AFRL-RI-RS-TR-2018-008

A NEXT GENERATION REPOSITORY FOR SHARING SENSITIVE NETWORK AND SECURITY DATA

UNIVERSITY OF MICHIGAN

JANUARY 2018

FINAL TECHNICAL REPORT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

STINFO COPY

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE

■ **AIR FORCE MATERIEL COMMAND** ■ **UNITED STATES AIR FORCE** ■ **ROME, NY 13441**

NOTICE AND SIGNATURE PAGE

AFRL-RI-RS-TR-2018-008 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE CHIEF ENGINEER:

/ S /
FRANCES A. ROSE
Work Unit Manager

/ S /
JOHN D. MATYJAS
Technical Advisor, Computing
& Communications Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* JAN 2018	**2. REPORT TYPE** FINAL TECHNICAL REPORT	**3. DATES COVERED** *(From - To)* SEP 2012 – SEP 2017

4. TITLE AND SUBTITLE

A NEXT GENERATION REPOSITORY FOR SHARING SENSITIVE NETWORK AND SECURITY DATA

5a. CONTRACT NUMBER
N/A

5b. GRANT NUMBER
FA8750-12-2-0314

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Michael Kallitsis

5d. PROJECT NUMBER
DHSP

5e. TASK NUMBER
MI

5f. WORK UNIT NUMBER
CH

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Regents of the University of Michigan
Division of Research Development Administration
503 Thompson St
Ann Arbor, MI 48109-1340

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory/RITE
525 Brooks Road
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/RI

11. SPONSOR/MONITOR'S REPORT NUMBER
AFRL-RI-RS-TR-2018-008

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited. This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Defending critical infrastructure from cyber-security threats, understanding macroscopic Internet and developing the next-generation tools for monitoring, maintaining and controlling Internet's complex ecosystem requires access to real-world data. Towards this goal, in the course of this project, our team has disseminated to the networking and security community, in a responsible and ethical manner, unique datasets collected and curated at Merit Network, Inc. Thus, researchers had the unique opportunity to get their hands on realistic data from a large Internet service provider, and utilize such data for experimentation, discovery of Internet trends, research and development of new software and tools, pedagogical purposes and others.

15. SUBJECT TERMS

Real-world data, cyber-security, flow data, Darknet data, routing data, IMPACT, network data, measurements

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** **FRANCES A. ROSE**
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	UU	16	**19b. TELEPHONE NUMBER** (*Include area code)* **N/A**

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

This material is based on research sponsored by Air Force Research Laboratory and Department of Homeland Security under agreement number FA8750-12-2-0314. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

1. Summary

Defending critical infrastructure from cyber-security threats (e.g., distributed denial of service (DDoS) attacks, advanced persistent threats (APT), malware, etc.), understanding macroscopic Internet phenomena (such as power outages or connectivity disruptions due to BGP misconfigurations or prefix hijacking attacks) and developing the next-generation tools for monitoring, maintaining and controlling Internet's complex ecosystem requires access to real-world data. Towards this goal, in the course of this project, our team has disseminated to the networking and security community, in a responsible and ethical manner, unique datasets including real-world traffic flow data, Internet Background Radiation data (aka Darknet or network telescope data), routing data (e.g., BGP routing tables and RADb data) and other material collected and curated at Merit Network, Inc. Thus, researchers had the unique opportunity to get their hands on realistic data from a large Internet service provider, and utilize such data for experimentation, discovery of Internet trends, research and development of new software and tools, pedagogical purposes and others. Overall, we have made available through the IMPACT program hundreds of terabytes of data, and delivered (or provided access to) data to more than 60 unique researchers from about 50 unique organizations (from industry, academia and the government, both nationally and internationally). The collected data has been leveraged in at least a dozen research publications performed by our team and collaborators, and has supported the work of several PhD students from our institution and elsewhere.

Specific work accomplishments:

1. Over 15 TB of longitudinal Netflow data, 50 TB+ of Darknet data, 1.5 TB of BGP data, 2 TB+ of DDoS attack snapshots in Netflow format, 3 TB+ of Mirai scanning data and 559 GB of RADb data.
2. Data dissemination. We have received more than 400 data requests and have delivered data to all researchers that have been approved for receiving the data and have completed all the necessary steps for receiving the data (e.g., MOAs, access keys, etc.). Overall we provided data to more than 60 unique researchers from about 50 unique organizations. We have made available more than 50 different datasets.
3. Administrative support. As members of the IMPACT team, we
    a. Attended 10+ onsite PI meetings
    b. Attended 50+ PI calls
    c. Performed 1 IRB submission, and 5 yearly IRB reviews
    d. Provided legal support for MOA data provider and host agreements and amendments
    e. Feedback and bug reporting for the new IMPACT portal

4. Research. The project supported through funds or data more than dozen scientific publications by authors at University of Michigan and collaborators.
5. Supporting research infrastructure and data analytics as a service (DASP). For the past few years, longitudinal data from Merit network are actively supporting:
    a. University of Michigan's Censys -- live BGP feed from Merit
    b. CAIDA's IODA -- live Darknet feed from Merit
6. Professional development. The project has supported the research of five graduate students in the field of cyber-security.
7. Outreach. We have supported the project in a number of different outreach activities including conference presentations and seminars, blog posts and contributions to the IMPACT forum.

2. Introduction

Internet's decentralized design, management and operation make it vulnerable to a plethora of cyber-security threats, including Distributed Denial of Service (DDoS) attacks, BGP route hijackings, malware propagation, scanning, data breaches, etc. High-value empirical data is critical to network and security research and development efforts that aim to protect critical infrastructure against cyber-threats. Our participation in the IMPACT project, has allowed our team to ethically and responsibly provide high-quality, real-world, operational and longitudinal data to the IMPACT community.

The provided datasets aimed at addressing various cyber-security R&D needs such as a) experimentation with high-volume, high-variety, high-velocity data in order to improve the design and evaluation of new intrusion detection methods and systems, b) event-reconstruction and evidence-based insights into global trends (e.g., DDoS attacks and malware propagation), and c) situational awareness (e.g., outage detection). We have leveraged IMPACT's policy and legal framework to minimize any risks associated with real-world data (e.g., mapping data or meta-data back to individuals or intervening and/or interacting with any humans or the devices thereof identified via the use of our data). Further, we have coupled this policy framework with additional technical controls in order to safeguard against any leakage of private information: we have anonymized internet protocol (IP) addresses from network traces as needed, and we have designed and developed a "secure enclaves" framework for sharing sensitive data with researchers in a "code-to-data" manner. Over the course of the past 5 years, an array of unique Internet datasets have been provisioned, including real-world traffic flow data, network telescope (darknet) data, routing and RADb data, among others. We next describe the datasets collected and our data sharing mechanisms.

3. Methods, Assumptions, and Procedures

The overarching objective of our project was to provide *real-world* data to researchers, collected and curated at Merit Network, Inc, a large research and education network (R&E) serving the State of Michigan. Merit has infrastructure that covers the entire state of Michigan, and serves an estimated population of about 1 million users. Merit serves higher education institutions, K-12 schools, the government and several other non-profit organizations. In this section, we describe our collection infrastructure and our methodology of sharing traffic flow data, Darknet data and routing data (BGP).

3.1 Traffic Flow Data

Figure 1 shows a simplified view of Merit's backbone network. We have deployed infrastructure to help us collect operational flow data for (approximately) *all* ingress and egress traffic at Merit.

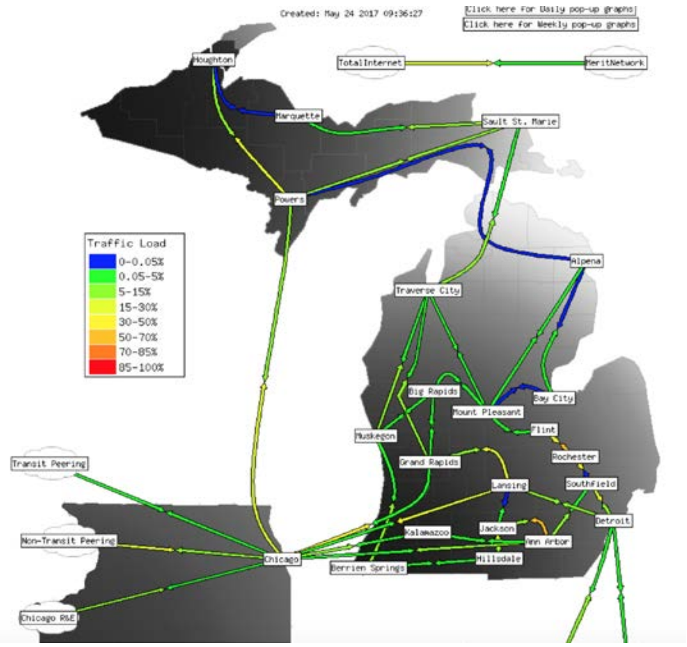

Figure 1. High-level overview of Merit's backbone network. We have installed measurement stations in Chicago and Detroit to collect unsampled Netflow data

To achieve this we have equipped all of our monitoring servers with PF_RING. PF_RING is an efficient packet capturing software that has helped us scale to multi-10 Gbps speeds. During high-peak school seasons, the monitored traffic was approaching 70 Gbps and we would collect about 200 GB of Netflow data per day. The collected flows were transmitted to a centralized server that would first *anonymize* the data (by stripping out the last 11 bits of both the source and destination

IPs) and then write the Netflow data on disks. We have used the nprobe software tool to construct the flows at the collection points and flow-tools to write the data on disks.

Longitudinal collection of traffic flow data at Merit from several vantage points has enabled our team to capture several interesting *snapshots* of internet traffic involving attacks such as volumetric DDoS attacks or state-exhaustion attacks (e.g., SYN flooding). We have been providing to the IMPACT community reflection and amplification attacks based on the NTP, DNS, SSDP and CHARGEN protocols. **Figure 2** shows a volumetric NTP-based DDoS attack captured at Merit. This dataset involved several insecure hosts within Merit being used as amplifiers, targeting various victims globally [1]. A snapshot of this event, collected and offered via IMPACT, has been frequently requested by researchers designing the next-generation intrusion detection and mitigation tools.

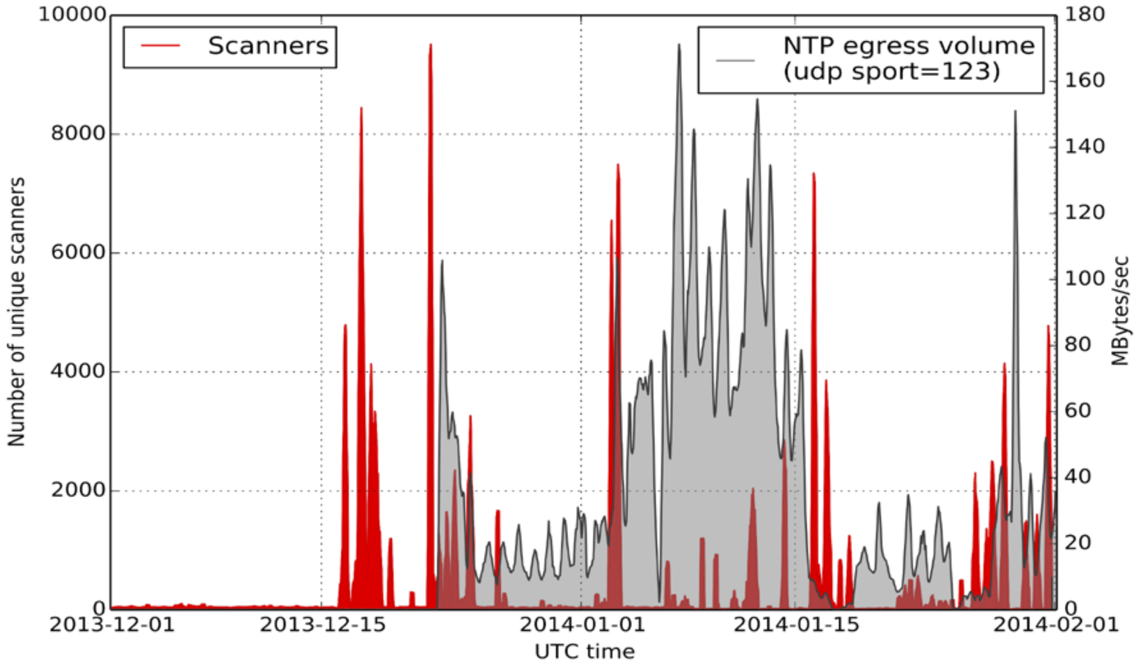

Figure 2. NTP DDoS attacks (and scanning activity appearing in our Darknet) captured at Merit early 2014.

The real-world flow datasets have supported IMPACT customer needs such as: a) experimentation, i.e., supporting researchers aiming to improve the design and testing of intrusion detection tools and methodologies. Researchers have utilized the data to assess the performance of their algorithms using a variety of case studies, e.g., distributed, high-volume, high-variety attacks (e.g., NTP DDoS), or low-volume, seemingly innocuous attacks (e.g., SYN flood attack or DDoS attacks based on protocols with low amplification factor such as SSDP, CHARGEN). The aforementioned datasets can also be "replayed" in a simulation/emulation environments such as DHS's DETER for educational/training purposes; b) event reconstruction, i.e., used as a “sensor” to support evidence regarding high-impact, global events such as DoS attacks, outages (due to natural phenomena, human-based nefarious activity, censorship or misconfigurations); c) evidence-based, time-series analyses to disclose emerging Internet trends, such as adoption of SSL/TLS protocols, trends regarding the Internet-of-Things (e.g., in terms of traffic growth) or malicious trends (e.g., scanning).

3.2 Darknet Data

Network telescopes (see **Figure 3**) receive real-world Internet data destined to an unused address space, and hence yield valuable information about global trends and activities [2]. For example, one can a) obtain evidence of spoofing-based DDoS activities manifested as backscatter, b) detect outages, c) observe global trends in scanning and worm propagation, and d) identify misconfigurations (e.g., routing). Merit has been operating one of the largest network telescopes still operational, and maintains historical, longitudinal darknet data that date back to 2005. Internet background radiation traffic arriving into our Darknet can provide unique insights into global / Internet-wide activities. In our recent work [3], Merit’s Darknet data was coupled with several other unique datasets (e.g., Censys data) to characterize the Mirai botnet, which crippled the Internet with DDoS attacks reaching 1 Tbps late 2016.

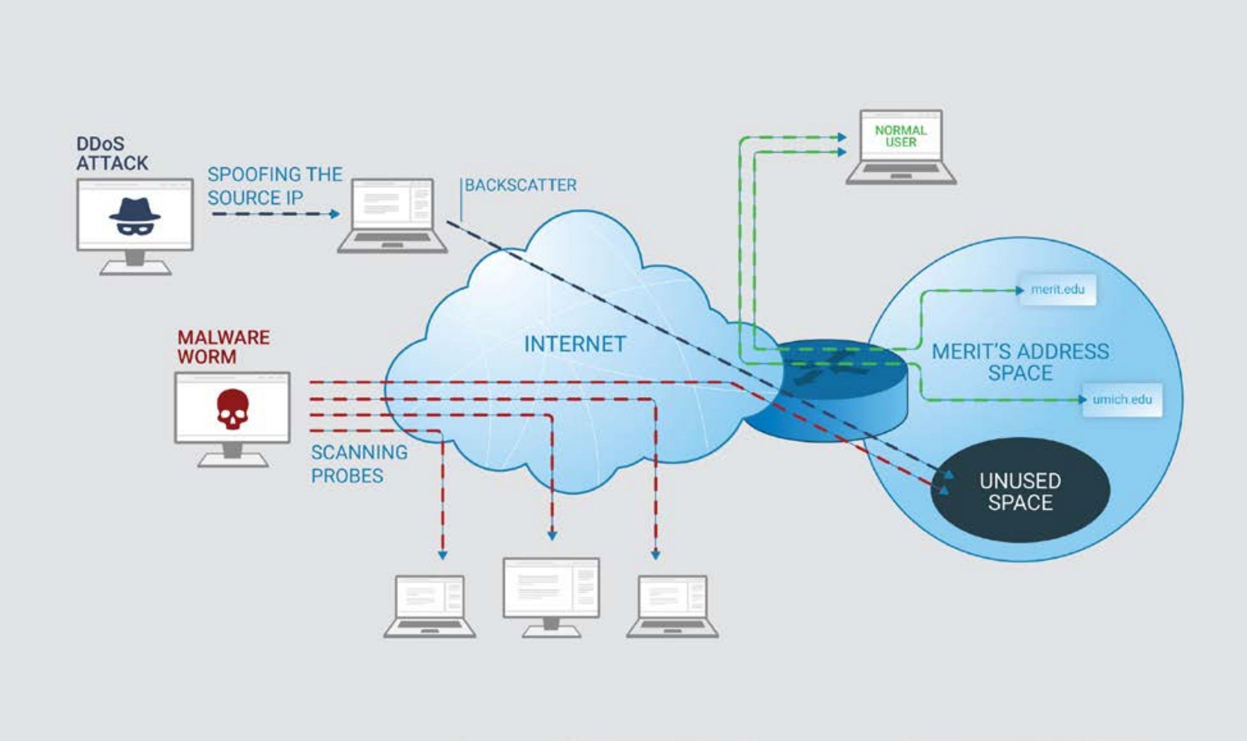

Figure 3. Network telescopes receive real-world Internet data destined to an unused address space

The darknet datasets have supported IMPACT customer needs such as: a) network event identification and monitoring: a prime example of this is the near-real-time identification of Mirai-infected IoT devices [3]. Further, darknet data can be used to detect outages due to natural disasters (eg., see the IODA tool by CAIDA that utilizes our Darknet data [3]), spoofing-based DDoS attacks, and several types of scanning. The ability to view data that capture Internet-wide scanning activities provides evidence-based insights into ongoing adversarial threats and hazards; b) time-series analyses to understand evolving trends: one can use darknet summaries of longitudinal data to understand global trends in malware propagation. For example, malware activities due to scanning on port 445 (attributed to Conficker and Sasser worms) were prevalent in previous years; however, nowadays, scanning activities to port 23 top the list; these are activities associated with the Mirai outbreak and IoT scanning; c) assess network reputation, cyber-risk and network hygiene: darknet data can be utilized as cyber-risk indicators or assessment tools for the reputation of networks [4]. E.g., ASNs that are consistently seen to be originators of malicious spoofing activities (such as bulletproof ASNs) can be associated with poor network hygiene.

Access to Darknet data were provided via our “secure enclaves”. Even though this is a dataset of minimal risk (traffic destined to a Darknet is inherently malicious and is directed to an unused, ungoverned space), we wanted to make sure 1) that researchers will not transfer any data outside our organization, and 2) that researchers would not probe or communicate with any hosts identified by our data. This technical provisions, along with IMPACT’s policy framework, allowed us to ethically share sensitive data while not deteriorating their scientific utility (i.e., Darknet data were not anonymized).

3.3 BGP Routing Data

The BGP protocol is a distributed protocol, characterized as the “glue of the Internet”, whose secure operation is critical for the Internet’s smooth and efficient operation. However, the lack of security considerations when BGP was originally designed, makes it still vulnerable to a large class of attacks known as *BGP prefix hijacking* attacks. In this project, we have been collecting longitudinal, global routing data, by peering (using a software router such as Quagga) with actual Merit core routers. This allowed us to collect global routing tables (dumps) and routing updates, collected from Merit’s vantage point. We have managed to detect a few BGP hijacking incidents involving Merit’s routing assets (i.e., prefixes) and made these datasets available to IMPACT researchers.

In addition, our BGP datasets have been continuously utilized for the past several years, in an operational manner, by University of Michigan’s Censys project (censys.io) [5].

3.4 Secure Enclaves

In order to enable sharing of sensitive data (such as our Darknet data) without sacrificing scientific utility (i.e., without applying brute anonymization schemes), we have designed, implemented and operated a “Secure enclaves” framework (see **Figure 4**) built on-top of VMWare’s virtualization platform. We have been offering our Darknet data, our scanners data and our Mirai/IoT data via the enclaves.

The secure enclaves system is a set of virtual machines (VMs) hosted at Merit's cloud computing infrastructure. These VMs have *read-only network* access (i.e., via an NFS network filesystem) to data being shared (such as our Mirai datasets). These allows researchers to *locally* process the data without having to transfer multiple terabytes of data to their organizations. Most importantly, though, the secure enclaves *would not allow* researchers to copy any data *outside* of the VM nor they would allow any *inbound* traffic to it; this "code-to-data" approach ensures that sensitive data are processed only in "safe places" (in our case, only within Merit) and are never copied without permission outside the data host. Researchers can access the system via a browser plugin by VMWare that allows secure communication (TLS) with the enclave VM. No data can be transferred through that channel, besides interacting with the VM via a remote terminal.

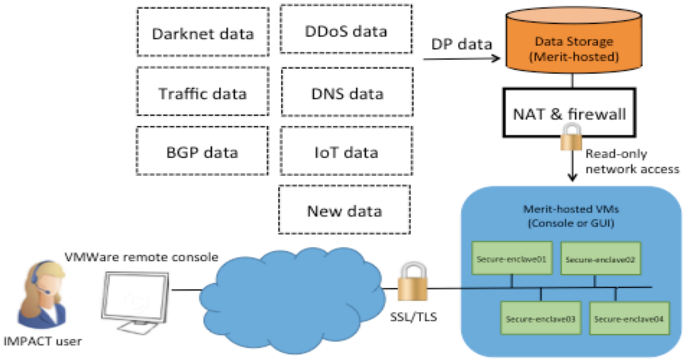

Figure 4. The secure enclaves data sharing system, hosted at Merit's cloud infrastructure.

4. Results and Discussion

4.1 Publications

Support from IMPACT and data collected from our team have enabled the following research papers:

1. **Understanding the Mirai Botnet**

Manos Antonakakis, Tim April, Michael Bailey, Matt Bernhard, Elie Bursztein, Jaime Cochran, Zakir Durumeric, J. Alex Halderman, Luca Invernizzi, Michalis Kallitsis, Deepak Kumar, Chaz Lever, Zane Ma, Joshua Mason, Damian Menscher, Chad Seaman, Nick Sullivan, Kurt Thomas, Yi Zhou. 26th USENIX Security Symposium, 2017

2. **Taming the 800 pound gorilla: The rise and decline of NTP DDoS attacks**

J Czyz, M Kallitsis, M Gharaibeh, C Papadopoulos, M Bailey, M Karir. Proceedings of the 2014 Conference on Internet Measurement Conference, 435-448

3. **Estimating internet address space usage through passive measurements**

A Dainotti, K Benson, A King, , KC Claffy, M Kallitsis, E Glatz, X Dimitropoulos. ACM SIGCOMM Computer Communication Review 44 (1), 42-49

4. **Understanding IPv6 internet background radiation**

J Czyz, K Lady, SG Miller, M Bailey, M Kallitsis, M Karir. Proceedings of the 2013 conference on Internet measurement conference, 105-118

5. **AMON: An Open Source Architecture for Online Monitoring, Statistical Analysis, and Forensics of Multi-Gigabit Streams**

M Kallitsis, SA Stoev, S Bhattacharya, G Michailidis. IEEE Journal on Selected Areas in Communications 34 (6), 1834-1848, July 2016

6. **Leveraging internet background radiation for opportunistic network analysis**

K Benson, A Dainotti, AC Snoeren, M Kallitsis, KC Claffy. Proceedings of the 2015 ACM Conference on Internet Measurement Conference

7. **Trimming the Hill estimator: robustness, optimality and adaptivity**

S Bhattacharya, M Kallitsis, S Stoev (submitted to Journal of Extremes). arXiv preprint arXiv:1705.03088

8. **An Internet-Wide View of Internet-Wide Scanning**

Z Durumeric, M Bailey, JA Halderman, 23rd USENIX Security Symposium (SEC'14)

9. **On the Mismanagement and Maliciousness of Networks**

J Zhang, Z Durumeric, M Bailey, M Liu, M Karir. 21st Network & Distributed System Security Symposium (NDSS'14)

10. **An Internet-Wide View of ICS Devices**

A Mirian, Z Ma, M. Bailey, M Tischer, T Chuenchujit, T Yardley, R Berthier, et al. 14th IEEE Privacy, Security and Trust Conference (PST'16)

11. **Cloudy with a Chance of Breach: Forecasting Cyber Security Incidents**

Y Liu, A Sarabi, J Zhang, P Naghizadeh, M Karir, M Bailey, M Liu. USENIX Security, 1009-1024

12. **On The Power and Limitations of Detecting Network Filtering via Passive Observation**

M Sargent, J Czyz, M Allman, M Bailey. International Conference on Passive and Active Network Measurement, 165-178

13. **Characterization of blacklists and tainted network traffic**

J Zhang, A Chivukula, M Bailey, M Karir, M Liu. International Conference on Passive and Active Network Measurement, 218-228

Merit’s network telescope data (and, in general, Darknet data) has played a critical role in supporting several of the abovementioned works. This unique data source offered the opportunity to understand macroscopic internet trends and/or obtain evidence-based analyses of high-impact events (e.g., see Understanding the Mirai Botnet, Taming the 800 pound gorilla: The rise and decline of NTP DDoS attacks, Understanding IPv6 internet background radiation, An Internet-Wide View of Internet-Wide Scanning, An Internet-Wide View of ICS Devices, Leveraging internet background radiation for opportunistic network analysis, On The Power and Limitations of Detecting Network Filtering via Passive Observation), and supported efforts for network reputation and risk assessment (see On the Mismanagement and Maliciousness of Networks, Cloudy with a Chance of Breach: Forecasting Cyber Security Incidents, Characterization of blacklists and tainted network traffic).

The infrastructure supported by IMPACT for monitoring **Merit’s ingress/egress traffic** has not only assisted in the preparation of academic publications that employed such real-world data (see Taming the 800 pound gorilla: The rise and decline of NTP DDoS attacks, Characterization of blacklists and tainted network traffic, and Trimming the Hill estimator: robustness, optimality and adaptivity) and in the curation of high-value operational data to be used by the IMPACT community (e.g., DDoS datasets), but has also supported Merit Network itself in its mission to protect/secure its network and customers. Flow data collected with support from IMPACT have been utilized *internally* at Merit for forensics analysis and identification of culprits when network attacks would occur. Further, darknet datasets have also been utilized to identify vulnerable hosts.

Furthermore, access to **multi-10 Gbps streaming data** has allowed Merit and our collaborators at University of Michigan to design and engineer next-generation monitoring systems, such as AMON (All-packet MONitor: An Open Source Architecture for Online Monitoring, Statistical Analysis, and Forensics of Multi-Gigabit Streams) [6]. AMON constructs, in an efficient manner, network summaries that can be used as proxies for online network anomaly detection, real-time visualizations/dashboards and searching for network heavy-hitters. Contrary to most commercial solutions that process Netflow data, AMON receives as input raw packet data; this means that high-impact network events can be detected at *sub-second time scales, rather than minutes*. It is built with broadly accessible hardware and is currently operational and evaluated at Merit. Via AMON we had the opportunity to detect and capture several interesting DDoS events that we have shared with the IMPACT community.

4.2 Data Dissemination

We have distributed data to more than 50 unique organizations and more than 60 unique researchers. We have made available more than 50 unique datasets. We have served data for a variety of “research purposes”: most researchers were interested in using our flow-based attack datasets (e.g., our DDoS traffic flow data) for experimentation and evaluation of new intrusion detection (e.g., based on advanced machine learning techniques) and mitigation systems. We have also received requests for analyzing botnet behavior (e.g., our Mirai data), requests for identifying scanners, requests for data to be used for educational purposes, requests for BGP data to identify outages and others.

In addition, our data has supported the following "Data Analytics as a Service" systems:

1. University of Michigan's Censys search engine [5] (censys.io): Censys integrates scanning/probing data from ZMap (e.g., Internet-wide TLS certificates) into a database that can be efficiently queried. Merit's BGP data provides information about the AS PATH between University of Michigan (the scan origin) and the scanned host.
2. CAIDA's Internet Outage Detection and Analysis (IODA, ioda.caida.org): CAIDA's IODA system leverages several distinct sources (BGP data, active probing and Darknet data) to provide a Web-based dashboard and analysis engine for Internet outages. Merit's Darknet data is one of the two large Darknets used in this service.

The list of organizations that have requested data from our team includes (note that this list might be incomplete; it only includes organizations requesting data after IMPACT has launched):

- *Academia*: CAIDA, umich.edu, mit.edu, illinois.edu, wustl.edu, isi.edu, purdue.edu, utoledo.edu, rutgers.edu, princeton.edu, ucla.edu, wsu.edu, morgan.edu, pitt.edu, cuny.edu, cmu.edu, masonlive.gmu.edu, iastate.edu, rpi.edu, u.nus.edu, umass.edu, my.fit.edu
- *Government*: lanl.gov, mail.mil, ornl.gov
- *Industry*: galois.com, area1security.com, crw-innovations.com, i-a-i.com, unisys.com, srcinc.com, appcomsci.com, ranksoftwareinc.com, parsons.com, ampcus.com, digitalharmonic.com, ltsnet.net, modusoperandi.com, bah.com, microsoft.com, vencorelabs.com
- *International*: cs.dal.ca, dsto.defence.gov.au, uq.edu.au, ens.etsmtl.ca, napier.ac.uk

4.3 Professional Preparation

Data provided by our team has supported professional preparation of the following Computer Science graduate students, conducting research in the areas of networking and cyber-security: Jake Czyz (University of Michigan, graduated), Jing Zhang (University of Michigan, graduated), Zakir Durumeric (University of Michigan, graduated), Karyn Benson (University of California, San Diego, graduated), Zane Ma (University of Illinois, ongoing PhD student).

5. Conclusions

Access to real-world networking and security data is a necessary condition for research and development in cyber-security, for understanding Internet trends, for forensics analysis and for defending complex cyber-physical systems, such as the Internet. Research infrastructure for performing longitudinal measurements is costly and labor intensive, but is fundamental to the Internet's secure operation: security flaws, misconfigurations, deviations from "best-practices", infected devices and vulnerabilities cannot be identified without security measurements and real data.

In the project, our team strived to provide high-quality data, collected at Merit Network, to the research and networking community. We had to balance between 1) providing high-value raw networking data without any filtering, and 2) sacrificing the scientific utility of the data and enforce crude anonymization schemes to protect user privacy. Finding the "golden section" between the two is no-easy task, but we believe that our chosen data curation process allowed us to provide quality data to the IMPACT community without endangering privacy and/or the safety of Merit's network and its customer base. Towards that goal, IMPACT's legal and policy framework was a catalyst in our aim for ethical data sharing.

Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of Air Force Research Laboratory and Department of Homeland Security or the U.S. Government.

6. References

[1] J Czyz, M Kallitsis, M Gharaibeh, C Papadopoulos, M Bailey, M Karir. Taming the 800 pound gorilla: The rise and decline of NTP DDoS attacks, Proceedings of the 2014 Conference on Internet Measurement Conference, 435-448

[2] Karyn Benson et al. Leveraging internet background radiation for opportunistic network analysis. IMC, 2015.

[3] Manos Antonakakis, Tim April, Michael Bailey, Matt Bernhard, Elie Bursztein, Jaime Cochran, Zakir Durumeric, J. Alex Halderman, Luca Invernizzi, Michalis Kallitsis, Deepak Kumar, Chaz Lever, Zane Ma, Joshua Mason, Damian Menscher, Chad Seaman, Nick Sullivan, Kurt Thomas, Yi Zhou. Understanding the Mirai Botnet, 26th USENIX Security Symposium, 2017

[3] Internet Outage Detection and Analysis (IODA). CAIDA. https://ioda.caida.org, 2017

[4] Jing Zhang, Zakir Durumeric, Michael Bailey, Mingyan Liu, and Manish Karir, *On the mismanagement and maliciousness of networks*. NDSS, 2013.

[5] Zakir Durumeric, David Adrian, Ariana Mirian, Michael Bailey, J. Alex Halderman, *A Search Engine Backed by Internet-Wide Scanning*. 22nd ACM Conference on Computer and Communications Security (CCS'15)

[6] M Kallitsis, SA Stoev, S Bhattacharya, G Michailidis, *AMON: An Open Source Architecture for Online Monitoring, Statistical Analysis, and Forensics of Multi-Gigabit Streams***,** IEEE Journal on Selected Areas in Communications 34 (6), 1834-1848, July 2016.

List of Symbols, Abbreviations and Acronyms

AMON – All-packet Monitor
APT – Advanced Persistent Threats
ASN – Autonomous System Number
BGP – Border Gateway Protocol
CAIDA – Center for Applied Internet Data Analysis
CHARGEN – Character Generation protocol
DDoS – Distributed denial of service
DETER – Defense Technology Experimental Research
DHS – Department of Homeland Security
DNS – Domain Name System
GB – Gigabyte
ICS – Industrial Control Systems
IMPACT – Information Marketplace for Policy and Analysis of Cyber-risk & Trust
IP – Internet Protocol
IODA – Internet Outage Detection and Analysis
IoT – Internet of Things
IRB – Institutional Review Board
MOA – Memorandum of Agreement
NFS – Network File System
NTP – Network Time Protocol
PI – Principle Investigator
RADb – Routing Assets Database
R&E – Research and Education
SSDP – Simple Service Discovery Protocol
SSL – Secure Socket Layer
SYN – Synchronized
TB – Terabyte
TLS – Transport Layer Security
VM – Virtual Machine