DeapSECURE: Empowering Students for Data- and Compute-Intensive Research in Cybersecurity through Training

Wirawan Purwanto
Information Technology Services
Old Dominion University
Norfolk, Virginia, USA
wpurwant@odu.edu

Hongyi Wu
Electrical and Computer Engineering
Old Dominion University
Norfolk, Virginia, USA
h1wu@odu.edu

Masha Sosonkina
Modeling, Simulation, and Visualization Engineering
Old Dominion University
Norfolk, Virginia, USA
msosonki@odu.edu

Karina Arcaute
STEM Education and Professional Studies
Old Dominion University
Norfolk, Virginia, USA
karcaute@odu.edu

ABSTRACT

As the volume and sophistication of cyber-attacks grow, cybersecurity researchers, engineers and practitioners rely on advanced cyberinfrastructure (CI) techniques like big data and machine learning, as well as advanced CI platforms, e.g., cloud and high-performance computing (HPC) to assess cyber risks, identify and mitigate threats, and achieve defense in depth. There is a training gap where current cybersecurity curricula at many universities do not introduce advanced CI techniques to future cybersecurity workforce. At Old Dominion University (ODU), we are bridging this gap through an innovative training program named DeapSECURE (**D**ata-**E**nabled **A**dvanced Training **P**rogram for Cyber **Secu**rity **R**esearch and **E**ducation). We developed six non-degree training modules to expose cybersecurity students to advanced CI platforms and techniques rooted in big data, machine learning, neural networks, and high-performance programming. Each workshop includes a lecture providing the motivation and context for a CI technique, which is then examined during a hands-on session. The modules are delivered through (1) monthly workshops for ODU students, and (2) summer institutes for students from other universities and Research Experiences for Undergraduates participants. Future plan for the training program includes an online continuous learning community as an extension to the workshops, and all learning materials available as open educational resources, which will facilitate widespread adoption, adaptations, and contributions. The project leverages existing partnerships to ensure broad participation and adoption of advanced CI techniques in the cybersecurity community. We employ a rigorous evaluation plan rooted in diverse metrics of success to improve the curriculum and demonstrate its effectiveness.

PEARC'19, July, 2018, Chicago, Illinois USA

ACM ISBN 978-1-4503-7227-5/19/07...$15.00
https://doi.org/10.1145/3332186.3332247

CCS CONCEPTS

• **Social and professional topics** → Computing education programs → Computational science and engineering education

KEYWORDS

Training, Cyberinfrastructure, High-performance computing, Big data, Machine learning, Parallel computing, Cybersecurity, Workforce development

ACM Reference format:

Wirawan Purwanto, Hongyi Wu, Masha Sosonkina, Karina Arcaute. 2019. DeapSECURE: Empowering Students for Data- and Compute-Intensive Research in Cybersecurity through Training. In *Proceedings of ACM Practical Experience in Advanced Research Computing Conference (PEARC'19). ACM, New York, NY, USA, 8 pages.* https://doi.org/10.1145/3332186.3332247

1 INTRODUCTION

Cybersecurity focuses on protecting computers, networks, programs, and data from attack, destruction or unauthorized access. With the increasing reliance on computer systems and networks, cyberattacks are becoming more common,

sophisticated, and harmful in recent years. There is a surging demand for a well-trained cybersecurity workforce to address the multifaceted cybersecurity problems. According to CyberSeek [1], the U.S. employs nearly 716,000 people in cybersecurity positions in 2019, with approximately 313,000 current cybersecurity openings including over 33,000 in the Commonwealth of Virginia—the second highest in terms of number and the highest in terms of the demand concentration in the nation.

As the volume and sophistication of cyber-attacks grow, cybersecurity researchers, engineers and practitioners heavily rely on advanced cyberinfrastructure (CI) techniques such as Big Data (BD) and machine learning (ML), as well as advanced CI platforms, e.g., cloud and high-performance computing (HPC) to assess cyber risks, identify and mitigate threats, and achieve defense in depth [2, 3]. The vast majority of them extensively utilize computation, modeling and/or large amounts of data, leading to a surging demand for advanced CI techniques. The same CI techniques are also largely deployed by many corporations (IT, security, finance, merchants, etc.) to protect ever-growing and diverse digital infrastructure.

1.1 Cybersecurity Research at Old Dominion University

Old Dominion University (ODU) is located in the southeastern Virginia city of Norfolk, in the region named for the largest natural deep-water harbor on earth, Hampton Roads. Hampton Roads is home to the Naval Station Norfolk, the Newport News Shipbuilding, the Port of Virginia, and the Langley Air Force Base, the NASA Langley Research Center, numerous other federal facilities, as well as 164 international businesses representing 28 countries. This significant infrastructure represents a mosaic of assets and makes Hampton Roads an increasingly attractive target and particularly vulnerable to cyberattacks. The needs for cybersecurity workforce are particularly high and pressing in this region.

As the largest university in this region, ODU is strategically located for leading a strong cybersecurity program to address fundamental cybersecurity challenges and train the next generation of cybersecurity professionals. The University has recently made substantial investment in this highly demanding area, as evidenced by the establishment of the Center for Cybersecurity Education and Research (CCSER), which is housed in the Office of Academic Affairs, aiming to weave together disparate threads of programmatic and facility resources across the campus to create a strong education and research program focusing on cybersecurity. It represents an interdisciplinary effort bringing together about 30 faculty members, with a wide range of expertise in computer science and engineering, mathematics, information technology, modeling and simulation, criminal justice, psychology and philosophy, along with postdocs, staff, and undergraduate and graduate students, who work together to develop degree programs and conduct research in cutting-edge areas of cybersecurity. (Some research work done under the umbrella of CCSER are showcased in our training program, as elaborated in Sec. 2.2.)

1.2 Educational Gap in Advance CI

Despite the fact that cybersecurity focuses on the protection of "*cyber*" *infrastructure* in the most general sense, advanced CI techniques have not been widely introduced in undergraduate and graduate cybersecurity programs. In particular, the current cybersecurity curricula at ODU include core components of cybersecurity such as cryptography, network security, reverse software/hardware engineering, cyber threats and vulnerabilities, risk assessment and management, cyber defense and operation techniques, forensic investigation, and cyber laws and ethics, aiming to prepare students with the theory, technologies, skills, and practices necessary to safeguard critical cyberinfrastructure and protect confidential information against unauthorized access or corruption. The cybersecurity programs in other universities in the Commonwealth and across the country are structured in a similar way. Advanced CI skills have not been identified as a core component of nor integrated into cybersecurity curricula.

The lack of advanced CI skills, however, is becoming a hurdle for many senior undergraduates and early-stage graduate cybersecurity students who are keen to conduct cutting-edge cybersecurity research and/or participate in advanced industrial cybersecurity projects. For example, a graduate student is currently working with one of the authors (Wu) to develop privacy-preserved machine learning techniques by embedding efficient homomorphic cryptosystems in deep neural networks [4]. The student's preliminary research shows that the computation time would be increased by three orders of magnitude if the neural-network training is based on homomorphically encrypted data. Hence, her research will progress faster if the student masters the requisite CI techniques to effectively speed up the computation. Another example involves two students who studied hardware vulnerabilities in smart home devices while participating in ODU-led cybersecurity Research Experience for Undergraduates (REU) program. In ten weeks, they managed to gather data and run small-scale vulnerability test on a workstation, but the project was not completed in full-scale because they were not able to use an appropriate HPC platform. The interdisciplinary nature of cybersecurity further exacerbates the problem. A significant portion of cybersecurity students have never been exposed to advanced CI platforms throughout their entire education path. The abovementioned cases are illustrative of the gap between formal cybersecurity education and set of skills required for advanced cybersecurity research and projects. There is an urgent need to bridge this gap and train future workforce and researchers to be able to take advantage of state-of-the-art CI capabilities.

This paper introduces the **D**ata-**E**nabled **A**dvanced Training **P**rogram for Cyber **Secu**rity **R**esearch and **E**ducation (DeapSECURE), an innovative CI training program created at ODU. DeapSECURE aims to prepare undergraduate and graduate

students with advanced CI techniques and teach them to use CI resources, tools, and services to succeed in cutting-edge cybersecurity research and industrial cybersecurity projects. The project also aims to generate understanding about the types of knowledge that CI training needs across the broad swath of cybersecurity careers. We anticipate that DeapSECURE will promulgate knowledge of advanced CI techniques into future cybersecurity research workforce, leading to an increase in their utilization of advanced CI techniques and platforms for research, therefore accelerating the development of secure and resilient cyberinfrastructure. In this paper, we will first describe the background context of cybersecurity research and education at ODU. We will also cover the goal of the training program as well as the approach and philosophy we adopt to guide the development of the program, followed by the description of the training contents. We will describe the training participants and our experience with delivering the training to them. We employ an external assessor to help us receive constant feedback through the training; we will share some lessons we learn in this first year of the training program. Finally, we will present our future direction of the program.

2. THE DEAPSECURE TRAINING PROGRAM

The DeapSECURE training program aims to train current and future cybersecurity researchers, engineers and practitioners with advanced CI techniques and data analytic skills. The goals of the program are to provide cybersecurity students with:

1. Exposure to advanced CI techniques and platforms (collectively referred to as "CI technologies");
2. Exposure to the application of CI technologies in the state-of-the-art cybersecurity research;
3. Experience of these techniques via hands-on activities;
4. Increased computational competency for their degree curricula and research projects, which are especially important for advanced degrees (master and doctorate).

It is important to note that DeapSECURE is not a training in basic literacy in cybersecurity, nor it is about cyber operation techniques. The training is focused on the advanced CI tools to the extent of their applicability to cybersecurity research and practical problems. The exposure to cybersecurity research topics coupled with advanced CI may open up greater realm of possibilities to the students and attract them to pursue cybersecurity research.

2.1 Approach and Philosophy

Our preliminary studies have led to two observations that influence DeapSECURE's design philosophy. First, the training program targets students majoring in cybersecurity or interested in cybersecurity topics. They come from different educational backgrounds, representing diverse competence in computing abilities, ranging from those who are well-versed in computer hardware and programming to those who never interacted with a command-line interface. Students are also diverse in age, ethnic background, and work experience. Most of them are new to CI. Second, cybersecurity research often calls for data-intensive techniques, including big data analytics (for handling large volume, variety and velocity of data streams), machine learning (for extracting insight and making predictions from the data), and traditional high-performance programming paradigms, such as OpenMP and Message Passing Interface (MPI). To this end, we adopt the following approach and philosophy in training:

1. Motivate the CI technique to the students by drawing real-world examples from state-of-the-art cybersecurity research;
2. Leverage real data sets, such as data generated by cybersecurity researchers and practitioners;
3. Emphasize practical tools and adopt a "concept-by-example" approach; this is in contrast to the ground-up approach usually employed in standard curriculum courses.

In DeapSECURE, we do not attempt to present a comprehensive introduction of the CI technologies to students. Given the limited interaction time with students, we provide "just enough" exposure and practical experience of each CI technique. In so doing, the key concepts and issues must be communicated clearly to the students. Note that this approach is very similar to the one adopted by the Carpentries community [5, 6] for their lessons because this is also a non-degree training to be delivered to "complete novices" that are interested in the subject and very motivated. We also adopt elements and ideas from XSEDE Monthly Workshops [7, 8] conducted by Pittsburgh Supercomputing Center as follows: many example codes are provided, which will also serve as templates for the trainees later when they embark their own research or project; we also pose a number of "challenge homework problems" for students to pursue in their own time to solidify their skills. Finally, we are providing numerous pointers to (carefully selected) external resources for trainees to pursue the topics further on their own. By providing a concise introduction with pointers to additional learning resources and many sample codes, students can minimize the "wandering time" before they can begin to be productive. That is, that students would devote their time learning the "right things" rather than trying to bump here and there to figure out the correct things.

In our vision, DeapSECURE will combine some face-to-face time with the trainees (students) and a rich set of online resources: The face-to-face time is used to motivate and introduce them to the basics of CI techniques; the "depth" of the training will be provided via online means, which include additional learning resources, access to HPC resources, and a learning community. We plan to set up the online resources in the coming years, as elaborated in Sec. 4.

2.2 Contents of the Training

In this project, we are developing six new CI training modules in the context of cybersecurity research. Each training module follows a storyline and includes three sections:

1. A research presentation by an invited cybersecurity faculty, concluding with a research problem that heavily depends on CI techniques;
2. An introduction of corresponding CI skills, tools and platforms;
3. A guided hands-on lab session where students will apply the CI techniques to solve the research problem formerly introduced by the cybersecurity faculty.

Based on the typical cybersecurity computing needs mentioned in Sec. 2.1, we elect to cover the following CI topics for the six modules:

1. Introduction to standard HPC platform and elementary parallel processing
2. Big data analytics
3. Machine learning
4. Deep learning
5. Advanced cryptography (homomorphic encryption)
6. High-performance programming, including introduction to parallel programming

These topics guide the selection of the research presentation (first section) and hands-on activities (third section) of every module.

ODU's strong cybersecurity research program gives ample research topics for the first part of each module. For 2018–2019 academic year, the following are the research topics, the presenting faculty and their departmental affiliations (all from ODU):

1. Big Data and Cybercrime: "An Ounce of Prevention is Worth a Pound of Cure", by Roderick Graham (Sociology and Criminal Justice)
2. Cybersecurity of Internet-of-Things [9], by Li Da Xu (Information Technology and Decision Sciences); From Big Data to Knowledge, by Jingwei Huang (Engineering Management and Systems Engineering)
3. GNU Radio Empowered Radio Frequency Signal Classification for Unmanned Aircraft Systems Detection and Identification, by Sachin Shetty

Hands-on Activity	Module	CI Technologies	Toolkit Introduced	Dataset/Program
Extraction of originating IP addresses and countries from a massive set of spam emails	1	Standard HPC platform	UNIX shell, SLURM job scheduler; GNU parallel	(1) Spam collection of untroubled.org (1998-2018, >8.5 million emails, >45 GB in total size); (2) IP2Location LITE free dataset for geolocation mapping
Analytics on the collected spam activities by the country and year	2	Big data analytics	PySpark, ipython	
Analytics on a large dataset of network traffic flows	2	Big data analytics	PySpark, ipython	CICIDS2017 Intrusion Detection Evaluation Dataset
Classification of threats based on network traffic flows	3	Machine learning	Scikit-learn; Pandas; ThunderSVM	
Classification of drone based on RF signal features	3	Machine learning	Scikit-learn; Pandas; ThunderSVM	Research data from Sachin Shetty and Abdulkabir Bello
Featureless deep learning for malicious URL classification	4	Deep learning	KERAS; Pandas	Hands-on and URL dataset provided by Melissa C. Kilby
Privacy-preserving computation with homomorphic encryption	5	Homomorphic encryption	Python-paillier	(no dataset)
Parallelization in data encryption	6	Parallel programming with MPI	Python-paillier, mpi4py	(no dataset)
Simulation of physical unclonable functions (PUFs)	6	Parallel programming with OpenMP	(GNU) C/C++ compiler	PUF model from Weize Yu and Yiming Wen

Table 1: Hands-on activities, CI technologies (techniques and platforms), toolkits, programs and datasets used in DeapSECURE modules.

(Modeling, Simulation, and Visualization Engineering; and Virginia Modeling, Analysis, and Simulation Center)

4. Virtual MAC Spoofing Detection through Deep Learning [10], by Chunsheng Xin (Electrical and Computer Engineering)
5. Privacy-Preserving Neural Networks [4], by Cong Wang (Computer Science)
6. Hardware Security through Physical Unclonable Functions [11], by Weize Yu (Electrical and Computer Engineering)

Each module includes one or two hands-on activities that are either directly drawn from the faculty's work, or from a cybersecurity problem that are very close to that work. The details of the hands-on activities are shown on Table 1. We design the modules to introduce students to a set of CI techniques and platforms through hands-on experiences using carefully chosen toolkits, all of which are publicly available, open source, and widely supported in the computational communities. In addition, we intentionally choose tools that are at the highest programming level so as not to bog trainees with cumbersome technical details. For example, this leads to the choice of KERAS instead of TensorFlow or Caffe for deep learning. We use Python as the programming language in most of the modules, as it is easy to learn and can be used in interactive mode, which provides learners an opportunity of quick iterative learning through trial and error. Similarly, we select realistic datasets that are publicly available without encumbrance (for example, not behind a paywall nor under strict licensing requirements). Thus, students gain hands-on experience of working with real data, many of which are not small in size. Finally, the hands-on problems must be sufficiently interesting and challenging without being excessively complicated.

Most of the hands-on activities were newly created or based on actual research activities at ODU. In our effort to build out our modules, we found out that there are not that many appropriate hands-on cybersecurity-focused activities that utilize CI techniques and tools. For example, in the area of big data analytics for cybersecurity application (for module 2), while there are many research papers and white papers on this very topic [12, 13, 14, 15, 16], hardly one provides the datasets alongside a clear description on their methodologies suitable for novice learners. We believe that the hands-on activities developed in this module can be useful resources for (1) people who are embarking cybersecurity research on advanced CI, and (2) instructors who want to develop cybersecurity-focused courses adopting advanced CI.

2.3 Modes of Delivery

The modules will be delivered via two distinct means: monthly workshops and summer institutes. Six monthly workshops are conducted during academic year (roughly a month apart between two consecutive workshops), primarily targeting students enrolled at ODU. The summer institutes present these six modules to students from local community colleges, Research Experiences for Undergraduates program at ODU, and other Virginia universities; they will also include special activities such as field trips, open house for K-12 students, Cyber Night events, cybersecurity career panels, and student competitions. In both the monthly workshop and summer institute formats, each module is delivered as a three-hour hands-on workshop; the first research talk is about 30 minutes long, and the remainder is divided judiciously between the CI introduction and hands-on activities. In practice, we interleave the activities into the CI introduction. All the workshop sessions are recorded. The recordings are available for students to review the workshop materials. When the program is completed, workshops and institutes will eventually serve as a gateway to a wealth of carefully curated learning resources (our own and external resources) that interested trainees should pursue on their own.

Students obtain experience of using a real supercomputer through DeapSECURE. The hands-on portions of the training modules are carried out on ODU's Turing cluster [18]. Turing is a Linux HPC cluster consisting of over 250 compute nodes, 16–32 Intel Xeon CPU cores per node, at least 128 GB RAM per node, 56 Gbps FDR Infiniband interconnect fabric, 1.2 PB shared home and long-term storage, as well as 180 TB Lustre scratch space. Turing also has over 30 graphical processing units (GPUs) from NVIDIA, ranging from K40 to V100. Turing is currently the main computational workhorse that support diverse research activities at ODU.

3 WORKSHOP EXPERIENCE, EVALUATION, FINDINGS

3.1 Workshop Development

The 2018-2019 academic year is the first cycle of DeapSECURE, during which time the training materials are also being constructed from ground zero. The Principal Investigation (PI) team consists of two faculty members (Wu and Sosonkina) and one computational scientist (Purwanto). Together, they provide complementary expertise needed to build and deliver this training program. Wu is the Batten Chair of Cybersecurity and the Director of the Center for Cybersecurity Education and Research (CCSER); CCSER directly supports ODU's cybersecurity initiative through education (workforce development), research, and strategic partnership with industry, business, and military. Sosonkina has solid expertise in HPC, including the research area of power and resilience to achieve reliable exascale computing. Purwanto's area of expertise includes computational science (physics) and HPC technologies; he also has extensive experience in HPC user support, facilitation, and training, which include organizing ODU satellite site for XSEDE Monthly HPC workshops [7, 8] since 2016 and spearheading Carpentry and Carpentry-like workshops at ODU. Arcaute provides her expertise as an independent assessor for the program. Three graduate research assistants (GRAs) support the PIs by co-developing and testing the hands-on exercises as well as

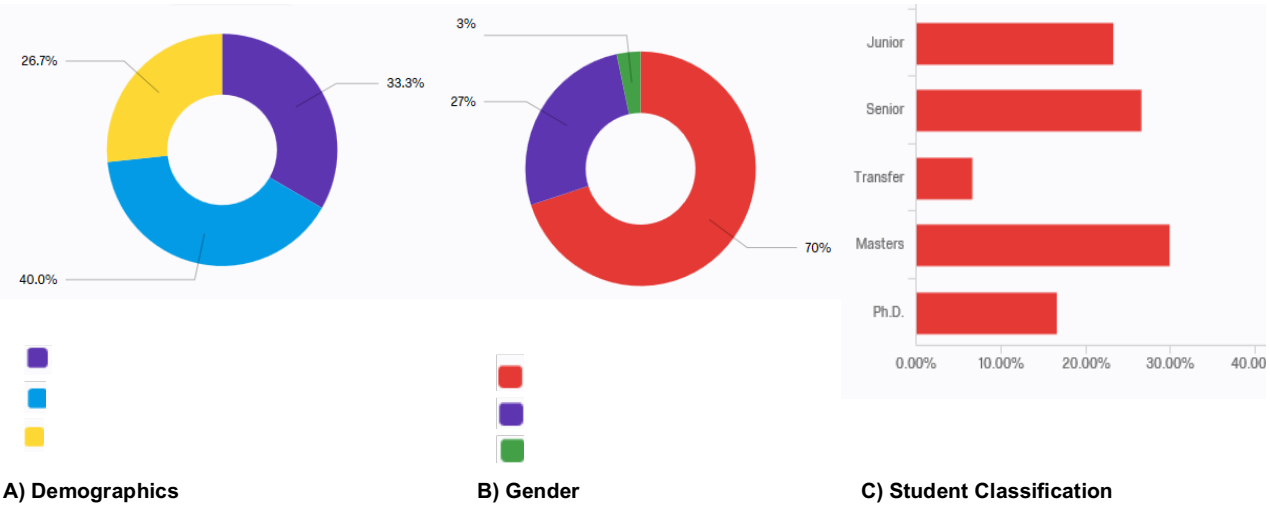

Figure 1. Participants' profile based on responses to survey.

assisting the trainees during the hands-on sessions of the workshops.

The DeapSECURE team uses collaborative technologies to build out the training materials: Gitlab for code and module contents; Google docs for sharing key documentations among the team members; Slack for responsive coordination during work sessions; as well as email communications. We use the Carpentries website template to create our training materials [13]. We work closely with the cybersecurity faculty members who present the research talks to align our CI techniques and hands-on exercises. The PIs and GRAs meet regularly on weekly basis to discuss the module contents and strategies; additional work sessions are utilized to build and test the hands-on exercises. The PIs also meet about once a month with the assessor to discuss the feedback received from workshop participants and make adjustment for the subsequent workshops.

3.2 Learner's Profile

ODU cybersecurity programs are unique in that it draws students from very diverse backgrounds: there is diversity in ethnic backgrounds (cybersecurity major enrollment includes 48% minority) as well as in academic preparation (for example: 44% military; there are also many students who are working professionals and returning to school at the same time).

For this year's training, we open the training widely to all interested participants. We received nearly 50 signups from different fields of study: nearly 50% of them from cybersecurity major, followed by electrical/computer engineering, modeling and simulation, and computer science. About 30% work full-time and another 30% works part-time, with few of them (~20%) being professionals in the information technology or cybersecurity; they are enthusiastic about learning new CI skills. About 30 actually attended one or more of the workshops.

3.3 Assessment; Student's Reaction and Feedback

We continuously improve our training, both contents and format, in response to feedback given by students. Karina Arcaute from ODU's department of STEM Education and Professional Studies is performing independent assessment of the workshops. This includes pre- and post-workshop surveys. For every workshop, the surveys consist of both knowledge assessment to evaluate students' understanding of the topics before and after the workshop, and workshop evaluation questionnaire to obtain student's feedback regarding the workshop itself. Focus group interviews will also be conducted to obtain a more in-depth feedback.

Overall, the four trainings that have taken place have been rated as excellent or very good by the majority of the participants, and the hands-on session is considered the most valuable part of the training. The response rate has ranged from 100% in the pre-training survey for the first training, to 40% in post-training evaluations. The number of attendees has ranged from 12–30 participants per workshop. As shown in Figure 1, the demographics of participants (based on 100% response survey) are as follows: 40% African-American, 33% Asian, and 27% White. Regarding gender, 70% of participants are male, 27% are female, and 3% preferred not to answer. The age distribution of participants ranges from 19 to 53, with the majority of the participants (60%) in the 18–22 and 23–27 age range. At least 45% of the participants are pursuing a graduate degree (30% Masters, 15% Ph.D.), 27% are seniors, 23% are juniors, and 7% are transfer students. The formative evaluation after each workshop has allowed us to continuously improve subsequent trainings. For example, based on the feedback from the survey, we relocated the training to an auditorium-style classroom with web-based learning capabilities, since the original location of the training was too small. We also gave student access to the

workshop slides and instruction websites, which greatly facilitated their learning both in-class and after the workshop.

The evaluation of the workshop has been very positive, with at least half of the participants considering the training very useful and worth their time. After the location change, the positive responses have increased, for example, in the first workshops only 47% of participants would definitely recommend the training to others and 27% indicating they might or might not, probably not, or definitely not recommend the training to others. In latter workshops the number of participants who would definitely recommend the workshop to others increased to 75%, with the other 25% probably recommending the training to others. The same observation is seen regarding the overall rating of the workshop: in the first workshop only 33% of participants were extremely satisfied while in the latter workshop 50% of participants were extremely satisfied and the other 50% were satisfied with the workshop. Overall in all workshops, the comments about the most valuable aspect of the training has been the hands-on session, and there are positive responses even for the open evaluation question of "*What is the least valuable aspect of the training?*" with responses being either "All is valuable" or "N/A."

There are other assessment metrics which are not discussed here, because at the time this paper was written, DeapSECURE has yet to finish its first cycle. These metrics include things such as: the effectiveness of the training method; student's learning outcomes; the training's impact on student's academic, research, as well as professional performance. We plan to delve into these metrics in another publication in near future.

3.4 Broader Benefits

From this year's training, we already observed many broader benefits to this training program in addition to the core mission of teaching CI skills to the cybersecurity student population as follows: In-person workshops afforded opportunities for the students to interact and network among themselves and with the training team. Several trainees decided to pursue HPC- and CI-related curriculum courses. For example, two of them are currently taking Sosonkina's High-Performance Computing and Simulation graduate course in Spring 2019. The GRAs for this project received cross-training in a wide variety of CI techniques by being part of the team. They also gained valuable skills in collaborative development of software and training materials. Project team's engagement with cybersecurity faculty and their graduate students provided ample cross-fertilization research opportunities. For example, the DeapSECURE team members helped one research group to discover and remove a major computational bottleneck in their research code. We believe that having this kind of informal trainings, even if the students will not eventually pursue research in cybersecurity, has the benefit of bringing campus communities together, spreading the awareness of cybersecurity research as well as CI techniques to the students.

4 FUTURE DIRECTIONS

Monthly workshops are useful to kickstart and further students' learning. We realized that for those who have no prior knowledge of advanced CI techniques this is a steep learning curve. Hence, in the future we will consider providing two broad workshop sections or instances, one for students without prior CI knowledge and another for students already equipped with some programming skills. In education, the skills being taught must be reinforced through intentional practice and repetition, as well as more engagement with the learning materials, ideally on a constant basis. Being aware of this, in the following years we will set up an online continuous learning community to complement the workshops and summer institutes. This includes a "learning tree"-style repository of additional training materials (our own materials and carefully curated pointers to external resources) for students to further pursue topics of their interest, a virtual computer lab (which in our case is the Turing cluster at ODU), a student forum, as a place for students to continue their learning engagement after the face-to-face sessions. Workshops will eventually serve as a gateway to the wealth of learning resources that interested trainees can and should pursue on their own. Students in the cohort will become part of a learning community, where they can further learn in different modes (whether moderated or independent). Online Q&A and/or discussion forum will be leveraged to facilitate communication. In a moderated approach, for example, a moderator can lead students to look at a challenging cybersecurity problem together, propose and attempt an approach to solve the problem as a group. The discussions generated will draw upon the things they have learned in the workshop as well as from the online resources. In independent mode, a learner may already have a particular problem to solve, and pick certain modules to learn particular a CI technique that can be used to address his/her specific problem. They will have freedom as to what topics they will learn, and how much. Trainees who have real-world project or research to tackle should become motivated to learn, as this repository will become resources available at their fingertips. Archived workshop materials as well as additional learning materials are also posted on this online platform as open educational resources, to be made available to the research and education communities at large. The open-source style development of the learning modules facilitates a wide range of adoption, adaptations, and contributions in an efficient manner.

5 SUMMARY

In this paper we described DeapSECURE, a non-degree training program developed at Old Dominion University (ODU) that introduces CI technologies in the context of cybersecurity research. We developed six modules to expose cybersecurity students to advanced CI platforms and techniques rooted in big data, machine learning, neural networks, and high-performance programming. Each workshop includes a cybersecurity research presentation which provides the motivation and context for a CI technique, which is then examined during a hands-on session. The modules are delivered through (1) monthly workshops for

ODU students, and (2) summer institutes for students from nearby universities and Research Experiences for Undergraduates participants. The modules adopt a hands-on approach based on real-world cybersecurity problems. The materials developed by this project will be made available using an open-source license to encourage a wide range of adoption, adaptation, and contributions. The project leverages existing and new partnerships to ensure broad participation, and accordingly broaden the adoption of advanced CI techniques in the cybersecurity community.

ACKNOWLEDGMENTS

The development of DeapSECURE training program is supported by NSF CyberTraining grant #1829771. We thank our graduate research assistants for their hard work to support the success of this project: Issakar Doude (Modeling, Simulation, and Visualization Engineering), Qiao Zhang (Electrical and Computer Engineering), Rui Ning (Electrical and Computer Engineering). DeapSECURE training workshops utilize the Turing HPC cluster provided by ODU Research Computing Services, part of Information Technology Services. We also thank the ODU Distance Learning for their support in recording the workshop sessions.

REFERENCES

[1] National Initiative for Cybersecurity Education. 2019. Cyberseek heatmap. http://cyberseek.org/heatmap.html. Retrieved April 5, 2019.

[2] Jason R. Hamlet, Curtis M. Keliiaa. 2010. Assessment of current cybersecurity practices in the public domain: cyber indications and warnings domain. Sandia National Laboratory Technical Report SAND2010-4765. https://doi.org/10.2172/992337.

[3] Jason R. Hamlet, Curtis M. Keliiaa. 2010. National cyber defense high performance computing and analysis: concepts, planning and roadmap. Sandia National Laboratory Technical Report SAND2010-4766. https://doi.org/10.2172/992325.

[4] Qiao Zhang, Cong Wang, Hongyi Wu, Chunsheng Xin and Tran V. Phuong. 2018, GELU-Net: A Globally Encrypted, Locally Unencrypted Deep Neural Network for Privacy-Preserved Learning. In Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI), 3933–3939 (2018). https://doi.org/10.24963/ijcai.2018/547.

[5] Greg Wilson. 2006. Software Carpentry: Getting Scientists to Write Better Code by Making Them More Productive. Computing in Science & Engineering 8(6), 66–69 (Nov–Dec 2006). https://doi.org/10.1109/MCSE.2006.122.

[6] Software Carpentry. 2019. https://software-carpentry.org/. Retrieved April 8, 2019.

[7] Pittsburgh Supercomputing Center. 2019. XSEDE HPC Workshop Series. https://psc.edu/xsede-hpc-series-all-workshops. Retrieved April 8, 2019.

[8] John Urbanic and Thomas Maiden. 2018. Evaluating the Wide Area Classroom After 10,500 HPC Students. Presented at EduHPC-18: Workshop on Education for High-Performance Computing (Dallas, TX). https://grid.cs.gsu.edu/~tcpp/curriculum/sites/default/files/JohnUrbanic_1.pdf.

[9] Yang Lu, Li Da Xu. 2018. Internet of Things (IoT) Cybersecurity Research: A Review of Current Research Topics. In IEEE Internet of Things Journal (Early Access). https://doi.org/10.1109/JIOT.2018.2869847. Retrieved April 8, 2019.

[10] Peng Jiang, Hongyi Wu, Cong Wang, Chunsheng Xin. 2018. Virtual MAC Spoofing Detection through Deep Learning. In Proceedings of 2018 IEEE International Conference on Communications (ICC), 1–6 (2018). https://doi.org/10.1109/ICC.2018.8422830.

[11] Weize Yu, Selçuk Köse. 2017. A Lightweight Masked AES Implementation for Securing IoT Against CPA Attacks. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 64(11), 2934–2944 (2017). https://doi.org/10.1109/TCSI.2017.2702098.

[12] Raghul Gunasekaran, Sarp Oral, David Dillow, Byung Park, Galen Shipman, Al Geist. 2011. Real-Time System Log Monitoring/Analytics Framework. Presented at Cray User Group Conference (Fairbanks, AK). https://www.osti.gov/biblio/1056901 and https://pdfs.semanticscholar.org/9efa/4559c6a3ccf899bb49f668f67e1214c54e0f.pdf.

[13] Cloudera, Inc. 2018. FireEye: uncovering zero-day and advanced persistent threats more quickly. https://www.cloudera.com/about/customers/fireeye.html. Retrieved December 25, 2018.

[14] Min Du, Feifei Li, Guineng Zheng, Vivek Srikumar. 2017. DeepLog: Anomaly Detection and Diagnosis from System Logs Through Deep Learning. In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, 1285–1298 (2017). https://doi.org/10.1145/3133956.3134015.

[15] Faheem Ullah, M. Ali Babar. 2018. Architectural Tactics for Big Data Cybersecurity Analytic Systems: A Review. https://arxiv.org/abs/1802.03178. Retrieved April 8, 2019.

[16] Alvaro A. Cárdenas, Pratyusa K. Manadhata, Sreeranga P. Rajan. 2013. Big Data Analytics for Security. IEEE Security & Privacy, 11(6), 74–76 (Nov–Dec 2013). https://doi.org/10.1109/MSP.2013.138.

[17] The Carpentries. 2018. Styles for the Carpentries lessons. https://github.com/carpentries/styles . Retrieved October 10, 2018.