CYBER SECURITY DIVISION

2014 R&D SHOWCASE AND TECHNICAL WORKSHOP

Internet Mapping Primitives

Naval Postgraduate School

Robert Beverly

December 17, 2014

Homeland Security

Science and Technology

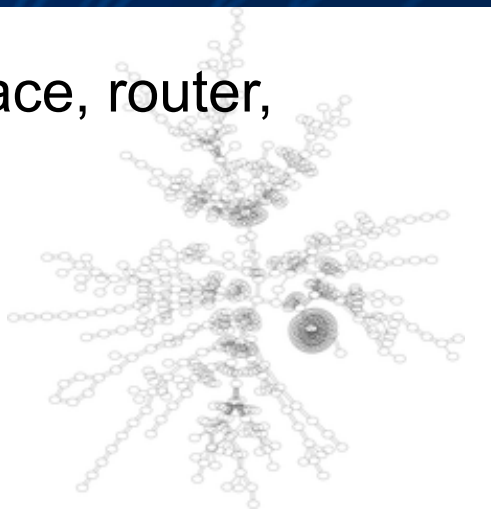
Team Profile

- **Naval Postgraduate School**:
  - US Navy's Research University
  - Located in Monterey, CA
  - ~1500 students (all services, civilians, foreign military)

- **Center for Measurement and Analysis of Network Data**:
  - **PI**: Robert Beverly
  - **Faculty**: Geoffrey Xie, Justin Rohrer (NPS CS), Ralucca Gera (NPS Math), Arthur Berger (Akamai)
  - **Students**: Lance Alt, Guillermo Baltra, Billy Brinkmeyer, Blake Lafever, Daryl Lee, Erik Rye, Sam Trassare

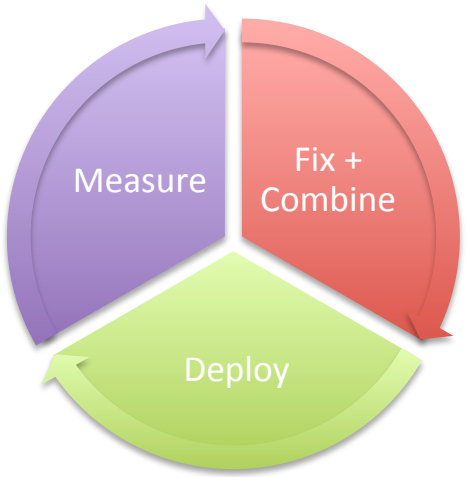
Customer Need

- Efficiently gather accurate network maps (interface, router, autonomous system) amid:
  - Network dynamics (long and short-lived)
  - Vantage points with different views
  - Topological sparsity (e.g. IPv6)
  - Potential deception/fakery (discussed later)
- Such maps are used for:
  - Security (e.g. route hijacking, weak/strong network points)
  - Optimizing content delivery (e.g. CDNs)
  - Geolocation
  - Network management/debugging (e.g. reverse traceroute)
- DHS BAA:
  - "*…identify infrastructure components in greatest need of protection*."



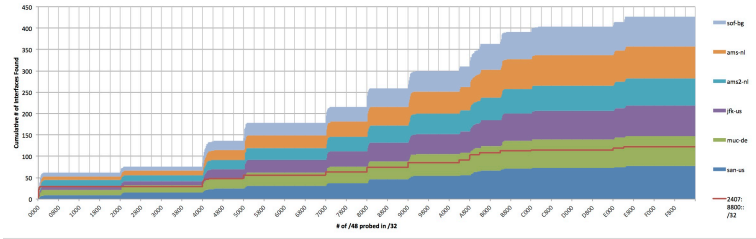
Approach (prior, year 1 work)

- Started with primitives we proposed in [BBX10]:
  1. Utilize available external knowledge
  2. Maintain state over prior rounds of probing
  3. Adaptively sample to discover subnet structure
- Implement on CAIDA's Archipelago (Ark) infrastructure
- Real-world implementation and deployment led to:
  - Fixing primitives
  - Combining primitives
  - Ingress Point Spreading [BBX14] (rank order vantage points in order of expected utility for a destination)

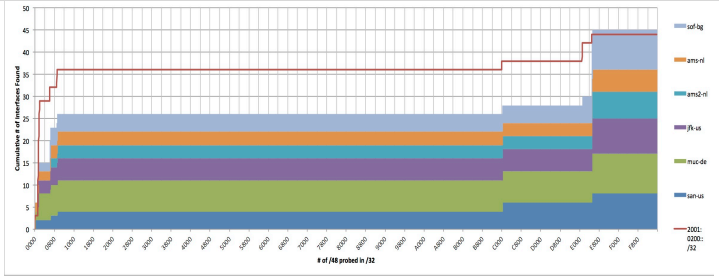


Approach (this year – what's new)

- Active collaboration with CAIDA
- Technology transfer plan includes deployment as part of CAIDA's production (Ark) IPv6 mapping:
  - In practice, did not obtain efficiency/coverage gain
  - Due to common IPv6 subnetting practices
    - Visualizing all /48's in a subnetted /32 IPv6 prefix:



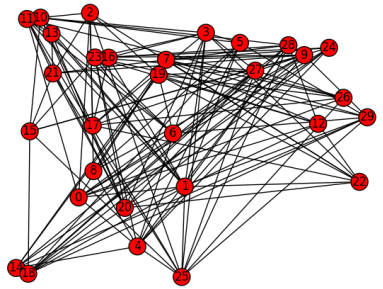
    - But, most /32 IPv6 prefixes have very little subnetting:



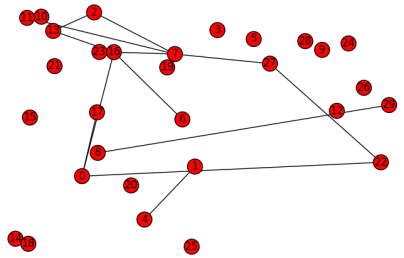
- Implication: different probing strategy required for IPv6

Approach (this year -- what's new)

- No ground-truth is a common measurement obstacle:
  - Evaluate relative benefit of approaches
  - Or, limited validation
- Want more concrete understanding of mapping abilities
- So, create own ground-truth:
  - Emulate real router images (IOS/JunOS)
  - Emulate complex topologies and dynamics
  - On a 96GB 16-core machine, can emulate 300 Cisco 7200 routers implementing customer, provider, and peering policies with > 150k BGP routes injected
- Understand limits of tools on a variety of topologies
- Automation: explore lots of topologies, results
- Emulation reveals artifacts simulation cannot



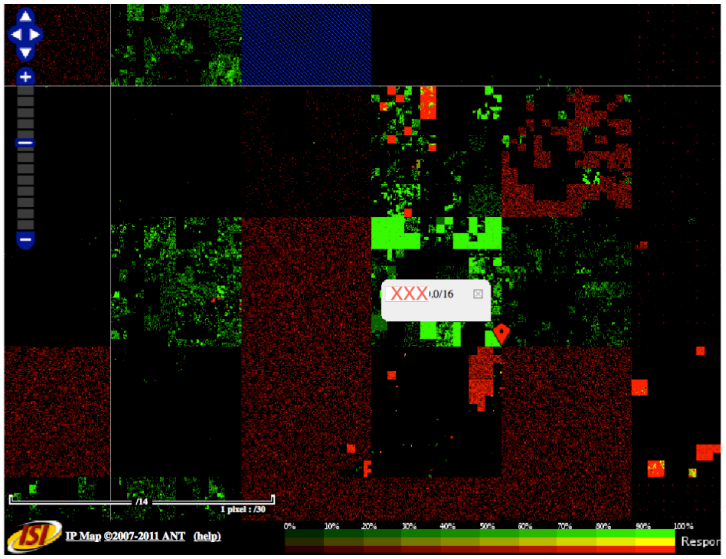
inferred topology



difference from known ground truth

Approach (this year -- what's new)

- Not only is ground-truth elusive, people lie
- An unexpected insight of our research: making Internet measurements more robust to deception
- ACSAC 2014 "Uncovering Network Tarpits with Degreaser"
- Developed a tool to detect tarpits
- Randomly scan all /24s on Internet
- Found > 215k (fake) IPs
- As large as /16's!
- Synergistic w/ DHS-funded ISI census work



Data from USC/LANDER
internet_address_census_it58w-20140122
Visualized with their IPv4 browser

Current Status

- Have met DHS year-2 deliverables:
- Produced working implementations of:
- Recursive Subnet Inference (RSI)
- Ingress Point Spreading (IPS)
- Bonus work in:
  - IPv6
  - Emulation/ground-truth
  - Deception/tool robustness
- Working on:
  - Gathering more topology snapshots using methodology
  - Technology transfer

Benefits

- Developed and integrated new topology primitives into a cohesive mapping system
- Real-world implementation permits adoption
- Demonstrated the utility of using vantage points intelligently
- Demonstrated ability to discover more topology with half the load (amount of probing) and time
- First large-scale experimentation with new IPv6 mapping techniques
- Demonstrated ability to detect fake host responses and generate more accurate maps

Next Steps

- Final year thrusts:
  - Refine primitives for IPv6
  - Gather and characterize more topologies, especially IPv6
  - More detection/analysis in space of measurement adversaries (deception)
  - Utilize emulation testbed to evaluate tools, especially under various topology dynamics
  - Deploy primitives in production

Transition

- Software contributions:
  - Running/working implementation on CAIDA's Ark
  - Native python scamper controller
  - Native python scamper warts binary parser
  - Dynamips patches (our automation is triggering bugs no one else is encountering)
  - Degreaser publicly available on github
- Academic papers:
  - **ACSAC2014**: "Uncovering Network Tarpits with Degreaser"
  - **PAM2014**: "Ingress Point Spreading"
  - **PAM2013**: "IPv6 Alias Resolution via Induced Fragmentation"

Contact Information

- Center for Measurement and Analysis of Network Data @NPS: http://www.cmand.org
- Contact:

Robert Beverly
Assistant Professor
http://rbeverly.net/research
rbeverly@nps.edu
831-656-2132