Human Rights Protocol Considerations Research Group
Internet-Draft
Intended status: Informational
Expires: May 16, 2018

S. Bortzmeyer
AFNIC
N. ten Oever
ARTICLE 19

November 12, 2017

Anonymity, Human Rights and Internet Protocols
draft-tenoever-hrpc-anonymity-01

Abstract

Anonymity is less discussed in the IETF than for instance security [RFC3552] or privacy [RFC6973]. This can be attributed to the fact anonymity is a hard technical problem or that anonymizing user data is not of specific market interest. It remains a fact that 'most internet users would like to be anonymous online at least occasionally' [Pew].

This document aims to break down the different meanings and implications of anonymity on a mediated computer network.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 16, 2018.

Table of Contents

1. Introduction

There seems to be a clear need for anonymity online in an environment where harassment on the Internet is on the increase [Pew2] and the UN Special Rapporteur for Freedom of Expression calls anonymity 'necessary for the exercise of the right to freedom of opinion and expression in the digital age' [UNHRC2015].

Nonetheless anonymity is not getting much discussion at the IETF, providing anonymity does not seem a (semi-)objective for many protocols, even though several documents contribute to improving anonymity such as [RFC7258], [RFC7626], [RFC7858].

There are initiatives on the Internet to improve end users anonymity, most notably [torproject], but these initiatives rely on adding encryption in the application layer.

This document aims to break down the different meanings and implications of anonymity on a mediated computer network and to see whether (some parts of) anonymity should be taken into consideration in protocol development.

2. Vocabulary Used

Concepts in this draft currently strongly hinges on [AnonTerm]

Anonymity A state of an individual in which an observer or attacker cannot identify the individual within a set of other individuals (the anonymity set). [RFC6973]

Linkability Linkability of two or more items of interest (IOIs - Items Of Interest, e.g., subjects, messages, actions, ...) from an attacker's perspective means that within the system (comprising these and possibly other items), the attacker can sufficiently distinguish whether these IOIs are related or not. [AnonTerm]

Pseudonymity Derived from pseudonym, a persistent identity which is not the same as the entity's given (or official) name. For most (TODO all?) IETF protocols, pseudonimity is a given: protocols don't care whether the identity is an official one or not. But it should be noted that, if the user cannot create new pseudonyms easily, pseudonyms suffer from linkability. Unlikability depends on this ability to create new pseudonyms. TODO: or decide that pseudonyms _require_ this ability to be created at will?

Unlinkability Unlinkability of two or more items of interest (IOIs, e.g., subjects, messages, actions, ...) from an attacker's perspective means that within the system (comprising these and possibly other items), the attacker cannot sufficiently distinguish whether these IOIs are related or not. [AnonTerm]

Undetectability The impossibility of being noticed or discovered

Undetectability of an item of interest (IOI) from an attacker's perspective means that the attacker cannot sufficiently distinguish whether it exists or not [AnonTerm]

Unobservability

Unobservability of an item of interest (IOI) means: undetectability of the IOI against all subjects uninvolved in it and

anonymity of the subject(s) involved in the IOI even against the other subject(s) involved in that IOI. [AnonTerm]

It should be noted that the word "anonymity" is both very loaded politically (witness all the headlines about the "darknet") and poorly understood. Most texts talking about anonymity actually refer to pseudonymity (for instance, when people say that "Bitcoin is

anonymous"). This confusion is even in the example given in [RFC4949] definition of anonymity.

Anonymity is strongly linked to unlinkability: if your actions are linkable, it suffices that one of them is tied to your identity, and anonymity is over.

It should be noted that anonymity is not binary: there have been these recent years a lot of progress of desanonymisation techniques. Data is never fully "anonymous", it is only more or less anonymous. [RFC6235] [MITdeano] [Utexas] [Article29]

3. Should protocols promote anonymity?

The amount of data that is generated by and about individuals is growing exponentially. This can be attributed to the fact that an ever increasing number of actions is digitally mediated, and the increase of connected sensors in the every day environment. Even though these two causes do not fully fall within the scope of the IETF, there is a significant part of these two examples that do.

With the increase of data there is also an increasing ability for third parties to analyze human behaviour. It should be noted that any data that could identify an individual is personally identifiable information (PII). This means that information which can be used to distinguish an indivual from other individuals can be considered as personally identiable information. The access and control of personally identifiable information by a third party is a (potential) liability for both the third party and the individual. This liability could for example translate into a physical risk for the individual or into a legal risk for the third party under information security and privacy laws.

Some network operators argue that without the opportunity to persistently identify individual users it becomes harder to thwart attacks and troubleshoot network issues. Whereas identification might be helpful to address issues in some cases, it poses an inherent threat to the anonymity of users. Not protecting the anonymity of users leads to a deterioration of the right to privacy, and the right to freedom of opinion and expression. There can be limitations the right to privacy and freedom of expression, but these should always be provided by law and necessary and proportionate to achieve one of a handful of legitimate objectives.

Anonymity will always be a balancing act between user protection (which requires a high level of anonymity) and other requirments for operations and research, such as routing information. Anonymity is by no means achieved by default in an online environment, nor has it

been a strong consideration in protocol development in the development of the Internet. Increasing anonymity in the digital environment is not an easy task, exactly because the ubiquity of data that is generated and stored. But exactly the fact that we generate so much data urges us to address this issue.

4. Example of use cases

4.1. Simultaneous use

One user may use concurrently several identities, mixing them in operations, while wanting to keep them distinct. The protocol and its implementations should not preclude this use.

4.2. Successive use

One user may switch from one identity to another. In that case, it must be doable without a "bleedover" from the old identity to the new one.

5. Practical advices

5.1. Protocol developers

First, the protocol should avoid to have mandatory persistent identifiers.

Even without persistent identifiers, anonymity could be broken by examining the patterns of access. If an user visits each morning the three same Web sites, always in the same order, it will be easy to identify them even without persistent identifier. Protocol designers should therefore ask themselves if patterns are easily visible, or obfuscated in some way.

If the protocol collects data and distributes it (see [RFC6235]), "anonymizing" the data is often suggested but it is notoriously hard. Do not think that just dropping the last byte of an IP address "anonymizes" data.

Pay attention to the fact that Internet actors do not all see the same thing. Consider the anonymity of the user with respect to:

- local network operator
- other networks you connect to
- your communications peer on the other end of the pipe

- intermediaries ([RFC6973])

- enablers ([RFC6973])

- someone who is in several roles, for instance a big state surveillance agency

5.2. Protocol implementors

Avoid adding options or configurations that create or might lead to patterns or regularities that are not explicitely required by the protocol.

An example is DHCP where sending a persistent identifier as the client name was not mandatory but, in practice, done by many implementations, before [RFC7844].

If an implementation allows for identity management, there should be a clear barrier between the identities to ensure that they cannot (easily) be associated with each other.

If there are anonymization option for the protocol, these should be enabled by default.

6. Open Questions

While analyzing protocols for their impact on users anonymity, would it make sense to ask the following questions:

1. How does the protocol impact pseudonymity? If the protocol limits the creation of new pseudonyms, it can limit their usefulness to "hide" an user's identity. For instance, IP addresses are pseudonyms but, since they are not under end users's control, they have strong linkability. That's why they are rightly regarded as personal identifiers [EUcourt]. On the other hand, Bitcoin addresses are pseudonyms with limited linkability, since the user can always create a lot of them.

2. Could there be more advice for protocol developers and implementers to improve anonymity? (Besides the ones in Section 5.)

7. Security Considerations

As this draft concerns a research document, there are no security considerations.

8. IANA Considerations

This document has no actions for IANA.

9. Research Group Information

The discussion list for the IRTF Human Rights Protocol Considerations proposed working group is located at the e-mail address hrpc@ietf.org [1]. Information on the group and information on how to subscribe to the list is at https://www.irtf.org/mailman/listinfo/hrpc

Archives of the list can be found at: https://www.irtf.org/mail-archive/web/hrpc/current/index.html

10. References

10.1. Informative References

[AnonTerm] Pfitzmann, A. and M. Hansen, "A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management", 2010, <http://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.34.pdf>.

[Article29] Article29, ., "Opinion 05/2014 on Anonymisation Techniques", 2014, <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf>.

[EUcourt] "EUCJ Case C-70/10: Scarlet Extended SA vs. Societe belge des auteurs, compositeurs et editeurs SCRL (SABAM)", 2011, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:62010CJ0070:EN:HTML&lipi=urn%3Ali%3Apage%3Ad_flagship3_pulse_read%3BSFHas%2FXMRHeHVu46775ezw%3D%3D>.

[MITdeano] de Montjoye, Y., Hidalgo, C., Verleysen, M., and V. Blondel, "Unique in the Crowd: The privacy bounds of human mobility", 2013, <https://www.nature.com/articles/srep01376>.

[Pew] Rainie, L., Kiesler, S., Kang, R., and M. Madden, "Anonymity, Privacy, and Security Online", 2013, <http://www.pewinternet.org/2013/09/05/anonymity-privacy-and-security-online/>.

[Pew2] Duggan, M., "Online Harassment", 2014, <http://www.pewinternet.org/2014/10/22/online-harassment/>.

[RFC3552] Rescorla, E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP 72, RFC 3552, DOI 10.17487/RFC3552, July 2003, <https://www.rfc-editor.org/info/rfc3552>.

[RFC4949] Shirey, R., "Internet Security Glossary, Version 2", FYI 36, RFC 4949, DOI 10.17487/RFC4949, August 2007, <https://www.rfc-editor.org/info/rfc4949>.

[RFC6235] Boschi, E. and B. Trammell, "IP Flow Anonymization Support", RFC 6235, DOI 10.17487/RFC6235, May 2011, <https://www.rfc-editor.org/info/rfc6235>.

[RFC6973] Cooper, A., Tschofenig, H., Aboba, B., Peterson, J., Morris, J., Hansen, M., and R. Smith, "Privacy Considerations for Internet Protocols", RFC 6973, DOI 10.17487/RFC6973, July 2013, <https://www.rfc-editor.org/info/rfc6973>.

[RFC7258] Farrell, S. and H. Tschofenig, "Pervasive Monitoring Is an Attack", BCP 188, RFC 7258, DOI 10.17487/RFC7258, May 2014, <https://www.rfc-editor.org/info/rfc7258>.

[RFC7626] Bortzmeyer, S., "DNS Privacy Considerations", RFC 7626, DOI 10.17487/RFC7626, August 2015, <https://www.rfc-editor.org/info/rfc7626>.

[RFC7844] Huitema, C., Mrugalski, T., and S. Krishnan, "Anonymity Profiles for DHCP Clients", RFC 7844, DOI 10.17487/RFC7844, May 2016, <https://www.rfc-editor.org/info/rfc7844>.

[RFC7858] Hu, Z., Zhu, L., Heidemann, J., Mankin, A., Wessels, D., and P. Hoffman, "Specification for DNS over Transport Layer Security (TLS)", RFC 7858, DOI 10.17487/RFC7858, May 2016, <https://www.rfc-editor.org/info/rfc7858>.

[torproject] The Tor Project, ., "Tor Project - Anonymity Online", 2007, <https://www.torproject.org/>.

[UNHRC2015]
Kaye, D., "Anonymity, Privacy, and Security Online (A/HRC/29/32)", 2015, <http://www.ohchr.org/EN/HRBodies/HRC/RegularSessions/Session29/Documents/A.HRC.29.32_AEV.doc>.

[Utexas] Narayanan, A. and V. Shmatikov, "Robust De-anonymization of Large Sparse Datasets", 2008, <http://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pd>.

10.2. URIs

[1] mailto:hrpc@ietf.org

Authors' Addresses

Stephane Bortzmeyer
AFNIC

EMail: bortzmeyer+ietf@nic.fr

Niels ten Oever
ARTICLE 19

EMail: niels@article19.org