Network Working Group
Internet-Draft
Intended status: Informational
Expires: November 17, 2019

I. Learmonth
Tor Project
May 16, 2019

Guidelines for Performing Safe Measurement on the Internet
draft-learmonth-pearg-safe-internet-measurement-02

Abstract

Researchers from industry and academia will often use Internet measurements as a part of their work. While these measurements can give insight into the functioning and usage of the Internet, they can come at the cost of user privacy. This document describes guidelines for ensuring that such measurements can be carried out safely.

Note

Comments are solicited and should be addressed to the research group's mailing list at 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

Copyright Notice

1. Introduction

When performing research using the Internet, as opposed to an isolated testbed or simulation platform, means that you research co-exists in a space with other users. This document outlines guidelines for academic and industry researchers that might use the Internet as part of scientific experiementation.

1.1. Scope of this document

Following the guidelines contained within this document is not a substitute for any institutional ethics review process you may have, although these guidelines could help to inform that process. 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 privacy.

1.2. Active and passive measurements

Internet measurement studies can be broadly categorized into two groups: active measurements and passive measurements. Active measurements generate traffic. Performance measurements such as TCP throughput testing [RFC6349] or functional measurements such as the feature-dependent connectivity failure tests performed by [PATHspider] both fall into this category. Performing passive measurements requires existing traffic. Passive measurements can help to inform new developments in Internet protocols but can also carry risk.

The type of measurement is not truly binary and many studies will include both active and passive components. Each of the considerations in this document must be carefully considered for their applicability regardless of the type of measurement.

2. Consent

Ideally, informed consent would be collected from all users of a shared network before measurements were performed on them. In cases where it is practical to do so, this should be done.

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.

2.1. 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.

2.2. 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 you asked for their informed consent for the measurements you'd like to perform.

For example, 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.

If practical, attempt to obtain informed consent or proxy consent from a sample of users to better understand the expectations of other users.

3. Safety Considerations

3.1. Use a 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.

3.2. Only record your own traffic

When performing 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.

3.3. Be respectful of other's 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.

3.3.1. 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 oppertunity 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.

3.4. Only collect data that is safe to make public

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

3.5. Minimization

For all data collected, consider whether or not it is really needed.

3.6. Aggregation

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

3.7. Source Aggregation

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.

4. Risk Analysis

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

5. Security Considerations

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

6. IANA Considerations

This document has no actions for IANA.

7. 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].

8. Informative References

[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/>.

[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>.

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

[PATHspider]
Learmonth, I., Trammell, B., Kuehlewind, M., and G. Fairhurst, "PATHspider: A tool for active measurement of path transparency", DOI 10.1145/2959424.2959441, July 2016, <https://dl.acm.org/citation.cfm?doid=2959424.2959441>.

[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>.

[RFC6349] Constantine, B., Forget, G., Geib, R., and R. Schrage, "Framework for TCP Throughput Testing", RFC 6349, DOI 10.17487/RFC6349, August 2011, <https://www.rfc-editor.org/info/rfc6349>.

[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>.

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

Author's Address

Iain R. Learmonth
Tor Project

Email: irl@torproject.org