

Workgroup: Network Working Group  
Internet-Draft:  
draft-irtf-pearg-safe-internet-measurement-08  
Published: 10 July 2023  
Intended Status: Informational  
Expires: 11 January 2024  
Authors: I. R. Learmonth    G. Grover  
          HamBSD                Centre for Internet and Society  
          M. Knodel  
          Center for Democracy and Technology

## **Guidelines for Performing Safe Measurement on the Internet**

### **Abstract**

Internet measurement is important to researchers from industry, academia and civil society. While measurement of the internet can give insight into the functioning and usage of the Internet, it can present risks to user privacy. This document describes briefly those risks and proposes guidelines for ensuring that internet measurements can be carried out safely, with examples.

### **Note**

This document is a draft. It is not an IETF product. It does not propose a standard. Comments are solicited and should be addressed to the research group's mailing list at [pearg@irtf.org](mailto:pearg@irtf.org) and/or the author(s).

The sources for this draft are at:

<https://github.com/irl/draft-safe-internet-measurement>

### **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 <https://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 11 January 2024.

## Copyright Notice

Copyright (c) 2023 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

## Table of Contents

- [1. Introduction](#)
  - [1.1. Scope of this document](#)
  - [1.2. Terminology](#)
  - [1.3. User Impact from Measurement Studies](#)
- [2. Safe Internet Measurement Guidelines](#)
  - [2.1. Obtain consent](#)
    - [2.1.1. Informed Consent](#)
    - [2.1.2. Proxy Consent](#)
    - [2.1.3. Implied Consent](#)
  - [2.2. Isolate risk with a dedicated testbed](#)
  - [2.3. Be respectful of others' infrastructure](#)
  - [2.4. Maintain a "Do Not Scan" list](#)
  - [2.5. Minimize data](#)
    - [2.5.1. Discard it](#)
    - [2.5.2. Mask it](#)
    - [2.5.3. Aggregate it](#)
  - [2.6. Reduce accuracy](#)
  - [2.7. Analyze Risk](#)
- [3. Security Considerations](#)
- [4. IANA Considerations](#)
- [5. Acknowledgements](#)
- [6. Informative References](#)
- [Authors' Addresses](#)

### 1. Introduction

Measurement of the internet provides important insights and is a growing area of research. Similarly the internet plays a role in enhancing research methods of different kinds.

Performing research using the Internet, as opposed to an isolated testbed or simulation platform, means that experiments co-exist in a space with other services and end users. This document outlines guidelines for academic, industry and civil society researchers who

might use the Internet as part of scientific experimentation to mitigate risks to the safety of other users.

### **1.1. Scope of this document**

These are guidelines for how to measure the Internet safely. When performing research on a platform shared with live traffic from other users, that research is considered safe if and only if other users are protected from or unlikely to experience danger, risk, or injury, now or in the future, due to the research.

Following the guidelines contained within this document is not a substitute for institutional ethics review processes, although these guidelines could help to inform that process. It is particularly important for the growing area of research that includes Internet measurement to better equip review boards to evaluate Internet measurement methods[[SIGCOMM](#)], and we hope that this document is part of that larger effort.

Similarly, these guidelines are not legal advice and local laws must also be considered before starting any experiment that could have adverse impacts on user safety.

The scope of this document is restricted to guidelines that mitigate exposure to risks to Internet user safety when measuring properties of the Internet: the network, its constituent hosts and links, or its users traffic.

For the purpose of this document, an Internet user is an individual or organisation whose data is used in communications over the Internet, most broadly, and those who use the Internet to communicate or maintain Internet infrastructure.

### **1.2. Terminology**

**Threat model:** A threat is a potential for a security violation, which exists when there is a circumstance, capability, action, or event that could breach security and cause harm [[RFC4949](#)]. Every Internet measurement study has the potential to subject Internet users to threat actions or attacks.

**Measurement studies:** Internet measurement studies can be broadly categorized into two groups: active measurements and passive measurements. The type of measurement is not truly binary and many studies will include both active and passive components.

**Active measurement:** Active measurements generate or modify traffic.

**Passive measurement:** Passive measurements use surveillance of existing traffic.

On/off-path: A measurement, or attack, that is on-path happens on the network. Off-path indicates activity in a side-channel, end-point or at other points where the user, their connection, or their data can be accessed.

One-/two-ended:

### **1.3. User Impact from Measurement Studies**

Any conceivable Internet measurement study might be considered an attack on an Internet user's safety. The measurement of generated traffic may also lead to insights into other users' traffic indirectly as well.

It is always necessary to consider the best approach to mitigate the impact of measurements, and to balance the risks of measurements against the benefits to impacted users, which we get to in the following section. But first, we describe the consequences of attacks that are made possible by measurement.

Surveillance: An attack whereby an Internet user's information is collected. This type of attack covers not only data but also metadata.

Inadequate data protection: An attack where data, either in transit or at rest, is not adequately protected from disclosure. Failure to adequately protect data to the expectations of the user is an attack even if it does not lead to another party gaining access to the data.

Traffic generation: An attack whereby undue traffic is generated to traverse the Internet.

Traffic modification: An attack whereby on-path Internet traffic is nonconsensually modified.

Breach of privacy: User privacy can be violated in the context of data collection. This impact also covers the case of an Internet user's data being shared beyond that which a user had given consent for.

Impersonation: An attack where a user is impersonated during a measurement.

Legal: Users and service providers are bound by a wide range of policies from terms of service to rule of law, each according to context and jurisdiction.

Other Retribution: A user may be actioned by a service or provider during a measurement or as a consequence of measurement behaviour over the network.

System corruption: An attack where generated or modified traffic causes the corruption of a system. This attack covers cases where a user's data may be lost or corrupted, and cases where a user's access to a system may be affected.

Data loss, corruption: An attack where the integrity of user data is violated during or as a result of a measurement.

Denial of Service (by which self-censorship is covered): An attack introduced during measurement that overwhelms the user client or service with excessive traffic, thus resulting in a denial of service.

Emotional trauma: An attack by which either a measurement of or exposure to content or behaviour in an internet measurement causes a user harm.

## **2. Safe Internet Measurement Guidelines**

### **2.1. Obtain consent**

Accountability and transparency are fundamentally related to consent. As per the Menlo Report, "Accountability demands that research methodology, ethical evaluations, data collected, and results generated should be documented and made available responsibly in accordance with balancing risks and benefits."[\[MenloReport\]](#) A user is best placed to balanced the risks and benefits for themselves therefore consent must be obtained. From most transparent to least, there are a few options for obtaining consent.

#### **2.1.1. Informed Consent**

Informed consent should be collected from all users that may be placed at risk, no matter how small a risk, by an experiment, however it may be impractical to do so at scale. In cases where it is practical to do so, this should be done.

However for consent to be informed, all possible risks must be presented to the users. The considerations in this document can be used to provide a starting point although other risks may be present depending on the nature of the measurements to be performed.

Example: A researcher would like to use volunteer owned mobile devices to collect information about local Internet censorship. Connections will be attempted by the volunteer's device with

services and content known or suspected to be subject to censorship orders.

This experiment can carry substantial risk for the user depending on the circumstances, from disciplinary action from their employer to arrest or imprisonment. Fully informed consent ensures that any risk that is being taken has been carefully considered by the volunteer before proceeding.

### **2.1.2. Proxy Consent**

In cases where it is not practical to collect informed consent from all users of a shared network, it may be possible to obtain proxy consent. Proxy consent may be given by a network operator or employer that would be more familiar with the expectations of users of a network than the researcher.

In some cases, a network operator or employer may have terms of service that specifically allow for giving consent to third parties to perform certain experiments.

Example: A researcher would like to perform a packet capture to determine the TCP options and their values used by all client devices on an corporate wireless network.

The employer may already have terms of service laid out that allow them to provide proxy consent for this experiment on behalf of the employees, in this case the users of the network. The purpose of the experiment may affect whether or not they are able to provide this consent. Say, performing engineering work on the network may be allowed, whereas academic research may not be already covered.

### **2.1.3. Implied Consent**

In larger scale measurements, even proxy consent collection may not be practical. In this case, implied consent may be presumed from users for some measurements. Consider that users of a network will have certain expectations of privacy and those expectations may not align with the privacy guarantees offered by the technologies they are using. As a thought experiment, consider how users might respond if asked for their informed consent for the measurements you'd like to perform.

Implied consent should not be considered sufficient for any experiment that may collect sensitive or personally identifying information. If practical, attempt to obtain informed consent or proxy consent from a sample of users to better understand the expectations of other users.

Example: A researcher would like to run a measurement campaign to determine the maximum supported TLS version on popular web servers.

The operator of a web server that is exposed to the Internet hosting a popular website would have the expectation that it may be included in surveys that look at supported protocols or extensions but would not expect that attempts be made to degrade the service with large numbers of simultaneous connections.

Example: A researcher would like to perform A/B testing for protocol feature and how it affects web performance. They have created two versions of their software and have instrumented both to report telemetry back. These updates will be pushed to users at random by the software's auto-update framework. The telemetry consists only of performance metrics and does not contain any personally identifying or sensitive information.

As users expect to receive automatic updates, the effect of changing the behaviour of the software is already expected by the user. If users have already been informed that data will be reported back to the developers of the software, then again the addition of new metrics would be expected. There are risks in pushing any new software update, and the A/B testing technique can reduce the number of users that may be adversely affected by a bad update.

The reduced impact should not be used as an excuse for pushing higher risk updates, only updates that could be considered appropriate to push to all users should be A/B tested. Likewise, not pushing the new behaviour to any user should be considered appropriate if some users are to remain with the old behavior.

In the event that something does go wrong with the update, it should be easy for a user to discover that they have been part of an experiment and roll back the change, allowing for explicit refusal of consent to override the presumed implied consent.

## **2.2. Isolate risk with a dedicated testbed**

Wherever possible, use a testbed. An isolated network means that there are no other users sharing the infrastructure you are using for your experiments.

When measuring performance, competing traffic can have negative effects on the performance of your test traffic and so the testbed approach can also produce more accurate and repeatable results than experiments using the public Internet.

WAN link conditions can be emulated through artificial delays and/or packet loss using a tool like [netem](#). Competing traffic can also be emulated using traffic generators.

### **2.3. Be respectful of others' infrastructure**

If your experiment is designed to trigger a response from infrastructure that is not your own, consider what the negative consequences of that may be. At the very least your experiment will consume bandwidth that may have to be paid for.

In more extreme circumstances, you could cause traffic to be generated that causes legal trouble for the owner of that infrastructure. The Internet is a global network crossing many legal jurisdictions and so what may be legal for you is not necessarily legal for everyone.

If you are sending a lot of traffic quickly, or otherwise generally deviate from typical client behaviour, a network may identify this as an attack which means that you will not be collecting results that are representative of what a typical client would see.

### **2.4. Maintain a "Do Not Scan" list**

When performing active measurements on a shared network, maintain a list of hosts that you will never scan regardless of whether they appear in your target lists. When developing tools for performing active measurement, or traffic generation for use in a larger measurement system, ensure that the tool will support the use of a "Do Not Scan" list.

If complaints are made that request you do not generate traffic towards a host or network, you must add that host or network to your "Do Not Scan" list, even if no explanation is given or the request is automated.

You may ask the requester for their reasoning if it would be useful to your experiment. This can also be an opportunity to explain your research and offer to share any results that may be of interest. If you plan to share the reasoning when publishing your measurement results, e.g. in an academic paper, you must seek consent for this from the requester.

Be aware that in publishing your measurement results, it may be possible to infer your "Do Not Scan" list from those results. For example, if you measured a well-known list of popular websites then it would be possible to correlate the results with that list to determine which are missing.

### **2.5. Minimize data**

When collecting, using, disclosing, and storing data from a measurement, use only the minimal data necessary to perform a task.

Reducing the amount of data reduces the amount of data that can be misused or leaked.

When deciding on the data to collect, assume that any data collected might be disclosed. There are many ways that this could happen, through operation security mistakes or compulsion by a judicial system.

When directly instrumenting a protocol to provide metrics to a passive observer, see section 6.1 of RFC6973 [[RFC6973](#)] for the data minimalization considerations enumerated below that are specific to the use case.

### **2.5.1. Discard it**

Discard data that is not required to perform the task.

When performing active measurements be sure to only capture traffic that you have generated. Traffic may be identified by IP ranges or by some token that is unlikely to be used by other users.

Again, this can help to improve the accuracy and repeatability of your experiment. [[RFC2544](#)], for performance benchmarking, requires that any frames received that were not part of the test traffic are discarded and not counted in the results.

### **2.5.2. Mask it**

Mask data that is not required to perform the task. Particularly useful for content of traffic to indicate that either a particular class of content existed or did not exist, or the length of the content, but not recording the content itself. Can also replace content with tokens, or encrypt.

### **2.5.3. Aggregate it**

When collecting data, consider if the granularity can be limited by using bins or adding noise. XXX: Differential privacy.

Do this at the source, definitely do it before you write to disk.

[[Tor.2017-04-001](#)] presents a case-study on the in-memory statistics in the software used by the Tor network, as an example.

## **2.6. Reduce accuracy**

Binning, categorizing, geoip, noise.

## 2.7. Analyze Risk

The benefits should outweigh the risks. Consider auxiliary data (e.g. third-party data sets) when assessing the risks.

## 3. Security Considerations

Take reasonable security precautions, e.g. about who has access to your data sets or experimental systems.

## 4. IANA Considerations

This document has no actions for IANA.

## 5. Acknowledgements

Many of these considerations are based on those from the [TorSafetyBoard] adapted and generalised to be applied to Internet research.

Other considerations are taken from the Menlo Report [MenloReport] and its companion document [MenloReportCompanion].

## 6. Informative References

[netem] Stephen, H., "Network emulation with NetEm", April 2005.

[RFC2544] Bradner, S. and J. McQuaid, "Benchmarking Methodology for Network Interconnect Devices", RFC 2544, DOI 10.17487/RFC2544, March 1999, <<https://www.rfc-editor.org/info/rfc2544>>.

[TorSafetyBoard] Tor Project, "Tor Research Safety Board", <<https://research.torproject.org/safetyboard/>>.

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

[Tor.2017-04-001] Herm, K., "Privacy analysis of Tor's in-memory statistics", Tor Tech Report 2017-04-001, April 2017, <<https://research.torproject.org/techreports/privacy-in-memory-2017-04-28.pdf>>.

[MenloReport] Dittrich, D. and E. Kenneally, "The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research", August 2012, <<https://>

[www.caida.org/publications/papers/2012/menlo\\_report\\_actual\\_formatted/](http://www.caida.org/publications/papers/2012/menlo_report_actual_formatted/)>.

**[MenloReportCompanion]** Bailey, M., Dittrich, D., and E. Kenneally, "Applying Ethical Principles to Information and Communication Technology Research", October 2013, <[https://www.impactcybertrust.org/link\\_docs/Menlo-Report-Companion.pdf](https://www.impactcybertrust.org/link_docs/Menlo-Report-Companion.pdf)>.

**[RFC6973]** Cooper, A., Tschofenig, H., Aboba, B., Peterson, J., Morris, J., Hansen, M., and R. Smith, "Privacy Considerations for Internet Protocols", RFC 6973, July 2013, <<https://www.rfc-editor.org/info/rfc6937>>.

**[SIGCOMM]** Jones, B., Ensafi, R., Feamster, N., Paxson, V., and N. Weaver, "Ethical Concerns for Censorship Measurement", August 2015, <<http://conferences.sigcomm.org/sigcomm/2015/pdf/papers/nsethics/p17.pdf>>.

#### Authors' Addresses

Iain R. Learmonth  
HamBSD

Email: [irl@hamsbd.org](mailto:irl@hamsbd.org)

Gurshabad Grover  
Centre for Internet and Society

Email: [gurshabad@cis-india.org](mailto:gurshabad@cis-india.org)

Mallory Knodel  
Center for Democracy and Technology

Email: [mknodel@cdt.org](mailto:mknodel@cdt.org)