SaTC: Small: New Encryption Paradigms to Circumvent Internet Censorship

ts says: Expand to accomodate applications of HE, in particular.

Nation-states and other organizations increasingly seek to censor Internet communications. These censorship efforts have expanded to include detection of circumvention tools using deep-packet inspection (DPI) technologies. These look past network headers to identify application layer protocols and contents, and have been successfully deployed in countries such as China, Iran, Ethiopia, Kazakhstan, and others in order to prevent use of tools such as Tor. Social activists have responded with ad-hoc fixes and a call to build new obfuscation tools that provide more principled prevention of DPI systems.

This proposal lays out a plan to provide them. We will use a new methodology of empirically-driven provable-security to develop new security goals for encryption and the new primitives needed to meet them. By mixing theory with empiricism, we will be able to provide strong guarantees that our primitives treat the right problems in a deployable and formally sound manner. We will in particular introduce and explore two new cryptographic primitives:

- **Format-transforming encryption**: In ongoing work we have suggested the use of encryption primitives that allow fine control over the format of ciphertexts. In the proposed work, we will formalize this new primitive, explore security notions for it, and build practical schemes that will enjoy proofs of security. We will evaluate their security as well via empirical means, and use this to bootstrap an iterative process for refinement of the security goals and, in turn, the schemes.
- **Distribution-matching encryption**: We introduce a new primitive that goes beyond formats to also allow control over the distribution of ciphertexts generated by encryption. Like previous stegonagraphic approaches, we will explore security notions that measure security against powerful attackers, but, as educated by our empirical experiences, we will look at new weaker notions as well. We will provide new schemes based on techniques such as rejection sampling.
- **Honey encryption**:

  ts says: add

Throughout we will develop new attacks and test them using datasets collected from public sources such as the top Alexa websites, as well as private ones such as university network packet captures. We will explore attacks based on the use of machine-learning tools and statistical tools such as entropy estimation, as well as more conventional approaches.

ts says: drop transitions reference, but perhaps keep text re: Tor.

In our transitions supplement, we discuss how we will transfer these new technologies from the lab to the users targeted by nation-state censorship. This will involve building into the Tor pluggable transports framework new mechanisms based on format-transforming and distribution-matching encryption, as well as an outreach agenda that will seek to engage and educate the broader public on the topic of censorship and its circumvention.

Intellectual Merit: This work will require developing new methodologies to integrate advantageously the use of empiricism with formal security analysis, the latter in the vein of modern cryptography's provable security. We suspect that this will form the basis of a new approach to development of secure cryptographic protocols. We will have to advance theory to realize these new primitives in a secure manner. This will include interaction with other topic areas such as language ranking algorithms, rejection sampling, and hypothesis tests.

Broader Impact: This proposal will positively benefit:

- *Students* who participate in the research through new cross-disciplinary research opportunities on security, cryptographic theory, and networking.
- *Science of cybersecurity research* by introducing and exercising a new methodology of empirically-driven provable security that tightly couples data-driven experiments with theory, all within an iterative refinement process.

- *People targeted by censorship* by providing new, secure techniques for ensuring access to information via the Internet.
- *Society* through the use of outreach, education, and engagement on the tricky issues surrounding Internet censorship and circumvention tools.

Key words: censorship, security, cryptography, encryption.