The Role of the Internet Engineering Task Force (IETF) in Improving Privacy on the Internet

Jon Peterson*, Hannes Tschofenig†, Bernard Aboba‡, Karen Solins§
*Neustar, Email: jon.peterson@neustar.biz
†Nokia Siemens Networks, Email: Hannes.Tschofenig@nsn.com
‡Microsoft, Email:bernarda@microsoft.com
§Massachusetts Institute of Technology, Email: sollins@csail.mit.edu

I. Introduction

Who we are (e.g. our thoughts, dreams, feelings, DNA sequence), what we own (such financial property), what we have experienced (audio/visual/olifactory transcripts), and how we can be reached (location, endpoint identifiers) are among the most private pieces of information about us. More and more of this information is being digitized and made available electronically.

As this information becomes more available, it gets exposed in unpredicable and surprising ways: health record breaches are commonplace today (see, for example, [1]), quasi-public information (e.g. birthdates, social security numbers) can be used to obtain financial data or even impersonate others, casual sharing of personal experiences is now a common activity (see, for example, [2]) and the Web increasingly extends into our private lives via webcams, microphones, etc., (as it was seen with the recent incident at a middle school [3]). Learning the geographical whereabouts of a person is possible - see problems with the W3C Location API [4] - and the same is true obtaining reachability information, such as Caller-ID ordinarily concealed by anonymous calls [5]. Unfortunately, we are certain to see more of these incidents due to the implementation of increasingly more powerful technical mechanisms, including the W3C Device APIs [6] as currently specified, which simplifies, and perhaps even encourages, intrusions by third parties who have no relationship with the end user.

What can be done about this widespread erosion of privacy? From the long experience of the IETF, the authors believe that an important initial step is to consider privacy while designing protocols and architectures, rather than as something to "bolt on" as an afterthought. The IETF has successfully applied this method with a a number of design criteria in its process, most notably security. As we illustrate later in this article, the IETF has been considering privacy in our protocol designs already for many years, although often implicitly and without thorough documentation. In order to optimize technical designs for privacy, in addition to thinking early about privacy design there are further challenges: first, the technical work needs to be backed-up by laws and appropriate disincentives to violate them; second, the best technology will not help end users if it does not get implemented properly and deployed in a privacy-friendly manner; third, few privacy experts are available in some communities; and finally, the organizational structures of some standards development organizations lack the checks and balances that might favor privacy controls for users over implementation expedience. While these other challenges are important, we focus the remainder of this paper on designing protocols from the ground up with privacy in mind.

Advocacy for privacy must not, however, devolve into dogmatism. In a USENIX Security 2003 presentation [7], Eric Rescorla investigated the regrettable status of security on the Internet, which is something of a curiosity given the great variety of tools (such as TLS, SSH, and IPsec) available. He argued that the reason that Internet communication remained insecure is largely due to the focus on a mistaken threat model:

- We worry about all known threats.
- Too good security is trumping deployment.
- Practical security is not glamorous.

We believe that the above statements are also applicable to privacy. In line with a quote by McGeorge Bundy - 'If you guard your toothbrushes and diamonds with equal zeal, you'll probably lose fewer toothbrushes and more diamonds' - we argue that the work on privacy has to be guided by the following principles:

- One size does not fit all.
- Protect what's important to users.
- Do not over-design in areas that yield no practical benefit.

We will explain these three guiding principles in light of completed as well as ongoing work in the IETF.

II. Completed IETF Work

The IETF is known for its seminal contributions to the design of the Internet and its applications. The "Request for Comment" (RFC) specifications that IETF participants produce divide into several categories, such as technical specification, best current practice descriptions, and architectural

Position paper for the W3C Workshop on Privacy for Advanced Web APIs; 12/13 July 2010, London. This paper represents the views of the authors in their role as individual contributors to the Internet standards process. They do not reflect the consensus output of any IETF working group or the IAB. Referenced documents may, however, reflect IETF consensus. Jon, Hannes, and Bernard are members of the Internet Architecture Board (IAB) and have contributed to the standardization efforts discussed in this document. Karen co-chairs the Privacy and Security working group of the MIT Communications Futures Program (see http://cfp.mit.edu/).

writeups. While these RFCs do not mandate a specific implementation they are often, if not always, guided by different architectural design decision; to illustrate the history of the IETF's commitment to work on privacy, we briefly discuss a selected list of completed specifications:

A. Presence

A presence service, as defined in the abstract in RFC 2778 [8], allows users of a communications service to monitor one another's availability and disposition in order to make decisions about communicating. Presence information is highly dynamic, and generally characterizes whether a user is online or offline, busy or idle, away from communications devices or nearby, and the like. Necessarily, this information has certain privacy implications, and from the start the IETF approached this work with the aim to provide users with the controls to determine how their presence information would be shared.

The Common Profile for Presence (CPP) [9] defines a set of logical operations for delivery of presence information. This abstract model is applicable to multiple presence systems. The SIP-based SIMPLE presence system [10] uses CPP as its baseline architecture, and the presence operations in the Extensible Messaging and Presence Protocol (XMPP) have also been mapped to CPP [11].

SIMPLE [10], the application of the Session Initiation Protocol (SIP) to instant messaging and presence, has native support for subscriptions and notifications (with its event framework [12]) and has added an event package [13] for presence in order to satisfy the requirements of CPP. Other event packages were defined later to allow additional information to be exchanged. With the help of the PUBLISH method [14] clients are able to install presence information on a server, so that the server can apply access-control policies before sharing presence information with other entities.

The integration of an explicit authorization mechanism into the presence architecture has been a major improvement in terms of involving the end users in the decision making process before sharing information. Nearly all presence systems deployed today provide such a mechanism, typically through a reciprocal authorization system by which a pair of users, when they agree to be "buddies," consent to divulge their presence information to one another.

B. Extending Presence: The Support for Location

Sharing location based information was always seen as a very attractive feature for application designers. With the desire to standardize protocols for systems sharing geolocation IETF work was started in the GEOPRIV working group [15]. During the initial requirements and privacy threat analysis in the process of chartering the working group, it became clear that the system would an underlying communication mechanism supporting user consent to share location information. The resemblance of these requirements to the presence framework was quickly recognized, and this design decision was documented in RFC 4079 [16].

While presence systems exerted influence on location privacy, the location privacy work also influenced ongoing IETF work on presence by triggering the standardization of a general access control policy language called the Common Policy (defined in RFC 4745 [17]) framework. This language allows one to express ways to control the distribution of information as simple conditions, actions, and transformations rules expressed in an XML format. Common Policy itself is an abstract format which needs to be instantiated: two examples can be found with the Presence Authorization Rules [18] and the Geolocation Policy [19]. The former provides additional expressiveness for presence based systems, while the latter defines syntax and semantic for location based conditions and transformations.

As a component of the prior work on the presence architecture, a format for presence information, called Presence Information Data Format (PIDF), had been developed. For the purposes of conveying location information an extension was developed, the PIDF Location Object (PIDF-LO). With the aim to meet the privacy requirements defined in RFC 2779 [20] a set of usage indications (such as whether retransmission is allowed or when the retention period expires) in the form of the following policies have been added that always travel with location information itself.

We believes that the standardization of these meta-rules that travel with location information has been a unique contribution to privacy on the Internet, recognizing the need for users to express their preferences when information travels through the Internet, from website to website. This approach very much follows the spirit of Creative Commons [21], namely the usage of a limited number of conditions (such as 'Share Alike' [22]). Unlike Creative Commons, the GEOPRIV working group did not, however, initiate work to produce legal language nor to design graphical icons since this would fall outside the scope of the IETF. In particular, the GEOPRIV rules state a preference on the retention and retransmission of location information; while GEOPRIV cannot force any entity receiving a PIDF-LO object to abide by those preferences, if users lack the ability to express them at all, we can guarantee their preferences will not be honored.

While these retention and retransmission meta-data elements could have been devised to accompany information elements in other IETF protocols, the decision was made to introduce these elements for geolocation initially because of the sensitivity of location information.

C. The Role of Intermediaries

Open Pluggable Edge Services (OPES) are services that would be deployed at application-level intermediaries in the network, for example, at a web proxy cache between the origin server and the client. These intermediaries would inspect and transform or filter content, with the explicit consent of either the content provider or the end user.

In the process of chartering work in the IETF around OPES various communities familiar with privacy laws voiced their concerns and led the Internet Architecture Board to publish a document [23] demanding basic privacy principles to be respected with the work in the OPES working group.

This effort showed that the transparent process, the open nature of the IETF to involve a diversity of stakeholders, and

the strong commitment of the IAB to address architectural challenges to help with the development of high-quality Internet standards.

D. Wiretapping

Many communication systems have been subject to laws requiring the exposure of details on the communication, including the content of the communication itself, to law enforcement agents. From a technology perspective, the telephony system is one such example where wiretapping has been built-in by design.

Together with the IAB, the IETF investigated the implications of wiretapping for the design of Internet protocols, and concluded (see RFC 2804 []RFC2804) as part of the RAVEN process that IETF standardization will not incorporate requirements for wiretapping a part of the process for creating and maintaining IETF standards. The requirement that a third-party be able to eavesdrop on communications without the awareness of end users simply fell outside the acceptable range of behavior for end-to-end security.

RFC 2804 has particularly shown the willingness of the IETF community to go against the mainstream thinking prevailing in traditional standards organizations in order to ensure better end-to-end security, as well as better privacy protections for end users. The IETF, however, is not the sole arbiter of the Internet, however, and other organizations share a similar responsibility for the stewardship of critical user privacy rights on the Internet.

III. Ongoing IETF Work

In addition to those examples of completed work, there are a number of ongoing activities with respect to privacy worth mentioning:

A. Open Web Authentication

Use of distributed "web services" to authorize access to resources by third parties is increasing. Such resources are protected and require those who desire access to first authenticate so that the services can evaluate their permissions in its authorization policies.

These web-services are compelling, but users cannot be required to share their long-term credentials with third-party applications, as the security and privacy implications would be devasting. Instead, a user must have some other means of granting explicit permission for a particular service to act as their agent. Furthermore, users have to be given the ability to restrict access of these agents to a limited subset of their resources, limit the duration of the access grant, and impose various other restrictions (e.g. only certain HTTP methods).

The IETF has started work on OAuth, see [24], with the publication of the community specification (OAuth 1.0, see RFC 5849 [25]) and related work OAuth 2.0 [26]. Quite naturally, the work on delegated authentication for access to resources on the Internet is impacted by privacy considerations, and thisus the group will ensure that IETF quality expectations are met.

B. Clarification of the Geolocation Privacy Architecture

While the work of the GEOPRIV working group was well-known and accepted within the IETF, the discussions surrounding the W3C geolocation API made it clear that standards development organizations outside the IETF did not share the same understanding of the importance of considering privacy early in their design process.

The GEOPRIV working group had decided to clarify the architecture to make it more accessible to those outside the IETF, and also provides a more generic description applicable beyond the context of presence. [27] shows the work-in-progress writeup.

The clarification of the architectural document in in keeping with the group's strong and longstanding emphasis on security and privacy matters.

C. Session Recording Protocol (SIPREC)

The Session Recording Protocol (SIPREC) working group [28] is chartered to define a SIP-based protocol for controlling a session (media) recorder.

Session recording is a critical requirement in many business communications environments such as call centers and financial trading floors. In some of these environments, all calls must be recorded for regulatory and compliance reasons. In others, calls may be recorded for quality control, business analytics, or consumer protection. Recording is typically done by sending a copy of the media to the recording devices.

Media recording is a sensitive issue when it comes to privacy and for this reason the charter spells-out the need to address privacy in the output of deliverables of the group, highlights the requirement for end-user notification of any recording action, and the need to conform to RFC 2804 [29]. This working group very well illustrates the challenges of balance conflicting requirements of user privacy with regulation, and demonstrates how, simply through ensuring consent, a compromise can be found.

D. Application-Layer Traffic Optimization (ALTO)

In May 2008 the Real-time Applications and Infrastructure Area Directors of the IETF organized a workshop to discuss network delay and congestion issues resulting from increased Peer-to-Peer (P2P) traffic volumes. The report of the workshop was published as RFC 5594 [30]. The outcome of the workshop lead to a number of IETF and IRTF activities, including the establishment of the ALTO working group [31] and the very recent formation of the Congestion Exposure (conex) working group [32].

In contrast to client/server architectures, P2P applications often must choose one or more suitable candidates from a selection of peers offering the same resource or service. The goal of the working group effort is to provide applications with information to perform better-than-random initial peer selection.

Although this working group at first sight might have very little relationship with privacy the members of the working group frequently bring up privacy aspects in their contributions [33], [34], [35]. While the used terminology might not

always be correct but the basic understanding for the topic is certainly pointing in the direct direction.

E. IAB Privacy Program

The IAB has a number of tasks and among them is the architectural oversight function. As part of this task, IAB members investigates challenges the Internet faces and to determine what the IETF, other standards organizations, and the broader Internet community can do to help making the Internet work better.

Some of the activites span multiple IAB generations. The name for such a longer term effort is called 'program' and the current IAB members have decided to create a privacy program to improve the quality of the IETF specifications with regard to privacy, to reach out to other organizations, and to engage with regulators. The work also includes the development of guidelines with respect to the handling privacy sensitive information in IETF protocols, and to improve the general awareness of privacy issues in protocol development.

The IAB program descriptions can be found at [36].

IV. Conclusion

The IETF together with the IAB have worked on privacy and privacy-related topics over a long timeframe. As is the case with security, privacy-sensitive work is of higher quality if privacy is considered early in its design process. The IETF achieved these results by involving a broader community, including privacy experts with technical as well as legal background, from early on. The open nature of the IETF makes participation for a wide range of experts easy, and keeps the barrier for participation low. The organizational structure and the detailed review process of documents creates an ideal environment for developing high-quality work in a way that organizations focused on implementation expedience might find difficult to replicate.

The IETF has not yet developed a global policy on privacy, but it has arrived at more or less the same position in several work areas. Part of the ingoing IAB privacy program is to sort out if a generalized IETF position can developed into a set of principles and guidelines to be followed by future IETF specifications and, hopefully, to inspire other bodies that deal regularly with privacy-sensitive information.

While the IETF has made important strides forward, the challenge to make the Internet work better is ongoing. The recently published analysis [4] of the W3C geolocation API by focusing only on a design-by-policy [37] argues (we believe persuasively) that the W3C current approach does not work in practice. We maintain that the W3C, the IETF and the Internet community of privacy experts must work together to provide an online experience that conforms with user expectations of privacy and the emerging regulatory environment.

V. Acknowledgements

We would like to thank Alissa Cooper and Martin Euchner for an early review of this document.

References

[1] M. Neil, "Celebrity Medical Files Breached at UCLA," Apr. 2009, http://tinyurl.com/2v24hhb.
[2] "Privacy Concerns Hit Facebook, Google," May 2010, http://www1.voanews.com/learningenglish/home/business/Privacy-Concerns-Facebook-Google-95063839.html.
[3] C. Doctorow, "School used student laptop webcams to spy on them at school and home," Feb. 2010, http://www.boingboing.net/2010/02/17/school-used-student.html.
[4] N. Doty, D. Mulligan, and E. Wilde, "Privacy Issues of the W3C Geolocation API," Feb. 2010, http://escholarship.org/uc/item/0rp834wf.
[5] K. Poulsen, "Anonymous Caller? New Service Says, Not Any More," Feb. 2009, http://www.wired.com/threatlevel/2009/02/trapcall/.
[6] W3C, "Device APIs and Policy Working Group Charter," Jun. 2010, http://www.w3.org/2009/05/DeviceAPICharter.
[7] E. Rescorla, "The Internet is Too Secure Already," 2003, http://www.rtfm.com/TooSecure-usenix.pdf.
[8] M. Day, J. Rosenberg, and H. Sugano, "A Model for Presence and Instant Messaging," Feb. 2000, RFC 2778, Request For Comments.
[9] J. Peterson, "Common Profile for Presence (CPP)," Aug. 2004, RFC 3859, Request For Comments.
[10] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP: Session Initiation Protocol," May 2002, RFC 3261, Request For Comments.
[11] P. Saint-Andre, "Mapping the Extensible Messaging and Presence Protocol (XMPP) to Common Presence and Instant Messaging (CPIM)," Oct. 2004, RFC 3922, Request For Comments.
[12] A. Roach, "Session Initiation Protocol(SIP)-Specific Event Notification," Jun. 2002, RFC 3265, Request For Comments.
[13] J. Rosenberg, "A Presence Event Package for the Session Initiation Protocol (SIP)," Aug. 2004, RFC 3856, Request For Comments.
[14] A. Niemi, "Session Initiation Protocol (SIP) Extension for Event State Publication," Oct. 2004, RFC 3903, Request For Comments.
[15] "Geographic Location/Privacy (geopriv) - Charter," Jun. 2010, http://datatracker.ietf.org/wg/geopriv/charter/.
[16] J. Peterson, "A Presence Architecture for the Distribution of GEOPRIV Location Objects," Jul. 2005, IETF 4079, Request For Comments.
[17] H. Schulzrinne, H. Tschofenig, J. Morris, J. Cuellar, J. Polk, and J. Rosenberg, "Common Policy: A Document Format for Expressing Privacy Preferences," Feb. 2007, IETF 4745, Request For Comments.
[18] J. Rosenberg, "Presence Authorization Rules," Dec. 2007, RFC 5025, Request For Comments.
[19] H. Schulzrinne, H. Tschofenig, J. Morris, J. Cuellar, and J. Polk, "An Architecture for Location and Location Privacy in Internet Applications," Jul. 2009, http://tools.ietf.org/html/draft-ietf-geopriv-policy-21.
[20] M. Day, S. Aggarwal, and J. Vincent, "Instant Messaging / Presence Protocol Requirements," Feb. 2000, RFC 2779, Request For Comments.
[21] "Creative Commons," Jun. 2010, http://creativecommons.org/about/licenses.
[22] "Creative Commons - Licenses," Jun. 2010, http://creativecommons.org/about/licenses.
[23] S. Floyd and L. Daigle, "IAB Architectural and Policy Considerations for Open Pluggable Edge Services," Jan. 2002, IETF 3238, Request For Comments.
[24] "Open Authentication Protocol (oauth) - Charter," Jun. 2010, http://datatracker.ietf.org/wg/oauth/charter/.
[25] E. Hammer-Lahav, "The OAuth 1.0 Protocol," Apr. 2010, IETF 5849, Request For Comments.
[26] E. Hammer-Lahav, D. Recordon, and D. Hardt, "The OAuth 2.0 Protocol," Jun. 2010, IETF draft (work in progress), http://datatracker.ietf.org/doc/draft-ietf-oauth-v2/.
[27] R. Barnes, M. Lepinski, A. Cooper, J. Morris, H. Tschofenig, and H. Schulzrinne, "An Architecture for Location and Location Privacy in Internet Applications," May 2010, IETF draft (work in progress), http://datatracker.ietf.org/doc/draft-ietf-geopriv-arch/.
[28] "SIP Recording (siprec) - Charter," Jun. 2010, https://datatracker.ietf.org/wg/siprec/charter/.
[29] IAB and IESG, "IETF Policy on Wiretapping," May 2000, IETF 2804, Request For Comments.
[30] J. Peterson and A. Cooper, "Report from the IETF Workshop on Peer-to-Peer (P2P) Infrastructure, May 28, 2008," Jul. 2009, IETF 5594, Request For Comments.
[31] "Application-Layer Traffic Optimization (alto) - Charter," Jun. 2010, https://datatracker.ietf.org/wg/alto/charter/.
[32] "Congestion Exposure (conex) - Charter," Jun. 2010, http://datatracker.ietf.org/wg/conex/charter/.

[33] Y. Wang, H. Song, and M. Chen, "Analysis for ALTO privacy and load issues," Oct. 2009, IETF draft (work in progress),https://datatracker.ietf.org/doc/draft-wang-alto-privacy-load-analysis.
[34] S. Kiesel, S. Previdi, M. Stiemerling, , R. Woundy, and Y. Yang, "Application-Layer Traffic Optimization (ALTO) Requirements," Jun. 2010, IETF draft (work in progress), https://datatracker.ietf.org/doc/draft-ietf-alto-reqs/.
[35] J. Seedorf and E. Burger, "Application-Layer Traffic Optimization (ALTO) Problem Statement," Oct. 2009, IETF 5693, Request For Comments.
[36] "Internet Architecture Board," Jun. 2010, http://www.iab.org/.
[37] S. Spiekermann and L. Cranor, "Engineering Privacy," Sep. 2008, http://ssrn.com/abstract=1085333.