[BY EMAIL (techconnect@sfgov.org)]

February 21, 2006

Chris A. Vein
Acting Executive Director
Department of Telecommunications and Information Services
City & County of San Francisco
875 Stevenson Street, 5th Floor
San Francisco, CA 94103-0948

Re: TechConnect RFP 2005-19 / Privacy and Municipal Broadband

Dear Mr. Vein,

On October 19, 2005, the ACLU of Northern California, Electronic Frontier Foundation (EFF), and Electronic Privacy Information Center (EPIC) submitted comments to TechConnect concerning privacy issues raised by municipal broadband access.[1] In that letter, we raised a series of privacy issues that sought to focus attention on whether uses of the municipal broadband network will have secure and private access to the Internet. We applaud TechConnect for including the privacy issues we raised in RFP 2005-19.

At section 2.11 of the RFP, TechConnect requested proposers to provide a copy of their privacy policy, to certify that it complies with applicable law, and to explain how it will communicated to users. TechConnect also requested proposers to explain how they will address a series of privacy issues raised in our October letter.

In this letter, we stress that the city should consider minimum standards for the privacy issues raised by the RFP. Privacy notices are not enough. The short history of E-commerce has shown that companies often issue privacy policies that are substantively weak and extend to users few legal rights to redress privacy violations. Minimum standards are necessary for each of the privacy questions posed to proposers in order to guarantee respect for users' rights.

To assist TechConnect in this process, we suggest model minimum standards to each of the questions included in the RFP. We also urge TechConnect to consider the safeguards recommended in EFF's "Best Practices for Online Service Providers," which describes legal policies and technical procedures for protecting privacy. [2]

[1] Letter from Nicole A. Ozer, Technology and Civil Liberties Policy Director, ACLU of Northern California; Kurt Opsahl, Staff Attorney, EFF; & Chris Jay Hoofnagle, Senior Counsel, EPIC West Coast Office, to San Francisco TechConnect, Oct. 19, 2005, available at http://epic.org/privacy/internet/sfws10.19.05.html and attached as Appendix A.

[2] Attached as Appendix B. These guidelines were developed by technical and legal experts for service providers that wish to handle user data ethically. They are available at http://www.eff.org/osp/.

• What personal information is collected about users?

Providers should take all reasonable steps to enable use of the network without the collection of personal information. Data collection should accommodate the individual's right to communicate anonymously and pseudonymously through the service.

"Operation of the network" refers to actions necessary to technically run the network. This includes actions necessary for guaranteeing service availability, billing, network testing, and reasonable security measures.

• How is this information used?

Providers should use information for purposes necessary to operation of the network.

• How long is this information stored?

Providers should specify a data retention schedule for all information collected. Providers should store information only for so long as needed to operate the network. In no event should data be kept for more than a few weeks. Information that needs to be kept to provide enhanced services should be the minimum necessary to provide the service, be deleted as soon as operationally possible, and providers should employ technical measures to shield this information including obfuscation or aggregation.[3]

• With whom is this information shared?

Providers should only share information for purposes necessary to operate the network. Entities that receive personal information should be held to the same privacy standards as the provider.

• Is this information commercialized in any way?

Providers should not commercialize personal information collected in the course of operating the network unless the user opts in to such uses of data.

"Opt in" refers to affirmative consent, a situation where the user can employ the network for basic services, and affirmatively choose to enroll in additional services. That is, a user does not "opt in" to the service by simply using the network. Providers should obtain affirmative consent again where there is a material change to information collection or use policies. Furthermore, an expression of affirmative consent should only be effective for one year.

• Is this information correlated to a specific user, device or location?

Providers should correlate information to specific users, devices, or locations only to the extent necessary to operate the network.

[3] See Appendix B.

• Are mechanisms available to allow users to opt in or opt out of any service that collects, stores, or profiles information on the searches performed, websites visited, e-mails sent, or any other use of the Network?

Opt in should be the standard for services that exceed the basic function of providing individuals with Internet access.

• Are mechanisms available to allow users to opt in or opt out of any service that tracks information about the user's physical location?

Providers should take all reasonable steps to enable location-based services without creating a tracking or logging mechanism that will create records of individuals' location.

• Are users enumerated or assigned any unique number that can be used to track them from session to session?

Providers should take all reasonable steps to design the system to prevent enumeration from session to session.

Providers should obtain a user's affirmative consent before enumerating users across sessions.

• Are policies in place to respond to legal demands for users' personal information in accordance with applicable laws?

Providers should comply with legal demands for users' personal information only after verifying the legal sufficiency of the request, and notify the subject of the request as quickly as possible before providing information to the requestor. A good model is set forth by the Cable Communications Policy Act (47 USC § 551). That act, which also applies to satellite television providers, specifies a procedure where individuals are notified before their information is revealed to others pursuant to legal process. It was passed to protect individuals' television viewing habits from disclosure, information that is at least as sensitive as e-mail and web browsing records. It has been in effect since 1984, and accordingly many companies have processes to comply with its standards.

• Are users allowed access to all information collected about them?

Users should be able to access personal information collected and maintained by the provider and its affiliates or partners.

• Are users provided with a mechanism to review this information and to correct inaccuracies or delete information?

Providers should extend reasonable opportunities for users to correct or delete personal information collected and maintained by the provider and its affiliates or partners.

Thank you for considering our comments. If we can be of further help, please feel free to contact us.

Nicole A. Ozer
Technology and Civil Liberties Policy Director
ACLU of Northern California
nozer@aclunc.org
415-621-2493

Kurt Opsahl
Staff Attorney
Electronic Frontier Foundation (EFF)
kurt@eff.org
415-436-9333

Chris Hoofnagle
Senior Counsel and Director, West Coast Office
Electronic Privacy Information Center (EPIC)
hoofnagle@epic.org
415-981-6400

ELECTRONIC PRIVACY INFORMATION CENTER

Joint Letter on San Francisco Wireless Internet Access

[BY MAIL AND EMAIL (techconnect@sfgov.org)]

October 19, 2005

TechConnect RFI/C 2005-07
Dept. of Telecommunications and Information Services
City and County of San Francisco
875 Stevenson St., 5th Floor
San Francisco, CA 94103

Re: Privacy Issues Associated with Municipal Wireless Internet Access

The American Civil Liberties Union of Northern California (ACLU), Electronic Frontier Foundation (EFF), and Electronic Privacy Information Center West Coast Office (EPIC West) submit these comments on TechConnect RFI/C 2005-07 in response to information received by the City concerning municipal wireless Internet access.

The ACLU is a nonprofit, nonpartisan organization dedicated to the defense and promotion of the civil liberties and civil rights secured by the state and federal constitutions and related statutes. The ACLU of Northern California, based in San Francisco, is the largest ACLU affiliate in the nation, with 50,000 members spanning communities from Crescent City to Fresno.

EFF is a nonprofit donor-supported membership organization working to protect fundamental rights regardless of technology; to educate the press, policymakers, and the general public about civil liberties issues related to technology; and to act as a defender of those liberties. Among its various activities, EFF opposes misguided legislation, initiates and defends court cases preserving individuals' rights, launches global public campaigns, introduces leading edge proposals and papers, hosts frequent educational events, engages the press regularly, and publishes a comprehensive archive of digital civil liberties information on the most linked-to web sites in the world at www.eff.org.

EPIC is a not-for-profit research center founded in Washington, DC in 1994 to focus public attention on privacy and open government. EPIC's West Coast office is based in San Francisco, and concentrates on consumer privacy issues.

Municipal wireless offers our society an opportunity to address digital divide issues, to give more individuals access to more information, to keep San Francisco competitive with other cities offering free or low-cost wireless, and many other valuable social ends.

We are heartened that the City has already recognized the profound importance of proper privacy protections for the municipal wireless system by stating in the RFI that:

The City anticipates a Network that protects the privacy of users, respects consumer choice, and fosters diversity of information and ideas.

Additionally, by asking vendors to specify the privacy policies and security standards that will be put in place "to protect the privacy of--and information transmitted by--users," the City has wisely made privacy a key policy standard for municipal wireless Internet access.

We have surveyed the privacy and free speech issues raised by the proposals and have provided some

Appendix A

concrete questions to assist the City in addressing these issues in a meaningful manner.

The Importance of Privacy

Privacy is an inalienable right under the California State Constitution. As an inalienable right, a citizen's privacy is not to be bought, sold, or bargained away.[1] Proposition 11, which added the privacy right to the State Constitution recognized that both the government and the private sector pose risks to information privacy:

The right of privacy is the right to be left alone. It is a fundamental and compelling interest. It protects our homes, our families, our thoughts, our emotions, our expressions, our personalities, our freedom of communion, and our freedom to associate with the people we choose. It prevents government and business interests from collecting and stockpiling unnecessary information about us and from misusing information gathered for one purpose in order to serve other purposes or to embarrass us.[2]

As the ballot proposition recognized, privacy is important because it gives individuals a zone of autonomy in which they can explore intellectual interests, personal relationships, and other socially valuable ends without fear of intrusion and oversight.[3] The "ability to speak one's mind without the burden of the other party knowing all the facts about one's identity can foster open communication and robust debate."[4]

San Franciscans have the right to a network that respects privacy and autonomy, allowing users to explore what the Internet has to offer, including information about medical conditions and the use of online banking, without fear of surveillance or intrusion.

We note that these principles cannot be viewed as mere aspirations. In general, when a government entity establishes and assumes responsibility for a system that provides public electronic communications services, that constitutes "state action" for constitutional purposes and requires the City to comply with the dictates of the state and U.S. Constitutions, including the First and Fourth Amendments.

Comment on Question 8

Question 8 from the RFI solicits comment on how to implement both privacy and freedom of expression on the network:

What privacy policies and security standards will you put in place to protect the privacy of--and information transmitted by--users?

We wish to emphasize that this question raises two important issues: first, how will the network protect the privacy of users. Second, how will the network protect information transmitted by users? These two questions, while they sound similar, are different. Many of the commercial responses to the RFI focus exclusively on the second question, emphasizing how their approach will protect against malicious users of the system. Such protection is critical to operation of the network. But both must be addressed to fully serve the City's policy standard of developing a network that protects the privacy of users and fosters diversity of information and ideas.

Protecting the Privacy of Users

A dialogue on how to protect users' information must encompass the following issues:

· Will users be enumerated, that is, assigned a unique number that can be used to track an individual from session to session?

Computers accessing the Internet must be identified in order to route content to the appropriate user. Computers must also be identified when they "host" or provide resources to other users. However, in most situations, there is no requirement that a unique identifier be employed to keep track of what an Internet user

does in a previous session. Linking session activity and creating a log of activities creates a profile of a user's activity. It is well settled that the First Amendment protects privacy of association, such as the sanctity of group membership lists, as well as the right to speak anonymously. Accordingly, it must be permissible for system users to use technical measures that shield their identities.

Special attention must be paid to whether users will be tracked by identifiers that are unchangeable, such as the "MAC" identifier embedded in network cards or by "usernames" assigned by the service. Such vendor plans can lead to a significant reduction in privacy.

- **Will the service attempt to commercialize data?**

A main goal of municipal wireless is to bridge the digital divide. Much of the population affected by the divide cannot exercise choice in the marketplace and choose a privacy-sensitive service provider. We therefore think it especially important that the city not bargain away privacy by choosing a service provider that commercializes users' data. In addition, we have specific privacy concerns with several of the proposals that include commercialization of the data.

For example, we are skeptical of claims that systems that use transactional logs to target advertising are truly anonymous. Any system that scans users' Internet usage for content can be tweaked to serve other purposes, or altered to track specific individuals. Furthermore, such targeting could lead to harm where, for instance, a family computer is used to research a sensitive and very private issue such as health concerns or political activity, and a later user of the same computer is presented with advertising pertaining to that earlier user's browsing.

We are similarly skeptical of bids where the service provider seeks to commercialize user or transactional data through affiliate or non-affiliate sharing agreements. If such a provider is chosen, the standard should be opt-in. Affirmative consent should be obtained before data is used for marketing by affiliates or non-affiliates.

- **Will the service provider resist legal demands for users' personal information?**

Because service providers are the vital link between individuals and Internet resources, they face legal pressures from other network users, industries, and governments to disclose personal information. As courts have noted, users "who have committed no wrong should be able to participate online without fear that someone who wishes to harass or embarrass them can file a frivolous lawsuit and thereby gain the power of the court's order to discover their identities."[5] Typically, when user information is sought, the service provider is the first entity informed of the request.

This issue is especially sensitive when the service provider is, as here, a state actor, and may therefore face additional pressures from government to provide information about individuals' Internet use. Except in circumstances where law enforcement presents a court order binding the service provider to secrecy, the service provider should inform the user of the request as soon as possible, and, in any event, the service provider should be prepared to litigate to avoid disclosing data if the request is legally insufficient.

The City should discuss procedures and policies for protecting users' personal information in the hands of vendors. Specifically, to protect and preserve users' rights to speak freely, the City should:

(1) ensure that the service provider will provide notice, within no more than seven days of receipt of a subpoena, to each person whose personal information is sought;

(2) allow the user at least fourteen days from the time notice was received to file a motion to quash; and

(3) prohibit any disclosure pending the disposition of any motion to quash.

- **How long will server logs be maintained?**

As mentioned above, service providers can be the focus of extraordinary requests for users' data. As an

intermediary, a service provider finds itself in a position to collect and store detailed information about its users and their online activities that may be of great interest to third parties. As a result, any municipal wireless service provider must deal with requests from law enforcement and lawyers to hand over private user information and logs. Yet, compliance with these demands takes away from the City's goal of providing users with reliable, private and secure network services.

Reducing the amount of time that the system stores user and transactional data will enhance privacy and reduce the costs and burdens of responding to requests for user data. [6] Personal information about users should be kept only as long as it is operationally necessary, and in no event for more than a few weeks. Aside from reducing retention, privacy risks can be managed by eliminating or obscuring personally identifiable information or by tracking usage in the aggregate rather than by personal identifiers.

We urge the City to ensure that its municipal wireless vendor adopt procedures along the lines of EFF's "Best Practices for Online Service Providers," which describes legal policies and technical procedures for protecting privacy.[7] Clear policies will conserve resources, help safeguard private data, and preserve freedom of expression online.

Protecting Information Transmitted by Users

The question of how to protect information transmitted by users can be addressed in a number of ways, and this list is not comprehensive. A dialogue on these issues should include the following considerations:

- **Will data be protected from interception by others?**

There must be measures to protect information transmitted by users from interception by others. A municipal wireless network will not be usable for personal activities, such as medical and banking activities if data can be intercepted and understood by others.

- **Will data be authentic? Will it be protected from corruption by others?**

There must be measures to ensure that the data flowing between the user and service provider is authentic. That is, there must be measures to shield users from being sent data that appears to be legitimate, but is really sent by a malicious actor. A typical example of this is the "man-in-the-middle" attack, where a malicious actor inserts himself between the service provider and the user in order to defraud one or both of the parties.

- **Will there be balance in addressing unlawful users?**

Malicious hackers and other bad actors will attempt to use the system. The City should strive to address these issues without punishing all users through identification requirements, such as the enumeration methods mentioned above. A few bad apples should not limit the network's ease of use for everyone else.

Where possible, unlawful uses should be addressed though techniques that do not involve identification. The service provider should track MAC addresses or usernames only after it determines that a specific computer is being used for unlawful purposes.

- **Will users have access to true end-to-end encryption?**

True end-to-end encryption allows communication that is shielded by mathematical algorithms from the user's computer to an online resource. It is not clear whether commercial commentators are proposing to offer true end-to-end encryption, or simply user-to-client encryption. In user-to-client encryption, the information is decrypted and sent "in the clear" after it reaches the service provider. Where possible, the system should employ true end-to-end encryption in order to properly protect user privacy.

Thank you for considering our comments. If we can be of further help, please feel free to contact us.

Nicole A. Ozer

Technology and Civil Liberties Policy Director
ACLU of Northern California
nozer@aclunc.org
415-621-2493

Kurt Opsahl
Staff Attorney
Electronic Frontier Foundation (EFF)
kurt@eff.org
415-436-9333

Chris Hoofnagle
Senior Counsel and Director, West Coast Office
Electronic Privacy Information Center (EPIC)
hoofnagle@epic.org
415-981-6400

---

[1] California law restrains the alienability of privacy rights in many respects. *See e.g.* Cal. Civ. Code § 1798.84(a) (making waivers of a variety of California-specific privacy protections inalienable by contract); Consumer Credit Reporting Agencies Act, Cal. Civ. Code § 1785.36.

[2] Proposed Amendments to Constitution, California Office of the Secretary of State, Nov. 7, 1972, available at http://library.uchastings.edu/ballot_pdf/1972g.pdf.

[3] Julie E. Cohen, *Examined Lives: Informational Privacy and the Subject as Object*, 52 Stan. L. Rev. 1373 (2000).

[4] *Columbia Insurance Co. v. Seescandy.com*, 185 F.R.D. 573, 578 (N.D. Cal. 1999).

[5] *Columbia Insurance Co. v. Seescandy.com*, 185 F.R.D. at 578.

[6] Because of Constitutional and statutory regulations limiting government access to user data, we assume that the City itself will not have access to personal data collected by the service provider absent appropriate legal process.

[7] These guidelines were developed by technical and legal experts for service providers that wish to handle user data ethically. They are available at http://www.eff.org/osp/.

EPIC Privacy Page | EPIC Home Page

Last Updated: October 19, 2005
Page URL: http://www.epic.org/privacy/internet/sfws10.19.05.html

Best Data Practices for Online Service Providers from the Electronic Frontier Foundation

Introduction

Online service providers (OSPs) are vital links between their users and the Internet, offering bandwidth, email, web and other Internet services. Because of their centrality, however, OSPs face legal pressures from all sides: from users, industry, and government. As an intermediary, the OSP finds itself in a position to collect and store detailed information about its users and their online activities that may be of great interest to third parties. The USA PATRIOT Act also provides the government with expanded powers to request this information. As a result, OSP owners must deal with requests from law enforcement and lawyers to hand over private user information and logs. Yet, compliance with these demands takes away from an OSP's goal of providing users with reliable, secure network services. In this paper, EFF offers some suggestions, both legal and technical, for best practices that balance the needs of OSPs and their users' privacy and civil liberties.

Are you an OSP?

If you think you might be an OSP, you probably are. As defined by the Digital Millennium Copyright Act (DMCA)[1], an OSP is any "entity offering the transmission, routing, or providing connections for digital online communications" or any "provider of online services or network access, or the operator of facilities therefor." The Electronic Communications Privacy Act (ECPA) defines two subcategories of OSPs: "electronic communication services"[2] and "remote computing services."[3] Access to users' information under ECPA is determined in large part by which of these subcategories fits your OSP. As a general rule, email and connectivity services would be electronic communication services, while website hosting would be considered a remote computing service. This means that virtually *any* website or access intermediary, not just established subscriber-based businesses, can be considered an OSP under the law. Indeed, even individuals may be "accidental OSPs" if they set up WiFi access points to share Internet connectivity with friends and neighbors.

How can OSPs develop sane network policies to protect themselves from legal liability and respond to subpoenas and court orders?

A key strategy is to minimize the amount of information OSPs collect and store in the first place. Unless they are in a specially regulated industry (finance or health care, for example), no law requires OSPs to collect and store information about their users. This means that OSP owners and operators are free to develop and implement reasonable data retention policies. Our suggestions for these policies, elaborated below, are for

---

1 See http://www4.law.cornell.edu/uscode/17/512.html

2 "any service which provides to users thereof the ability to send or receive wire or electronic communications…" http://www4.law.cornell.edu/uscode/18/2510.html

3 "the provision to the public of computer storage or processing services by means of an electronic communications system." http://www4.law.cornell.edu/uscode/18/2711.html

Appendix B

informational uses only. If you have any specific questions or concerns about your OSP, please consult an attorney. (EFF contacts are listed below.)

Legal Issues with Requests for User Data or Transactional Information

When law enforcement officers conduct civil or criminal investigations, they must obtain subpoenas, warrants or court orders to retrieve personal information from OSPs. The government may obtain basic subscriber information[4] with only a subpoena, but generally needs a warrant or a court order for more detailed records. These court orders might request the identity of the user, email message content, visited URLs, search queries, or any other kind of recorded information.

While the ECPA requires OSPs to disclose information in response to a legal process, it also prohibits certain disclosures without a proper request. For example, the ECPA prohibits an electronic communications service provider from producing the contents of electronic communications (i.e. the body of an email message or arguments in a URLs query string), even if served with a subpoena, except in limited circumstances. Thus, the OSP must evaluate the legal process carefully before retrieving the information and furnishing it to law enforcement. Often, this takes a great deal of time and resources, and the OSP should consult an attorney.

An OSP can keep its costs and risks down by setting clear policies about data retention. There are no laws that require OSPs to retain personally identifiable information (PII) or activity logs about users, unless this information is subject to other government regulation (such as financial transactions) or the OSP has received a backup preservation request from the government.[5] EFF believes that PII about users should be kept only so long as it is operationally necessary, and in no event for more than a few weeks. (We explore this issue in more detail in the technical section below.)

OSPs cannot be forced to provide data that does not exist. EFF suggests that OSPs draft an internal policy that states that they collect only limited information and do not retain any logs of user information on their networks for more than a few weeks. If a court order requests data that is more than a few weeks old, the OSP owner can simply point to the policy and explain that he cannot furnish the requested data. This saves the OSP time and money, while also providing the OSP with a X-week long cushion to examine their own logs.

Civil or criminal subpoenas may also be issued for identifying information called "subscriber information." This includes name, address, phone number and any other personal information that the OSP has collected from the user. Subpoenas for subscriber information are usually aimed at uncovering the identities of people who are posting anonymous comments. A typical scenario would be someone posting negative comments about a company. The company lawyer sends a subpoena for subscriber information about the poster, perhaps to determine whether it is an employee who can be fired or sued. Sometimes, these demands are simply used as a form of harassment, without any sound legal basis or intent to follow through with the legal process. In many cases, once

---

[4] Such as the user's name, address, records of session times and duration, IP or other network address.

[5] See http://www4.law.cornell.edu/uscode/18/2704.html.

the user's identity has been forcefully revealed, the requesting company takes extra-legal action against the user by firing or taking other forms of retribution against him.

Another common civil subpoena is a DMCA "Subpoena To Identify Infringer," which requires an OSP that hosts allegedly infringing material to disclose "information sufficient to identify the alleged infringer … to the extent such information is available to the service provider." Unlike an ordinary subpoena, the DMCA subpoena does not require a lawsuit to be filed first, but it must be accompanied or preceded by a notification of alleged infringement that has specific requirements. However DMCA subpoenas only apply to OSPs that actually host a work; not ISPs that merely provide connectivity, such as in the case of peer-to-peer filesharing. DMCA subpoenas also only apply to claims of copyright infringement.

In other circumstances, individuals may request information about a particular user, complaining that the user has engaged in harassment or other bad acts. In such cases, the OSP may be sympathetic to the alleged victim and be tempted to provide the information directly. However, an OSP has no way to verify the truth of the story and providing this information without legal process could subject the OSP to liability from the user. The safest course is to require a subpoena or other legal process before providing user information to anyone.

Remember, Internet users have a right to anonymous free speech under the First Amendment. An OSP receiving one of these subpoenas should notify the user as quickly as possible before responding to it.[6] This will give the user an opportunity to object to disclosure of his or her identity (technically, by filing a "motion to quash the subpoena"). Both Virginia and Arkansas currently require OSPs to give notice to users prior to turning over PII; California is considering a similar bill. Similar laws may soon be enacted in other states. Giving notice may also protect the OSP against lawsuits from users. It is important to set a data retention policy in place now that will protect your users' privacy and your own legal liability.

Technical Issues

Up until now, we have discussed EFF's recommendations for best practices to help OSPs minimize the cost of legal overhead. There is also a technical side to this issue. By being consumer-conscious about logging PII, network administrators can proactively save company resources and protect the privacy of their users at the same time. Upon receipt of a court order, OSPs are compelled by law to comb through their logs to extract the requested data using their own resources.[7] Thus, the cost of handling court orders scales proportionally with the retention of user traffic logs.

A general best practice to mitigate this problem is to log only enough information to maintain and upkeep the OSP's intended services—no more, no less. Logs should be

---

[6] On occasion, court orders to provide user information to the government may be accompanied with a request not to notify the user. In such circumstances, OSPs should consult with an attorney.

[7] In some cases, OSPs can seek reimbursement for the costs of compliance. See e.g. http://www4.law.cornell.edu/uscode/18/2706.html. However, reimbursement may not capture all the costs associated with legal compliance.

stored for a minimal amount of time. The "correct" strategy for a particular OSP will depend on the services they provide to their users. We outline some possible strategies below.

OSPs must first pinpoint, on every server, all logs where PII is being recorded. It's important to remember that IP addresses and MAC addresses are crucial sources of identity-revealing information, and they are often requested in court orders. The most common locations for PII include:

- DHCP logs (IP address-to-MAC address assignments, session times)
- RADIUS logs (user name, IP address assignment, callback telephone number, session time, etc.)
- Web and FTP server logs (client IP address, files accessed, request time, query string, etc.)
- Email server logs (sender/recipient addresses, message date and time, relay hostnames, etc.)
- Firewall and IDS logs (IP addresses, packet payloads, date and time of connections, protocol used, etc.)
- User contact information databases (mailing address, phone number, billing information, etc.)

For each piece of PII being recorded, it is imperative that network administrators justify why they are keeping the information and consider a realistic time limit for retaining the information. These decisions should be recorded in an internal data retention policy. We outline three possible methods for PII-elimination below: these are obfuscation, aggregation and deletion.

Obfuscation

The easiest, but least protective, strategy is to periodically scrub the logs to obfuscate all explicit or deducible PII. Since virtually all OSPs maintain multiple logs and user information databases, providers must ensure that user identity cannot be gleaned when matching two or more processed logs. Setting a reasonable time duration before PII obfuscation allows OSPs to administer and troubleshoot their networks in real-time. The amount of time PII-exposed logs are stored will depend on the service requirements, but of course PII should never be kept any longer than necessary.

Key solutions to wipe PII from logs include:

- Obscuring the last octet of all IPv4 addresses by either using a randomly seeded one-way hash, or replacing it with an arbitrary integer (between 1 and 254).
- Obscuring the third, fourth and sixth octets of all MAC addresses in the same way as above. This will obfuscate both the exact manufacturer ID (first three octets) and the specific device ID (last three octets) being used.
- Obscuring the last four digits of phone numbers, or replacing it with '0000', but keeping the area code and exchange.
- Obscuring or deleting all usernames in e-mail addresses.

- Obscuring or deleting all query strings in URLs (http://www.google.com/search?~~q=electronic+frontier+foundation~~).
- Obscuring or deleting all filenames from URLs (http://www.eff.org~~/IP/DMCA/unintended_consequences.pdf~~).

Some tactics that should not be used include:

- Encrypting PII with either symmetric or asymmetric keys: Any subpoena or court-order can still force OSPs to turn over the encryption keys along with the encrypted data.
- Hashing PII with a non-random, well-known one-way hash: Using trial-by-error, one could match hashed candidate IP addresses with the encrypted IP address to reveal the original data.

When implemented in a timely fashion, obfuscation gives OSPs the flexibility to glean general usage patterns without retaining PII; implemented poorly, OSPs will continue to be subject to the legal consequences of information requests.

Aggregation

A better strategy is to use aggregation techniques to compile general usage statistics followed immediately by log deletion. This allows OSPs to fully discard all logs, including PII-obfuscated logs, after a specified duration of time, but still keep tabs on network access patterns. OSPs can save a substantial amount of resources using this technique, since aggregation requires minimal hard disk space. It also ensures that no specific PII will be retained on OSP servers in the long term.

Consider an OSP which hosts an Internet search engine and wants to track popular search queries. Obfuscation of the query string would not work because it would mask the data the OSP wants to track. Obfuscation of only the IP address (while exposing the query string) could still lead to potential IP address matches and PII leaks. Using aggregation techniques, the OSP can simply extract the query strings from the log file, tally the number of times each query was made, and then delete the file completely. One OSP reported to us they automatically aggregate their web server logs every night, then immediately delete the previous day's logs. This method fully decouples users' identities from their search queries while allowing the OSP to keep track of popular search topics.

Deletion

Obfuscation and aggregation are only effective when used in tandem with log deletion. A strict policy which dictates when the OSP should fully purge logs from hard drives is a mandatory step in minimizing the potential challenges of legal compliance. Decisions on log retention time intervals will vary drastically. Free, open WiFi providers may delete connection logs immediately after log-off, while pay-per-use WiFi providers must keep logs for weeks until billing and collection have been completed. OSPs should note that different types of log files may have different data retention intervals.

Even after logs have been deleted from disk, the PII may still reside on the disk

until that memory segment is reused and written over. Even then, advanced forensic searches of server hard drives could still reveal past data stored on them. These processes may cause OSPs significant disruptions. If possible, you should use strong deletion utilities to fully scrub the hard drives containing deleted logs. This will ensure the removal of all sensitive PII.

The best way to protect against the risk of log artifacts on disk is to never create any user logs in the first place. This is the ideal and safest solution even though it is often impractical. By reconfiguring the logging preferences in server applications, one can easily change the log level to record nothing about network events. But for most OSPs, these logs are necessary for network troubleshooting and security precautions. This is also virtually impossible for large, for-profit providers that need to maintain billing and subscriber contact information. Thus, the best tactic for an OSP is to come up with a safe and sane network policy in which logs are retained for the shortest possible time.

Summary of Recommendations

a. Develop procedures for dealing with legal information requests and providing notice to users.
b. Collect the minimum amount of information necessary to provide OSP services.
c. Store information for the minimum time necessary for operations.
d. Effectively obfuscate, aggregate and delete unneeded user information.
e. Maintain written policies addressing data collection and retention.

Conclusion

OSPs need to understand their legal risks and obligations when codifying their logging practices. They must adopt a reasonable internal data retention policy and follow this policy consistently. Being strict about deleting all PII on servers will protect OSPs from many hidden costs. By taking proactive technical steps, and knowing their legal rights and obligations, OSPs can simultaneously maximize the privacy of users and protect themselves from the damaging effects of the DMCA, the ECPA and other data disclosure laws.