Multiple Objectives of Lawful-Surveillance Protocols

Joan Feigenbaum[1] and Bryan Ford[2]

[1] Computer Science Department, Yale University, New Haven CT 06520 USA,
Joan.Feigenbaum@yale.edu
[2] Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland,
Bryan.Ford@epfl.ch

Abstract. In recent work on *open, privacy-preserving, accountable surveillance*, we have proposed the use of cryptographic protocols that enable law-enforcement and intelligence agencies to obtain actionable information about targeted users of mass-communication systems without intruding on the privacy of untargeted users. Our suggestion that appropriate technology, combined with sound policy and the rule of law, can afford typical users significantly more privacy than they have now without hindering lawful and effective actions by law-enforcement and intelligence agencies has met with considerable skepticism. In this paper, we summarize the principal objections to our approach and address them.

1 Introduction

As networked devices become more available, more capable, and more ubiquitous in everyday life, tension mounts between users' desire to safeguard their personal information and government agencies' desire to use that personal information in their pursuit of criminals and terrorists. Since the Snowden revelations began in June of 2013, many people have asserted that society faces an unpleasant, stark choice: Citizens can either have control over their personal information, or they can have law-enforcement and intelligence agencies with the tools needed to keep the country safe. Others regard this stark choice as a false dichotomy and assert that, by deploying appropriate cryptographic protocols in the context of sound policy and the rule of law, citizens can have both user privacy and effective law enforcement and intelligence.

In this paper, we begin by briefly recapping our recent work on lawful, privacy-preserving surveillance, in which we adopt the second point of view and demonstrate its technical feasibility. We then present and address the principal objections to this viewpoint that we have heard from members of the cryptography, security, and privacy research communities.

2 Overview of previous work

In our work on *open, privacy-preserving, accountable surveillance*, we distinguish between *targeted users* (*i.e.*, those who are under suspicion and the subjects of

properly authorized warrants) and *untargeted users* (everyone else, *i.e.*, the vast majority of the users of any general-purpose, mass-communication system). We also distinguish between *known users* (*i.e.*, those for whom the relevant agency has a name, phone number, or other piece of personally identifying information (PII)) and *unknown users* (*i.e.*, those for whom no PII is available but who might nonetheless be legitimate targets of investigation). At first glance, it may seem nonsensical to describe a user as both "unknown" and lawfully "targeted," but it is not. For example, a "John Doe warrant" [1] might be issued for persons of interest for whom no PII is known but for whom relevant times and locations are known and for which the warrant can adequately demonstrate to a judge that reasonable suspicion is attached to that particular combination of times and locations.

We have explored the design and implementation of open processes and procedures for bulk surveillance that protect the privacy of all untargeted users but reveal information about lawfully targeted users, both known and unknown. Here, an "open" process or procedure is one that is unclassified and laid out in public laws that all citizens have the right to read, to understand, and to challenge through the political process. Our solutions make essential use of computation on encrypted data; roughly speaking, they enable agencies to obtain a large set of encrypted data about both targeted and untargeted users, feed it into a cryptographic protocol that winnows it down to the records of users targeted by a John Doe warrant, and decrypt only those records. Protocol-design principles include division of trust, limitations on scopes of individual warrants, sealing times and eventual target notifications for all warrants, and publicly reported statistics about the use of warranted-access mechanisms.

We have provided experimental evidence that actionable, useful information can indeed be obtained in a manner that preserves the privacy of innocent parties and that holds government agencies accountable. In particular, we have presented practical, privacy-preserving protocols for two operations that law-enforcement and intelligence agencies have used effectively: *set intersection* and *contact chaining*. Experiments with our protocols suggest that privacy-preserving contact chaining can perform a 3-hop privacy-preserving graph traversal producing 27,000 ciphertexts in under two minutes. These ciphertexts are usable in turn via privacy-preserving set intersection to pinpoint potential unknown targets in a set of 150,000 ciphertexts within 10 minutes, without exposing personal information about non-targets. Details of these experiments can be found in [7,8], along with a comprehensive overview of our approach to openness and accountability in lawful surveillance.

Other researchers have addressed privacy and accountability in government surveillance; a full review of the literature is beyond the scope of this paper. Most closely related to ours is the work of Kamara [3] and Kroll *et al.* [5], who propose cryptographic protocols that achieve privacy and accountability in the surveillance of *known* targets, and that of Kearns *et al.* [4], who propose differential-privacy-based, graph-search algorithms that distinguish targeted users from untargeted users.

3 Principal objections and responses

Our proposal for open, privacy-preserving, accountable surveillance is tantamount to an endorsement of a *social contract* that binds the cryptography, security, and privacy research communities together with the law-enforcement and intelligence communities. The contract requires us to provide technology that enables government agents to identify and pursue criminals and terrorists with minimal (if any) intrusion upon innocent users of information and communication systems. It requires democratic governments to conduct their pursuit of criminals and terrorists in a truly democratic fashion (employing *open processes*, as explained in Section 2) and a technologically sound fashion. We now summarize and respond to the wide range of objections to such a social contract that we have encountered since we first presented these ideas in [9].

3.1 The "don't be evil" objection

Unsurprisingly, we have encountered members of the cryptography, security, and privacy research communities who believe that our communities should not work with law-enforcement and intelligence agencies at all. They believe that the communities' goal should be "no surveillance" – of anyone by anyone ever for any reason.

This view is "unsurprising," because it exemplifies the cyber-libertarian tendency that has always been present in our communities. We anticipated this objection and pre-emptively responded to it in [9]:

> Before proceeding, we wish to address the question of why "privacy-preserving, accountable surveillance" is an appropriate topic for a workshop on "free and open communications on the Internet." While it may be interesting and appealing to contemplate an Internet in which there is little or no surveillance, it would not be an effective way to increase the degree to which "Internet freedom" is a lived experience for ordinary people. Law-enforcement and intelligence agencies have been and currently are active in every national- or global-scale mass-communication system, and the Internet will be no exception. The Snowden revelations may have provided an opportunity to design protocols that allow government agencies to collect and use data that are demonstrably relevant to their missions while respecting the privacy of ordinary citizens and being democratically accountable. The FOCI community should seize that opportunity.

3.2 The "political and social infeasibility" objection

Many have objected to our proposals simply on the grounds that they are politically unrealistic and will never be adopted. The law-enforcement and intelligence communities will not enter into a social contract of the type we support. Division

of trust, scope limitations, mandatory statistical reporting, *etc.*, are incompatible with "the way surveillance works," and thus even democratic governments will never commit to them. A very closely related objection is that such principles are vacuous: A scope limit, for example, could be set so high as to allow the decryption of all records obtained in a cell-tower dump or other act of bulk collection (and would be by a FISA court or equivalent "rubber-stamp" judicial system).

We acknowledge that this is a reasonable point of view. However, it is a description of "the way surveillance *currently* works" rather than an essential feature of the way it must work. To date, citizens of democratic countries have not demanded that their governments respect their privacy, autonomy, and other individual rights online as well as offline. On the contrary, citizens have been quite vocal in their demands that their governments stop criminals and terrorists from using the Internet in pursuit of violent aims, and many seem unconcerned ("I have nothing to hide") about whether their own civil rights would be trampled if governments heeded their demands.

This state of affairs could change. Political and social reality has changed drastically just in the last few years; for better and for worse, the range of feasible government policy has expanded. Citizens who once seemed complacent about (or even oblivious to) important societal problems have started to demand that their governments take action. Courts have ruled inadmissible some fruits of warrantless electronic searches, and presidential commissions have rejected blanket collection of call records. In time, law-enforcement and intelligence agencies may demand that we provide them with technology that has been thoroughly vetted by independent experts, that produces evidence that will not be ruled inadmissible, and that need not be keep secret and hence unavailable to prosecutors.

In summary, we believe that it would be foolish to abandon the study of open, privacy-preserving, accountable surveillance protocols simply because their adoption will take time.

3.3 The "technical infeasibility" objection

We have heard several times that, although secure, multiparty computation is very interesting theoretically, it is not usable in practice. It is described as too hard for software developers to understand and implement, too slow even when implemented well, or to hard to explain to our target users (law-enforcement and intelligence agencies). Sadly, this dismissive attitude is on display even in the cryptographic-research community, members of which have told us that they think "fancy crypto" or "exotic protocols" are ill-suited for this problem domain.

It is simply not true that secure, multiparty protocols for specific problems of interest in this context are too hard to implement or too slow to use on realistic-sized data sets; for example, the experimental results that we reported in [7] refute such criticisms. In general, there has been great progress in recent years on implementation and application of privacy-preserving computational techniques, including secure, multiparty computation, homomorphic encryption, and private

information retrieval. An overview of DARPA and IARPA[3] efforts in this area can be found in [2,6]. Whether the fruits of this research can be adequately explained to our target users is an empirical question, and we remain optimistic.

3.4 The "lack of generality" objection

Use of a privacy-preserving protocol for set intersection, contact chaining, or any particular computation requires an upfront commitment to the design and implementation of not only the protocol itself but also the necessary data infrastructure. The data that may be input to such a protocol, *e.g.*, phone-call records or IP-packet headers, must be formatted appropriately, encrypted under multiple public keys using the cryptosystem that is used in the protocol, and stored by an approved data custodian that may or may not be the communications-service provider whose system originally produced the data. Some people have rejected our proposals on the grounds that it does not make sense to create a data infrastructure to support only one operation (or even a small number of operations). Their claim is that government agencies would be willing to fund the creation and maintenance of such an infrastructure only if it were fully general-purpose, *i.e.*, if the encrypted data that it contained could be fed into *all* surveillance and data-mining protocols that the agencies use now or may use in the future.

Although a general-purpose data infrastructure may be a good long-term goal, we disagree that it is an appropriate goal at this time. In order to promote the use of privacy-preserving protocols in law enforcement and intelligence, we believe that the best starting point is a specific operation (or small number of them) that government agencies use routinely (and admit to using routinely) and that we know, based on rigorous experimental research, can be done efficiently, in a privacy-preserving manner, with current technology. Given that set intersection is a standard tool of law enforcement and intelligence (used, *e.g.*, in the NSA CO-TRAVELER program [10]) and that it is a well studied problem for which there are mature and practical privacy-preserving protocols that require only modest infrastructural investment, why would government agencies *not* be willing to compute set intersections in a privacy-preserving manner? We would be entirely justified, both as technologists and as citizens, in demanding that they do.

3.5 The "don't give aid and comfort to the enemy" objection

Finally, some people readily agree that particular cryptography-based solutions that we have proposed are clearly technically feasible and that they would enable government agencies to conduct in a privacy-preserving manner surveillance operations that they currently conduct in a privacy-invasive manner. Nonetheless, they believe that these solutions should not be adopted and that, merely by proposing them, we may be causing harm.

[3] The Defense Advanced Research Projects Agency and the Intelligence Advanced Research Project Activity are technology-research organizations within the US Department of Defense and the US Office of the Director of National Intelligence, respectively.

Essential to this objection is the belief that our proposals will be overinterpreted and/or misinterpreted by pro-surveillance zealots. Although we have clearly stated that we are proposing solutions to very specific problems, *e.g.*, how to find the records in the intersection of multiple cell-tower dumps without exposing the records that are not in the intersection, some critics claim that law-enforcement and intelligence agencies will, because they either don't understand or deliberately misrepresent our proposals, claim that we've provided fully general solutions. These agencies could assert that "academic cryptographers have shown that data-mining and surveillance operations can be done without compromising the privacy or security of innocent parties" and then interpret this assertion to mean that there would be no harm in their conducting whatever warrantless mass-surveillance operations they wish to conduct. Technically informed people who are paying attention will see immediately that the implied universal quantifier is not in fact present in what "academic cryptographers have shown," but the government officials who could grant a broad mandate for mass surveillance will not, in general, be technically informed and may not realize that they need expert advice (or may be convinced by the wrong "expert").

Another way in which we could do harm by proposing technically workable solutions that would provide privacy protection for untargeted users is by creating *function drag*. This term was coined by Paul Syverson to describe a situation in which it is preferable (for security, performance, or other reasons) to migrate to a new technology, but a particular function of the status quo technology appeals very strongly to a powerful constituency and thus exerts a drag on migration. Our existing communication infrastructure creates and stores a great deal of metadata, including phone-call records and IP-flow statistics, that is useful to law-enforcement and intelligence agents but potentially destructive of users' privacy. Infrastructure evolves, however, and we may someday be faced with the opportunity to route phone calls and IP packets without creating massive amounts of privacy-destructive metadata. Government agencies may resist the adoption of such a surveillance-resistant communication infrastructure, because they are increasingly dependent on communications metadata for their investigations. If we provide them with techniques for accessing those metadata in a privacy-preserving manner, we may make it easier for them to block desirable evolution of communications systems, because we will erode one of the reasons (*i.e.*, lack of privacy) that current systems are undesirable.

No doubt, these are reasonable concerns. Taken to their logical conclusions, they vitiate the very notion of a social contract that binds the cryptography, security, and privacy research communities together with the law-enforcement and intelligence communities. While acknowledging the risks of misinterpretation and function drag, we believe that research into privacy-preserving surveillance is still worth pursuing and that researchers should advocate for deployment of whatever workable solutions we obtain. As explained in Section 3.1, we simply don't see a better alternative. Democratic governments will continue to seek access to private information that they believe will enable them to catch criminals and terrorists, because their citizens will continue to demand that they do so.

Currently, it is fairly easy for law-enforcement and intelligence agencies to collect large amounts of information in plaintext form, most of which will prove to be irrelevant to their investigations. The cryptography, security, and privacy research communities have been saying for decades that our techniques can be used to compute a particular fact about a large, distributed data set without revealing anything about the data except what is implied by that fact and prior knowledge. It now behooves us to deploy these techniques in order to ensure that large-scale surveillance operations *of the sort that are routinely done now and that will continue to be done for the foreseeable future* are conducted in as privacy-respectful a manner as possible.

4 Conclusion

Stepping back from the specific points discussed in Section 3, we sense that much of the resistance to our notion of a social contract boils down to skepticism about whether government agencies should be trusted with technically sophisticated surveillance tools. More accurately, there is deep skepticism about whether they should be trusted with a larger arsenal of such tools than they already have. Obviously, the cryptography, security, and privacy research communities cannot stop government agencies from developing their own tools or from contracting with technology companies to develop them, but we could decide not to participate in such development efforts. If it were clear that our efforts would do more harm than good, then refusal to participate would be the only honorable choice.

The social contract that we envision would eliminate the need for trust without verification. Laws and processes governing surveillance would have to be open, as explained in Section 2, and would apply to everyone, including government officials. Users of surveillance technology in law enforcement and intelligence would have to abide by that part of the contract *and to show that they are abiding by it.* They would have to come out of the shadows, submit their needs to public scrutiny, and accept that one of the worthwhile prices of democracy is that rule of law will occasionally enable a criminal to evade surveillance who otherwise might not have. Crucially, their surveillance tools would have to be publicly proposed, publicly debated in a technically informed fashion, embodied in open-source designs and implementations, analyzed in public by technology and privacy experts, and verifiably deployed in configurations that technically enforce proper division of trust and rule of law.

We would sign that contract if they would.

Acknowledgements

This work was supported by US National Science Foundation grants CNS-1407454 and CNS-1409599, William and Flora Hewlett Foundation grant 2016-3834, and the AXA Research Fund.

We are grateful to our collaborator and former student Aaron Segal for all of his good work in this area and for helpful discussions.

References

1. Bieber, M.A.: Meeting the statute or beating it: Using John Doe indictments based on DNA to meet the statute of limitations. University of Pennsylvania Law Review 150(3), 1079–1098 (2002)
2. Greenberg, A.: DARPA will spend $20 million to search for crypto's holy grail. Forbes (Apr 6, 2011)
3. Kamara, S.: Restructuring the NSA metadata program. In: Proceedings of the 2nd Workshop on Applied Homomorphic Cryptography (held in conjunction with Financial Cryptography and Data Security). pp. 235–247. Springer, Berlin, Germany (March 2014)
4. Kearns, M., Roth, A., Wu, Z.S., Yaroslavtsev, G.: Private algorithms for the protected in social network search. Proceedings of the National Academy of Sciences 113(4), 913–918 (2016)
5. Kroll, J.A., Felten, E.W., Boneh, D.: Secure protocols for accountable warrant execution. `http://www.cs.princeton.edu/~felten/warrant-paper.pdf` (2014)
6. Lohr, S.: With 'Brandeis' project, DARPA seeks to advance privacy technology. The New York Times (Sep 14, 2015)
7. Segal, A., Feigenbaum, J., Ford, B.: Open, privacy-preserving protocols for lawful surveillance. `https://arxiv.org/abs/1607.03659` (2016)
8. Segal, A., Feigenbaum, J., Ford, B.: Privacy-preserving contact chaining [preliminary report]. In: Proceedings of the 15th Workshop on Privacy in the Electronic Society. pp. 185–188. ACM, New York NY, USA (October 2016)
9. Segal, A., Ford, B., Feigenbaum, J.: Catching bandits and *only* bandits: privacy-preserving intersection warrants for lawful surveillance. In: Proceedings of the 4th Workshop on Free and Open Communications on the Internet. USENIX, Berkeley CA, USA (August 2014)
10. Soltani, A., Gellman, B.: New documents show how the NSA infers relationships based on mobile location data. The Washington Post (Dec 10, 2013)