastly more capable of

gathering evidence against a suspect and constructing a narrative of guilt once a crime has been committed. Therefore, the NSA has failed to prevent several catastrophes by missing obvious signs in cases like the Boston Marathon Bombing perpetrated by Dzhokhar and Tamerlan Tsarnaev in April 2013 and the shooting of two New York City police officers by Ismaaiyl Brinsley in December 2014, but was quickly able to sift through their data landfill after the incidents to identify evidence linking the suspects to the acts.

Despite the initial outcry both within America and abroad after the Snowden revelations of 2013 about the scope of the NSA's surveillance activities, broad reform of the NSA has barely progressed and was definitively blocked by Senate Republicans in fall 2014.[1] Given this atmosphere, future whistleblowers are going to be increasingly reluctant to share information that is crucial to the public for fear of political, professional, or personal repercussions.

II. The Project

For my project, I will be working with Daniel Jackowitz, a graduate student in the department, to develop a private watermark circumvention protocol for the use of networks of whistleblowers who want to disclose information without compromising privacy or personal security. The protocol will allow a network of users who each have an individual copy of a potentially-watermarked, sensitive document that they desire to release to collude to ensure that the released document cannot be traced back to any individual participant.

The protocol is based on the idea that sensitive documents may be watermarked when distributed to individuals, such that if the document is released, the watermark on the specific released document can be traced back to the owner of the individual copy. The watermarks are generally invisible to human inspection – a watermark could be as simple as a few pixels an imperceptibly different shade of white from the rest – but this protocol will allow a group of users with their own individual copies to anonymously compute any hidden differences between their copies, and ensure that the final released document does not contain any information that can be traced back to one of the whistleblowers. Danny's work on the protocol specifically focuses on an image-based front end, such that a network of users can share copies of the same image and have the protocol determine any areas in the image which are not "safe" for release because they may contain watermarking information.

The project builds on the work of Kissner and Song as described in their 2005 paper, "Privacy-Preserving Set Operations," as well as the work of the more recent SEPIA (Security through Private Information Aggregation) project which provides a faster implementation of

[1] Ramsey, Nick. "Senate Republicans Block Broad NSA Reform Bill."*Msnbc.com*. NBC News Digital, 21 Nov. 2014. Web. 04 Feb. 2015.

similar cryptographic primitives to the ones described by Kissner and Song in the form of an open-source Java library complete with a high-level interface to secure, multi-party computation protocols.

I hope to implement a text-based front end for the same protocol, such that sensitive, potentially-watermarked text files can be shared and analyzed by the protocol as well as image files. I will also have to figure out a common way to reconstruct the output of the protocol – the parts of the input file deemed un-manipulated – into a file in the input format. Depending on the time required to complete such a front end for this protocol, the rest of my project will involve assisting the DeDiS group with the development of its next generation release of Dissent.

III. Deliverables

The deliverables for this project will consist of source code for the text-based front end of the private watermark circumvention protocol along with any appropriate documentation, both in the form of comments within the code and README files as appropriate. Depending on what other sub-tasks I am able to take on within the DeDiS group, my deliverables will include additional source and documentation files corresponding to the code I write for those tasks. At the end, I will also produce a write-up describing the functionality of the text-based front end as well as a summary of the additional work I've accomplished for the DeDiS group in its development of the next generation Dissent protocol.

IV. Timeline

Week of	**Work to Accomplish**
2/2	Meet with Danny to discuss project status, get familiar with Dissent code base
2/9	Get SEPIA library running and meet with Danny to discuss first steps
2/16	Begin designing the text-based front end based on existing code & SEPIA library
2/23	Produce a first-cut design of the text-based front end
3/2	Refine the design based on feedback from group members and produce first-cut code
3/9	*<SPRING BREAK; catch up as necessary>*
3/23	Attempt integration of front end with SEPIA library and existing code
3/30	Produce a second-cut of the code to refine for integration with the back end
4/6	Begin final integration
4/13	Complete final integration and testing of protocol with live users
4/20	Refine front end as necessary based on feedback, refine documentation
4/27	Produce final report and documentation