12th USENIX Workshop on Cyber Security Experimentation and Test (CSET '19)

PETER A. H. PETERSON AND ROB G. JANSEN

Peter A. H. Peterson is an Assistant Professor of Computer Science at the University of Minnesota, Duluth, where he teaches and researches operating systems and security, with a focus on R&D to make security education more effective and accessible. He received his PhD from the University of California, Los Angeles, for work on "adaptive compression"—systems that make compression decisions dynamically to improve efficiency. pahp@d.umn.edu

Dr. Rob Jansen is a Computer Scientist and a Jerome and Isabella Karle Distinguished Scholar Fellow at the US Naval Research Laboratory with research expertise in the areas of computer security and privacy, distributed systems, and parallel and distributed simulation. His research results have been published in the top peer-reviewed computer security and privacy conferences and workshops, and his work demonstrates an ability not only to develop and apply theoretical concepts, but also to build, evaluate, deploy, and measure real-world systems. When not in the lab, Rob enjoys jogging around the National Mall and through the streets of DC. rob.g.jansen@nrl.navy.mil

On Monday, August 12, 2019, 55 attendees joined us for the 12th USENIX Workshop on Cyber Security Experimentation and Test (CSET '19) in Santa Clara, California. CSET, one of the USENIX Security Symposium's co-located workshops, welcomes work in the broad categories of "cybersecurity evaluation, experimentation, measurement, metrics, data, simulations, and testbeds"—that is, research about research tools, data, and methods. The purpose of this article is to share our experience chairing CSET '19 and to highlight this year's papers.

Changes to the CSET PC

We made some experimental changes to the call for papers (CFP) and program committee (PC) this year, and we wanted to share them in the hope that they might be useful for other organizers. One of our main goals was to increase the community reach of the PC and the submission count, while reducing the PC review burden. To do this, we doubled the size of the PC to 46, inviting both established CSET community members and new people, including both junior and senior researchers. We also explicitly invited broad interpretations of the topics list. Additionally, we solicited a variety of paper lengths and types: traditional research papers, position papers, experience papers, preliminary work, and extended work. These could be long papers (eight pages), short papers (four pages), or extended abstracts (two-page talk proposals).

We explicitly invited preliminary work papers because CSET is a workshop; we wanted to encourage the lively discussion of new ideas, even if they were not fully developed. "Extended work" papers were meant to be expansions of security experimentation results, approaches, or tools developed in the course of other research (e.g., papers published at USENIX Security or elsewhere). Our rationale for soliciting these papers was that all security research requires an experimental approach; this often includes the development of tools, data, or knowledge that could be useful to the community. Unfortunately, these details are often drastically reduced in published papers due to space constraints. This cut material is often squarely in CSET's bailiwick, and we hoped that papers like this would be relatively easy for authors to prepare, interesting for attendees to discuss, and of service to the research community.

We are also happy to report that women comprised 46% of the CSET '19 PC, up from the recent peak of 32% in 2015. Women are in high demand and may already be committed to a full slate of PCs; to find the 21 women who were able to join the PC this year, we invited approximately double that number. Our takeaway was that it is absolutely possible to improve gender representation on PCs, but until the underlying diversity in our field improves, doing so may take a little time and effort.

Overall, our changes seemed to work well; we received 61 submissions, more than doubling 2018's submission count of 27. Each reviewer had approximately four papers to review. (We had wanted to limit each PC member to three reviews, but the volume of submissions precluded that.) Ultimately, we accepted 19 papers (31%). For more information about our process and statistics this year, please see our slides on the workshop site.

Sessions and Presentations

The 19 accepted papers this year were arranged into five sessions. The first session was "Cyberphysical and Embedded Testbeds and Techniques," chaired by Eric Eide (University of Utah). First, Paul Pfister (Iowa State University) presented a cyber-physical system (CPS) extension to ISEAGE, an event simulator used for cyberdefense competitions that included a physical model of a city, complete with LEDs representing system status. Next, Woomyo Lee (The Affiliated Institute of ETRI) presented a system for automatic generation of CPS research data about a power plant featuring a GE turbine, an Emerson boiler, and a FESTO water treatment system. After this, Sam Crow (UC San Diego) told us about Triton, a configurable testbed for avionics security research. Triton is, in the words of Crow, "real hardware from a real airplane that thinks it's running on an actual airplane in flight." Finally, we heard from Zachary Estrada (Rose-Hulman Institute of Technology) about CAERUS, a framework that is able to identify, through automated testing, timing sensitivities of undocumented embedded systems that can interact negatively with add-on security components.

Elissa Redmiles (Microsoft Research/Princeton University) chaired our "Data and Metrics" session. Michael Brown (Georgia Institute of Technology) described how debloaters can improve security by reducing the number of ROP gadgets through eliminating unimportant code, but also how they can accidentally introduce new high-quality gadgets. Instead of focusing on gadget count as the key metric, Brown proposes metrics based on gadget quality. Next, Aniqua Baset (University of Utah) discussed SecPrivMeta, an interactive website (secprivmeta.net) that provides visualizations of topic modeling on 36 years of security and privacy publications. After this, Josiah Dykstra (US Department of Defense) described how the NSA uses the Innovation Corps (I-Corps) methodology to improve the sharing of Cyber Threat Intelligence (CTI). Last, Jim Alves-Foss (University of Idaho) gave an entertaining talk containing a variety of cautionary tales of problematic data analysis and experimentation to admonish the community to use care and best practices in research.

"Usability, Effects, and Impacts" was chaired by Heather Crawford (Florida Institute of Technology). Zane Ma (University of Illinois at Urbana-Champaign) gave the first talk, which was about the effect of TLS and browser presentation on the success of phishing attacks in an A/B test on 266 users. Next, Victor Le Pochat (KU Leuven) described the design and evaluation of Tranco, a "top sites" ranking that aggregates Alexa, Majestic, Quantcast, and Umbrella, to create a stable and robust list for use by researchers. Third, Xiaodong Yu (Virginia Tech) presented work investigating how seven cache configuration parameters affected timing-based side-channel attacks; their talk included suggestions for improving security while minimizing performance impact. Last, Ildiko Pete (University of Cambridge) presented preliminary results from the Cambridge Cybercrime Center's analysis of usability issues with the data sets they share.

David Balenson (SRI International) chaired "Problems and Approaches," which began with Qiao Kang's (Rice University) presentation of their work automating the detection of attacks against the data planes of programmable routers. Our next presenter, Fatima Anwar (UCLA, now University of Massachusetts at Amherst) described how the timing capabilities of trusted execution environments (TEEs) can be vulnerable to timing attacks in realistic scenarios, and provided requirements for securing time facilities in these environments. Next, Sri Shaila G (University of California, Riverside) presented results of a study using IDAPro to reverse engineer the binaries of real-world IoT malware samples as compiled with various options, finding that, while unstripped binaries are amenable to analysis, performance on stripped binaries is generally poor. Last, Jonathan Crussell (Sandia National Laboratories) talked about their analysis of 10,000 experiments comparing differences between virtual and physical testbeds for research.

The final session of the day was "Testbeds and Frameworks," chaired by Jelena Mirkovic (University of Southern California Information Sciences Institute). Aditya Ashok (Pacific Northwest National Laboratory) described PACiFiC, a sufficiently realistic campus microgrid testbed model to allow a phish-to-blackout attack simulation. Second, Russell Van Dam (Sandia National Laboratories) presented Proteus, an emulation framework that supports the analysis of a wide variety of peer-to-peer distributed ledger technologies against different types of automated scenarios. Finally, Ryan Goodfellow (Information Sciences Institute) described the DComp Testbed, an open-source testbed using EVPN routing, a set of independently useful tools, and featuring a high level of abstraction and isolation.

For more detail, please see the workshop program online at www.usenix.org/cset19/program.

We would like to offer our sincere thanks to the fantastic USENIX staff, CSET's program and steering committees, authors, session chairs, shepherds, presenters, and attendees. The 13th CSET will once again be co-located with USENIX Security 2020 in Boston, with papers due in spring 2020. If you're interested in research around security experimentation, please consider submitting to and/or attending CSET next year!