MATT BLAZE

UNIVERSITY OF PENNSYLVANIA[1]

US HOUSE OF REPRESENTATIVES
COMMITTEE ON ENERGY AND COMMERCE
SUBCOMMITTEE ON OVERSIGHT AND INVESTIGATIONS
HEARING ON "DECIPHERING THE DEBATE OVER ENCRYPTION"

APRIL 19, 2016

Thank you for the opportunity to offer testimony on the important public policy issues raised by cryptography and other security technologies. Since the early 1990's, my research has focused on cryptography and its applications for securing computing and communications systems, especially as we rely for increasingly critical applications on relatively insecure platforms such as the Internet. My work has focused particularly on the intersection of this technology with public policy issues. For example, in 1994, I discovered some fundamental technical flaws with the ill-fated "Clipper Chip", an encryption system designed by the National Security Agency intended to provide a government backdoor to encrypted communications.

I am currently an associate professor in the computer and information science department at the University of Pennsylvania. From 1992 until I joined Penn in 2004, I was a research scientist at AT&T Bell Laboratories. However, this testimony is not offered on behalf of any organization or agency.

I. ROBUST DIGITAL SECURITY TECHNOLOGIES ARE VITAL TO PROTECTING OUR NATIONAL AND CRITICAL INFRASTRUCTURE

It is difficult to overstate the importance of robust and reliable computing and communications to our personal, commercial, and national security today. Virtually every aspect of our lives, from our health records to the critical infrastructure that keeps our society and economy running, is reflected in or supported in some way by increasingly connected digital

[1] University of Pennsylvania Computer and Information Science, 3330 Walnut Street, Philadelphia, PA 19104. *mab@crypto.com.* Affiliation for identification only.

technology. The influx of new communications and computing devices and software over the last few decades has yielded enormous benefit to our economy as well as to our ability to connect with one another. This trend toward digital systems, and the benefits we reap from them, will only accelerate as technology continues to improve. Preventing attacks against our digital infrastructure by criminals and other malicious actors is thus now an essential part of protecting our society itself.

Unfortunately, modern computing and communications technologies, for all their benefits, are also notoriously vulnerable to attack by criminals and hostile nation-state actors. And just as the benefits of increased connectivity and more pervasive computing will continue to increase as technology advances, so too will the costs and risks we bear when this technology is maliciously compromised. It is a regrettable (and yet time-tested) paradox that our digital systems have largely become *more* vulnerable over time, even as almost every other aspect of information technology has (often wildly) improved. New and more efficient communication technologies often have *less* intrinsic security than the systems they replaced, and the latest computers and similar devices are regularly found to suffer from unexpected vulnerabilities that can be exploited remotely by malicious attackers. Large-scale data breaches and similar security failures have so become commonplace that they now only make the news when their consequences are particularly dramatic.

Serious security failures have become literally a daily occurrence, and it is not an exaggeration to characterize this situation as a national crisis.

Modern digital systems are so vulnerable for a simple reason: computer science does not yet know how to build complex, large-scale software that has reliably correct behavior. This problem has been known, and has been a central focus of computing research, literally since the dawn of programmable computing. As new technology allows us to build larger and more complex systems (and to connect them together over the Internet), the problem of software correctness becomes exponentially more difficult.[2] Worse, as this insecure technology becomes more integrated into the systems and relationships upon which society depends, the consequences become increasingly dire.

While a general solution to the problem of software reliability and

[2] That is, the number of software defects in a system typically increases at a rate far greater than the amount of code added to it. So adding new features to a system that makes it twice as large generally has the effect of making it far more than twice as vulnerable. This is because each new software component or feature operates not just in isolation, but potentially interacts with everything else in the system, sometimes in unexpected ways that can be exploited. Therefore, smaller and simpler systems are almost always more secure and reliable, and best practices in security favor systems the most limited functionality possible.

correctness has eluded us (and will continue to do so absent some remarkable and unexpected breakthrough), there are two tried-and-true techniques that can, to some extent, ameliorate the inherent vulnerability of software-based systems. One is the use of encryption to protect data stored on or transmitted over insecure media. The other is to design systems to be as simple as possible, with only those features needed to support the application. The aim is to minimize the "attack surface" that any software vulnerabilities would expose.

Neither the use of encryption nor designing systems to be small and simple are perfect solutions to the software security problem. Even carefully designed, single-purpose software that encrypts data whenever possible can still harbor hidden, exploitable vulnerabilities, especially when it is connected to the Internet. For this reason, software systems must be exposed to continual (and resource intensive) scrutiny throughout their lifecycle to discover and fix flaws before attackers find and exploit them. But these approaches, imperfect and fragile as they might be, represent essentially the only proven defenses that we have.

II. LAW ENFORCEMENT ACCESS REQUIREMENTS INTRODUCE GREAT RISKS

U.S. law enforcement agencies have for at least two decades been warning that wiretaps and other forms of electronic evidence gathering are on the cusp of "going dark". These fears have been focused chiefly on the potential for criminal use of encryption (which, properly used, can prevent eavesdroppers from recovering communications content), as well as on emerging decentralized communications paradigms, such as peer-to-peer communication, that are not easily intercepted with the same techniques that were used to wiretap traditional telephone calls. They call for developers to incorporate "lawful access"[3] features into products and services in order to facilitate wiretapping.

At first blush, a "lawful access only" mechanism that could be incorporated into the communications systems used by criminal suspects might seem like an ideal technical solution to a difficult policy problem. Unfortunately, harsh technical realities make such an ideal solution

[3] These law enforcement access features have been variously referred to as "lawful access", "back doors", "front doors", and "golden keys", among other things. While it may be possible to draw distinctions between them, it is sufficient for the purposes of the analysis in this testimony that all these proposals share the essential property of incorporating a special access feature of some kind that is intended solely to facilitate law enforcement interception under certain circumstances.

effectively impossible, and attempts to mandate one would do enormous harm to the security and reliability of our nation's infrastructure, the future of our innovation economy, and our national security.

A. Access Requirements Make Encryption Vulnerable and Expensive

Let us consider first the relatively narrow problem of ensuring law enforcement access to encrypted communication.[4] This is perhaps the simplest part of the law enforcement access problem, but it is dauntingly – and fundamentally – difficult to solve in practice without creating significant risk.

Encryption systems encode messages in a way that prevents their decryption without knowledge of a secret, called a *key*. Ordinarily, only the parties to the communication know the key, which can be destroyed and forgotten as soon as the communication has ended and need never be sent to anyone else. In most well designed encrypted communications systems, third parties – including the developer of the software used to perform the encryption and the service providers who operate the infrastructure through which it traverses – do not know or have copies of these keys; the encryption is said to be *end-to-end*, meaning it is conducted entirely between the communicating parties. End-to-end encryption is an important simplifying principle that allows for secure communication even over insecure media. It means that only the endpoints (the computers or devices being directly used by the parties) need to have access to and protect the keys, and the compromise of any other part of the system has no effect on the security of the messages. Securing the endpoints can sometimes be perilously difficult in practice, but it is a much simpler problem than securing the entire path over which messages are transmitted.

Any law enforcement access scheme of the kind apparently envisioned by the FBI would, necessarily, involve a mechanism for the transmission and storage of sensitive secret keys to a third party (whether the government or some other entity that holds it). This approach is sometimes called *key escrow*, *key recovery* or *trusted-third party* encryption; the secret is held "in escrow" by a third party. Key escrow was the widely criticized approach incorporated into the Clipper Chip in the early 1990's. It destroys the end-to-end design of robust encryption systems without any benefit to the application.

There are several fundamental problems with such schemes.

The most basic problem with third-party access cryptography is simply

[4] Decrypting encrypted communication is only one aspect of the law enforcement access problem as posed by law enforcement, but any access design mandate would, at a minimum, introduce the problems and risks discussed here, as well as others.

that we do not fully understand how to design it securely. Any key escrow or lawful access cryptography system, by its very nature, increases its number of points of failure. Unfortunately, we do not understand the problem well enough to even precisely quantify how this reduces security, let alone identify a safe level for this reduction.

The design and implementation of even the simplest encryption systems is an extraordinarily difficult and fragile process. Very small changes frequently introduce fatal security flaws. Ordinary (end-to-end, non-escrowed) encryption systems have conceptually rather simple requirements and yet, because there is no general theory for designing them, we still often discover exploitable flaws in fielded systems. Adding key escrow renders even the specification of the protocol itself far more complex, making it virtually impossible to assure that any systems using it will actually have the security properties that these systems are intended to have. It is possible, even likely, that lurking in any key escrow system will be one or more design weaknesses that allow recovery of data by unauthorized parties. The commercial and academic world simply does not have the tools to analyze or design the complex systems that arise from key recovery.

This is not simply an abstract concern. Virtually all law enforcement key recovery or key escrow proposals made to date, including those designed by the National Security Agency (the Clipper Chip[5]), have had unanticipated, serious design weakness discovered after the fact.

Frequently, subtle but devastating weaknesses in cryptographic systems and protocols are only discovered long after they are deployed in products and services, which means that sensitive data was at risk from their very first day of use. Law enforcement access requirements make such hidden flaws far more likely to exist.

Aside from cryptographic weaknesses, there are significant operational security issues. Third-party access, by its nature, makes encrypted data less secure because the third party itself creates a new target for attack.

The FBI has not stated whether the cryptographic access mechanisms they desire would be operated centrally or by the vendors of individual products. Either approach creates its own inherent risks and costs. A centralized system becomes a large and highly attractive target, while leaving the task to individual product vendors introduces the likelihood that some vendors will be lack the resources to securely mange the keys for their customers or will be specialty targeted for attack by national adversaries.[6]

---

[5] See M. Blaze. "Protocol Failure in the Escrowed Encryption Standard". *ACM Conference on Computer and Communications Security*, 1994.

[6] An alternative, but equivalently risky, design approach involves incorporating a law enforcement access mechanism into the end-user devices that would respond to remote commands from law enforcement to reveal its keys. In this case, managing and securing the

Importantly from a business perspective, the infrastructure to properly support any scheme of this kind would be very expensive to operate.

Even more significant risks arise from the *operational complexity* of managing access to the access keys. Key access centers must presumably be prepared to respond to law enforcement requests for key data on an emergency basis, completing transactions within a short time of receiving each request and without alerting the target of the investigation. There are thousands of law enforcement agencies in the United States authorized to perform electronic surveillance; the escrow centers must be prepared to identify, authenticate and respond to any of them within a short time frame. Even if we imagine relaxing these requirements considerably (e.g., one day or perhaps one week response time), there are few existing secure systems that operate effectively and economically on such a scale and under such tightly constrained conditions.[7] It is simply inevitable that lawful access systems that meet the government's requirements will make mistakes in giving out the wrong keys from time to time or will be vulnerable to unauthorized key requests. Nation-state adversaries could be expected to be particularly interested in, and adept at, fraudulent access to our law enforcement access services.[8]

B. Access Requirements Make Critical Software Vulnerable to Attack

The vulnerabilities introduced by the cryptographic and operational complexity of introducing law enforcement access are significant; by itself, this should be sufficient reason to render any policy that requires access unacceptably risky. But these are not the only problems. Even more serious, subtle, and difficult to prevent risks arise from the process of integrating the mechanism into the end-user software itself.

As noted above, computer science does not, in general, have the tools to

---

secret required to remotely issue such commands is essentially an equivalent problem to managing and securing cryptographic keys. The same risks and costs are present in either design.

[7] Perhaps the closest existing analog to such a system can be found in the law enforcement service centers operated by telephone companies to service wiretap and pen register requests. But these operations do not hold sensitive cryptographic keys of their customers or similar data. They simply act as a clearinghouse and point of contact to which law enforcement agencies serve legal processes. They do not have the problem of managing, controlling access to, or distributing any data as sensitive as cryptographic keys.

[8] In fact, there have already been several cases where hostile intelligence services have exploited the "lawful access" interfaces in telephone switches. The most famous published case involved the (still unsolved) compromise of a Greek mobile phone carrier. See V. Prevelakis and D. Spinellis, "The Athens Affair". *IEEE Spectrum*. July 2007.

build reliably correct software at scale, and any added requirements or features will increase the likelihood that the system as a whole will suffer from unintended, exploitable, vulnerabilities. Law enforcement access requirements are especially problematic in this regard because of their inherent interaction with the most security-sensitive aspects of the systems that would use them.

As of the time of this writing, the most specific proposal for access mandates is the recently circulated Feingold-Burr "Compliance with Court Orders" discussion draft. It is exceptionally broad, and would appear to implicate the design of virtually all computing and communications software and hardware. But even under a much more narrowly tailored mandate, ensuring law enforcement access in this way would necessarily add complex requirements to a broad range of consumer, business, and infrastructure-support software. We enjoy today flourishing, heterogeneous software and service marketplace. Everything from small mobile apps that provide instant messaging services to large-scale communication and data storage platforms routinely process communication and stored data that might potentially serve as evidence in criminal cases at some point.

The design approach advocated in such proposals would affect software across the full range of modern computing, from small systems built by startups and entrepreneurs to large platforms managed by multinational corporations, be engineered to incorporate the law enforcement access features, from decentralized and standalone application to centralized, cloud-based services. In small systems, the law enforcement access mechanism could be expected to represent almost as much design and development effort as the underlying function of the software itself. In larger systems, depending on the specifics of the software architecture, the law enforcement access function would have to be designed around and interact with a large number of data management, security, and communications functions.

Compounding the difficulty is the range of different application and service architectures whose designs would have to accommodate integration with the law enforcement access features. Each application would require significant engineering effort, much of which would be highly specific to the particular piece of software. That is, much of engineering effort required to put applications in compliance would not be able to be re-applied to other systems, because each system has its own particular architectural and design constraints. And because the access features are so security sensitive, this engineering work will require the highest quality assurance, testing, and validation, making it a difficult, slow and very expensive process. Doing this properly (to the extent it can be done safely at all) will make the access feature a significant bottleneck to many projects. Given the time and budget

pressures under which many software projects operate, and because the access feature is not directly useful to users, many developers will be able to devote only the minimum engineering resources possible to meet the requirements. The result will be that while the features might work in the sense that they allow law enforcement access, they can also be expected to account for a large proportion of the potentially exploitable defects in the system as a whole.

Incorporating law enforcement access features across even a subset of the most widely used software systems is an extraordinary engineering task, the correctness of which would be crucial for the security and integrity of any data that the software might handle and of the environment in which it will run.

In other words, the risks here come not just from the potential for direct misuse or abuse of the law enforcement access mechanism itself, but from the inevitable introduction of unintentional software bugs that can be exploited by bad actors to bypass the "front door" of the access mechanism entirely and gain access to sensitive user data.

An alternative approach to requiring each software developer to design its own access mechanism is also possible, but would have even more negative effects on the software ecosystem. This would involve the government developing approved software libraries that implement the access mechanism and requiring software developers to incorporate them in their systems. Unfortunately, this scheme would have the effect of essentially outlawing software whose design and architecture is incompatible with the standard official libraries. It would hugely attenuate the innovation that has driven the software economy, and it would still carry most of the risks discussed above.

C. These Risks Would Cut Across Our Nation's Infrastructure

An important task for policymakers in evaluating the FBI's proposal is to weigh the risks of making software less able to resist attack against the benefits of more expedient surveillance. It effectively reduces our ability to prevent crime (by reducing computer security) in exchange for the hope of more efficient crime investigation (by making electronic surveillance easier). Unfortunately, the costs of the FBI's approach will be very high. It will place our national infrastructure at risk.

This is not simply a matter of weighing the desires for personal privacy and for safeguards against government abuse against the need for improved law enforcement. That by itself might be a difficult enough balance for policymakers to strike, and reasonable people might disagree on where that balance should lie. But the risks here go far beyond that, because of the

realities of how modern software applications are integrated into complete systems.

Vulnerabilities in software of the kind likely to arise from law enforcement access requirements can often be exploited in ways that go beyond the specific data they process. In particular, even small hidden vulnerabilities often allow an attacker to effectively take control over an entire system, injecting its own software and compromising the platform as a whole.[9] The unintended defects inevitably introduced by access mandates such as those discussed in the previous section are especially likely to include vulnerabilities in this category. They are difficult to defend against or contain, and they current represent perhaps the most serious practical threat to networked computer security.

For better or worse, ordinary citizens, large and small business, and the government itself all depend on the same software platforms that are used by the targets of criminal investigations. It is not just potential terrorists, members of the Mafia and local drug dealers whose software would be weakened, but everyone's, including the systems used at almost all levels of government. The stakes involve not just the potential for unauthorized leaks of inconsequential personal chitchat, but also exposure of personal financial and health information, disclosure of proprietary corporate data, and compromises of the platforms that manage and control our national critical infrastructure.

These risks are not merely speculative concerns. There is overwhelming consensus in the technical security community that requirements for "exceptional access" mechanisms such as those being advocated for by law enforcement "open doors through which criminals and malicious nation-states can attack the very individuals law enforcement seeks to defend."[10]

III. THE FOCUS ON DESIGNED-IN ACCESS IGNORES ALTERNATIVES

The cryptography debate is sometimes characterized as a stark, zero-sum choice between privacy and security on the one hand and effective law enforcement and evidence gathering on the other. Fortunately, there appear to be viable alternatives to that permit law enforcement to continue without weakening security.

First, much user data today is stored a multitude of places, typically

[9] Such vulnerabilities, for example, are how so-called "botnets" used by criminals are able to take control over large numbers of computers on the Internet for sending spam and other fraudulent messages.

[10] See Abelson, et al, "Keys Under Doormats". *Oxford Journal of Cybersecurity*, 2015. http://cybersecurity.oxfordjournals.org/content/early/2015/11/17/cybsec.tyv009.article-info

creating multiple copies of evidence in the hands of third parties, such as at "cloud" services that provide backups and remote computing services. When evidence relevant to an investigation is stored in this way, it generally can be obtained by law enforcement under conventional legal processes.

Furthermore, as noted above, the systems we use today, including those protected by cryptography, are not impenetrably secure against sophisticated attack. Indeed, they are often woefully insecure, and are frequently compromised by criminals, which is why access mandates that would make them less secure would be so dangerous. However, this inherent insecurity can, under some circumstances, create opportunities for targeted evidence collection by law enforcement by exploiting preexisting security flaws (which are virtually always present) in the devices used by investigative subjects. With sufficient resources (perhaps beyond those currently available, but well within the potential resources of a national law enforcement agency), such weaknesses can often be exploited to obtain evidence.

An example of the fruitfulness of such approaches can be found in the recent San Bernardino shooting case, in which the FBI sought to unlock an Apple iPhone model 5c used by one of the shooters. Initially, the FBI believed that the device could not be unlocked, but some time after the initial court filings in the case, a targeted technical solution was discovered that enabled the agency to obtain the data stored on the phone without assistance from Apple.

Neither the use of third-party cloud data nor the use of targeted technical attacks against devices will be "one stop shopping" solutions for law enforcement. Each technology and product will be different, and in some cases considerable resources may be required to develop a particular solution. But a systematic, broad, and up-to-date arsenal of technical forensic capabilities, while costly, can be expected to provide a viable alternative to "going dark" in many cases, even as strong cryptography (without any explicit access mechanism) is increasingly used.[11]

Alternative approaches such as those discussed here have been largely absent from the "going dark" debate.

[11] See Steven M. Bellovin, Matt Blaze, Sandy Clark, and Susan Landau, "Lawful Hacking: Using Existing Vulnerabilities for Wiretapping on the Internet," *12 Nw. J. Tech. & Intell. Prop. 1* (2014). http://scholarlycommons.law.northwestern.edu/njtip/vol12/iss1/1

IV. CONCLUSION

The technical vulnerabilities that would inevitably accompany design requirements for law enforcement access being proposed will harm our security far more than they will help law enforcement. They will provide rich, attractive targets not only for relatively petty criminals such as identity thieves, but also for organized crime, terrorists, and hostile intelligence services. It is not an exaggeration to understand these risks as a significant threat to our economy and to national security.