# Attribution on the Internet

David Clark
MIT CSAIL
Internet Policy Research Initiative
April 2018

# Accountability

- The range of misbehavior and malicious activity on the Internet has triggered calls for “an accountable Internet”.
  - Is this a good idea?
  - What would that concept actually mean?
- Should we add some sort of identity tracking to the Internet?

# For example

- In February 2010, former Director of the National Security Agency Mike McConnell wrote, "We need to develop an early-warning system to monitor cyberspace, identify intrusions and locate the source of attacks with a trail of evidence that can support diplomatic, military and legal options — and we must be able to do this in milliseconds. More specifically, we need to reengineer the Internet to make attribution, geolocation, intelligence analysis and impact assessment — who did it, from where, why and what was the result — more manageable."
- What, exactly, might he have meant?

# Some questions

- At what layer of the Internet would such a mechanism be added?
- What entity should issue the identities?
- How would they be verified?
- What purpose would these identities actually serve?
  - Put differently, what problem are we solving?

# At what layer?

- To users of the Internet, the term describes the overall experience.
  - "I was on the Internet for 3 hours today."
- To the designers, the Internet is a packet carriage platform on which apps run.
  - Facebook is just an app.
- Should the packet carriage layer of the Internet concern itself with tracking identity?
- Or should management of identity be a higher-level (application layer) concern?

# Identity at the packet layer

- Perhaps add some field to the IP header, or to a new layer just above IP.
- Would imply a single, uniform mechanism applied to all applications.
  - Could it be used to correlate behavior of users across multiple apps?
- Would imply that the goal is that the identity credential be visible (and meaningful?) to third-party observers (e.g., ISPs, law enforcement).
  - If the credentials are only used at the end-point, no need to put them into the IP layer.

# Some social questions

- Who would issue these identities?
  - Governments? Google?
  - Could a government ban citizen from the Internet by refusing to issue an identity?
- If useful to third parties, under what circumstances can they be resolved to find the actual identity?
- What sort of tracking would they facilitate?
- How can forged identities be prevented?
  - Should routers check that they are valid?
    - Seems much too costly, by far.
  - Hold ISPs accountable for verifying their users?

# In the real world

- We do not expect to have the same knowledge of identity in all contexts.
  - Sometimes interacting parties have a strong need to confirm the identity of the opposite parties.
  - Sometimes assured anonymity is desirable.
  - Sometime it does not matter.
- We do not demand that everyone we meet give us a validated identity number.

# Wrong idea

- In my view, adding an identity mechanism to the packet layer of the Internet is a Bad Idea.
  - The need for identity should be defined by the application—the context in which actors are interacting.
- Identity at the packet level would not work.
  - Too easy to forge, too hard to check, would not actually enhance security. (Holding identified parties accountable.)
- Not clear how this would address McConnell's concern.

# Identity at the app layer

- Better match to the real world.
  - Works at the end-points, not “in” the network.
- To the extent that the application implements a strong identity scheme, it can use this information for attribution.
- But puts a burden onto the application developer.
- In my view, we need a set of supporting services that provide identity management tools for apps, and which can be tailored to different needs.

# Addresses as identity

- Today, sources are identified by their IP address.
  - They can forge these addresses if they don't want an answer back.
  - We should clean this up.
  - BCP 38 calls for ISPs to verify source addresses.
  - For current status of this effort, see https://www.caida.org/projects/spoofer/

# Enforcement by AS

- For all the AS that have been measured:
  - Degree of compliance with the checking requirement of BCP 38

| Status | Count |
|---|---|
| Spoofable | 132 |
| Mostly spoofable | 28 |
| Partly spoofable | 31 |
| Blocked | 392 |

# Why do forged addresses matter?

- What harm are we trying to prevent?
  - Forged source addresses are only useful for simple DDoS attacks.
  - In any case where the attacker needs to get anything back, the source address must be valid.
- For any attack “at the application layer”, the IP address will be valid.
  - So need a map of different sorts of attacks.
  - (That is another talk…)

# The real issue with identity

- Multi-stage attacks.
  - Attackers first penetrate “helper” machines by exploiting security vulnerabilities.
  - They then use these machines as the platform from which to launch the attack.
    - Bots and botnets
    - Intermediate nodes for data exfiltration and espionage.
  - The IP address is that of the intermediate, not the attacker.
    - Consider the difficulties today in taking down botnets.
    - Must disrupt the command and control mechanisms.

# Goals

- If the goal is prevention or blocking of attack, the source IP address may be useful.
- If the goal is to identify the attacker, we need identity of the person.
  - There is no way “the Internet” can (as a technical capability) map IP addresses to people.
  - Needs the help of ISPs and other actors.
  - Becomes a cross-state, political challenge, not a technical challenge.
  - What is “beyond a reasonable doubt?”
- Unless the event is state-sponsored.
  - Then we need attribution at the state level.
    - But what about “patriotic hackers”. Hold the state responsible?

# Conclusions

- The real challenge with attribution is the multi-stage nature of today's attacks.
- Identity management needs to be tailored to the differing requirements of different apps.
- Making IP addresses harder to forge is a good objective, but does not solve the attribution problem.
- Further reading: Clark, D. D. , Landau, S. *Untangling Attribution*. Harvard National Security Journal, Vol. 2, Issue 2 (2011)