Cybersecurity Experimentation Workshop Report (2022)

January 2, 2023

Jelena Mirkovic, David Balenson, Srivatsan Ravi,
Luis Garcia, Terry Benzel
USC Information Sciences Institute

USC Viterbi
School of Engineering
Information Sciences Institute

Introduction

This report summarizes the discussion and key findings from a three-hour workshop on improving cybersecurity experimentation, held virtually on Wednesday, December 14, 2022. Additionally, the report summarizes the findings from a follow-up survey. The workshop and survey served as forums for learning about and understanding of obstacles in cybersecurity experimentation, what is needed to overcome them, and how to improve reproducibility and reuse of cybersecurity artifacts. Both benefited from the degree of openness and interaction displayed by the participants, while discussing cybersecurity experimentation in their research. By openly sharing their experiences and knowledge, insights were gained which are documented in this report and should provide value to all participants and the broader community. Furthermore, the workshop was the first in what will hopefully be an ongoing series of discussions and interactions around the cybersecurity experimentation and cybersecurity artifacts.

Participants

A large number of researchers from different areas of cybersecurity and privacy were invited via email to participate in the workshop (see appendix). Out of these, around 30 joined the workshop and remained engaged throughout the discussion. The majority were faculty or PhD students at universities in the U.S. and abroad. Some researchers had experience in building and operating research infrastructure and spoke from that perspective too.

Goals

The workshop had several goals:

1. Learn about researcher needs around cyber experimentation and how they meet those needs today
2. Learn about researcher needs with regard to common research infrastructure, and about any obstacles researchers face when using common research infrastructure
3. Revisit and extend recommendations from previous Cybersecurity Experimentation of the Future (CEF) workshops [1]
4. Understand the obstacles around experiment artifact sharing and reuse

Format

The workshop was broadly structured into three discussion sessions: (1) cyber experimentation needs, (2) research infrastructure needs and experiences and (3) artifact sharing and reuse. After very brief introductory presentations, all participants were called upon to speak (some also volunteered multiple times) in the sessions, and to share their opinions on all three topics, regardless of the session. A copy of the agenda and links to the slides used for the introductory presentations are included in appendices.

Reception

Participants were very engaged in this workshop, and had a lot to share. They came from different research fields (CPS, automotive security, IoT, human-centric security, internet measurement, social network measurement, embedded systems, smart phones, hw in the loop, detailed simulation of power systems) and thus had unique views around cyber experimentation. There was a lot of enthusiasm expressed for discussing the workshop's topics. A list of workshop participants and their affiliations is included in an appendix.

Follow-up Survey

After the workshop, around 500 researchers in cybersecurity and privacy were invited via email to fill out a short questionnaire around the same main questions that were discussed in the workshop (see appendix). A total of 58 participants filled out the questionnaire within five days of being invited. Their responses are summarized below. **The questionnaire remains open at** https://forms.gle/PtSyNw9UeWbfvuwt6 **and we encourage any interested researcher to participate. We will revisit the responses every six months to glean new insights.**

Findings

We discuss our findings separately around the three main themes of the workshop.

Cybersecurity and privacy experimentation

Need for datasets and common evaluation environments: Each research area has its own needs. Abstracting from these two main topics emerged. First, *researchers in a given area need a common evaluation environment* so they can compare their work with prior work in the field. One example of this was adversarial machine learning (ML), where many attacks exist on many different systems, but there is no way to compare them with each other. Second, *researchers need high-quality, standardized and labeled datasets*. In many research areas datasets are hard to find (private, because they include some human actions), they are insufficient in size and improperly labeled, and they also may be synthetic or outdated.

Need for modeling or including human users: In many real cybersecurity and privacy scenarios, humans play a crucial role. There is a need to include humans in cyber experimentation, but we do not yet know how to do this in a systematic, reproducible, and scalable manner. If we were able to include humans in experiments this would improve chances of research systems transitioning to practice. *Opportunity:* Research infrastructure may be able to help here – automate setup of human user studies and preserve it for future use by other researchers.

Cybersecurity and privacy research infrastructure

Need for representative experimentation environments: Representative (realistic) environments and datasets are very important for publications. Researchers often struggle to publish, because

their experiments are deemed to be “toy” experiments – too simplistic. Researchers need pre-set, representative environments (i.e., topology, applications, traffic) and datasets, which will be accepted as “real enough” by program committees. Such environments also need to evolve - be updated periodically as trends in the Internet change. In some areas, like CPS, there is the notion of digital twins, which is well-accepted.. For datasets, there are a few challenges: how to remove noise and how to achieve accurate labels on large datasets. There was a significant discussion about the need for representative traffic, both legitimate and malicious. *Opportunity:* a shared research infrastructure could host complex experimentation environments, used by many researchers. *Opportunity:* experimental traffic could be captured and labeled automatically, to be used as a quality dataset in other research or in education. *Opportunity:* a shared research infrastructure with ability to recreate representative legitimate and malicious traffic would significantly aid researchers.

Need for user-friendly interfaces: Researchers invest a lot of effort to learn how to use new infrastructure. User-friendly interfaces are needed, which enable quick learning, such as Jupyter notebooks and pre-set experiments. Researchers also need programmability – they need a way to automate the entire experiment lifecycle (execution, post-processing, storage, sharing) in addition to experiment setup. Researchers further need a way to manage large-scale experiments.

Need for inclusion of third-party devices: Many researchers currently buy and instrument their own devices. They feel ownership over them and the cost is not too large. However, research evolves and these devices may become redundant in a given lab. *Opportunity*: Research infrastructure could help include these third-party devices in experiments, for reproducibility or just to utilize them once they have become redundant for their lab.

Need for a variety of devices and experimentation modes: There are many special devices - embedded devices, FPGA, IoT devices - which should be included in common research infrastructure. There are many different manufacturers, chip vendors, and devices that can have their own security features. In some cases these devices are difficult to instrument. Researchers also need a way to combine emulation and simulation, and potentially even include external elements (e.g., third-party devices, external Web sites, etc) in their experimentation.

Difficulty in environment setup and reuse: Setting up an experimental environment in any infrastructure (shared or in a private lab) takes time and is a lot of work. Researchers struggle how to amortize this time and effort. Often students from the same lab may end up exploring different research questions and they need different experimental environments. *Opportunity:* a shared research infrastructure could help package environments for reuse by others. This way someone would benefit from the invested work, albeit not the same researcher that did the work.

Difficulty in understanding research infrastructure’s limitations: Researchers need to understand limitations and/or assumptions in research infrastructure’s architecture. For example, what kind of experiments is the infrastructure meant to support, and with what fidelity. Researchers also need to understand the cost of using research infrastructure; both monetary and time/effort cost.

Cybersecurity and privacy artifact sharing and reuse

Difficulty - incomplete artifacts: Artifacts are sometimes published with insufficient documentation or missing scripts. *Opportunity:* Standardize artifact format and metadata, enforce this standard during artifact submission and evaluation.

Difficulty - non-portable artifacts: Some artifacts work in a given setting but not in a different environment. Being able to prove that an artifact works in different environments is necessary to make research relevant.

Need - more artifact evaluation and research reproduction: According to Security Venues with No Page-Limit - a short list [2], out of approximately 96 cybersecurity and privacy conferences there are only six that have artifact evaluation. We need to extend this practice to more venues. ML conferences have established datasets and benchmark tracks. This could be useful to cybersecurity and privacy research. Research reproduction papers (especially negative results) also need a venue where to get published and receive recognition.

Need - large storage for ML models. Many ML models require days to train on powerful, but expensive GPUs. ML researchers need pretrained models, space where these can be stored, and access to GPU infrastructure.

Questionnaire Findings

The questionnaire was administered online to 58 anonymous participants between December 17 and December 21, 2022. We summarize the findings below.

Participants

Participants were asked to self-declare their field of research in a free-form answer. Participants' field of research varied broadly across the following categories: software security, network security, systems security, web security, AI security, psychological aspects of cyber defense, CPS security, information integrity, data-driven security, IoT security, security education, Internet security, hardware security, distributed systems security, vehicular security, privacy-preserving computation, data privacy, human factors in security and privacy, routing security, side channels, cloud security, usable privacy and security, enterprise security, and user authentication.

Cybersecurity and privacy experimentation

Next, participants were asked to discuss what is needed for cyber experimentation in their field of research. We have grouped the answers into categories and we list the categories here:

- Representative datasets and applications, representative experimentation environments
- GPUs for model training
- Large storage for model-saving and sharing
- Inclusion of humans in cybersecurity scenarios

- Special devices (programmable logic controllers, mobile phones, IoT devices, cars, drones, cellular network, oscilloscopes or spectrum analyzers, Arduino, Raspberry Pi)
- Extensive logging (data collection) to package and release artifacts
- Computational resources
- Virtual machines with full (superuser) control
- Internet access for measurement or third-party datasets
- Domain names that can be attached to research infrastructure
- Ability to run large-scale experiments
- Routing experiment support (ASN, peering relationships, ability to emulate large BGP experiments, vantage points to advertise Internet routes)
- Data visualization
- High-fidelity simulators, e.g., for CPS or wireless

Compared to our virtual workshop findings, these answers provide a new input around the *needs for simulators and for support for routing experiments.*

Next, participants were asked how they experimented today. We grouped similar responses and summarize them here:

- University cluster (2 responses)
- In their lab (35 responses)
- Simulator (6 responses)
- Measurement on the Internet (25 responses)
- Testbed (18 responses, including DeterLab, CloudLab, Chameleon, PEERING, EC2 and private commercial testbeds)

The response numbers add up to more than 58 participants, because some participants listed multiple experimentation modes. In fact, experimentation in a lab, coupled with measurement on the Internet was the most frequent selection.

Cybersecurity and privacy research infrastructure

Next, participants were asked to describe an ideal, hypothetical research infrastructure for their use. We categorized their responses and summarize them here:

- A dataset aligned with the researcher's research focus (e.g., ML models with backdoors, large open-source software repositories)
- A representative environment/scenario for a given cybersecurity research focus
- An orchestration software for executing large-scale experiments
- Virtual machines with the user having full control
- Ability to mix simulation and experimentation
- Ease of use: automated saving of experiment results and configuration (packaging), support for the entire experiment life-cycle
- Access to special devices (mobile phones, laptops, oscilloscopes, cellular equipment, drones, cars)
- Full Internet access, ability to include external systems in the experiment
- Ability to run experiments at different locations in the real Internet
- Ability to run routing experiments

- Ability to run large-scale experiments
- Reliable and encrypted storage
- No wait time for resources, no limits on how long they can be held
- Trusted computing hardware
- Ability to attract human users and engage them in behaviors, which can then be monitored and logged for analysis. Ability to include humans in experiments.

Overall, these responses echoed the findings from the virtual workshop and expanded on the need for *usability, ease of access, full control, diversity and representativeness of research infrastructures.*

Cybersecurity and privacy artifacts

Finally, participants were asked to comment on how the research community can improve quality of shared artifacts and promote artifact reuse. Answers were categorized and are summarized below:

- Venues should encourage/require artifacts and run artifact evaluation
- Venues dedicated to cybersecurity artifacts and research reproduction (positive or negative results)
- Recognition by peers/promotion bodies that artifact publishing is a large contribution
- Standards for packaging artifacts and common place to store them (with limited access by others if desired)
- Better artifact documentation and test data and scripts
- Automated artifact packaging by research infrastructures
- Easy way to run a packaged experiment
- Common research infrastructure where to share/reuse artifacts
- Ability to rate artifact quality

Overall, these responses echoed the findings from the virtual workshop. The overwhelming majority of answers talked about *venues requiring artifacts for publication and the need for standards around artifact packaging*, which clarify requirements for a high-quality artifact. *Research infrastructures* were seen as key players here, which could *enable sharing and reuse of artifacts via automated artifact packaging support, and via storage and reuse of artifacts on the infrastructure.*

Acknowledgements

The workshop organizers thank everyone who took the time to attend the workshop and for their open and collegial participation. We also thank the anonymous participants who completed the online questionnaire. Your feedback will help inform future efforts to support cybersecurity and privacy experimentation, infrastructure, and artifact sharing and reuse.

The workshop was supported, in part, by the Sharing Expertise and Artifacts for Reuse through Cybersecurity Community Hub (SEARCCH) project which is supported by the National Science Foundation under grant numbers 1925773, 1925616, 1925588, 1925564. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Conclusion and Next Steps

Together, the workshop discussion and the questionnaire responses identified many needs and challenges as well as opportunities for improved cyber experimentation, infrastructure, and artifact sharing and reuse. We now reflect back on the workshop goals:

1. Learn about researcher needs around cyber experimentation and how they meet those needs today. ***We learned that most researchers experiment in their own lab, using limited resources they have, and may combine this with measurement on live Internet. Many researchers expressed unmet needs around cyber experimentation, revolving around the needs for common environments and datasets, and inclusion of humans.***
2. Learn about researcher needs with regard to common research infrastructure, and about any obstacles researchers face when using common research infrastructure. **We learned that researchers need user-friendly research infrastructure, which helps them perform experimentation in representative environments on a variety of devices. Research infrastructure should aid researchers publish (by providing diverse devices and representative environments and datasets, and ability to mix experimentation modes - emulation, simulation and Internet interaction) and share artifacts (by providing support for artifact packaging).**
3. Revisit and extend recommendations from previous Cybersecurity Experimentation of the Future (CEF) workshops. **Our current findings validated all the findings from CEF workshops.**
4. Understand the obstacles around experiment artifact sharing and reuse. **We learned that researchers would benefit from standardization of requirements around artifact sharing, from research infrastructure aiding in artifact packaging and from incentives (at publication venues and in the promotion process) for artifact sharing and reuse.**

Furthermore, the workshop was the first in what will hopefully be an ongoing series of discussions and interactions around the use of cybersecurity experimentation and cybersecurity artifacts. We hope to hold additional workshops and conduct additional surveys in the future.

References

[1] David Balenson, Laura Tinnel, and Terry Benzel, Cybersecurity Experimentation of the Future: Catalyzing A New Generation of Experimental Cybersecurity Research, Final Report, July 31, 2015. https://cef.cyberexperimentation.org/

[2] Security Venues with No Page-Limit. https://secnopagelim.github.io/

Appendix: Workshop Invitation Email

Dear <FIRST NAME>

We would like to invite you to a three-hour virtual workshop on improving cybersecurity experimentation on Wed, Dec 14, from 8 to 11am PST. The workshop will focus on discussion of obstacles in cybersecurity experimentation, what is needed to overcome them (e.g., more diverse hardware, labeled datasets, reusable experiments, etc.), and how to improve repeatability and reuse of cybersecurity artifacts. Your feedback will inform our future endeavors in supporting cybersecurity experimentation through public testbeds we host at USC Information Sciences Institute.

We encourage you to attend even if all cybersecurity experimentation you have done were in your lab, or on your laptop/desktop. We would like to hear from everyone what are obstacles researchers face in experimentation and what can be done to overcome these obstacles. A tentative agenda is pasted below. Please come prepared to offer your point of view, no feedback is too small.

Wed, Dec 14, all times PST
8:00 - 8:20 Introductions
8:20 - 8:40 Main questions for discussion
8:40 - 9:10 Cybersecurity experimentation needs
9:10 - 9:20 Break
9:20 - 9:50 Testbeds for cybersecurity experimentation
9:50 - 10:00 Break
10:00 - 10:30 Cybersecurity artifacts
10:30 - 11:00 Final remarks and wrap-up

Please register at this Zoom link:
https://usc.zoom.us/meeting/register/tJcucu6grzMvH9AaMSNff2Xm9aMZgKEy8n0L

If your schedule does not permit you to attend all three hours, please feel free to drop in when you can. In case you cannot attend, we would appreciate it very much if you offered your thoughts via email. Simply reply to this email and share your feedback around the following questions:
(1) what is needed for cyberexperimentation in your field of research?
(2) how do you experiment today (e.g., in a lab)?
(3) what would a testbed have to offer to entice you to experiment there?
(4) how can we as a community improve the quality and increase reuse of cybersecurity artifacts (code and datasets, experiment scenarios, etc. in published papers)?

If you would rather offer your thoughts anonymously please use this Google form:
https://forms.gle/GsBaKVbtkhoLmoih7

We look forward to your participation in the workshop and hope that you will be able to participate as much as possible!

Best regards,
Jelena Mirkovic, Terry Benzel, David Balenson, Srivatsan Ravi, Daniel Massey
USC Information Sciences Institute

Appendix: Workshop Agenda

Wed, Dec 14 (all times PST)
8:00 - 8:20 Introductions
8:20 - 8:40 Main questions for discussion (slides)
8:40 - 9:10 Cybersecurity experimentation needs (slides)
9:10 - 9:20 Break
9:20 - 9:50 Testbeds for cybersecurity experimentation (slides)
9:50 - 10:00 Break
10:00 - 10:30 Cybersecurity artifacts (slides)
10:30 - 11:00 Final remarks and wrap-up

Appendix: Presentation Slides

Presentation slides for each session can be found at the following links:

Main questions - https://docs.google.com/presentation/d/1G1w4_1jTIZE0EdWWK-2_pdg31LGLufBz-y3Apg2ISDY/edit?usp=sharing

Cybersecurity experimentation - https://docs.google.com/presentation/d/1zk4I4PS3kbTIfzHnyzys5TbCDtyUGWLvYzxPrGFB8Ok/edit?usp=sharing

Cybersecurity testbeds - https://docs.google.com/presentation/d/1GWYu-CFUZz3Lb9tJ7f7Z15bZTEtPd6LjC74_6IPFzwE/edit?usp=sharing

Cybersecurity artifacts -

https://docs.google.com/presentation/d/1O8kOj-lrmONCj_72uSRyLmvjeilEDgrMU3tE-1hEZzs/edit?usp=sharing

Appendix: Workshop Participants

Afsah Anwar – Postdoc, Northeastern University
Anita Nikolich – Director of Research and Technology Innovation and Research Scientist, School of Information Sciences (iSchool), University of Illinois, Urbana-Champaign
Brendan Saltaformaggio – Assistant Professor, School of Cybersecurity and Privacy and the School of Electrical and Computer Engineering, Georgia Institute of Technology
Christos Papadopoulos – Professor of Computer Science, University of Memphis
Daphne Yao – Professor, Computer Science, Virginia Tech
David Balenson – Associate Director Networking and Cybersecurity Division and Senior Computer Scientist, USC Information Sciences Institute
David Nicol – Herman M. Dieckamp Endowed Chair in Engineering; Director, Advanced Digital Sciences Center; and Director, Critical Infrastructure Resilience Institute, University of Illinois, Urbana-Champaign
Deepankar (Deep) Medhi – Program Director, Division of Computer and Network Systems, National Science Foundation
Efren Lopez Morales – PhD student, Texas A&M University
Eugene Vasserman – Associate Professor, Computer Science, Kansas State
Gang Wang – Assistant Professor, Computer Science, University of Illinois
Giovanni Apruzzese – Assistant Professor, Computer Science, University of Liechtenstein
Guillermoi Francia – Director, Research and Innovation, University of West Florida
Ilkan Esiyok – PhD student, CISPA
Jelena Mirkovic – Research Faculty, USC Information Sciences Institute
Luis Garcia – Research Assistant Professor, Computer Science, USC; and Research Lead, USC Information Sciences Institute
Nik Sultana – Assistant Professor, Computer Science, Illinois Institute of Technology
Niklaus Kang – Infrastructure Assistant Manager, National University of Singapore
Sascha Fahl – Professor, Computer Science, CISPA
Shanchieh (Jay) Yang – Professor, Computer Engineering, Rochester Institute of Technology; and Director Global Outreach for Global Cybersecurity Institute
Shuang Hao – Assistant Professor, Computer Science, UT Dallas
Srivatsan Ravi – Assistant Professor of Research, USC Dept. of Computer Science; and Research Scientist, USC Information Sciences Institute
Stefanos Koffas – PhD student, TU Delft
Terry Benzel – Associate Director ISI; and Director Networking and Cybersecurity Division, USC Information Sciences Institute
Xenofan Koutsokos – Chair, Department of Computer Science; Professor of Computer Science; and Professor of Electrical and Computer Engineering, Vanderbilt University
Ximing (Simon) Ou – Professor, Computer Science and Engineering, University Southern Florida
Zhenkai Liang – Associate Professor, School of Computing, National University of Singapore
Ziming Zhao – Assistant Professor, Computer Science and Engineering, University Buffalo

Appendix: Cybersecurity Experimentation Survey

All questions allowed for a paragraph-style response and all questions were optional and anonymous.

1. What is your broad field of research (e.g.,, IoT security)
2. What is needed for cyber experimentation in YOUR field of research?
3. How do you experiment today (e.g., in a lab, in a testbed, in real Internet)? More details are very helpful.
4. What would a testbed have to offer to entice you to experiment there? Imagine an ideal testbed for your research, then describe it to us.
5. How can we as a community improve the quality and increase reuse of cybersecurity artifacts (code and datasets, experiment scenarios, etc. in published papers)? Any ideas are welcome here.
6. If there is anything else you want to share, feel free to use the space here.