Guidelines for Performing Safe Measurement on the Internet

Authors

- Iain Learmonth, Tor Project
- Gurshabad Grover, Centre for Internet and Society India
- Mallory Knodel, Center for Democracy and Technology
Goal

For industry and academia using measurements to research the functioning and usage of the Internet, this document describes guidelines for ensuring that such measurements can be carried out without violating user privacy.
Scope

● Not a substitute for any institutional ethics review process.
● Not legal advice.
● Restricted to guidelines for measurement of: the network, its constituent hosts and links, or its users traffic.
● 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.
Structure

- Consent: informed, proxy, implied
- Safety considerations:
  - Isolate risk with a dedicated testbed
  - Be respectful of others’ infrastructure
  - Data Minimization
  - Masking Data
- Risk analysis
1. Introduction
   1.1. Scope of this document
   1.2. Threat Model
   1.3. Measurement Studies
   1.4. User Impact from Measurement Studies

2. Consent
   2.1. Informed Consent
   2.2. Informed Consent: Case Study
   2.3. Proxy Consent
   2.4. Proxy Consent: Case Study
   2.5. Implied Consent
   2.6. Implied Consent: Case Study 1
   2.7. Implied Consent: Case Study 2

3. Safety Considerations
   3.1. Isolate risk with a dedicated testbed
   3.2. Be respectful of other's infrastructure
       3.2.1. Maintain a "Do Not Scan" list
   3.3. Data Minimization
       3.3.1. Discarding Data
       3.3.2. Masking Data
       3.3.3. Reduce Accuracy
       3.3.4. Data Aggregation

4. Risk Analysis

5. Security Considerations
Changes

Since -04

- Add section on responsible disclosure
- Add recent research, discuss IP address
- Safety != Ethics?

Since -05

- Nits and edits
- Accountability and transparency
Future of the draft (open issues)

- Responsible disclosure
- Talk about availability, not just disclosure
- Talk about IP addresses
- Future computing capabilities
- Look at CAIDA's Promotion of Data Sharing Webpage
- Look at Workshop on Ethics in Networked Systems Research
PRs, reviews welcome

https://github.com/IRTF-PEARG/draft-safe-internet-measurement