Towards Designing Effective Visualizations for DNS-based Network Threat Analysis

Rosa Romero-Gómez, Yacin Nadji, and Manos Antonakakis

{rgomez30,yacin,manos}@gatech.edu

Georgia Tech School of Electrical and Computer Engineering
College of Engineering

What is Network Threat Analysis?

Analyst

The Domain Name System (DNS) is an essential protocol used by both legitimate Internet applications and cyber attacks

[Building a Dynamic Reputation System for DNS, Antonakakis et al. 2010]

Challenges

Threat Intelligence Acquisition

"Security analysts are still collecting threat intelligence via email, spreadsheets, and cutting/pasting information from web-based sources. Obviously, these manual processes don't scale"

[Enterprise Strategy Group (ESG) Research Report: Threat Intelligence and Its Role Within Enterprise Cybersecurity Practices,

June 2015]

Analytics

"Threat intelligence may offer clues but human beings are left to do the heavy lifting by investigating and analyzing the data on their own"

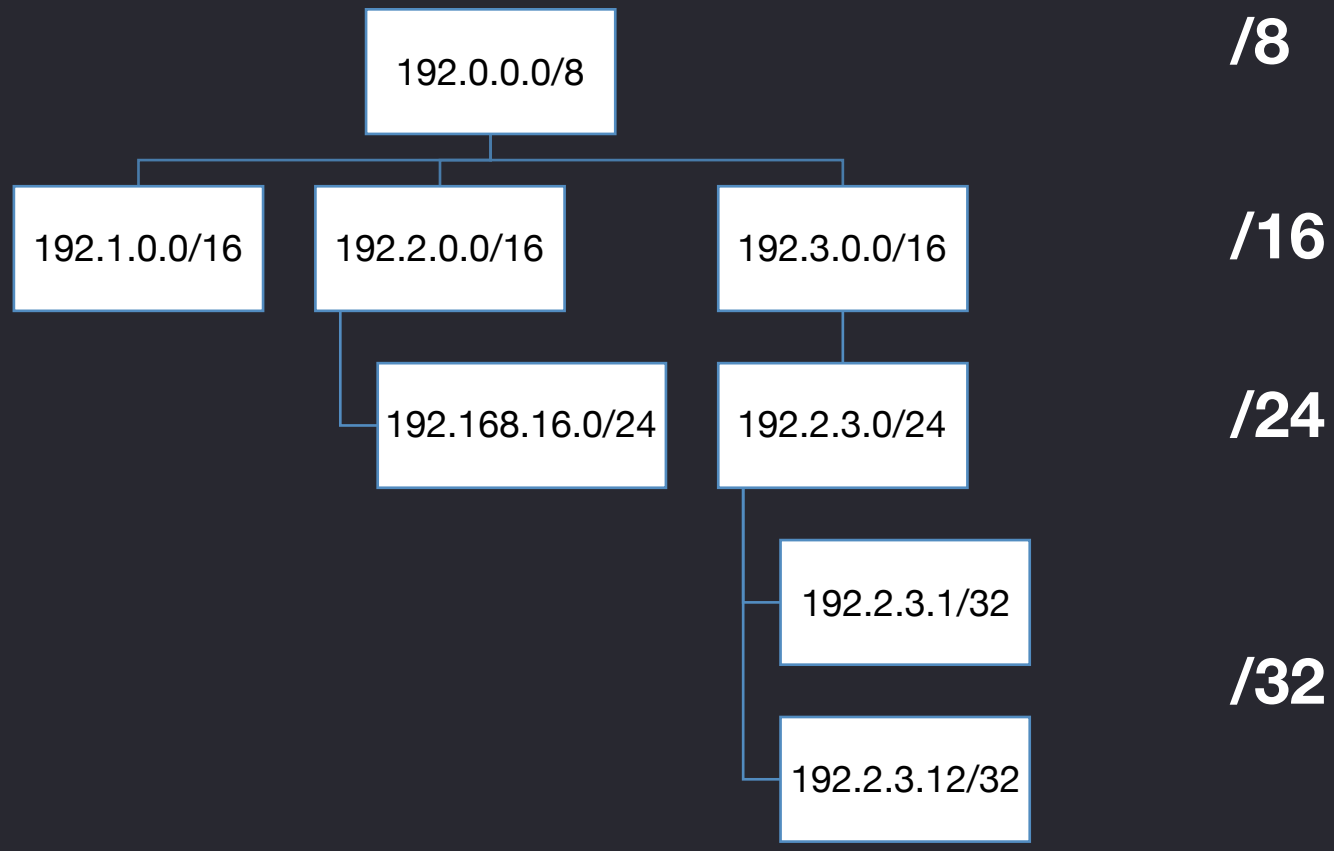
[ESG Research Report: Threat Intelligence and Its Role Within Enterprise Cybersecurity Practices, June 2015]

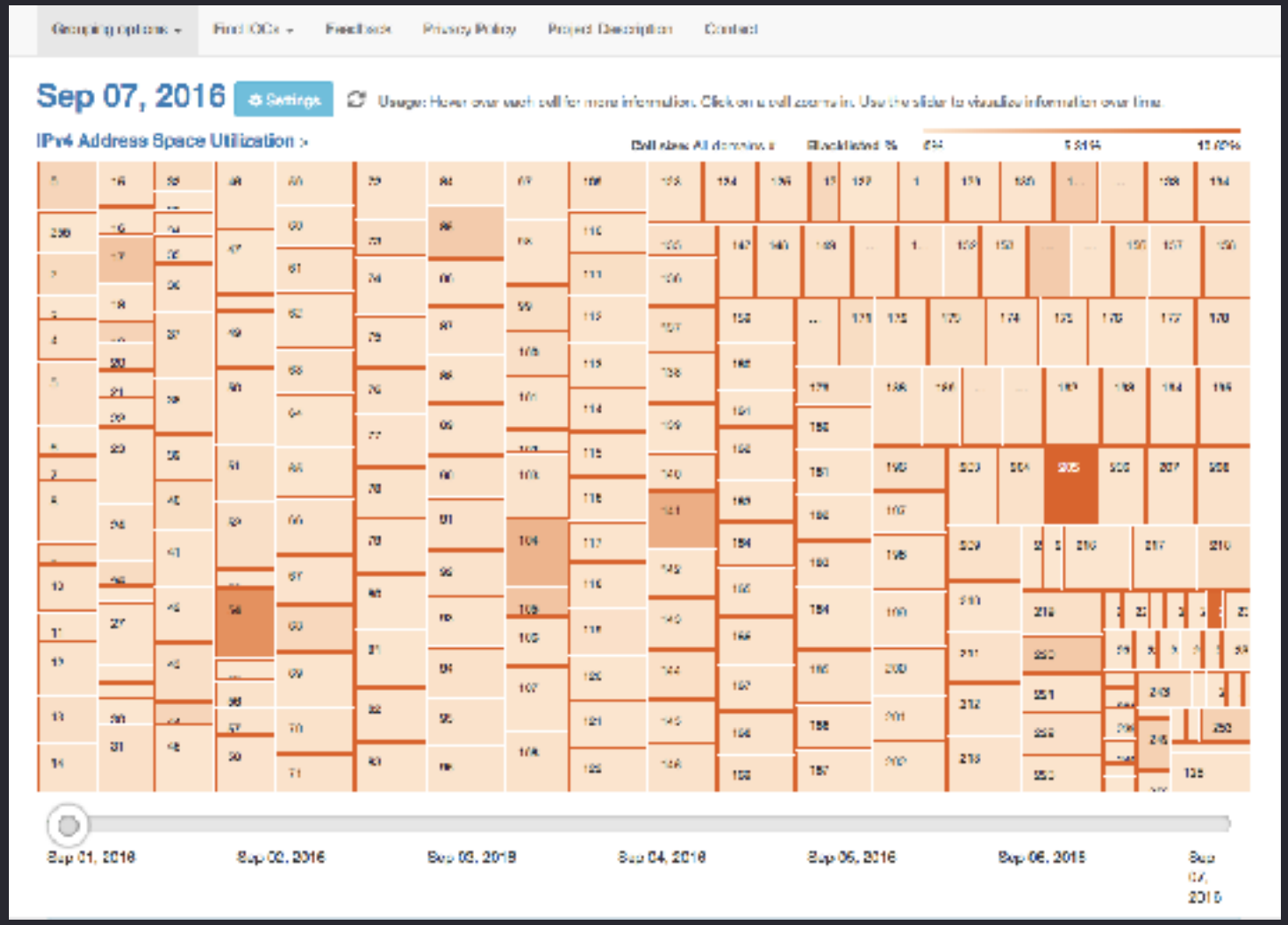
Our Approach:
Open Source THreat Analysis COnsole (THACO)

Open Datasets

Visualization Techniques

Classless Inter-Domain Routing (CIDR) Notation





Access to THACO live demo: https://ipviz.gtisc.gatech.edu/

User-centered Visualization Design

Domain Problem / Data Characterization

Design

Prototype

Evaluation

Domain Problem & Data Characterization

- Procedure: informal interviews with two domain experts in network threat analysis during two months
- Output: two main high-level categories of tasks and data requirements:

Top-down network threat analysis or Threat Hunting

Bottom-up network threat analysis or Incident Response

Domain Problem & Data Characterization

- Effective threat intelligence involves the combination of multiple data sources:
  - Active DNS datasets (https://www.activednsproject.org/)
  - Public Domain Blacklists such as abuse.ch
  - Malware Traces (https://www.virustotal.com)
  - Domain WHOIS records (https://www.threatminer.org/)

Design

- Design Goals
    1. Multiple views
    2. Different levels of detail
    3. Scalability

Multi-grouping, zoomable treemap

Evaluation

- Participants:
  - Network threat analysts are hard to find
  - Seven in-situ and thirty-one online network threat analysts from both academia and industry
  - Years of experience ranging < *1 year* to > *10 years*
- Procedure:
  - In-situ evaluation: tasks scenarios and semi-structured interviews
  - Online evaluation: web-based survey (SUS, System Usability Scale)

Evaluation

- Main Results:
  - Threat analysis experience of participants affects neither task completion rates nor task completion times using THACO
  - Experience analysts satisfaction garnered THACO an “A” grade in usability
- Limitations:
  - THACO could be improved for tasks involving keeping track of different pieces of information over time

Want data?

Active DNS datasets: https://www.activednsproject.org/

Want a demo?

THACO live demo: https://ipviz.gtisc.gatech.edu/

Want code?

Source code on GitHub: https://github.com/Astrolavos/THACO

Questions?

Towards Designing Effective Visualizations for DNS-based Network Threat Analysis

Rosa Romero-Gómez, Yacin Nadji, and Manos Antonakakis

{rgomez30,yacin,manos}@gatech.edu

Georgia Tech School of Electrical and Computer Engineering
College of Engineering