Risking Communications Security: Potential Hazards of the "Protect America Act"

Steven M. Bellovin, Columbia University
Matt Blaze, University of Pennsylvania
Whitfield Diffie, Sun Microsystems
Susan Landau, Sun Microsystems
Peter G. Neumann, SRI International
Jennifer Rexford, Princeton University

September 30, 2007

Abstract

The Protect America Act passed in August 2007 changes U.S. law to allow warrantless foreign-intelligence wiretapping from within the U.S. of any communications believed to include one party located outside the United States. U.S. systems for foreign intelligence surveillance located outside the United States minimize access to the traffic of U.S. persons by virtue of their location. The new law does not — and could lead to surveillance on a unprecedented scale that will unavoidably pick up some purely domestic communications.

The civil-liberties concern is whether the new law puts Americans at risk of spurious — and invasive — surveillance by their own government. The security concern is whether the new law puts Americans at risk of illegitimate surveillance by others. We focus on security. How will the collection system determine that communications have one end outside the United States? How will the surveillance be secured? We examine risks and put forth recommendations to address them.

1 Introduction

The U.S. Foreign Intelligence Surveillance Act (FISA), first adopted in 1978, governs electronic surveillance of communications within the United States for foreign intelligence purposes. It permitted surveillance with warrants in three basic cases:

- any person in the United States communicating via wire
- a U.S. person [1] in the United States whether communicating via wire or radio
- any person in the United States communicating via radio with people all of whom are in the United States [2].

The law was also clear in its exception: no warrant was required to intercept radio communications between persons in the U.S. and persons abroad unless the government was intentionally targeting a particular known U.S. person who was in the U.S. This exception was viewed as a temporary one; the Senate Judiciary Committee Report on the FISA legislation makes it clear that interception of radio communications was to be considered separately [1, p. 34]. But that separate legislation never came to pass, and so the warrantless exception continued.

At the time that FISA was written, communications satellites (radio) had revolutionized international communications. In subsequent decades there was a major shift to fiber optic cables with a decreasing percentage of foreign communications that travel by radio. Thus the exemption allowing warrantless interception became increasingly less applicable. In recent years the National Security Agency (NSA), the U.S. signals intelligence agency, began pressing to have it updated. While many in the field agreed that there was plausibly a problem as a result of fiber-optic cables, the Protect America Act (PAA)[3], passed in August 2007, was an entirely different matter. At issue was the dropping of the warrant requirement for communications

[1] U.S. citizens, permanent residents, and U.S. corporations, per 50 U.S.C. §1801 (i).

[2] 50 U.S.C. §1801 (f); the rules are, in fact, even more complicated, but this is sufficient for our purposes.

[3] P.L. 110-55.

(in any medium) of U.S. persons located in the United States with persons "reasonably believed to be located outside the United States"[4].

The "reasonably believed to be located outside the United States" aspect of the PAA arguably changes the rules on using Call Detail Records (CDRs). CDRs are records of such transactional information as calling and called numbers for phone calls, IP addresses and user URI in the case of VoIP, SMTP headers for email, etc., time and date of communications, etc. They can be surprisingly revelatory of relationships and organizational structure (although, as we explain below, this data does not always reveal where the parties to a communication are physically located). CDRs can, in particular, be used for targeting further surveillance, i.e., wiretapping. The effect is to enable warrantless wiretapping of communications based on a "nexus" of suspicion. The Protect America Act thereby creates serious civil liberties concerns. The PAA would appear to require access to CDRs, records that are surprisingly revelatory of relationships and organizational structure. The act removes the telecommuncations carriers, natural guardians of their customers' interests, from being in a position to defend their customers against abuses of the wiretap process. In attempting to collect communications with one end outside the United States, the new law allows the development of a system that will probably pick up many purely domestic communications. By removing the requirement to specify exactly who is being tapped, the new law allows surveillance on an unprecedented scale.

This new law threatens not only the privacy but the security of American communications. Recent events in Europe demonstrate reasons for our concern. For almost a year beginning in April 2004, over one hundred phones belonging to members of the Greek government, including the prime minister, ministers of defense, foreign affairs, justice and public order, and opposition members in the Greek parliament, were wiretapped through surreptitious software that turned official built-in tapping capabilities —capabilities to be invoked *only* with legal authorization — to the advantage of as yet unknown parties. What is known is that private communications of the highest levels of the Greek government were wiretapped for ten months [2].

The United States has also experienced difficulties in building communications surveillance systems. Under the Communications Assistance for Law Enforcement Act (CALEA), the FBI was responsible for determining tech-

[4]Protect America Act, §105(a), 2007.

nical specifications for wiretapping built into switches of digital telephone networks, and DCS 3000 was designed according to those requirements. Recently-released FBI documents reveal serious problems in the system's implementation [5].

The system has had trouble providing auditing, a situation that is quite surprising for a system intended for evidence collection. It has no unprivileged userids, relies on passwords rather than token-based or biometric authentication, and even uses an outdated hashing algorithm[6]. Most seriously, the system relies on a single shared login, rather than a login per authorized user. The ability to audit user behavior depends entirely on people following proper process, including using a manual log sheet to know who has used the system at a given time. Since remote access — in an insecure fashion — is permitted from other DCS 3000 nodes, this is a serious issue indeed. These issues make the system vulnerable to insider attacks. It is worth noting that Robert Hanssen — the most damaging spy in FBI history — abused his authorized access to the FBI's internal computer systems, both to steal information and to track progress of the investigation aimed at finding the leak.

The serious problems that occurred in the implementation of DCS 3000 illustrate the risks in building a communications surveillance system. We do not know whether the DCS 3000 was merely poorly implemented or whether it was poorly specified. What were the requirements on the FBI system? Did these requirements include full auditing and full user identity? What were the project's goals? Were the designers of the FBI system required to meet *all* requirements or goals? These are questions that should have been asked of the DCS 3000 designers. If there have been answers to them, these answers have not been made public.

While NSA has great expertise in building surveillance systems, that does not mean that things cannot go wrong. In fulfilling the Protect America Act's goals, what techniques will intelligence agencies use to determine that the communications involve someone located outside the United States? How will the surveillance be secured? Can this surveillance be done without creating security risks to Americans? What type of surveillance architecture will be

[5]See http://www.eff.org/flag/061708CKK/

[6]MD5 appears in a 2007 "system security plan,"[3, p. 32] several years after Chinese researchers found serious problems with this already weak hashing algorithm.

used? Is there historical precedent to guide us? Have unauthorized parties been able to subvert NSA intelligence systems and turn them to their own advantage? It is critical that these questions be asked, that answers be given to Congress and the citizenry, and that these answers give sufficient assurance that by allowing our own government to spy on American communications, we have not given that power to others as well.

In this paper we discuss potential risks in building a surveillance system to satisfy the Protect America Act. We provide background on FISA, and then discuss a likely surveillance architecture. We next discuss Call Detail Records and determining location of international Internet traffic. We enumerate security risks and conclude by putting forth recommendations.

2 Background

Prior to the Protect America Act, 2001 U.S. wiretapping law was essentially governed by Title III of the Omnibus Safe Streets and Crime Control Act of 1968[7], which regulated the procedure for wiretaps in criminal investigations, and FISA, which did the same for foreign intelligence surveillance. These laws — and their later derivatives — laid out clear and specific procedures for obtaining wiretap warrants, which, with very minor exceptions, specified the particular line (or particular IP address, email account, etc.) on which the tapping was to occur [4, pp. 323-332, 338-341]. Law enforcement obtained a warrant and communicated this information to the communications provider, which installed the tap.

In the United States, government officials learned the hard way that oversight was critical if surveillance technologies were to be kept within legal bounds. During the Watergate era, a special Senate Committee (the Senate Select Committee to Study Governmental Operation with respect to Intelligence Activities) investigated thirty-five years of government electronic surveillance in the United States and uncovered many abuses. These included wiretaps on Congressional staff and Supreme Court Justices as well as the wiretapping of Martin Luther King for many years[8] and govern-

[7] 18 U.S.C. 2510 et seq.

[8] One incident involved wiretapping during the 1964 Democratic Party convention when President Lyndon Johnson was supporting the seating of one delegation from Mississippi and King, was supporting another; the tapping enabled Johnson to learn about King's

ment investigations of such decidedly non-violent groups as the American Friends Service Committee, the National Association for the Advancement of Colored People, and the Women's Strike for Peace. It was clear that the "national-security" grounds for many of the wiretaps were not justified. The requirements governing FISA wiretapping were lifted almost verbatim from the carefully-crafted recommendations of the Senate committee report [5, pp. 292-330]. Some of these safeguards delimiting government surveillance were removed by the USA PATRIOT Act (arguably the most important change the PATRIOT Act made in wiretapping law was modifying the requirement that foreign intelligence be the "primary" purpose of a FISA tap to a "significant" purpose; see [6, pp.280-285]).

Some might argue that the excesses of surveillance in the nineteen-sixties and seventies were long ago, occurring during a period of domestic unrest and international tension. But government excesses in this realm continue. A recent report by the FBI Inspector General, for example, sharply criticized the bureau over the FBI's abuse of the National Security Letters, "administrative" subpoenas that are issued with *no* judicial oversight and that require the recipient to turn over certain records. The IG concluded that FBI agents may have violated the law 3,000 times since 2003 in their collection of telephone and financial records of U.S. citizens and foreign nationals [7].

The U.S. is a major communications hub in our communication-centered world, giving NSA significant opportunities for access to traffic. There are numerous reasons for U.S. centrality in the world's communication's systems. One is cost: the U.S. is the world's leading economy and fiber optic cables — which is how modern wired communications travel — have been built to the United States. With their economies of scale, these cables enable U.S. providers to underbid regional carriers. For example, much of South America transits its traffic through Miami. Another is politics, which can lead to strange communication paths. For many years communications could not travel directly between Taiwan and the People's Republic of China: calls traveled by way of Sacramento over AT&T lines. A third reason is the Internet. Many of the servers that are the very reason for communication — yahoomail, hotmail, gmail, etc. — are in the U.S. (though this is an ever decreasing percentage of the world's mail servers, especially as China comes

strategy and counter it.

online).

NSA argued that, as communications moved from satellite to fiber, the lack of a warrant exemption for fiber communications created an impediment (this was occurring just as the events of September 11, 2001 put increased pressure on the intelligence agency). However, as we have already notedd, the PAA broadened the warrantless exemption from communications by radio between persons in the U.S. and persons abroad[9] to communications where at least one end was "reasonably believed" to be outside the U.S. Modern communications technology — mobile phones, WiFi, and the Internet — often make it difficult to discern whether communication is from a location inside or outside the U.S. The technological issue is whether the NSA can determine, in real time, from where a cell call originates, and what is the origin of an IP transmission. The security issue is whether the NSA do so without jeopardizing the U.S. communications infrastructure. Knowing answers to these questions will help assure Congress and the citizenry that the consequent NSA wiretapping can be done without increasing the risk that communications of Americans will be exploited by others who are not authorized to do so.

3 A Likely Architecture

We start with a brief primer on current communications networks.

International communications enter the United States by satellite, terrestrial microwave, older copper cable, and newer fiber-optic cable. There are about twenty-five cableheads in the United States. The cableheads are directly connected to gateways, that is, network nodes configured to support a multiplicity of connections and a variety of protocols. Increasingly communications are IP-based. The Internet is the interconnection of many networks (this is the origin of the term) and the connections occur in various ways. For the largest networks, they occur at peering connections: interconnections between administratively separate domains.

Since information about the design of the NSA surveillance architecture has not been made public, it is impossible to know exactly what this architecture might be. However, a current court case gives hints. In late 2005 and

[9]Unless the person in the U.S. was a particular known U.S. person, in which case a warrant was needed.

spring 2006 the *New York Times* and *USA Today* revealed post-September 11th warrantless wiretapping by the NSA. Shortly afterwards, civil-liberties groups and individuals sued AT&T over the "illegal spying of telephone and Internet communications." Affidavits filed in *Tash Hepting et al. v. AT&T Corporation et al.,*[10] describe the architecture for NSA surveillance at the AT&T switching office in San Francisco. Although the administration initially denied the warrantless wiretapping, Director of National Intelligence Mike McConnell revealed in late August 2007, that the NSA had begun a program of warrantless wiretapping sometime after September 11th [8]. Meanwhile AT&T has acknowledged the authenticity of the documents describing the layout and configuration for the secure room of the AT&T San Francisco office in which the electronic surveillance took place[9].

Our discussion is based on these documents and on affidavits submitted by two expert witnesses, Mark Klein (a technician in the AT&T San Francisco office) [10] and J. Scott Marcus (a designer of large scale IP-based data networks, former CTO at GTE Internetworking and at Genuity, and former senior advisor for Internet Technology at the Federal Communications Commission) [11]. Optical fiber carrying inter-ISP peering traffic associated with AT&T's Common Backbone[11, p. 15] was "split," dividing the signal so that 50% went to each output fiber (the weakened signal on each output fiber still contained sufficient information to allow reading the communications) [11, pp. 12-14]. One of the output fibers was diverted to the secure room; the other carried the communications on to the AT&T switching equipment. The secure room contained traffic analyzers and logic servers made by Narus Inc.; Narus states that such devices are capable of doing real-time data collection (recording data for consideration) and capture at high data rates. Certain traffic was selected and sent over a dedicated line to a "central location." The setup in the San Francisco office was one of many throughout the country, including in Seattle, San Jose, Los Angeles, and San Diego [10, p. 7]. According to Marcus's affidavit, the diverted traffic "represented all, or substantially all, of AT&T's peering traffic in the San Francisco Bay Area," [11, p. 24] and thus, "the designers of the ... configuration made no attempt, in terms of location or position of the fiber split, to exclude data sources comprised primarily of domestic data" [11, pp. 24-25].

[10]United States Second District Court for Northern California, Case 3:06-cv-0672-vrw.

4 Call Detail Records

To understand how intelligence agencies determine which calls to wiretap, we begin with a brief discussion of signals intelligence and then consider the role of CDRs.

Signals intelligence is organized into a seven-step process: access, collection, processing, exploitation, analysis (intelligence analysis), reporting, and dissemination. The first three are of particular concern here. Access is what happens at a radio, a fiber splitter, a tap on a wire, or a tap in a telephone switch. Collection is the process of recording signals for consideration. Recorded signals may be kept briefly or for very long periods.

Processing is shorthand for selecting the information you want (and filtering out the information you don't). As in any learning process, if you can find information at all, you often have too much of it and must sort what interests you from what doesn't. This is where CDRs come in; they provide an amazingly effective guide to communications activity but have historically been viewed as much less deserving of privacy protection than call content.

Modern telecommunications allow the construction of smooth-running organizations that span the globe; telecommunications are the nervous systems of these organizations. To listen to an organization's communications is to read its mind; following just the pattern of its communications, Call Detail Records, is a long step in this direction. CDRs contain very complete call traffic data (calling and called numbers and location, time of day, call duration, etc.) and provide a window into the past. CDR data is used internally by the phone companies for billing, engineering, marketing, and fraud detection. Unlike a wiretap or pen register, which respectively provide real-time access to the content or number being dialed, a CDR database contains a wealth of data on past calls. Thus an interested government agency need not have the proper legal authorization or technology in place before a call is made but may search the call detail database later, once a suspect has been identified. For international calls and some purely domestic calls, two CDRs exist for each communication, one from the origination point — which may be an interface to another company — and one from the termination.

CDRs are a major privacy risk. Cortes et al. showed, for example, that, even though the calling number had changed, it was possible to identify an individual caller from a 300-terabyte CDR database by simply looking at called number patterns [12]. George Danezis relates a story in which Intel

Corporation researchers studying ambient Bluetooth activity to improve ad-hoc routing protocols issued its staff members Bluetooth devices. One of the discoveries was that a pair of researchers were meeting nightly, a relationship that had not been previously known to the other lab members[13, pp. 7-8].

If the telephone companies are providing real-time access to CDR — and the affidavits in the Hepting case indicates this is the case — NSA could conceivably learn of a call in progress in time to intercept the content. This means that real-time CDR could be used to target which people to wiretap and then do so without a court order. The more tightly-coupled CDR and content collection are, the more likely it is that content wiretapping will occur as a result of CDR information, without regard to the intentions of the parties involved.

5 Monitoring International Internet Traffic (or "Where is that call from?")

Monitoring international traffic requires an effective way to identify whether the communication starts or ends outside the United States. This is a surprisingly difficult problem to solve on today's Internet. Perhaps even more surprisingly, this is not an easy task on the telephone network either. According to a 1998 National Academies study, "the underlying telephone network is unable to provide [caller ID] information with high assurance of authenticity" [14, p. 36]. (Or, to put it another way, CDR is an amazingly effective guide to to communications activity, but the data can't always provide real-time answers to the location of a call.) NSA has worked on the problem, and the agency even has a patent for using time latency to determine a communication's location (U.S. Patent # 6,947,978: Method for geolocating logical network addresses).

Monitoring international traffic requires either (i) limiting monitoring to links that carry only international traffic or (ii) filtering out any traffic transferred between two domestic hosts. The first approach seems easy if monitoring is limited to the cableheads terminating links connecting the U.S. to other countries. The second approach also seems easy, by mapping the IP addresses of the sending and receiving hosts to their geographic locations. However, both approaches have serious technical problems.

While most traffic on international links travels to or from a foreign host, a small amount of *domestic* traffic traverses these links as well. For example, some domestic traffic travels through Canada and then back to the U.S., due to the vagaries of Internet routing[11]. As such, monitoring the links at the U.S. borders, with the goal of warrantless tapping of international traffic, could lead to unintentional tapping of domestic traffic. Because these links operate at very high speed, it is difficult to analyze the measurement data as they are collected. Furthermore, Internet traffic does not necessarily follow symmetric paths — the traffic from host A to host B does *not* necessarily traverse the same links as the traffic from B to A — monitoring both ends of a conversation sometimes requires combining data collected from multiple locations, making this type of monitoring difficult in practice.

Monitoring very close to the sending or receiving host ensures that (i) both directions of the traffic are visible and (ii) the link speeds are typically small enough for detailed data collection. But monitoring near the domestic end-point would almost certainly capture a large amount of traffic exchanged with other U.S.-based hosts. To identify and filter the domestic traffic, the NSA could map the remote host's IP address to a country using registries that identify the institution that owns the IP address block. The problem is that these registries are notoriously incomplete and out-of-date. Instead, the NSA could use existing IP geo-location services (such as Quova, www.quova.com). Although geo-location mapping services are often accurate to small tens of miles, errors of hundreds of miles or more are not uncommon. As such, a host might easily look as though it resides on the opposite side of the border with another country, such as Mexico or Canada. Geo-location services apply techniques to limit these errors, but the techniques are necessarily imperfect.

Even if the geo-location services are accurate, the source and destination addresess in the IP packet do not necessarily correspond to communicating hosts. Some VoIP services, such as Skype, routinely use relay nodes to enable calls between two hosts that could not otherwise communicate, due to a firewall or a Network Address Translator (NAT), a device that enables multiple hosts on a private network to access the Internet using a single public IP address. A relay node is a third machine that may reside in the

[11]This is partially a result of a 1940s AT&T master plan that made the U.S., Canada, and most of the Caribbean one integrated country, with no cableheads, or even international gateways, between them.

same country as one, or both, of the other hosts, or in yet a third country. Depending on where traffic is monitored, the source or destination address may correspond to the relay node, rather than one of the communicating end-points, complicating the efforts to determine whether both end-points are domestic. In addition, some users apply anonymization tools like Tor (The onion router) that intentionally hide the source and destination addresses from packet sniffers. Whether the traffic traverses a relay or an anonymizer, the monitor may capture erroneous IP addresses that do not correspond to the ultimate source and destination of the traffic.

Even if the traffic does not traverse a relay or anonymizer, real-time association of an IP address with a particular person-of-interest is a difficult task. For example, an IP address may correspond to a NAT box. Identifying the particular host responsible for the traffic requires access to transient information available only to the NAT box. Even in the absence of NAT boxes, the IP address of each end host may be assigned dynamically through the Dynamic Host Configuration Protocol (DHCP). Mapping the IP address to a particular host may require DHCP logs from the local site, and these logs are often incorrect [15]. Mapping from the host to a particular user is difficult if the machine is shared among many people, as in a cyber-cafe or an academic lab. In addition, mobile hosts such as laptops or PDAs acquire new IP addresses frequently (see e.g.,[16]).

Even if the communicating end-points can be appropriately identified, determining what application they are running is difficult. In the simplest case, applications are easily discernable from numerical identifiers (i.e., "port numbers") in the data packets. However, some applications do not use well-known port numbers, and others intentionally use port numbers normally reserved for other applications in order to evade detection; for example, some peer-to-peer file sharing applications use port 80. (There is active research in determining the type of traffic using other information.) Such analysis is difficult to perform in real time on high-speed links, such as the links connecting the U.S. to other countries. In addition, a malicious party trying to avoid detection might intentionally pad or jitter the packets to evade detection, adding further complexity to an already difficult problem. Finally, some applications, like Skype, encrypt the data, making it difficult to extract meaningful information about the content of the communication between the end hosts.

The real problem is that these difficulties are intrinsic to the basic design

of the Internet. Additional issues arise when interworking VoIP with other telephony services, such as the public-switched telephone network. For example, an international call may terminate in the U.S., and then use VoIP the rest of the way (and vice versa), requiring joint analysis across the two kinds of communication networks. The many difficulties in accurately distinguishing domestic and foreign communication make it extremely unlikely that an intelligence agency could avoid tapping domestic calls.

6 Risks

The change from a system that wiretapped particular lines upon receipt of a wiretap order specifying those lines to one that sorts through high-speed transactional data in real time and selects communications of interest is massive. Can such a selection process be built securely without risking exploitation of U.S. communications by others? We see a number of serious risks that must be addressed.

Risk 1: Risk of exploitation by opponents: A system that is accessing domestic communications necessarily poses a greater direct risk to communications of Americans than a surveillance system that is fielded overseas. While it was undoubtedly the case that, even prior to the PAA, U.S. systems were vulnerable to surveillance, building surveillance systems costs money. The system designed as a result of the PAA must not provide foreign powers an easier way to gain access to U.S. communications.

Risk 2: Removal of safeguards by communications carriers: Previous approaches to foreign intelligence surveillance of U.S. persons went through the communications carriers, who combine technical expertise regarding communications with responsibility for their customers' security and privacy. What risks are introduced by leaving a single entity in charge of selection and retention decisions? This process provides no recourse in cases where "mistakes were made."

Risk 3: Lack of inherent technical minimization of traffic: Intercepting at switches creates unnecessary risks because the switches handle domestic as well as foreign communications.

Risk 4: Remote control of filters; Who controls the filter? Is NSA designing sufficiently robust mechanisms to assure complete control?

Risk 5: Domestic traffic entering the NSA collection system: It is likely that the current surveillance architecture filters out most "US-person traffic" before such traffic gets to NSA headquarters at Fort Meade. Does the design of the expanded surveillance system take into account that much of the traffic entering the NSA system will be purely domestic?

Risk 6: Large attack surface: Because of the likelihood that similar NSA systems are being deployed around the globe, it is likely that the system design is well known. Thus risks include the ability by various enemies to modify content capable of evading the NSA's controls, thus reducing the value of the surveillance system — and increasing the risk of even further surveillance down the road.

Risk 7: Call Detail Record information: CDR systems were originally intended to be used by telephone company employees for billing; later uses include determining customer usage patterns and thus anticipating future needs. It is a truism in the security field that problems frequently occur when new uses are found for an old system, since the protection mechanisms and system architecture were never designed for such uses. Will new vulnerabilities be created when copies of this data are sent to law enforcement or intelligence agencies? It is impossible to give a definitive answer, but the past history of such changes do not leave us sanguine.

There are also ways in which the Protect America Act enables an architecture that may reduce risk. Being able to place equipment on U.S. soil reduces the need to place equipment abroad. Beyond the direct security risks to equipment, which could be alleviated by high quality shielding and tamper resistance, there is an intrinsic risk. When intercept capability is installed in other countries' communication systems the privilege must be paid for — often by sharing information. Host countries might demand not only a share of the intelligence take — whether this could ever pose a threat to U.S. communications is hard to assess — but inspection authority over the installation and information about the techniques. Intercept facilities hosted by foreign governments are expected not to spy on the host countries. However, the charge that they are doing so is often made, and the host countries quite reasonably insist on taking measures aimed at preventing this.

Note that we have not enumerated all possible future scenarios. In particular, the security risks will be exacerbated by the direction of the Internet's development. The Internet is currently a network with only millions of de-

vices connected to it, but the world is rapidly moving to a situation in which billions of small, resource-limited devices such as radio-frequency ID (RFID) tags and sensors will use networks for communication and control. While many of these devices will be on local-area networks, others will make use of the Internet. Any future surveillance architectures must take such growth and directions into account.

The Greek wiretapping case and the number of serious security problems in DCS 3000 — a system that attempts far fewer accesses to communications than the PAA envisions — heighten our concern about the security of the system that will implement the PAA. While we would expect an NSA-designed system to include better technical controls than DCS 3000 does, we know the importance of outside evaluation even in national-security systems. (Recall, for example, the 1994 attack by Matt Blaze against the viability of the Law Enforcement Access Field in the Clipper system [17]. This attack did not break the encryption algorithm, but instead enabled a user to use the encryption system while disabling law enforcement access to the keys. An outsider uncovered a flaw in the NSA-designed, publicly-fielded cyrptosystem.)

The Protect America Act, a law quickly proposed and enacted, potentially vastly increases the number of Americans whose communications and communication patterns will be studied. This sets up access to U.S. communications, a target of great value. The nation may build for its opponents something that would be too expensive for them to build for themselves: a system that allows them to see the intelligence interests of the U.S., a system that may tell them how to thwart those interests, and a system that might be turned to intercept the communications of American citizens and institutions. It is critical that the new surveillance system neither enable exploitation of U.S. communications by unauthorized parties nor permit abuse by authorized ones.

7 Recommendations

We see three serious security risks that have not been adequately addressed (or perhaps not even addressed at all): the danger that the system can be exploited by unauthorized users, the danger of criminal misuse by a trusted insider, and the danger of misuse by the U.S. government. Our recommendations are based on these concerns.

Minimization is critical. Allowing collection of calls on U.S. territory necessarily entails greater access to the communications of U.S. persons; the architecture must minimize collection of both the call details and the content of these communications. The best way to prevent problems is to intercept as early as possible: at the cableheads; such a solution, by decreasing the number of interception points will simplify the security problem. Surveilling at the cableheads will help minimize collection but it is not sufficient. Intercepted traffic should be studied (by geo-location and any other available techniques) to determine whether it comes from non-targeted U.S. persons and if so, discarded before any further processing is done.

The Need for Oversight

In 2002 Attorney General John Ashcroft proposed changing FISA procedures. The FISA Court, whose job it is to review FISA wiretapping warrant applications, was not pleased with this, in part because of mistakes that had occurred in earlier FISA applications. The court issued a report criticizing the proposal [18] and FBI mishandling of the *wall* between foreign intelligence cases and criminal investigations, "In virtually every instance, the government's misstatements and omissions in FISA applications and violations of the Court's orders involved information sharing and unauthorized disseminations to criminal investigators and prosecutors." An extremely important check on government abuse is oversight. As the founders of the United States knew, another branch of the government can provide the objectivity necessary for such an investigation. Public knowledge also matters. When the FISA court was dissatisfied with the administration in 2002, it declassified its opinion, helping to shape the later debate on the PATRIOT Act renewal and other administration requests for changes in the wiretap laws.

The architecture should be developed in collaboration with the communications carriers, organizations with long experience of responsibility for the privacy and security of their customers' communications. That responsibility should *not* be removed from the communication providers.

Oversight is necessary to prevent abuse and ensure information assurance. Independent oversight of operations is also essential and is a fundamental tenet of security. To assure independence the overseeing authority should be as far removed from the intercepting authority as practical.

To guarantee that electronic surveillance is effective and free of abuse *and* that minimization is in place and working appropriately, it is necessary

that there be frequent, detailed reports on the functioning of the system. Of particular concern is the real-time use of CDR for targeting content, which must neither be abused by our own government nor allowed to fall into unauthorized hands. For full oversight, such review should be done by a branch of government different from the one conducting the surveillance. We recommend frequent ex post facto review of the CDR-based real-time targeting.

The oversight mechanism must include outside reviewers who regularly ask, "What has gone wrong lately — regardless of whether you recovered — that you have not yet told us about?"

Security of U.S. communications has always been fundamental to U.S. national security. The surveillance architecture implied by the Protect America Act will, by its very nature, capture some purely domestic communications, risking the very national security that the act is supposed to protect. In an age so dependent on communication, the loss may be greater than the gain. To prevent greater threats to U.S. national security, it is imperative that proper security — including minimization, robust control, and oversight — be built into the system from the start.

References

[1] United States Senate, Committee of the Judiciary (1977), *Legislative History P.L. 95-511 Foreign Intelligence Surveillance Act,* Report 95-604, Ninety-Fifth Congress, First Session, November 15, 1977.

[2] Prevelakis, Vassilis and Diomidis Spinellis, "The Athens Affair," *IEEE Spectrum,* July 2007, pp. 18-25.

[3] Information Assurance Section, Federal Bureau of Investigation, "Controlled Interface 100 (CI-100) System Security Plan (SSP) DCS-3000 to EDMS," April 16, 2007.

[4] Solove, Daniel and Marc Rotenberg, *Information Privacy Law,* Aspen Publishers, 2003.

[5] United States Senate, Select Committee to Study Governmental Operations with Respect to Intelligence Activities (1976), *Intelligence Activi-*

ties and the Rights of Americans, Final Report: Book II, Report 94-755, Ninety-Fourth Congress, Second Session, April 23, 1976.

[6] Diffie, Whitfield and Susan Landau, *Privacy on the Line: The Politics of Wiretapping and Encryption, updated and expanded edition,* MIT Press, 2007.

[7] Inspector General, Federal Bureau of Investigation, *A Review of the Federal Bureau of Investigation's Use of the National Security Letters,* March 2007.

[8] Roberts, Chris, "Transcript: Debate on the foreign intelligence surveillance act" *El Paso Times Article,* August 22, 2007, http://www.elpasotimes.com/news/ci_6685679 (last viewed September 3, 2007).

[9] Exhibit A, *Tash Hepting et al. v. AT&T Corporation et al.,* United States Second District Court for Northern California, Case 3:06-cv-0672-vrw, June 8, 2006.

[10] Klein, Mark, affidavit in *Tash Hepting et al. v. AT&T Corporation et al.,* United States Second District Court for Northern California, Case 3:06-cv-0672-vrw, June 8, 2006.

[11] Marcus, J. Scott, affidavit in *Tash Hepting et al. v. AT&T Corporation et al.,* United States Second District Court for Northern California, Case 3:06-cv-0672-vrw, June 8, 2006.

[12] Cortes, Corinna, Daryl Pregibon and Chris Volinsky, "Computational Methods for Dynamic Graphs," AT&T Shannon Labs, January 9. 2004.

[13] Danezis, George, "Introducing Traffic Analysis: Attacks, Defences and Public Policy Issues."

[14] Schneider, Fred, *Trust in Cyberspace,* Computer Science and Telecommunications Board, National Research Council, 1999,

[15] Clayton, Richard,*Anonymity and Traceability in Cyberspace,* University of Cambridge Computer Lab, Technical Report Number 653, November 2005.

[16] Bellovin, Steve, Matt Blaze, Ernie Brickell, Clinton Brooks, Vint Cerf, Whitfield Diffie, Susan Landau, Jon Peterson, John Treichler, "Security Implications of Applying the Communications Assistance to Law Enforcement Act to Voice over IP," http://www.itaa.org/news/docs/CALEAVOIPreport.pdf, 2006.

[17] Blaze, Matt "Protocol Failure in the Escrowed Encryption Standard," *Proceedings of Second ACM Conference ony Computer and Communications Security,* Fairfax, VA, November 1994.

[18] United States Foreign Intelligence Surveillance Court, Memorandum Opinion (as Corrected and Amended,) May 17, 2002, in United States Senate, Committee on the Judiciary, 2002, *The USA PATRIOT Act in Practice: Shedding Light on the FISA Process,* Hearing on September 10, 2002, S. Hrg. 107-947, One Hundred Seventh Congress, Second Session.

[19] United States Foreign Intelligence Surveillance Court, *In Re All Matters Submitted to the Foreign Intelligence Surveillance Court, Memorandum Opinion,* May 17 2002.