Opt In Session

Discussion of experimentation with operational network traffic

$$2IRB \bigvee \neg 2IRB$$

R. R. Brooks

`rrb@acm.org`

Clemson University, ECE

Who I am

This session

- Simulated network traffic does not behave like operational traffic.
- People do not behave like bots.
- Answer - Use operational data.

This session

- Simulated network traffic does not behave like operational traffic.
- People do not behave like bots.
- Answer - Use operational data.
- **We can do that?**

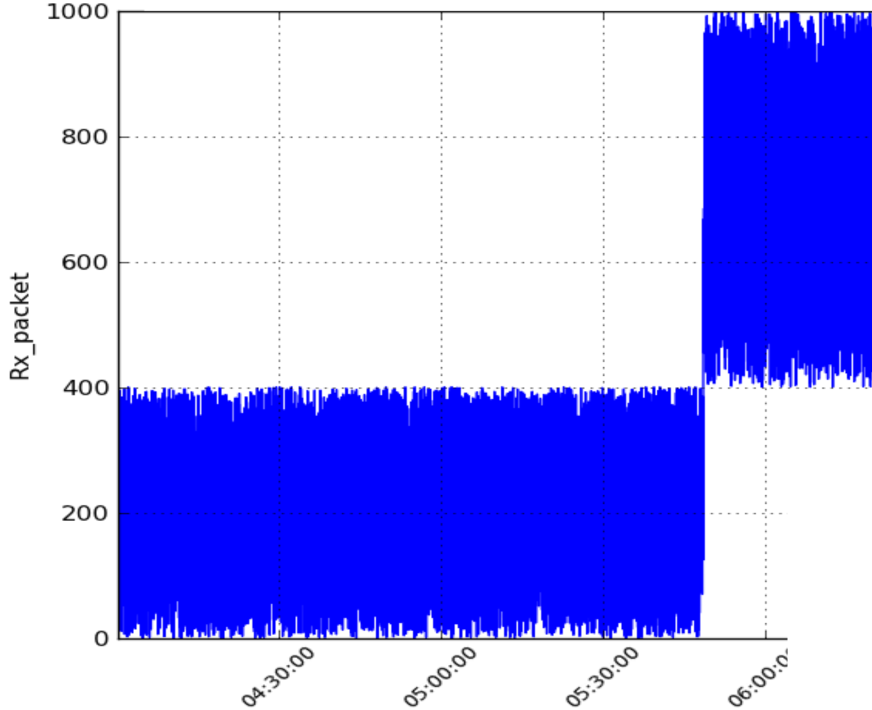
This session

- Simulated network traffic does not behave like operational traffic.
- People do not behave like bots.
- Answer - Use operational data.
- **We can do that?**
- *Is that ethical?*
- *But what about privacy?*
- *My IT group will never agree.*
- *What is an IRB? Aren’t they a pain?*

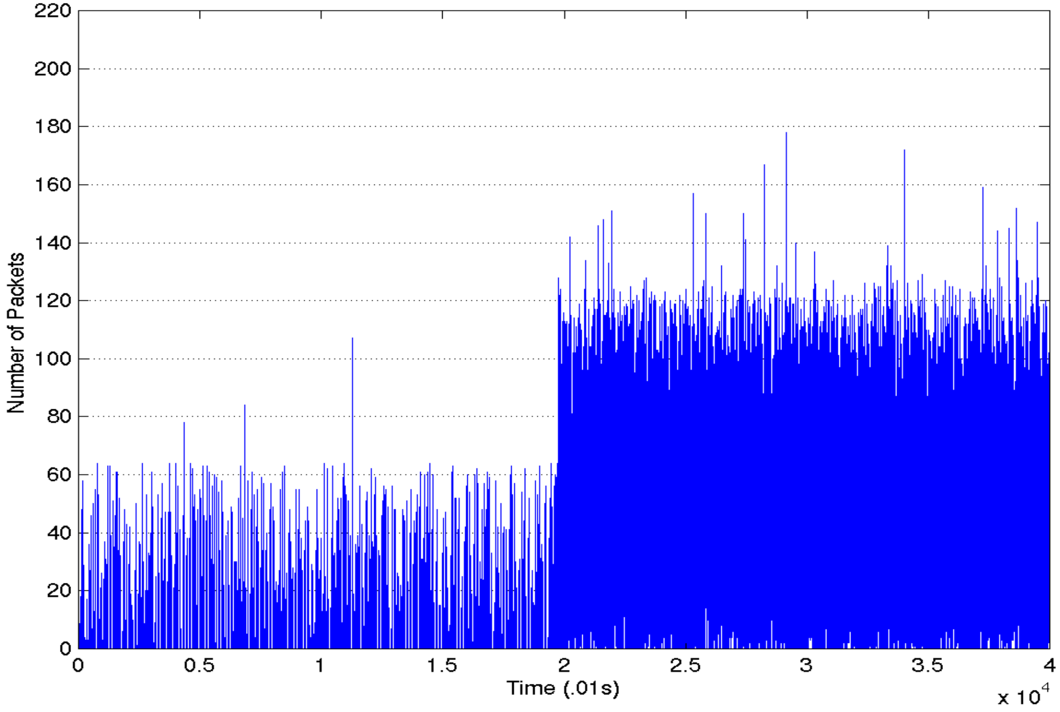
This session

- Simulated network traffic does not behave like operational traffic.
- People do not behave like bots.
- Answer - Use operational data.
- **We can do that?**
- *Is that ethical?*
- *But what about privacy?*
- *My IT group will never agree.*
- *What is an IRB? Aren't they a pain?*
- I have slides/discussion points to provoke comments.
- This should be group discussion *(therapy?)*

Artificial data

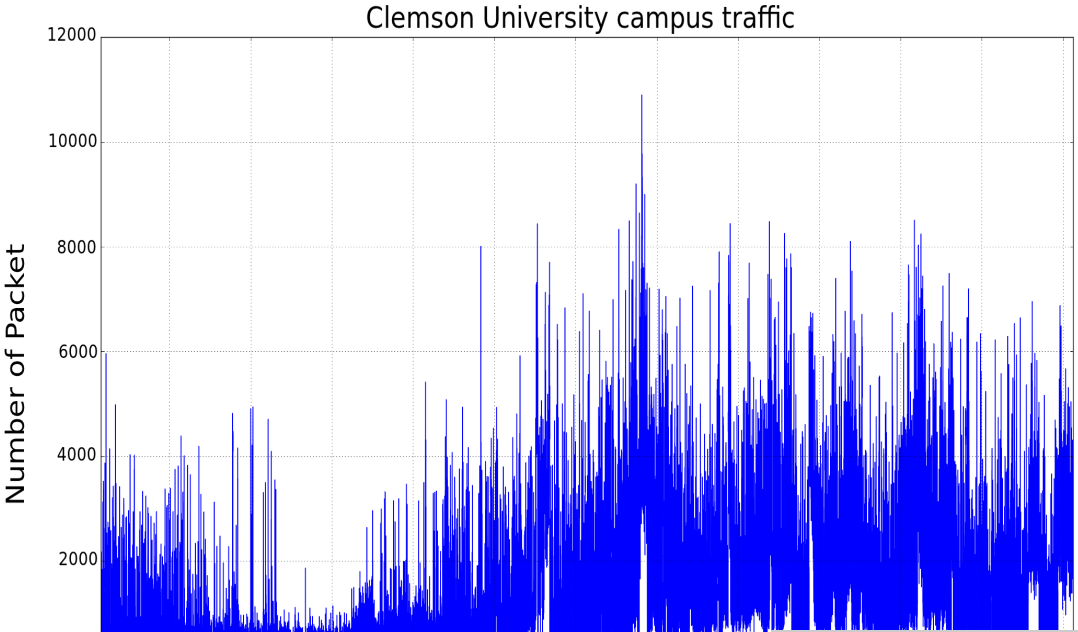


Matlab generated DDoS

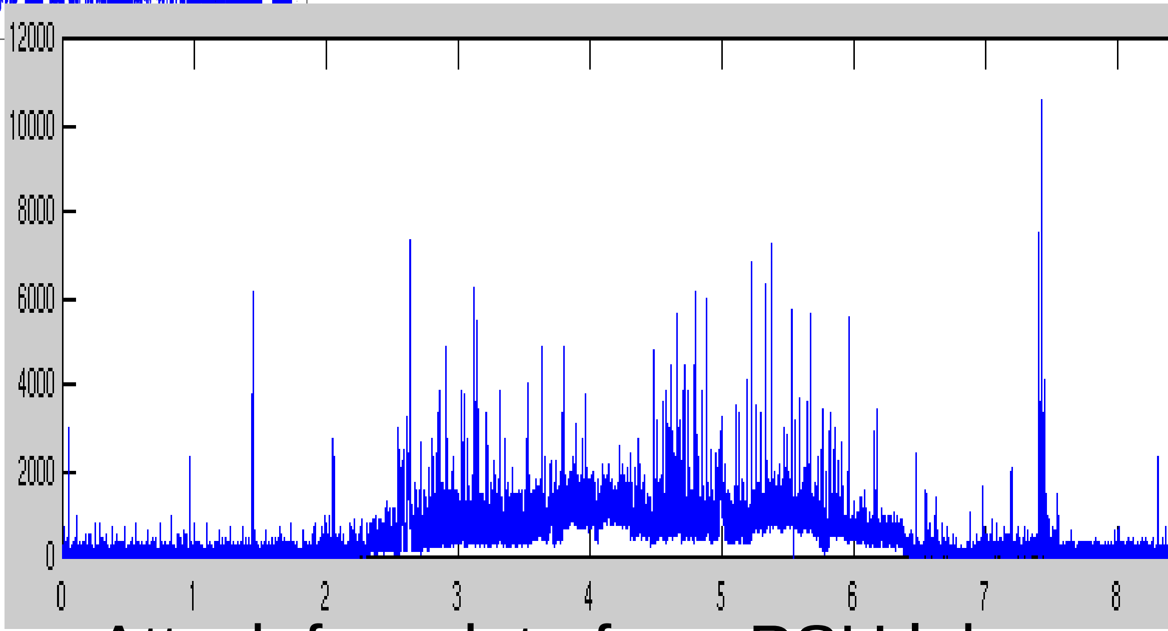


ns-2 generated DDoS

Ground truth data

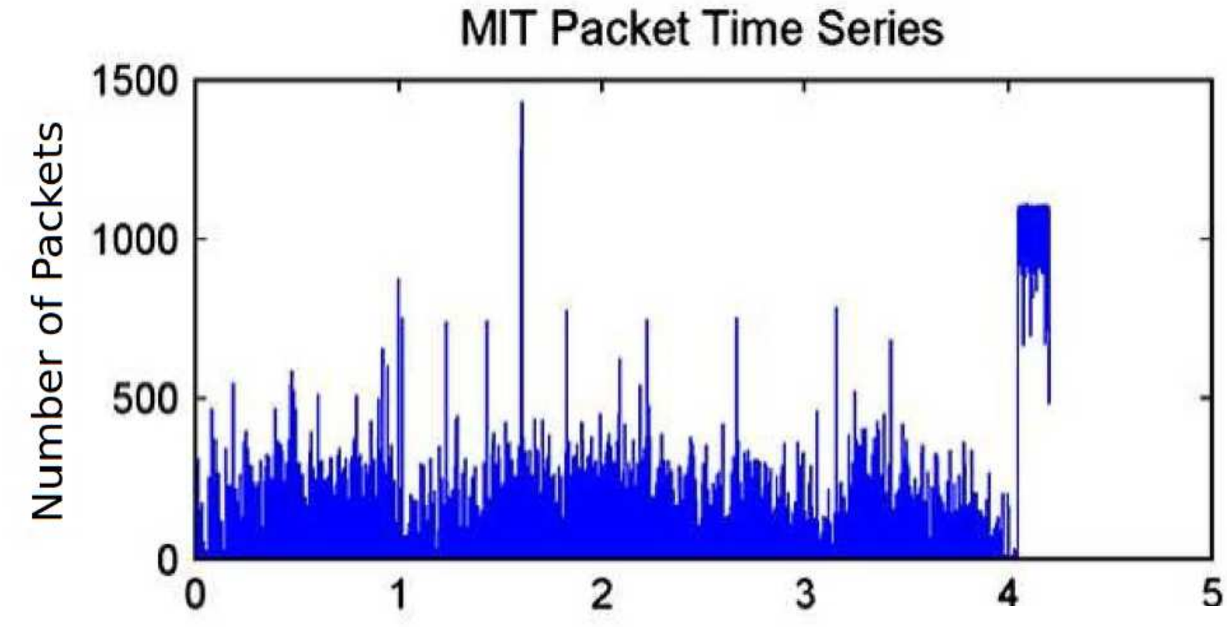


Clemson net w/o attack

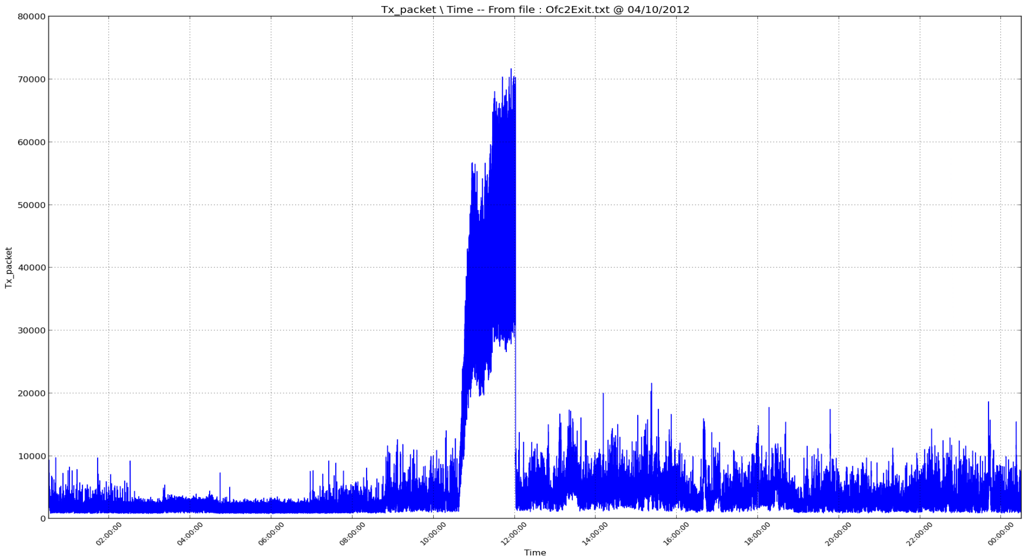


Attack free data from PSU lab

“Truthy” vs. real DDoS



MIT LL DARPA funded test



Clemson net with attack

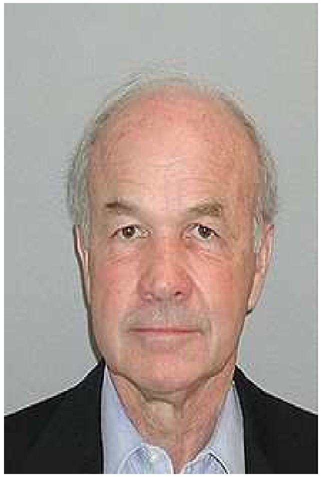
My work

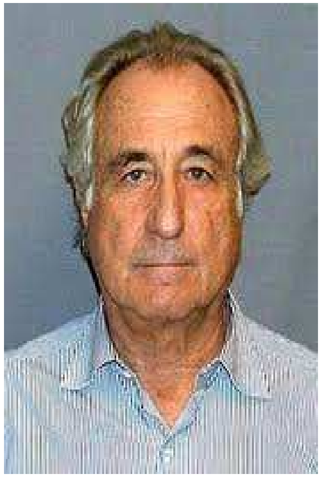
- DDoS detection/mitigation requires ground truth.
- We use operational traffic *without disturbing operations.*
- Buy in from CCIT, network security officer.
- Good news - Request for IRB review rejected.
- Collecting, analyzing, archiving traffic *statistics*
- No formal opt-in policy.
- Support from university CIO.

Ethics 1

- Highest ethical standards needed.
- We need to avoid temptation.
- People with strong moral foundation.
- Do the right thing.

Ethics 2





Ethics 3

- Consequences for all potential actions
  - Identify and evaluate consequences of actions
  - Compare the results from each
- Rights of others
  - Consider effects of actions on all participants
  - If anyone adversely affected, modify/avoid action
- Character perspective
  - What would a moral person do in this situation?
  - Try to emulate that approach
- Convergence
  - Iterate until these approaches converge

Privacy expectations

What privacy expectations do network users have?

- Are traffic statistics sensitive?

Privacy expectations

What privacy expectations do network users have?

- Are traffic statistics sensitive?
- Packet source/sink records?

Privacy expectations

What privacy expectations do network users have?

- Are traffic statistics sensitive?
- Packet source/sink records?
- Email traffic?
- *Email at public universities probably subject to FOIA*

Privacy expectations

What privacy expectations do network users have?

- Are traffic statistics sensitive?
- Packet source/sink records?
- Email traffic?
- *Email at public universities probably subject to FOIA*
- Google search queries?
- *AOL search query anonymization problems*

Privacy expectations

What privacy expectations do network users have?

- Are traffic statistics sensitive?
- Packet source/sink records?
- Email traffic?
- *Email at public universities probably subject to FOIA*
- Google search queries?
- *AOL search query anonymization problems*
- Packet headers OK as long as payload ignored?

Privacy expectations

What privacy expectations do network users have?

- Are traffic statistics sensitive?
- Packet source/sink records?
- Email traffic?
- *Email at public universities probably subject to FOIA*
- Google search queries?
- *AOL search query anonymization problems*
- Packet headers OK as long as payload ignored?
- **Action item –Classes of traffic not to analyze without opt-in**
- **Action item –Classes of traffic not to analyze with opt-in**

Privacy law

What is really private? *I am not a lawyer*

- Legal threshold is *reasonable expectation of privacy.*

Privacy law

What is really private? *I am not a lawyer*

- Legal threshold is *reasonable expectation of privacy.*
- Most traffic unencrypted, privacy expectation unreasonable.

Privacy law

What is really private? *I am not a lawyer*

- Legal threshold is *reasonable expectation of privacy.*
- Most traffic unencrypted, privacy expectation unreasonable.
- Information shared with third party not private.

Privacy law

What is really private? *I am not a lawyer*

- Legal threshold is *reasonable expectation of privacy.*
- Most traffic unencrypted, privacy expectation unreasonable.
- Information shared with third party not private.
- Employers can inspect traffic on their net.

Privacy law

What is really private? *I am not a lawyer*

- Legal threshold is *reasonable expectation of privacy.*
- Most traffic unencrypted, privacy expectation unreasonable.
- Information shared with third party not private.
- Employers can inspect traffic on their net.
- If corporate policy says no inspection, they still can.

Privacy law

What is really private? *I am not a lawyer*

- Legal threshold is *reasonable expectation of privacy.*
- Most traffic unencrypted, privacy expectation unreasonable.
- Information shared with third party not private.
- Employers can inspect traffic on their net.
- If corporate policy says no inspection, they still can.
- VOIP traffic more private than texting, email.

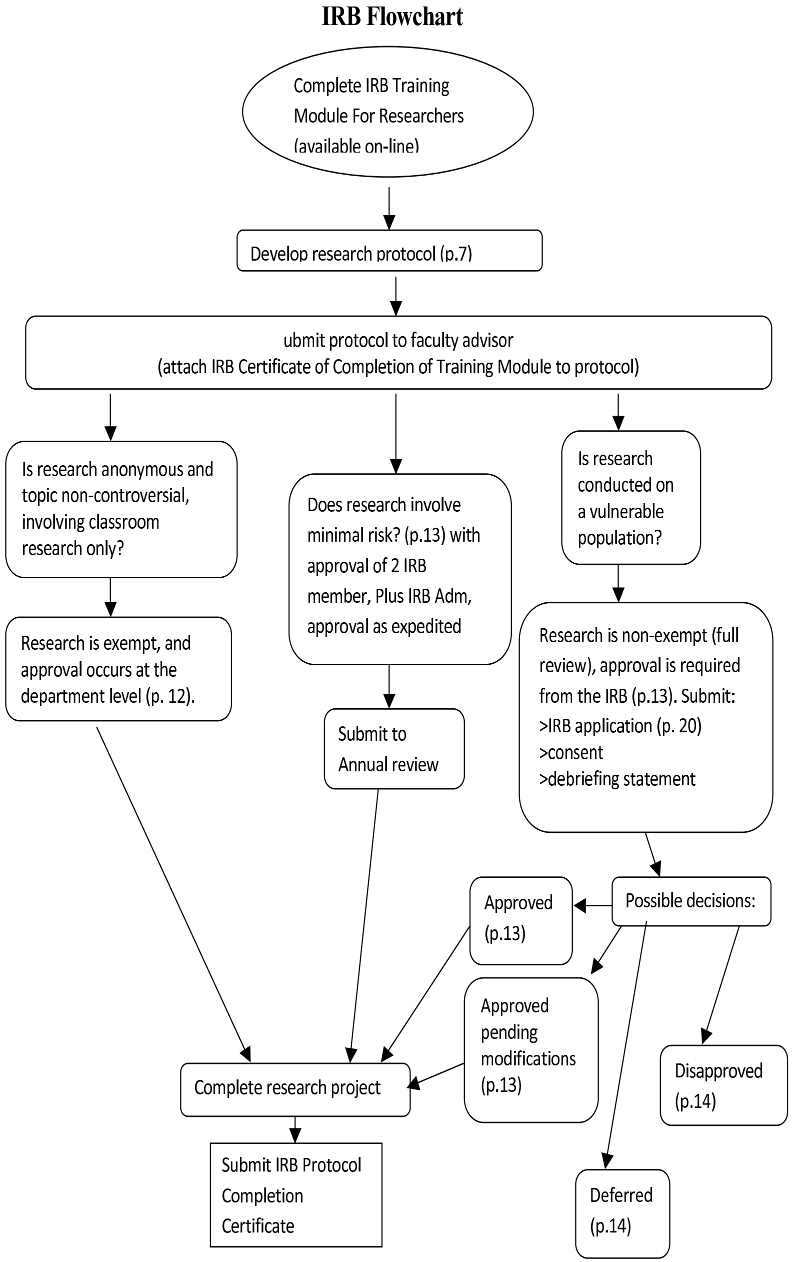
Privacy law

What is really private? *I am not a lawyer*

- Legal threshold is *reasonable expectation of privacy.*
- Most traffic unencrypted, privacy expectation unreasonable.
- Information shared with third party not private.
- Employers can inspect traffic on their net.
- If corporate policy says no inspection, they still can.
- VOIP traffic more private than texting, email.
- *“You have zero privacy...get over it.”* – Scott McNeally
- Orwell was an optimist. Zamyatin was a prophet.

IRB

- Institutional Review Board – DHHS mandated review of research involving humans.
- **Is it research?**
- **Are there human subjects**
- Training.
- Exempt, expedited, or full.
- $2IRB \bigvee \neg 2IRB$



IRB Flowchart

Complete IRB Training Module For Researchers (available on-line)

Develop research protocol (p.7)

ubmit protocol to faculty advisor
(attach IRB Certificate of Completion of Training Module to protocol)

Is research anonymous and topic non-controversial, involving classroom research only?

Research is exempt, and approval occurs at the department level (p. 12).

Does research involve minimal risk? (p.13) with approval of 2 IRB member, Plus IRB Adm, approval as expedited

Submit to Annual review

Is research conducted on a vulnerable population?

Research is non-exempt (full review), approval is required from the IRB (p.13). Submit:
>IRB application (p. 20)
>consent
>debriefing statement

Possible decisions:

Approved (p.13)

Approved pending modifications (p.13)

Disapproved (p.14)

Deferred (p.14)

Complete research project

Submit IRB Protocol Completion Certificate

Liability

- *I am not a lawyer.*
- As private enterprizes, companies and universities may have broad lee-way in handling data.
- As agents of the US government, they may have restictions in receiving and storing user data.
- For experimenters using the system, the rules seem opaque to me.
- Liability of GENI and hosting organizations for misuse of their systems. (i.e. GENI users attacking others.)
- Liability of GENI and hosting organizations for misuse of user data. (i.e. GENI abusing users.)
- **Action item: questions, discussion, future steps?**

Opt-in

- Action item: When should opt in be recommended/required?
- Action item: Could a boiler plate EULA and opt in framework/document be developed?
- Action item: Proper organizations to vet/review proposed?
- Action item: Other items?

Classes of users

- **Group exercise**
- Are there classes of users that need to be considered
  - students at the local university with IT approval,
  - users who agree to join in a prototype system,
  - faculty/staff,
  - the public at large?
- Foreign and domestic users?
- Anonymous users?

Banned

- What types of experiments on operational network traffic should not be performed, ever?

Banned

- What types of experiments on operational network traffic should not be performed, ever?
- Denial of service on GENI infrastructure.

Banned

- What types of experiments on operational network traffic should not be performed, ever?
- Denial of service on GENI infrastructure.
- Denial of service on external infrastructure.

Banned

- What types of experiments on operational network traffic should not be performed, ever?
- Denial of service on GENI infrastructure.
- Denial of service on external infrastructure.
- Port scans of external resources.

Banned

- What types of experiments on operational network traffic should not be performed, ever?
- Denial of service on GENI infrastructure.
- Denial of service on external infrastructure.
- Port scans of external resources.
- Spreading malware/spam...
- Others?

Prudence

Action items

- What steps should a prudent experimenter take before using operational data?
- What safeguards should be provided to prevent disturbing network operations?

Our approach

- Consider negative effects
- Discuss with local IT and testbed administrators
- Consider data storage/archival/privacy issues
- Prepare IRB information/application
- Verify need for IRB, (if necessary) get approval
- Run small test runs to verify lack of impact
- Do research

Action items

- Action items
- Comments
- Concerns