The 10th Workshop on Active Internet Measurements (AIMS-10) Report

kc claffy
UCSD/CAIDA
kc@caida.org

David Clark
MIT/CSAIL
ddc@csail.mit.edu

This article is an editorial note submitted to CCR. It has NOT been peer reviewed.
The authors take full responsibility for this article's technical content. Comments can be posted through CCR Online.

ABSTRACT

On 13-15 March 2018, CAIDA hosted its tenth Workshop on Active Internet Measurements (AIMS-10). This workshop series provides a forum for stakeholders in Internet active measurement projects to communicate their interests and concerns, and explore cooperative approaches to maximizing the collective benefit of deployed infrastructure and gathered data. An overarching theme this year was how to inform new legislation of communications policy in the U.S. Given the continued limited insight into Internet operations by researchers and policymakers, we tried to focus these discussions on what data is or could be measured to shape and support current and emerging policy debates. Materials related to the workshop are at http://www.caida.org/workshops/aims/1803/.

CCS CONCEPTS

• **Networks → Network measurement**; **Public Internet**; **Network dynamics**;

KEYWORDS

active Internet measurement, validation, policy

1 INTRODUCTION

On March 13-15, 2018 CAIDA hosted the tenth Workshop on Active Internet Measurements (AIMS-10) chaired by kc claffy (UC San Diego) and David Clark (MIT). Discussion topics this year included: how to make measurement results more accessible and informative to policy analysts and policymakers; existing and proposed active measurement platforms, architectures, methods, and tools; observations with policy implications; and novel measurement data to support innovative analyses.

We review highlights of discussions and consider insights that transpired from the workshop. This report does not cover each topic discussed; for more details examine workshop presentations linked from the workshop web page: http://www.caida.org/workshops/aims/1803/.

2 PLATFORMS TO SUPPORT INTERNET SCIENCE AND POLICY

A long-standing goal of the AIMS workshops series is to foster interaction in a community seeking to design and build complementary and cooperative measurement systems, including correlating data collected by disparate systems and groups. Scientific research communities are broadly benefiting from the exploding interest and investment into open source web development platforms, databases for storing and retrieving unstructured data, and availability of computing resources. In the case of Internet science, these evolving platforms enable rapid development of easy-to-use graphical systems that facilitate discovery, analysis, and navigation of data describing the Internet infrastructure. At the same time, there is growing support in the NSF-funded research community for building *science gateways* to expand access to high performance computations, analysis, data, and storage resources.

2.1 Integrated Platform for Applied Network Data Analysis (PANDA)

CAIDA is eager to apply these developments, undertaking a new large infrastructure effort to design and build a system that will allow scientists, operators, and eventually policy makers to ask high-level cross disciplinary questions about Internet structure, behavior, and performance. This recently funded NSF project, "DIBBs: Integrated Platform for Applied Network Data Analysis (PANDA)" [16], will integrate (in some cases first re-architect for higher performance) several data building blocks that CAIDA researchers have developed over the last twenty years: the Archipelago (Ark) Active Internet Measurement Platform and its supporting components and derivative data; ASRank, which computes and compares routing and economic relationships among ISPs; BGPStream, an efficient framework for routing (BGP) data analysis; MANIC: Mapping and Analysis of Interdomain Congestion; and Spoofer, assessment of IP source address validation best practices. The resulting

unified platform will support collaboration with multiple disciplines and increase community accessibility to the underlying components. The project targets research questions from network mapping interdomain Internet topology, security, economics and policy. Questions we are prioritizing include interconnection topology, path performance characteristics, and source address validation. During the workshop we discussed the current state of development PANDA components in detail: AS Rank, Archipelago (Ark) Internet measurement platform, and MANIC.

AS Rank. As the first of many examples of services that PANDA will integrate and offer public access via a web API, Bradley Huffaker (CAIDA) presented an update on rearchitecting of CAIDA's AS Rank service [4]. The new interactive service will offer researchers more user-friendly tools for collecting, analyzing, querying, and interpreting data on the Internet AS-level topology. We seek to optimize query efficiency to serve a larger user population and to provide the data to researchers in a useful format (JSON) via a new programmatic RESTful API.

BGPStream. Alistair King (CAIDA) gave an update on another PANDA component – BGPStream V2 [6] (in private beta testing at time of writing). BGPStream is an open-source software framework for live and historical BGP data analysis, supporting scientific research, operational monitoring, and post-event analysis. The new version includes native support for true real-time data access via BMP, local caching of dump files, high-level PyBGPStream API (prototype), a new filter interface, and miscellaneous performance improvements.

2.2 New measurement platforms

African BGP Route-Collector Platform. Roderick Fanou (CAIDA) et al. [28] worked with 42 IXPs in 30+ African countries (Af-IX) where 23 IXPs host 41 PCH Route collectors, only three have Route Views collectors, and less than 50% of IXPs provide publicly available data. Roderick contributed to the development of an open source web platform, ARDA [28], to assist researchers in the data collection and analysis of African route collectors. ARDA implements a common structure to store data, defines and periodically computes statistics with IXP, national, and regional views. It offers views for tracking IXP growth, interconnection development progress and gaps, and a locus for technical support and reporting. The source code been released and is publicly available on Github since March 2018 [27].

PacketLab. Kirill Levchenko (UCSD) described his idea for PacketLab, a universal measurement endpoint interface that could lower barriers faced by experimenters and measurement endpoint operators. Most research groups cannot afford to design, deploy, and maintain their own network of measurement endpoints, and thus rely on measurement infrastructure shared by others. Unfortunately, issues of compatibility, trust, and a lack of incentives hinder efficiency of such sharing. PacketLab is driven by two objectives: moving the measurement logic out of the endpoint to a separate experiment control server, thus making each endpoint a lightweight packet source/sink; and delegating access to measurement endpoints while retaining fine-grained control over how one's endpoints are used by others, thus allowing research groups to share measurement infrastructure with each other with little overhead. If successful, PacketLab could accommodate the research community's demand for future global-scale Internet measurement. This project received funding from NSF in late 2018.

Timing Verification as a Service. Darryl Veitch (U. of Technology Sydney) described the differences among various time "data types" and presented a model for how a timing verification service could work, the kind of services it could potentially provide and how, and their limitations. This includes separating the problems at the client side from those at the time-server present in the network, on which the client timekeeping generally depends. An example service is the following: an application could issue a "server audit request", which would provide continuous external monitoring of a time server(s) used throughout a measurement campaign. This monitoring data could be used subsequently for application timestamp quality validation, for troubleshooting, or simply stored to act as an authoritative timing audit trail if required. Darryl also summarized recent work on automated detection of time server anomalies that underpin many of the proposed services.

Fling: middlebox measurement. Middleboxes have been known to change packets in many ways, making it hard to design and deploy protocol extensions. Addressing the need to know what such middleboxes do, Ahmed Elmokashfi (Simula) presented a client/server tool called fling ("flexible ping") [2] that can detect which device along the path changed or dropped a packet. Example use cases include using fling to verify whether DSCP code points can be used by WebRTC for signaling its QoS expectations. fling also measures whether various transport protocols and their options are usable in an end-to-end fashion. These results are based on measurements from geographically distributed Ark, PlanetLab, and NorNet probes.

Auditing net neutrality violations globally. David Choffnes (Northeastern U.) and his team developed tools to detect DPI-based traffic differentiation, expose problematic ISP policies, and even to deploy countermeasures to work around them. In December 2017, David's team signed a contract with ARCEP (France's telecom regulator) to apply

their tool as a product that subscribers of French ISPs can use to audit whether they violate net neutrality laws in France. David talked about deployment and lessons learned along the way. He also encouraged the community to incorporate the Reverse Traceroute tool into their path measurement toolbox [13].

3 MEASURING SECURITY PROPERTIES OF NETWORKS

Marinho Barcellos (Federal U. of Rio Grande do Sul) discussed efforts to increase the visibility of networks lacking source address validation (SAV) with a new methodology and corresponding tools detecting spoofed traffic in network traces. The goals of the project are threefold: (i) investigate the prevalence, causes, and impact of IP source spoofing on the Brazilian IXP substrate; (ii) create a tool that enables IXPs to run compliance tests on source address validation; and (iii) measure trends in remediation after deployment of the tool. The project uses continuous flow data and other relevant information from large IXPs in Brazil.

Alberto Dainotti (CAIDA) described ARTEMIS, a new defense approach for an AS to detect and stop the hijacking of its own prefix, leveraging the pervasiveness of publicly available BGP monitoring services and their recent shift towards real-time streaming. Compared to previous work, this approach combines characteristics desirable to network operators such as comprehensiveness, accuracy, speed, privacy, and flexibility. Real-world experiments show that, with the ARTEMIS approach, prefix hijacking can be neutralized within a minute [29].

Mattijs Jonker (CAIDA/ U. Twente) described early results from an analysis of 1100 days of data (March 2015 - March 2018) that leverages his work to infer and characterize the practices and efficacy of blackholing events [14]. He used data from the UCSD Network Telescope amplification honeypots, and large-scale active DNS measurements. This work will appear at IMC2018.

4 DEVELOPMENTS IN ROUTER-LEVEL CARTOGRAPHY TECHNIQUES

Researchers exploring Internet topology data face the task of *IP address alias resolution*, which identifies IP interfaces belonging to the same router.

Alias resolution "on demand". Young Hyun (CAIDA) presented on-demand topology measurement service Vela [5]. Vela can feed millions of IPv4 addresses as input to the MIDAR [17, 18]. CAIDA alias resolution software system which executes on the Ark measurement platform. It infers which IPv4 addresses belong to the same routers, producing router-level and pop-level topologies for better understanding of AS peering arrangements and redundancy and resilience of ASes. MIDAR results can also help to identify traceroute path anomalies [19], decipher Multipath Detection Algorithm (MDA) traceroute results, and identify border links between ASes [20, 23]. Vela also provides *aliasq*, a web API for querying a database of MIDAR aliases harvested from CAIDA's Internet Topology Data Kits [7].

Mapping AS borders in the Internet. Alex Marder (U. of Pennsylvania) described early results from research into how to infer router operators and interdomain links from traceroute data. The bdrmap-IT technique synthesizes methods described in two IMC '16 papers [21, 24] into a single approach for identifying interdomain borders. Working with existing archives of traceroute data (for all ASes seen in the data), the algorithm creates a hybrid router-interface graph, establishes last-hop router owners, and uses a graph refinement loop to infer the AS ownership of routers and determine interdomain links. This work will appear in SIGCOMM IMC 2018 [23].

Capturing Layer 2 topology. Yves Vanaubel (U. Liege) presented recent research results improving the inferential power of basic traceroute by labeling IP addresses to reveal Multi-protocol Label Switching (MPLS) tunnels. (These tunnels route packets via path labels instead of network addresses.) He also described several tools and methods now implemented in scamper [22] for discovering more information about middleboxes, layer-2 devices, routers, subnets, improved alias resolution, and network fingerprinting. His implementation has enabled a longitudinal study of middleboxes in the wild [12].

Leveraging DNS naming conventions. Matthew Luckie (U. Waikato) presented his early work on examining router interface naming conventions to infer whether a set of hostnames belongs to the same router. The methodology seeks patterns in hostnames that could uniquely identify such a router in a domain. Using CAIDA's ITDKs [7] as input and testing data, Matthew developed an algorithm that automatically infers these patterns (regular expressions); he is exploring the use of this algorithm to find routers matching an organization's naming convention. Integration of this technique with MIDAR could potentially improve the coverage of alias resolution. Even more important, it could make a big difference for IPv6 alias resolution, because the best current technique for IPv6 alias resolution relies on routers sending fragments – which most networks do not.

One question emerged from this work: is there any approach to derive from this data a set of best practices for operators to use for router naming? For example, some operators (e.g., IIJ) use router naming conventions that may help alias resolution, yet provide no hints on geolocation or

link speeds. So far, the data appears to suggest no convergence toward any de facto standard practice. It is possible that router naming practices of many operators come from "grabbing a script at a NANOG meeting" and using it. Thus the research community might have an opportunity to provide a script that would establish common practice on router naming helping both operators and researchers. However, clearly some providers may not use naming conventions on purpose as it might reveal inefficiencies or odd paths in routing that operators would rather not make obvious. Yet the point was also made that good router naming practice creates a more "debuggable" network, and may reduce calls to the NOC. Matthew also posited that IPv6 naming may more likely come from automated provisioning systems because nobody wants to type out an IPv6 PTR record.

5 PATH MEASUREMENT: STANDARDIZING TOOLS AND DATA FORMATS

We had an extensive discussion about possible strategies to achieve more consistency across the three large-scale traceroute measurement infrastructure platforms serving academic researchers: RIPE Atlas, CAIDA's Ark, and Google's Measurement Lab.

Multipath discovery issues. One significant open issue is the interest some researchers have in expanding use of MDA traceroute. Kevin Vermeulen and Timur Friedman (Sorbonne U.) presented the challenges in deploying resource-intensive multipath traceroute probing on resource-limited RIPE Atlas probes. The Multipath Detection Algorithm (MDA) uses the Paris traceroute algorithm [1] to discover all paths between the source and the destination. The algorithm represents a tradeoff between improved efficiency (number of probes) and statistical guarantees of discovering complete topology. MDA has enabled discovery of odd cases of load balancing and evidence of asymmetric topologies that may cause performance asymmetries on parallel paths. Kevin and Timur ran a preliminary experiment using 10 PlanetLab nodes as sources and the IP Hitlist from USC/ISI as destinations, and developed heuristics to optimize probing for various scenarios of path asymmetry and load-balancing. They published their results on MDA-lite in the proceedings of ACM IMC [15]. RIPE NCC has deployed an MDA traceroutes implementation on RIPE Atlas nodes, and hopes to further contribute to developing and deploying smarter and more efficient MDA measurement techniques for future deployment on RIPE Atlas.

Stephen Strowes (RIPE NCC) is also investigating the data collected by RIPE Atlas traceroute measurements for conducting multipath discovery. The RIPE Atlas traceroute implementation (as of March 2018) modifies flow IDs on each iteration of ongoing measurements: by default a cycle consists of 16 iterations, one every 15 minutes, suggesting that once every four hours the Atlas platform completes an ersatz MDA traceroute measurement. RIPE NCC plans to analyze the data gathered by this implementation to understand the path variability made visible by Atlas, and compare it to that obtained by running full MDA traceroutes performed by researchers at Sorbonne University. Stephen's presentation inspired discussion on how the community might want to modify or standardize traceroute implementations, including algorithms and output formats.

Measurement Lab usage. Peter Boothe (Google) updated participants on Measurement Lab, a partnership of Google, Open Technology Institute, and Planet Lab with a mission to measure the Internet, save the data, and make it universally accessible and useful. open data and open science. Peter and Ya Chang (Google) proposed evolving the schema for Paris traceroute data, so that users will be able to reconstruct paths in BigQuery. Ya observed that the community does not seem to have a standard way to represent (IP traceroute) path data in SQL databases. She described some interesting use cases and proposed a few schema options to start the discussion.

6 PERFORMANCE AND AVAILABILITY

Detecting outages. Ramakrishna Padmanabhan (U. Maryland) presented a study on the role that weather plays in causing residential links to fail. Using a year-long dataset comprised of over 4 billion pings to 3.6 million IP addresses throughout the United States from before, during, and after periods of severe weather forecast by the National Weather Service, they introduced new techniques to (i) measure how weather correlates with failures across different geographic areas, link types, and providers; (ii) detect correlated failures that are caused, for instance, by power outages and network outages.

John Heidemann (USC/ISI) described a new clustering algorithm that scales to large datasets (millions of blocks by thousands of observations) to identify clusters of addresses that respond similarly. He successfully applied this technique to service outages and anycast catchment changes during DDoS and believes clustering is one step toward finding the "forest" in our trees and leaves of observations.

Geolocation. The community is interested in integrated geolocation information to enable regional analysis, mapping, and visualization. On Day 2, James Miller (FCC) described how the FCC has used some informal triangulation techniques to try to infer geolocation of MBA data, along with some hard-coded ground truth for known nodes such as Measurement Lab servers.

On Day 3, Emile Aben (RIPE NCC) (on behalf of Jasper den Hertog (RIPE NCC)) described challenges to geolocation of devices on the Internet and offered examples of sketches of regional Peer-to-Peer network fabrics in Denmark, South Korea, Ireland and the US, derived from RIPE Atlas measurement data, with the networks being classified by type (e.g., end-user, end-user/transit, local IXP, transit/external, network with less data). This work used RIPE Atlas data to infer the geolocation of Internet resources and make results queryable via a web API OpenIPMap [26]. The continued development of OpenIPMap aims to narrow a gap in knowledge of infrastructure topology geolocation by crowd-sourcing. Future work includes integration of RIR data, reverse DNS inferences, and automated IP address discovery for IPv6.

Measuring censorship. Roya Ensafi (U. Michigan) presented a recent collaboration intended to safeguard users from adversarial network interference by building tools to measure, understand, and defend against it. Censored Planet [25] uses novel measurement techniques that remotely detect instances of interference on a broad cross-section of the Internet. Compared to previous approaches, which relied on having volunteers in censored regions deploy special hardware or software, this approach yields significantly better coverage, lower costs, and reduced risk exposure for probing nodes. The system enables continuous monitoring of the deployment of network interference technologies, tracking policy changes in censoring nations, and better understanding of the targets of interference. (This work subsequently appeared at USENIX Security 2018 [30].)

Measuring bottlenecks. With a few hundred million users, Mozilla's Firefox browser offers an almost unprecedented opportunity to gather information about Internet connection quality. Saptarshi Guha (Mozilla) described a pilot experiment that aims to publish a continuously collected data set of "Internet connection quality" as experienced by Firefox clients across the world. Mozilla plans to provide open de-identified data sets to the research community, policy makers and the general public with transparent methodology for a set of performance measures around Internet quality, a public-facing report of globally distributed, representative longitudinal data for consumers and researchers (e.g. "network.metrics.mozilla.com" to mirror the hardware report at "hardware.metrics.mozilla.com"), and to provide a source of more relatable Internet quality measures for developers/marketers (e.g. Firefox updates are 3x slower in small towns in India).

Measuring interdomain congestion. There is significant interest in the technical and policy communities regarding the extent, scope, and consumer harm of persistent interdomain congestion. CAIDA hopes to provide empirical grounding for discussions of interdomain congestion by developing a system and method to measure congestion on thousands of interdomain links without direct access to them. Amogh Dhamdhere (CAIDA) described the Measurement and Analysis of Interdomain Congestion (MANIC) system, which uses the Time Series Latency Probes (TSLP) technique to identify links with evidence of recurring congestion suggestive of bandwidth under-provisioning. MANIC is currently deployed at 86 vantage points worldwide and has shown that congestion inferred using the lightweight TSLP method correlates with other metrics of interconnection performance impairment. CAIDA researchers and collaborators used this method to study interdomain links of eight large U.S. broadband access providers from March 2016 to December 2017, validating their inferences against ground-truth traffic statistics from two of the providers. They did not find evidence of widespread endemic congestion on interdomain links between access ISPs and directly connected transit and content providers, although some such links exhibited recurring congestion patterns. This work was published at ACM SIGCOMM 2018 [10] where it received the Best Paper Award.

Alex Marder (U. of Pennsylvania) presented his early study of relationships between parameters measured from the Internet edge, such as RTTs and loss, and severity of congestion. Using a small dedicated network, he ran controlled experiments with two primary goals: (i) classifying congestion based on the bandwidth available to each TCP flow, and (ii) exploring the possibility of identifying and classifying congestion for links subsequent to a congested link. He hopes to determine the average per flow throughput of TCP flows on persistently congested links, and ultimately find evidence of service degradation due to DDoS attacks.

Esteban Carisimo (U. de Buenos Aires) described early work using the Stable distribution for modeling distributions of latency measurements from the edge of the network. The study makes use of high-frequency ping data collected on the Ark platform [3]. Early results show that the Stable distribution is surprisingly effective at capturing characteristics of these latency distributions over 10-minute samples. Even more surprising, parametric fitting of the Stable distribution to such samples yields parameter values that effectively differentiate congested from uncongested links. Despite the complexity inherent to the Stable distribution and the resulting trade-offs, its ability to accommodate extreme values makes it a promising approach to analyzing Internet latency. (Esteban subsequently submitted this work to a conference based on feedback from workshop participants.)

Measuring video quality. Paul Schmitt (Princeton U.) presented a novel lightweight system running on a home gateway that integrates active and passive measurements to correlate video QoE with congestion events. The prototype system deployed on 20 hosts (7 in France and 13 in U.S.) separates video flows from other sources of traffic by mapping DNS requests to subsequent TCP flows, which allows the identification of active services, including those that utilize HTTPS encryption. It then tracks traffic patterns of video streams, to infer metrics such as average bitrate and re-buffering events. The system also uses lightweight probing, e.g., pings and traceroutes, to infer potential root causes.

Ricky Mok (CAIDA) described his new project to build a platform and framework to crowd-source measurements for assessing Quality of Experience (QoE). With support from NTT, the Quality of User Internet Customer Experience (QUINCE) project will examine the feasibility of gamifying various kinds of Internet measurement and subjective assessments of video streaming performance by using a web-based platform. The project will test the sustainability of performing these measurement tasks in a crowd-sourcing context and will collect data to study the topology and performance of the Internet, the streaming performance of video services, and subjective assessments of the quality of experience (QoE) of video streaming.

7 PUBLIC SECTOR INTERNET MEASUREMENT PRIORITIES

James Deaton offered some commentary on high performance broadband access to schools paid for with public funds (e.g. E-Rate subsidies, local funds). James asks researchers to consider, and study for the benefit of public sector networks, "what constitutes high performance?", "what should be measured?", "is bandwidth a good proxy for performance?"

DHS PM Erin Kenneally described how the DHS S&T IMPACT Program [11] can help companies address risks associated with data sharing for academic research, by vetting and providing standardized researcher data use agreements with customized additional restrictions per provider. Companies leveraging the IMPACT framework may, at their discretion, retain custody and control over the research data at all times. They employ rigorous data agreements to limit access to and use of shared data. IMPACT helps address risks by vetting researchers, providers, and data.

Through the Measuring Broadband America Project, the U.S. FCC has created a publicly accessible, consumer privacy-protecting, national database on broadband performance and characteristics based on standardized metrics and data formats, prioritizing open methodologies, open data, open source, and collaboration. The FCC has commissioned development of infrastructures for measuring both fixed and mobile broadband, and offers opportunities for academic collaboration via shared datasets. After several years they have still not released any of the data or a report on the mobile measurements. James also described the FCC's Transparency Privacy Rule that became effective 11 June 2018 [9] as required by the Restoring Internet Freedom Order [8]. Disclosure requirements include: network management practices, e.g., blocking, throttling, affiliated prioritization, congestion management, app-specific behavior, device attachment rules, and security; performance (e.g., service description and impact of non-broadband access); and commercial terms (e.g, price, privacy policies, and redress options).

8 RESEARCH REPRODUCIBILITY

Alberto Dainotti (CAIDA) started a conversation on hyperpapers & open co-authoring. The talk discussed the history of scientific publication and its current state today. Many in the community feel problems of secrecy, lack of reproducibility, and under-utilized data. In an environment of publish or perish Alberto asked whether the current publication process is optimized for the good of science, and whether we have struck a reasonable balance between secrecy and openness. New technologies such as iPython, Jupyter notebooks, and services such as BinderHub offer opportunities for improvement of reproducibility of work.

ACKNOWLEDGMENTS. The workshop was funded by the Department of Homeland Security (DHS) Science and Technology Directorate, Cyber Security Division (DHS S&T/CSD) contract FA8750-12-2-0326 and the National Science Foundation awards CNS-1513283 and OAC-1724853. The work represents the position of the authors and not necessarily that of DHS or NSF.

9 WORKSHOP PARTICIPANTS

The main reason we continue this workshop is the enthusiastic participation it attracts from some of the brightest and most productive people in the community. We are grateful for their engagement and insights, many of which are reflected in this report.

Emile Aben (RIPE NCC) (remote participant), Marinho Barcellos (UFRGS), Hrishikesh Bhatt Acharya (Rochester Institute of Technology), Peter Boothe (Google/M-Lab), M. Abdullah Canbaz (University of Nevada, Reno), Esteban Carisimo (CAIDA/Universidad de Buenos Aires), Ya Chang (Google), David Choffnes (Northeastern University) (remote participant), David Clark (MIT), Ann Cox (DHS S&T Cyber Security Division), James Deaton (Great Plains Network), Ahmed Elmokashfi (SIMETRIC/Simula), Roya

Ensafi (University of Michigan), Timur Friedman (Sorbonne U.), Saptarshi Guha (Mozilla), John Heidemann (USC/ISI), Mattijs Jonker (University of Twente), Scott Jordan (UC Irvine), Ginga Kawaguti (NTT), Erin Kenneally (DHS), Kirill Levchenko (UC San Diego), Matthew Luckie (University of Waikato), Alexander Marder (University of Pennsylvania), James Miller (FCC / OET), Lucas Muller (CAIDA / UFRGS), Hideki Nojiri (NTT), Ramakrishna Padmanabhan (University of Maryland), Paul Schmitt (Princeton University), Stephen Strowes (RIPE NCC), Yves Vanaubel (University of Liege), Darryl Veitch (University of Technology Sydney), Kevin Vermeulen (Sorbonne University), and from CAIDA: kc claffy, Alberto Dainotti, Amogh Dhamdhere, Roderick Fanou, Marina Fomenkov, Alex Gamero-Garrido, Shuai Hao, Bradley Huffaker, Young Hyun, Alistair King, Ricky Mok, and Joshua Polterock.

REFERENCES

[1] Augustin, B., Cuvellier, X., Orgogozo, B. Viger, F., Friedman, T., Latapy, M., Magnien, C., and Teixeira, Renata. 2006. Avoiding Traceroute Anomalies with Paris Traceroute. In *ACM SIGCOMM IMC*.

[2] R. Barik, M. Welzl, A. M. Elmokashfi, S. Gjessing, and S. Islam. 2017. fling: A Flexible Ping for Middlebox Measurements. In *2017 29th International Teletraffic Congress (ITC 29)*. https://doi.org/10.23919/ITC.2017.8064349

[3] CAIDA. [n. d.]. Archipelago Measurement Infrastructure. http://www.caida.org/projects/ark/.

[4] CAIDA. [n. d.]. AS-rank. http://as-rank.caida.org.

[5] CAIDA. 2016. Vela: On-Demand Topology Measurement Service. https://www.caida.org/projects/ark/vela/.

[6] CAIDA. 2018. BGPStream V2 Private Beta. https://bgpstream.caida.org/v2-beta.

[7] CAIDA's Macroscopic Internet Topology Data Kit (ITDK). [n. d.]. http://www.caida.org/data/active/internet-topology-data-kit/.

[8] Federal Communications Commission. 2018. Restoring Internet Freedom Order Taking Effect. https://transition.fcc.gov/Daily_Releases/Daily_Business/2018/db0510/DOC-350643A1.pdf.

[9] Federal Communications Commission. 2018. Transparency Disclosures Portal. http://www.fcc.gov/disclosures.

[10] A. Dhamdhere, D. Clark, A. Gamero-Garrido, M. Luckie, R. Mok, G. Akiwate, K. Gogia, V. Bajpai, A. Snoeren, and k. claffy. 2018. Inferring Persistent Interdomain Congestion. In *ACM SIGCOMM*.

[11] DHS Science & Technology. 2016. IMPACT: Information Marketplace for Policy and Analysis of Cyber-risk and Trust. https://impactcybertrust.org/.

[12] Korian Edeline and Benoit Donnet. 2017. A First Look at the Prevalence and Persistence of Middleboxes in the Wild. In *29th International Teletraffic Congress, ITC*. https://doi.org/10.23919/ITC.2017.8064352

[13] Ethan Katz-Bassett. 2010. Reverse Traceroute. http://research.cs.washington.edu/networking/astronomy/reverse-traceroute.html.

[14] M. Jonker, A. King, J. Krupp, C. Rossow, A. Sperotto, and A. Dainotti. 2017. Millions of Targets Under Attack: a Macroscopic Characterization of the DoS Ecosystem. In *Internet Measurement Conference (IMC)*.

[15] Vermeulen K., Strowes S.D., O. Fourmaux, and T. Friedman. 2018. Multilevel MDA-Lite Paris Traceroute. In *Proceedings of the 2018 Internet Measurement Conference (IMC '18)*.

[16] kc claffy. 2017. "DIBBs: Integrated Platform for Applied Network Data Analysis (PANDA)". http://www.caida.org/funding/dibbs-panda/.

[17] K. Keys and Y. Hyun. [n. d.]. MIDAR: Monotonic ID-Based Alias Resolution Tool. http://www.caida.org/tools/measurement/midar/.

[18] K. Keys, Y. Hyun, M. Luckie, and k. claffy. 2013. Internet-Scale IPv4 Alias Resolution with MIDAR. *IEEE/ACM Transactions on Networking* (Apr 2013).

[19] Matthew Luckie and kc claffy. 2014. A Second Look at Detecting Third-Party Addresses in Traceroute Traces with the IP Timestamp Option. In *PAM*, Michalis Faloutsos and Aleksandar Kuzmanovic (Eds.). https://doi.org/10.1007/978-3-319-04918-2_5

[20] M. Luckie, A. Dhamdhere, B. Huffaker, D. Clark, and k. claffy. 2016. bdrmap: Inference of Borders Between IP Networks. In *Proceedings of the ACM SIGCOMM Internet Measurement Conference (IMC)*.

[21] Matthew Luckie, Amogh Dhamdhere, Bradley Huffaker, David Clark, and kc claffy. 2016. bdrmap: Inference of Borders Between IP Networks. In *Internet Measurement Conference (IMC) (IMC)*. ACM.

[22] Matthew J. Luckie. [n. d.]. Scamper. http://www.caida.org/tools/measurement/scamper/.

[23] Alexander Marder, Matthew Luckie, Amogh Dhamdhere, Bradley Huffaker, kc claffy, and Jonathan M. Smith. 2018. Pushing the Boundaries with bdrmapIT: Mapping Router Ownership at Internet Scale. In *ACM SIGCOMM*.

[24] Alexander Marder and Jonathan M. Smith. 2016. MAP-IT: Multipass Accurate Passive Inferences from Traceroute. In *SIGCOMM ACM IMC*. http://dl.acm.org/citation.cfm?id=2987468

[25] University of Michigan. [n. d.]. Censored Planet. http://censoredplanet.com/.

[26] RIPE NCC. 1992-2018. RIPE NCC IPmap: A Collaborative Approach to Mapping Internet Infrastructure. https://openipmap.ripe.net/.

[27] Roderick Fanou. 2018. "(ARDA) African Route-collectors Data Analyzer Source Code". https://github.com/rodricfanou/African_Route-collectors_Data_Analyzer-ARDA.git.

[28] Roderick Fanou, Victor Sanchez-Aguero, Francisco Valera, Michuki Mwangi, Jane Coffin . 2017. (ARDA) African Route-collectors Data Analyzer. https://arda.af-ix.net/.

[29] P. Sermpezis, V. Kotronis, P. Gigis, X. Dimitropoulos, D. Cicalese, A. King, and A. Dainotti. 2018. *ARTEMIS: Neutralizing BGP Hijacking within a Minute*. Technical Report. Center for Applied Internet Data Analysis (CAIDA).

[30] Ben VanderSloot, Allison McDonald, Will Scott, J. Alex Halderman, and Roya Ensafi. 2018. Quack: Scalable Remote Measurement of Application-Layer Censorship. In *USENIX Security Symposium*.