DEAPSECURE COMPUTATIONAL TRAINING FOR CYBERSECURITY: THIRD-YEAR IMPROVEMENTS AND IMPACTS

Bahador Dodge
Jacob Strother
Rosby Asiamah
Karina Arcaute
Wirawan Purwanto
Masha Sosonkina
Hongyi Wu

Old Dominion University
5115 Hampton Boulevard
Norfolk, VA, USA
{bdodg001, jstro005, rasia003, karcaute, wpurwant, msosonki, h1wu}@odu.edu

ABSTRACT

The Data-Enabled Advanced Training Program for Cybersecurity Research and Education (DeapSECURE) was introduced in 2018 as a non-degree training consisting of six modules covering a broad range of cyber-infrastructure techniques, including high performance computing, big data, machine learning and advanced cryptography, aimed at reducing the gap between current cybersecurity curricula and requirements needed for advanced research and industrial projects. By its third year, DeapSECURE, like many other educational endeavors, experienced abrupt changes brought by the COVID-19 pandemic. The training had to be retooled to adapt to fully online delivery. Hands-on activities were reformatted to accommodate self-paced learning. In this paper, we describe and assess the third and fourth years of the project and compare them with the first half of the project, which implemented in-person instruction. We also indicate major improvements and present future opportunities for this training program to advance the cybersecurity field.

Keywords: big data, machine learning, neural networks, high performance computing training, cybersecurity

1 INTRODUCTION

Modeling and simulation (M&S) permeates the research topics investigated in this paper, such as training (Purwanto et al. 2021, Purwanto et al. 2019), cybersecurity, high-performance computing (HPC), decision-making (Collins and Etemadidavan 2021), and data science. M&S gives theoretical and practical insight to the corresponding fields of study. In this paper, we introduce a recently developed training program targeting M&S, HPC, and data science methodologies in the field of cybersecurity.

The age of cybersecurity goes back to the 1970s when researcher Bob Thomas created the first computer program called Creeper that could move across ARPANET's computers and leave a trail behind. Because of this program, Ray Tomlinson wrote the first example of an antivirus that chased the Creeper and deleted

MSVSCC'22, April 18-21, 2022, Suffolk, VA, USA; ©2022 Society for Modeling & Simulation International (SCS)

it (Chadd 2020). Before the late 1950s, when the first computer network was built for the U.S. military, most computers were huge in size, locked in secure places, and only a handful of scientists or computer programmers had access to them. Today, almost every computer and mobile phone in the world is connected to the Internet, and the amount of data being generated yearly is staggering. The security of these computer systems and the data contained therein becomes an urgency, as today's malicious attacks are becoming more pervasive and sophisticated.

There have been many records of cyberattacks after the 1970s in the age of the connected world. Some of them had dire consequences and major disruptions on daily life of people, organizations and governments. These attacks took many different forms, including Malware, Ransomware, Distributed Denial of Service (DDoS), Spam and Phishing, Corporate Account Takeover (CATO), Automated Teller Machine (ATM) cash out, etc. On May 7, 2021, the Colonial Pipeline, which carries gasoline, diesel oil and jet fuels from Texas to New York, suffered a ransomware attack by a group named DarkSide, causing a temporary but major economic disruption. The primary target of the attack was the billing system of the company. Although the pipeline itself was not the target, the company shutdown the pipeline due to inability to bill customers. The consequence of the ransomware attack on eastern shore of the U.S. was enormous and caused major disruptions from empty gas pumps to higher gas prices. The hackers did not have access to Colonial Pipeline's physical infrastructure locally; the hacking originated from outside of the U.S via Internet access to the company's major IT infrastructures. The attackers also stole 100 gigabytes of company data in the span of a two-hour window. Colonial Pipeline hacking shows the importance and vulnerabilities of the connected world.

In the era of the connected world, the importance of cybersecurity and its role in saving and protecting people, infrastructure, and countries cannot be underestimated. However, the task of defending cyber systems involves finding "needles in a haystack". Forbes estimated that the amount of data generated each day is about 2.5 quintillion bytes (about 1.7 megabytes of data per second per person). In contrast, many sophisticated computer viruses do not have more than 2 megabytes of size, like Stuxnet virus. Smithsonian magazine puts Stuxnet at number one on their list of the top ten most sophisticated and destructive computer viruses (Weinberger 2012). The phrase big data has been in use since the 1990s and is characterized by terms such as variety, veracity, and velocity. It is inefficient, if not impossible, to process data or events on the order of gigabytes, terabytes, or petabytes with the latest desktop and laptop computers; instead, processing big data requires advanced computing platforms such as HPC systems, as well as advanced computational tools and techniques. These are important constituents of the greater advanced cyberinfrastructure (CI) ecosystem, upon which state-of-the-art M&S and data science research takes place today (UITS 2019, Stewart et al. 2010).

At present, the cybersecurity curricula used in many colleges and universities lack inclusion of advanced CI techniques to strengthen cybersecurity analysis, research, and development. Recognizing this shortage, Old Dominion University (ODU), particularly the School of Cybersecurity, developed DeapSECURE (short of Data-Enabled Advanced Computational Training Platform for Cybersecurity Research and Education), an innovative, non-degree cybersecurity training program, funded by the National Science Foundation.

The DeapSECURE training program was created to address major curricular gaps in cybersecurity education in the areas of advanced computing. This non-degree training program consists of six modules that covers high performance computing (HPC), big data analytics (BD), neural networks (NN), machine learning (ML), parallel programming (PAR) and cryptography for privacy-preserving computation (CRYPT). These techniques are used extensively in state-of-the-art cybersecurity research and practice. Application examples in the modules are carefully selected to engage learners in the field of cybersecurity and aim to train current and future researchers, engineers and practitioners with advanced techniques and skills necessary to carry out cybersecurity research and industrial projects.

1.1 DeapSECURE Modules

Advanced CI techniques include processing, cleaning and visualization of big data, development of machine learning models as well as M&S algorithms, which are tuned for the best performance on the target architecture. They require the use of cloud and HPC with parallel programming. The DeapSECURE training program (DeapSECURE Team 2019) draws upon the importance of these CI techniques in state-of-the-art cybersecurity research, and develops lessons and hands-on sessions to introduce these topics to students and learners. This program aims at introducing undergraduate and graduate students to a well-rounded set of advanced CI techniques, primarily through hands-on learning activities, to address cybersecurity challenges in their respective fields of work, with the ultimate goal of introducing the developed methodologies and hands-on modules into formal curricula of cybersecurity, HPC, and M&S disciplines.

Table 1 outlines the six modules developed for DeapSECURE by describing their focus (in Column 3) and naming the software tools used in the hands-on exercises (in Column 4). In designing a sequence of workshops based on these modules, we may take a subset of these modules depending on the learners' interests and available time.

Table 1: DeapSECURE lesson modules.

Item	Module Name	Description	Toolkit Used
1	Introduction to HPC ("HPC")	Introduction to HPC systems and how to access, use and program such systems	UNIX shell commands, SLURM
2	Dealing with Big Data ("BD")	Processing, cleaning, analyzing, and visualizing big data sets	Year 1: Python, PySpark Year 2 onward: Python, Pandas, Matplotlib, Seaborn
3	Machine Learning ("ML")	Introduction to the concept of and building machine learning models, and analysis of the result	Python, Scikit-Learn library
4	Deep Learning using Neural Networks ("NN")	Introduction to advanced machine learning models using deep neural networks	Python, KERAS library
5	Cryptography for Privacy-Preserving Computation ("CRYPT")	Introduction to advanced, privacy-preserving encryption and decryption techniques	Python, AES and Python-Paillier libraries
6	Parallel and High-Performance Programming ("PAR")	Parallel programing utilizing HPC	Python, mpi4py library

Our objective behind varying the number of workshops and their focus in the training sequence is to imbibe each module with strong standalone characteristics, so that each may be more versatile and geared towards future usages in curricular and customized module assembly. DeapSECURE emphasizes hands-on experiences in the practical tools by which the CI techniques can be applied to real-world problems.

1.2 The First Year

In the first year (academic year 2018–2019, or “Y1”), the six modules were created from scratch. A significant effort was put into engaging cybersecurity researchers at ODU to compile a list of interesting cybersecurity problems and datasets as the target applications to teach the tools and techniques to the students. The training materials were developed collaboratively between faculty and teaching assistants (TAs) using Gitlab for codes, lessons, and data repositories, as well as Google Docs for workflow coordination among team members. Three Ph.D. students served as TAs in the development of the lessons and the hands-on parts of the workshops.

Each module was taught as a workshop with a three-hour duration. A workshop started with a 30-minute introduction of cyber research related to the CI topic, followed by the teaching of CI techniques necessary and hands-on work to further expose the topic of the workshop to students (Purwanto et al. 2019). The lessons taught in the workshop were published on the web on the Gitlab platform. Pre- and post-workshop surveys were conducted to generate feedback and statistical data for improvement of the materials and workshop delivery. As borne by the survey responses, the workshops were generally well-received by the students, and they were happy learning new computational techniques and tools beyond what they already knew (Purwanto et al. 2019). In the first year, however, many participants struggled with the hands-on part of the training. These activities assumed strong command-line and programming skills, which were lacking in many participants (Purwanto et al. 2021).

1.3 The Second Year

In the second year (2019–2020, “Y2”), major changes were done based on our assessment of Y1, and three modules were rewritten and/or retooled (Purwanto et al. 2021). As shown in Table 2, the modules were categorized as either compute- or data-intensive. The compute-intensive modules were taught in the Fall of 2019 and the data-intensive ones in the Spring of 2020. A second major change in Y2 is focusing just on three research themes: (1) spam processing, (2) homomorphic data encryption, and (3) mobile device security, instead of having a distinct storyline (theme) in each of the six modules as was done in Y1. With the new three main foci, it became easier to concentrate on the CI techniques rather than spending time introducing a research theme. For example, the three data-intensive modules were retooled with Pandas, a reputable Python library for data processing, and completely rewritten to utilize the SherLock dataset (Mirsky et al. 2016), and take the security of mobile devices as their overarching theme.

Table 2: Workshop Categories.

Compute-intensive	**Data-intensive**
Introduction to HPC	Dealing with Big Data
Cryptography for Privacy-Preserving Computation	Machine Learning
Parallel and High-Performance Computing	Deep Learning using Neural Networks

The third major change was the introduction of “hackshops” to further challenge the skills learned by students through the workshops. The hackshops were unstructured, additional hands-on sessions to solve predetermined challenge problems in small groups with guidance from the TAs.

Similar to the first year, survey outcomes were successful, rated “extremely good” or “good” by 80% of the attendants. The most valuable assets of the workshops based on the student feedback were students’ exposure to hands-on sessions. The new and reorganized modules enabled them to gain more understanding of cybersecurity and its importance from different perspectives.

2 YEAR THREE OF DEAPSECURE

The COVID-19 pandemic hit at the end of the second year; this compelled the team to rework the workshop materials for online delivery. The team conducted a pilot online workshop in August of 2020. The third year (Y3, academic year 2020-2021) of this training program built on the success of this pilot workshop. The online format used (1) Zoom, a video conferencing application, for face-to-face meeting and instruction; (2) Jupyter, a web-based interactive development environment for Python; and (3) Slack, a group-based messaging and communications platform. Zoom's breakout room feature was used to carry out the hands-on instruction in small groups, where the hands-on activities were prepared using Jupyter notebooks. Much effort was spent on preparing these notebooks, which were an adaptation of the web-based lesson modules. Slack was used as a hub that enabled the participants and project team members to directly communicate before, during, and after the workshop. The hackshops were not conducted in Y3 due to limited resources and time.

2.1 Lesson Development

The start of Y3 brought in two undergraduate students as new TAs to the DeapSECURE team, in addition to the more experienced graduate students. Initially, these undergraduate students took the role of new learners who had never seen the lesson materials; this allowed the team to receive valuable feedback from a learner's perspective before the actual workshop. While preparing for each workshop, the faculty exposed the undergraduate students not only to the actual contents, but to the development strategies. Eventually, this would graduate their role from learners to developers.

The primary development tools used were Git, Jekyll, Unix/Linux, and Jupyter—the same tools used in the first two years. Git was used to create and manage repositories for each lesson. The TAs would fork and clone each repository to allow for their own local changes, which could be merged to the master branch. Jekyll, a static-site generator tool, was used to generate the lesson websites. One main component of the training development was the use of the Linux cluster environment on ODU's Wahab cluster, which was essential for heavy computations via shell scripts and/or python programs.

The core of our hands-on development process relied heavily on the use of Jupyter notebooks. First, a master outline was constructed for each module to define its overall structure and goals. Two to three notebooks per module were prepared for learners based on the outline. (The HPC lesson was an exception, which was purely carried out on remote Linux terminal sessions.) These notebooks began as developer's notebooks, which contained experimental codes and research notes. These were further refined to become the hands-on codes and activities (which were intentionally made incomplete, to be worked out by the learners during and after the workshops), accompanied by appropriate instructional texts, and the worked-out solutions by the TAs. Typically, each TA was assigned to develop a particular notebook or lesson episode, after which one of the faculty would review the development done and make any polishing changes.

The worked-out solutions and extraneous research parts were removed from the developer's notebooks using a custom script to become the final notebooks for the learners. The research and development efforts to create the developer's notebooks, as well as to structure the materials for optimal teaching and learning experience, were the most time-consuming part of Y3's work. The texts, hands-on codes, etc. were mostly based on the materials published on the lesson websites, but they also closed the gaps that were still present after the second year of lesson development. The texts in the Jupyter notebooks were meant to convey only the most salient points from the lesson websites, not to replace them. Learners were still encouraged to review the lesson websites for more detailed information and hands-on activities/challenges.

2.2 Conducting Workshops

All six modules were offered as six workshops in Y3, just as in the previous years. The training program was announced widely to the University student body, and participants were required to register to attend the workshop series. Each workshop included a main presentation broken down into one-hour blocks, which included hands-on breakout sessions for the learners to work on the Jupyter notebooks. The cybersecurity research presentation was shortened and included as an optional part of the introduction to the CI method, instead of featured separately at the beginning of every workshop. These became a background introduction on the practical uses of the current module's topic. Breakout rooms were split into three main categories: beginner/novice, intermediate, and advanced. Initially, participants were assigned to each room based on the self-assessed skill level inferred from their registration data; however, this was later changed to allow participants the freedom to join any breakout room, with encouragement from the TAs to move to a different room if the pacing is too slow or too fast. The Kahoot platform was also used for interactive quizzes at the end of each session.

The hands-on learning in Y1 and Y2 was conducted by following the instructor's shared screen, while in Y3, learners were given Jupyter notebooks to work with independently. Initially, learners were left to explore on their own—resulting in slow progress and a potential of getting stuck on less essential issues. Later on, during the summer workshops, the TAs adopted a guided presentation approach by working out the notebook with the learners while highlighting the important parts. Learners were encouraged to work ahead if they desired, as there was usually not enough time to cover everything in the notebook; this left plenty of learning material for learners to try after the workshop if interested. Overall, it was difficult to get the participants to engage with others, instead they would prefer to work alone and in silence.

Figure 1 compares the attendance for each workshop among the three years. The workshops in Y1 and Y2 were carried out during the academic years (Fall and Spring semesters). "SI1" refers to the Summer Institute held at the end of Y1 as a weeklong event. The first three workshops in Y3 were carried out during the Fall/Spring semesters and the last three as a summer workshop series in three consecutive days. Generally, around 10 to 30 participants attended each workshop with the high points on the graph indicating when registration was open. In Y3, registration was open during three separate workshops: HPC, PAR, and BD, thus there was a spike in attendance at those points.

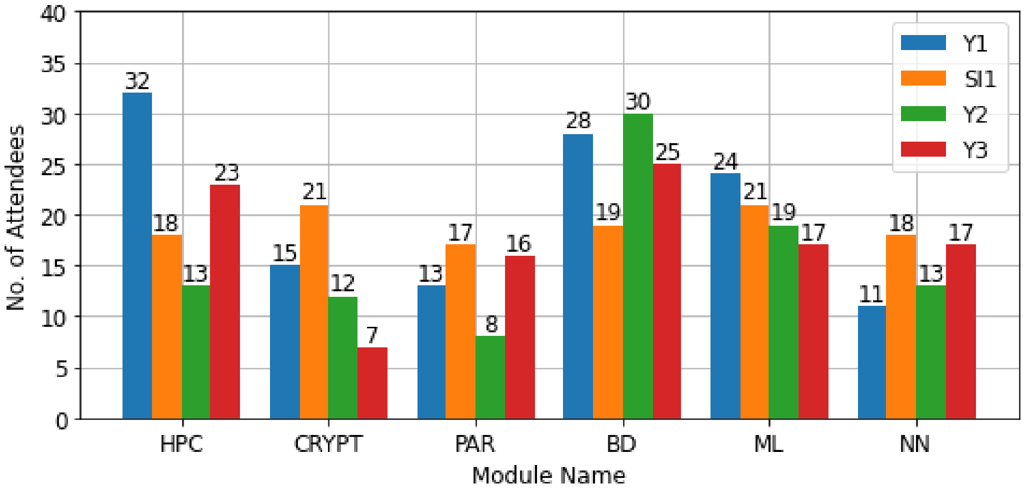

Figure 1: Number of attendees for each workshop during Y1, Y2, Y3 workshop series and SI1 (2019 Summer Institute).

A major problem throughout the three years was the attendance retention across the series in a year. Our data indicates that multi-day workshop series such as the Summer Institute format is better for attendance retention because they appear to provide continuity and focus of learning condensed in a shorter timespan. For example, Y3's summer workshop series (teaching BD–NN modules) started with 25 attendees, then decreased to a constant 17. Participants were more likely to attend the workshops during the summer compared to during the academic school year due, most likely, to having more free time for a non-degree training. There are also statistics shown in Figure 2 that include the total counts of the workshops attended by attendees, based on training year (higher participant counts for 5-6 workshops indicate better overall retention). Overall, the trend is the same: the summer institutes from Y1 and Y3 had high attendance rates (between 17–21).

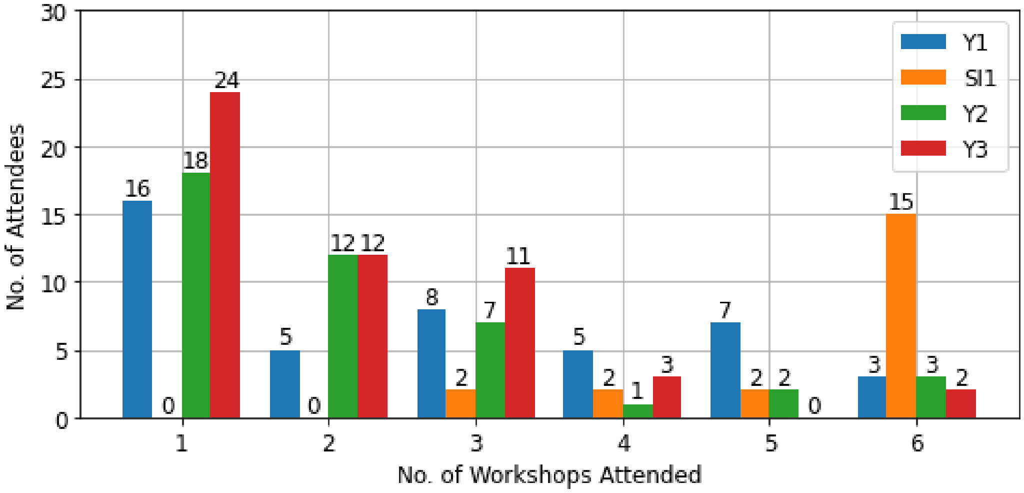

Figure 2: The distribution of participants based on the total numbers of workshops attended, separated for Y1, Y2, Y3, and SI1.

2.3 Accomplishments in Module Development

A major accomplishment in Y3 of DeapSECURE is the completion of the *deliverable contents* of each module. In the HPC module, there is now a sharp focus on the parallel processing of spam emails using a small set of basic Unix commands. In the cryptography module, a new challenge was introduced on brute-force decryption of various encrypted messages. For the parallel programming module, the conceptual differences between serial and parallel computing were identified; there was also an introduction to basic message-passing parallel programming building blocks with a demonstration of converting a computationally expensive serial program to a parallel program with improved timing performance. Next, in the machine learning module, the complete workflow of machine learning was introduced starting from the raw data to a deployable machine learning model capable of distinguishing two smartphone applications. Lastly, in the neural networks module, the basic concepts and implementation of neural network models were significantly elucidated.

2.4 Attendance Profile

Figure 3 shows the distribution of workshop participants based on their academic major. With no surprise, computer science (CS), electrical and computer engineering (ECE), and cybersecurity (CYSE) are the top three majors for the attendants at 35%, 29%, and 12% respectively. Other majors that are less prevalent include modeling and simulation engineering (MSIM) and data science (DATA). The OTHER category contains non-STEM majors and STEM majors representing less than 2% of participants like math and physics.

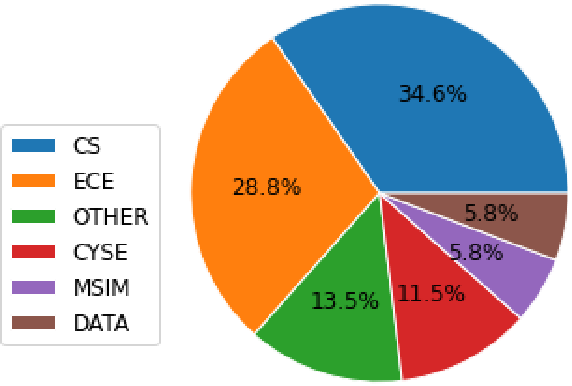

Figure 3: Distribution of Y3 workshop participants categorized by their academic majors.

During the registration process, participants answered many questions. They were asked to self-identify their skill levels (none, novice, intermediate, or expert) on Unix, Python, and C/C++. Figure 4 shows the results of this questionnaire. Many participants are novice or intermediate in each programming tool, with a considerably high number of intermediates for C/C++; this is likely due to C++ being taught in a required course for Engineering and computer-related majors at ODU.

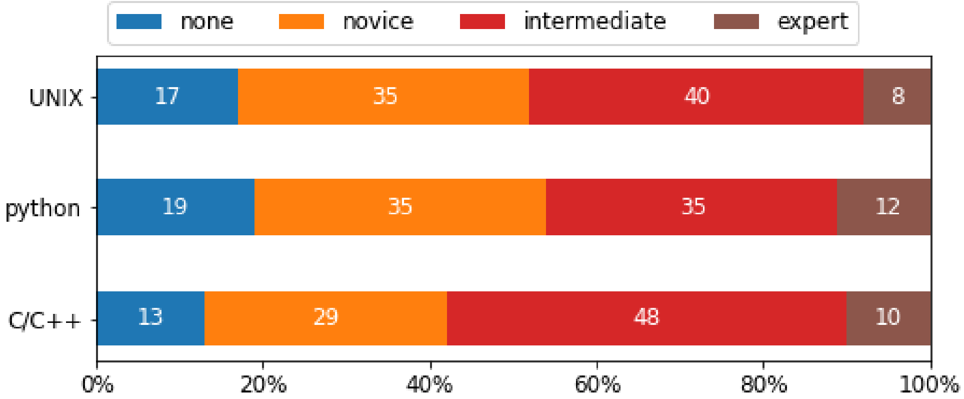

Figure 4: The distribution of the participants' self-assessed skills in key computing languages: Unix (shell), Python, and C/C++.

2.5 Feedback

We systematically gathered pre- and post-workshop evaluation and feedback from workshop participants. Besides the typical feedback such as ratings on various aspects of the workshop and materials, the team also utilized a series of knowledge questions to evaluate participants' understanding of the CI topics. Unfortunately, it was difficult to obtain a high response rate due to several factors, such as fully remote delivery. However, the responses gathered were helpful and instructive. In the HPC workshop, four participants noted in the post-survey that the workshop felt too slow while five participants noted that the pacing was too fast, with one of them stating that there should be a completely separate workshop for Unix. The CRYPT and PAR workshops also had similar constructive criticism: there needs to be more time and instruction allocated to the hands-on learning portion.

The three data-intensive workshops were held in June 2021 in a format very close to the 2019 Summer Institute; these workshops provided an opportunity to make iterative improvements based on the daily feedback in addition to those received from the previous three workshops. As a result, more instruction and guidance were given to the participants during the hands-on sessions. The way the notebooks were crafted and presented led to positive feedback from the participants, with one participant noting that "having a TA explain the steps and outputs" of the notebook was a key aspect that worked well in the training. Another participant provided some constructive feedback for what did not work well: "there was too much content in such a short amount of time." This suggests that the total workshop time may need to be increased, more time needs to be allocated for the hands-on materials, or the content amount needs to be decreased.

2.6 Lessons Learned by Teaching Assistants

TAs contributed significantly to the workshop from lesson planning and development to teaching in the hands-on sessions. During the course of the project, through the many challenges that required hard work and dedication to overcome, TAs acquired many skills that benefit them individually and professionally.

The TAs were trained to develop academic lessons in a research-oriented computational environment. The TAs learned all aspects of developing lessons from the planning and outlining stages through the proofreading and completion stages. Through the lesson development process, TAs learned three notable skills: (1) technical writing, (2) working as a team, and (3) software development. The TAs learned technical writing by contributing to the Jupyter notebooks and lesson websites. They also learned effective team work using Git version control system, shared Google workspace and other communication tools. There was also a peer editing strategy: one TA would write one part of the lesson and another TA would write a different part after which they would switch to review and edit each other's work. Lastly, due to the programming-intensive nature of the workshops, the TAs gained much experience in using programming tools like Git, Python, and shell. Almost all the lessons use Python and its major libraries, which are key components to the DeapSECURE's advanced training. TAs were also exposed to web development tools such as Markdown, Jekyll, HTML, and CSS. TAs also learned software development practices, which were taught mainly by examples and through much practice, such as: writing readable codes, adding meaningful comments, creating useful documentation, and writing meaningful commit messages. These skills are highly utilized in the tech industry today. Much of the software development was conducted in a supercomputing environment, which trained the TAs to be comfortable with the Unix terminal-based environment, scripting in various languages, and the Slurm job scheduler.

Another major skill is related to the core CI topics covered in the DeapSECURE lesson modules. Although all TAs had engineering or computer science backgrounds, the advanced nature of CI techniques requires significant time and effort to master the materials in order to help learners in the breakout sessions. Frequent regular meetings—twice a week—help reinforce the TAs' familiarity and understanding of the key concepts in CI topics such as big data, machine learning, etc. Eventually, TAs became proficient to teach parts of the lessons or hands-on activities to other students in the workshops.

2.7 Reinforcing CI Technique Competence

Through the experiences of teaching DeapSECURE lesson materials, both to the workshop participants and the teaching assistants, we discovered the best way to build up the computing skills in these cross-cutting areas of advanced CI and cybersecurity. Students are impacted by DeapSECURE in two different ways: as novice learners or as teaching assistants. New learners were initially exposed to the advanced CI topics by joining the workshop series taught by the DeapSECURE team. More importantly, there is a pathway for a subset of these students to gain further proficiency by being intensively trained to become TAs, or instructors, on the training project. In the lesson development process, the TAs are mentored by the faculty through regular interactions and direct work using the CI techniques. This leads to students' solid understanding of and competence in the CI topics, as well as skills in academic writing, team work, software development, and teaching (instruction). These two ways underline the need for students to pass through the CI topics more than once, with increasing level of intensity and practice, in order to gain CI skill proficiency that would become useful to advance state-of-the-art cybersecurity research.

3 FUTURE WORK

The DeapSECURE lesson and hands-on materials are available openly and freely on the Internet (DeapSECURE Team 2019). Currently, the team is in the process of finalizing and releasing all the materials as open-source educational resources which can be adapted and tailored to specific use cases in cybersecurity and other M&S or data-intensive disciplines. Currently, the big data lesson has been released as an open-source educational resource (DeapSECURE Team 2019). The other modules are being released in the same way. Given the extensive breadth of CI techniques in cybersecurity, we have not exhausted all the possible basic CI skills in the current incarnation of the lesson materials. For example, a dedicated module could be created with the focus on cybersecurity M&S (Kavak et al. 2021), which is important, for example, in the estimation of cyber risks.

The overwhelmingly positive feedback we received from the workshops indicates that DeapSECURE should continue beyond the initial funding period by NSF. Specifically, we consider that an advanced level of the training targeting instructors, teaching assistants, and lesson developers to be invaluable for their professional development, as well as producing the next-generation cybersecurity researchers and practitioners that are competent in taking advantage of M&S (Collins et al. 2020) and data science techniques to obtain insight about different scenarios and outcomes, especially when it is impossible to do so by traditional research methods (Etemadidavan and Collins 2021, Collins et al. 2021). For this reason, our future work will focus on this advanced training by leveraging the current teaching assistant training framework we have developed so far. We are also exploring ways to disseminate the awareness of cybersecurity research and education communities of the DeapSECURE training, to build community of users and contributors to further advance this training program to impact greater communities beyond the project's home institution (ODU). This will position DeapSECURE as a leader in the CI/M&S training for current and aspiring cybersecurity researchers. The hands-on focused training framework could also be extended to many other disciplines where M&S and data science play an indispensable role.

4 CONCLUSION

DeapSECURE is a non-degree training program that promulgates advanced CI techniques that are crucial to cutting-edge cybersecurity research. DeapSECURE was created to fill the gap in current cybersecurity curricula in the area of advanced CI techniques to strengthen cybersecurity analysis, research, and development. Six modules were created, covering high performance computing, big data analytics, neural networks, machine learning, parallel programming and cryptography for privacy-preserving computation. Application examples were carefully selected to engage learners in the field of cybersecurity and aimed to train current and future researchers, engineers and practitioners with advanced techniques and skills necessary to carry out cybersecurity research and industrial projects. DeapSECURE emphasizes the hands-on experience of these techniques. The lessons have undergone both major and iterative revisions through the three workshop series and one summer institute. The DeapSECURE lesson modules and hands-on materials are now available as open educational resources to the cybersecurity education and research communities. These modules can be used with or without adaptations for instructions (workshops, courses) as well as self-paced learning. By supplementing the existing university cybersecurity curricula, the lesson materials and hands-on training strategy developed by the DeapSECURE project may serve as an effective way to infuse literacy and competence in advanced CI and M&S techniques into current and future research workforce to advance state-of-the-art cybersecurity research and development programs. Efforts are underway to build a community around the training program, with a view to enhancing and expanding the impact of the training to benefit greater cybersecurity research and education communities beyond ODU.

ACKNOWLEDGMENTS

The development of DeapSECURE training program is supported by NSF OAC/CyberTraining grant #1829771. The workshops utilized Turing and Wahab HPC clusters provided by ODU Research Computing Services, part of Information Technology Services. The Wahab cluster was acquired in part using NSF Major Research Instrumentation grant #1828593.

REFERENCES

Chadd, Katie 2020. “The History of Cybersecurity”. https://blog.avast.com/history-of-cybersecurity-avast.

Collins, A. J., and S. Etemadidavan. 2021. “Interactive Agent-Based Simulation for Experimentation: A Case Study with Cooperative Game Theory”. *Modelling* vol. 2 (4), pp. 425–447.

Collins, A. J., S. Etemadidavan, and W. Khallouli. 2021. “Generating Empirical Core Size Distributions of Hedonic Games Using a Monte Carlo Method”. *International Game Theory Review* vol. 0 (0), pp. 2250001.

Collins, A. J., S. Etemadidavan, and P. Pazos-Lago. 2020. “A Human Experiment Using a Hybrid Agent-Based Model”. In *2020 Winter Simulation Conference (WSC)*, pp. 1016–1026.

DeapSECURE Team 2019. “DeapSECURE – Data Enabled Advanced Training Platform for Cybersecurity Research and Education – Project Website”. https://deapsecure.gitlab.io/.

Etemadidavan, S., and A. J. Collins. 2021. “An Empirical Distribution of the Number of Subsets in the Core Partitions of Hedonic Games”. *SN Operations Research Forum* vol. 2 (4), pp. 1–20.

Kavak, H., J. J. Padilla, D. Vernon-Bido, S. Y. Diallo, R. Gore, and S. Shetty. 2021. “Simulation for Cybersecurity: State of the Art and Future Directions”. *Journal of Cybersecurity* vol. 7 (1). tyab005.

Mirsky, Y., A. Shabtai, L. Rokach, B. Shapira, and Y. Elovici. 2016. “SherLock vs Moriarty: A Smartphone Dataset for Cybersecurity Research”. In *Proceedings of the 2016 ACM Workshop on Artificial Intelligence and Security*, AISec ’16, pp. 1–12. New York, NY, USA, Association for Computing Machinery.

Purwanto, W., Y. He, J. Ossom, Q. Zhang, L. Zhu, K. Arcaute, M. Sosonkina, and H. Wu. 2021. “DeapSECURE Computational Training for Cybersecurity Students: Improvements, Mid-Stage Evaluation, and Lessons Learned”. *The Journal of Computational Science Education* vol. 12, pp. 3–10.

Purwanto, W., H. Wu, M. Sosonkina, and K. Arcaute. 2019. “DeapSECURE: Empowering Students for Data- and Compute-Intensive Research in Cybersecurity through Training”. In *Proceedings of the Prac-*

tice and Experience in Advanced Research Computing on Rise of the Machines (Learning), PEARC '19. New York, NY, USA, Association for Computing Machinery.

Stewart, C. A., S. Simms, B. Plale, M. Link, D. Y. Hancock, and G. C. Fox. 2010. "What is Cyberinfrastructure". In *Proceedings of the 38th Annual ACM SIGUCCS Fall Conference: Navigation and Discovery*, SIGUCCS '10, pp. 37–44. New York, NY, USA, Association for Computing Machinery.

UITS 2019. "About Cyberinfrastructure". https://kb.iu.edu/d/auhf.

Weinberger, Sharon 2012, Mar. "Top Ten Most-Destructive Computer Viruses". https://www.smithsonianmag.com/science-nature/top-ten-most-destructive-computer-viruses-159542266/.

AUTHOR BIOGRAPHIES

BAHADOR DODGE is a Modeling and Simulation (M&S) graduate student at Old Dominion University. He has a Bachelor's degree in Electrical Engineering. His research interests are broad in the field of modeling and simulation tied to machine learning, data science, cybersecurity and engineering management. His email address is bdodg001@odu.edu.

JACOB STROTHER is an undergraduate student studying electrical engineering at Old Dominion University. His research interests and coursework are in computer hardware and digital design. His email address is jstro005@odu.edu.

ROSBY ASIAMAH is a recent graduate of Old Dominion University with a Bachelor's degree in Computer Science. He has a growing interest in cybersecurity topics like risk management and is proficient in problem solving, programming and web development. He currently serves as a reservist in the United States Air Force and is hoping to eventually merge his Computer Science career with his military career. His email address is rasia003@odu.edu.

KARINA ARCAUTE is a Senior Lecturer in the Engineering Fundamentals Division and Director of First Year Programs at Old Dominion University. Her research interests include STEM Education, First Year Engineering, and applications of Additive Manufacturing. Her email address is karcaute@odu.edu.

WIRAWAN PURWANTO is a Computational Scientist with the Information Technology Services at Old Dominion University. He holds a Ph.D. in Physics from the College of William and and Mary. His research interests lie in the applications of high-performance computing (HPC) in scientific simulation and modeling, and training scientific workforce to utilize HPC effectively. His email address is wpurwant@odu.edu.

MASHA SOSONKINA is a Professor of Modeling, Simulation Engineering at Old Dominion University. Her research interests include high-performance computing, large-scale simulations, parallel algorithms, and performance analysis. Her email address is msosonki@odu.edu.

HONGYI WU is the Batten Chair of Cybersecurity and the Director of the School of Cybersecurity at Old Dominion University. He is also a Professor in the Department of Electrical and Computer Engineering and holds joint appointments in the Department of Computer Science. His research focuses on networked and intelligent cyber-physical systems for security, safety, and emergency management applications. His email address is h1wu@odu.edu.