



Cyber Range and Cyber Defense Exercises: Gamification Meets University Students

Enrico Russo

DIBRIS - University of Genova
Genova, Italy
enrico.russo@unige.it

Marina Ribaudo

DIBRIS - University of Genova
Genova, Italy
marina.ribaudo@unige.it

Alessandro Orlich

DIBRIS - University of Genova
Genova, Italy
alessandro.orlich@dibris.unige.it

Giacomo Longo

DIBRIS - University of Genova
Genova, Italy
giacomo.longo@dibris.unige.it

Alessandro Armando

DIBRIS - University of Genova
Genova, Italy
alessandro.armando@dibris.unige.it

ABSTRACT

In the last decade, gamification has emerged as a valid alternative to more traditional learning processes both in academia and for professional training. Gamification has been successfully implemented in various disciplines to enhance the enjoyment and engagement of learning. This result can be achieved by providing challenges and quests, incentivizing task completion, and using role-playing games where learners assume different roles and perform tasks within a story format. In the case of cybersecurity, gamification can be introduced thanks to Capture The Flag (CTF) competitions or within virtual environments known as Cyber Ranges, where participants can test their skills on simulated networks, ICT systems, and other critical infrastructures. In this paper, we describe our experience with a cyber defender training activity proposed to computer science and computer engineering students. We organized lectures on cybersecurity, oriented towards developing problem-solving and practical skills. Then, we introduced gamification by running two on-site competitions: a Jeopardy CTF and a Cyber Defense Exercise. Here we will present the details of the second competition, discussing some results, the lessons learned, and the possible reuse of this experience in the broader target of IT professionals.

CCS CONCEPTS

• **Security and privacy** → *Human and societal aspects of security and privacy*; • **Applied computing** → *Education*.

KEYWORDS

Gamification, cybersecurity education, cyber range

ACM Reference Format:

Enrico Russo, Marina Ribaudo, Alessandro Orlich, Giacomo Longo, and Alessandro Armando. 2023. Cyber Range and Cyber Defense Exercises: Gamification Meets University Students. In *Proceedings of the 2nd International Workshop on Gamification in Software Development, Verification, and Validation*



This work is licensed under a Creative Commons Attribution 4.0 International License.

Gamify '23, December 4, 2023, San Francisco, CA, USA

© 2023 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0373-7/23/12.

<https://doi.org/10.1145/3617553.3617888>

(Gamify '23), December 4, 2023, San Francisco, CA, USA. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3617553.3617888>

1 INTRODUCTION

The presence of hardware devices and software applications is pervasive and affects every aspect of our daily lives. While this integration offers many benefits, it also has security implications for businesses and individuals, as neither hardware nor software is immune to vulnerabilities.

Preventing unauthorized access, alterations, or destruction of digital assets in the current digital landscape requires a twofold approach. First, we need dedicated *cybersecurity defenders* with the skills to identify attacks and implement effective countermeasures. Second, it is crucial to enhance cybersecurity awareness among software developers, system administrators, and even non-technical personnel. This diverse group of *cybersecurity stakeholders* benefits from gaining knowledge to develop secure software, configure systems, and raise awareness about potential threats. By doing so, a more comprehensive and robust approach to cybersecurity can be fostered within organizations.

Training is paramount in both scenarios, but unique challenges underscore its complexity. Cybersecurity training encompasses various technical topics, including operating systems, programming languages (low- and high-level), network protocols, web architecture, cryptography, and regulatory frameworks to mitigate cyber threats. Additionally, practical skills gained through hands-on activities are essential for effectively applying this knowledge.

To address the complexity mentioned earlier, creating engaging and practical learning environments widely exploits gamification techniques. Incorporating storylines and narratives that enable participants to play diverse cybersecurity roles, implementing leveling systems, scoring, leaderboards, and time constraints are some of the primary elements derived from the above techniques. These environments assist trainees in assimilating diverse concepts across multiple domains and enhance their problem-solving skills. Furthermore, they promote a competitive spirit that drives continuous improvement.

Capture The Flag (CTF) competitions and Cyber Defence exercises (CDX) are the most popular hands-on solutions [27] embodying the above features.

Creating and hosting Capture the Flag (CTF) events can often be managed with relative ease, and cybersecurity curricula in academic institutions widely adopted them among *non-traditional* approaches [17]. Conversely, organizing, managing, and executing CDXs require specialized systems [32], namely Cyber Ranges (CR), and demand significant human effort and financial resources [30]. As a result, CDXs are less appealing in higher education.

In this paper, we share our experience as educators, focusing on the use of a CR to run a CDX with university students. In particular, we created and managed the CR infrastructure, developed the exercise scenario (see Section 3.4), and acted as exercise managers and referees, i.e., played the roles of the Green and White teams (see Section 2.1).

While we have already disseminated some of the results of adopting CTF in our curricula [8], our current focus is on a recent CDX that we hosted, thereby extending our educational initiatives. Our experience encompasses the significant characteristics of CDXs and provides evidence that they can be affordable, even for less complex organizations. This inspiring example should encourage the broader use of CDXs, alongside other hands-on training methods, in academic education. It will be beneficial for courses aimed at cybersecurity defenders and those targeting potential cybersecurity stakeholders.

Structure of the paper. The paper is structured as follows. In Section 2, we recall some preliminary notions on hands-on exercises used in cybersecurity education and review the related work. In Section 3, we describe our whole training program, with a particular emphasis on the CDX. Section 4 and Section 5 present results and their discussion, respectively. Lastly, in Section 6, we draw the conclusions.

2 BACKGROUND AND RELATED WORK

In this section, we first introduce the foundational context of hands-on cybersecurity exercises. Then, we present the related work.

2.1 Hands-On Cybersecurity Exercises

CTF challenges are one of the prominent example of hands-on cybersecurity exercises. They cover various cybersecurity topics. Participants engage in activities that require problem-solving and vulnerability exploitation to obtain hidden *flags*, demonstrating successful task completion. Jeopardy and Attack/Defense (A/D) are two of the primary formats employed within the domain of CTFs. In Jeopardy CTF, participants face challenges covering diverse cybersecurity topics and must exploit vulnerabilities to find hidden flags. In A/D CTF, teams compete by attacking other teams' applications to acquire flags (points) while also defending their infrastructure from being hacked.

In addition to CTF competitions, another valuable training activity in cybersecurity education is represented by CDXs. The participants of a CDX become actors in a complex scenario and experience realistic cyber-attacks. They are organized into teams, each with a different focus. The White Team (WT) consists of organizers responsible for overseeing the correct execution of the CDX, ensuring rule compliance, and certifying the scores. Blue Teams (BTs) act as defenders and are tasked with protecting a simulated enterprise ICT infrastructure, namely *scenario*, from cyber attacks targeted at the

main assets. The attacks are executed through a team of specialists called the Red Team (RT). Additionally, the Yellow Team (YT) simulates legitimate user activities, which enhances the authenticity of the exercise and conceals the operations of the RT. Due to the higher complexity involved in the scenario setup and execution phase, CDXs need the capabilities of CRs. Responsibility for managing the CR infrastructure, as well as creating and implementing the scenario, lies with the Green Team (GT).

CDXs have Locked Shields (LS) [20] as a reference. LS is the cyber defense competition annually organized by NATO for member nations. As previously stated, those who have carried out similar exercises have generally aimed to maintain the complexity of LS also by leveraging capabilities of CRs [10, 18, 28, 32]. Furthermore, the above experiences underscore that CDXs are better suited than CTFs for training cybersecurity stakeholders as well. For example, including realistic applications in scenarios can require that cyber defenders should work with software engineers to review code. This cooperative approach allows those cybersecurity stakeholders to exercise and gain practical experience in secure software design, coding, testing, and bug fixing.

2.2 Related work

"Gamification as a winning cyber security strategy" is the title of the short paper by Wolfeden et al. [31] stating that gamification can be introduced through CRs to recruit and assess professionals. The author focuses on professional training, stating that the cybersecurity industry can evolve for the better through gamified learning.

As demonstrated in the literature below, the use of gamification for cybersecurity education has been thoroughly explored, involving different types of learners.

To raise cybersecurity awareness among non-experts, entry-level games with user-friendly interfaces can be employed. For example, Schofield et al. [19] discuss a role-playing quiz application to educate users about password security. Costa et al. [5] present an escape room where players, acting as on-site agents, must solve elementary cybersecurity challenges to block the company's ultimate malware with a secret code.

When dealing with university students, cybersecurity training heavily relies on various hands-on exercises [11]. The prominent ones in higher education focus on activities inspired by CTF competitions. Vigna et al. [26] and Tobey et al. [26] present early experiences introducing CTF-like competitions at the university level. Leune et al. [13] and Beltran et al. [2] explore the integration of gamification elements with traditional educational approaches as a means to enhance student motivation and yield improved outcomes in terms of skill acquisition. Also, Vykopal et al. [29] present the experience they learned from using Jeopardy CTFs as an assessment method for a university course, in contrast to the usual goal of showcasing the abilities of already skilled players. The nature of cybersecurity is adversarial, and Mirkovic et al. [15] discuss how this aspect can be included in classes through A/D exercises where students play both roles.

The first phase of our scholarship involves the organization of a Jeopardy CTF, which stems from similar reasoning and entails sharing and discovering comparable results. In contrast, the second phase of our initiative is built on CDXs, leveraging CRs for their

execution. CRs are mainly prevalent in academic settings as a research topic [4, 9], with a focus on proposing new architectures and enhancing their capabilities.

To the best of our knowledge, limited research explores the affordability of CDXs in the university context, discussing their feasibility with detailed examples, such as ours. Vielberth et al. [25] detail and discuss an exercise conducted with university students using a CR. However, their activity partially resembles a CDX exercise and only focuses on training Security Operation Center analysts. Leitner et al. [12] introduce the design and implementation of their CR. Although exercise and training courses are also promoted for educational purposes, they lack detailed explanations of their hands-on activities and their practicality for systematic use in the university setting.

3 ACME SCHOLARSHIP PROGRAM

The ACME¹ Scholarship Program is designed for students pursuing degrees in information technology (Computer Science, Computer Engineering, and Electronic and Telecommunication Engineering) at the University of Genoa. It aims to provide students with the knowledge and operational training to defend digital infrastructures against cyber threats. The program employs a blend of traditional training methods with hands-on sessions, and it is organized in different phases that can be seen as the *levels* of a game: students go through various levels, and those who perform best in one level are admitted to the next one. Below, we review all the phases of the training.

3.1 Level 1: Cyber learner

In the first level, participants are *learners* who need to acquire enough knowledge and skills to take part in the competitions. Activities start with meetings of 2 hours per week, and a typical meeting consists of the presentation of some introductory material so that students can understand (i) the topic, (ii) the context, and (iii) the main techniques and tools useful to detect and exploit possible vulnerabilities in the given context. Afterward, some exercises are proposed and solved during the meeting, and some others are left as (optional) homework.

The topics covered during the hands-on sessions are mainstream in ethical hacking, e.g., basics on Linux, programming, network protocols, Web security (client and server), basics on applied cryptography, and binary reverse engineering.

The platform chosen for the training is CTFd [6]. This solution is designed to support CTF organizers by handling the publication of exercises, registration of participants, flag submissions, and automatic scoring. Students can compare their progress during training thanks to the real-time leaderboard, which provides one of the primary ingredients often used to gamify education.

3.2 Level 2: Cyber beginner

At the end of the training, we organize a 4-hour long, on-site Jeopardy CTF; the 2022 edition took place in December, with 56 students enrolled, acting as *beginner* CTF players. The challenges were 18, organized in the classical categories: Web (4), Network (2), Binary

(2), Crypto (4), Crypto protocols (2), Misc (3), and Sanity (1). We will not go into the details of this Jeopardy CTF. The interested reader can refer to [8] for more information on previous editions.

The top-15 students in the CTF were invited to the next level of the project, described in the following section.

3.3 Level 3: Cyber pitch

Often, students with strong technical skills struggle to share their knowledge or explain the solutions they found to solve technical problems. For this reason, in the last two editions of the ACME Scholarship Program, we added a phase dedicated to assessing students' *soft skills*. The top-15 students on the leaderboard at the second level were invited to a meeting where they had the opportunity to introduce themselves in front of a committee formed by academic staff and ACME company human resource managers.

Very general indications were provided, asking students to prepare a presentation of about 10 minutes, briefly touching on the following points:

- Who am I? (name, year of study, a picture)
- What are my hobbies? (if any)
- Why am I interested in cybersecurity?
- How did I solve one challenge and find the flag?

Students made their *pitch* in front of the committee and received an evaluation (each member of the committee gave a vote out of 30 and the average value was computed). This score was combined with the CTF ranking to finalize the leaderboard and select the top-8 participants who were invited to participate in the next training level, e.g., the second on-site competition, detailed in Section 3.4.

3.4 Level 4: Cyber defender

The Cyber Defense eExercise was preceded by some preparatory meetings on ICT architectures, Linux administration, network security, and firewalls.

During the game, each student assumed the role of one BT (BT1, BT2, ...BT8). It is worth noting that even if CDXs are typically played in teams, we had the constraint of assigning four grants to the top-4 students. Therefore, students competed individually.

The exercise was conducted on an experimental CR based on a lightweight framework that we are developing inspired by [14]. Each BT was assigned a Linux Fedora 37 virtual machine hosting an instance of the scenario and configured with 4 virtual CPUs (Intel Xeon Silver 4310) with 4 cores, 16GB of RAM, and 160GB of storage.

As previously mentioned, we assumed the role of the GT by managing the CR infrastructure and developing and deploying the exercise scenario. Moreover, we acted together with cybersecurity experts of ACME as the WT by supervising the exercise and analyzing BT reports for drawing up the final ranking.

Below, we detail the internals of the exercise.

Objectives and rules. The exercise takes three hours, comprising three attack campaigns of one hour each. In the first campaign, the initial half hour is allocated for familiarization with the scenario, allowing the eight BTs to become acquainted with the assets they need to defend. They can analyze them, identify weaknesses and vulnerabilities, and potentially patch them before the RT launches

¹This program is organized with a company that funds four grants for the best-performing students. The name of the company has been changed to ACME.

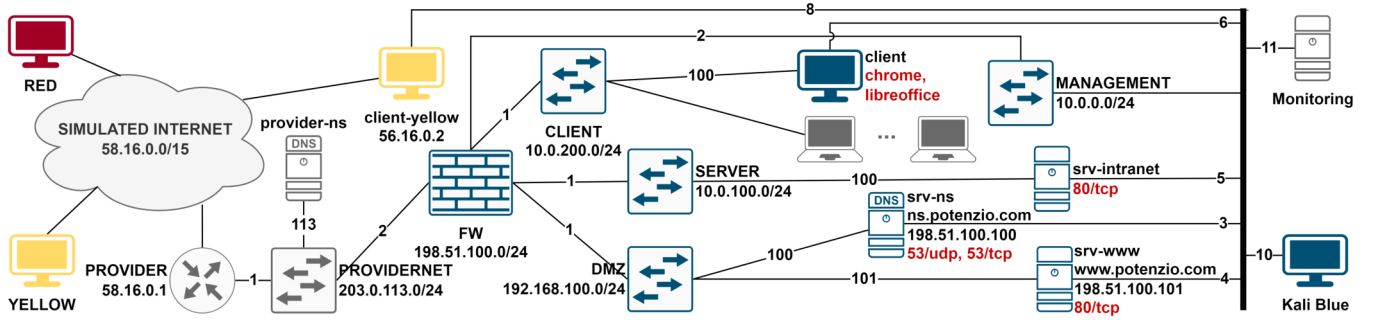


Figure 1: The scenario for the cyber defense competition.

the first attack. The familiarization time was intentionally made short because, similar to the approach taken in the LS exercise, we wanted to test the ability to respond to and analyze attacks rather than prevent them. In the second half hour, the RT carries out the first assault.

In the subsequent two campaigns, the RT conducts two attacks each, for a total of five for the entire CDX. It is noteworthy that BTs are subjected to the same attacks in the same order, allowing for a consistent evaluation of the exercise.

Each BT starts the exercise with 5000 points. The score changes according to (i) the integrity of the services to be protected, e.g., the availability of the corporate website, (ii) a successful attack executed by the RT, and (iii) the quality of the reports that BTs can fill in after each attack (see Section 3.4). In detail, a BT loses one point for each minute of unavailability of corporate services. The score is decremented by 100 points whenever the RT performs a successful attack. On the other hand, points can be awarded to reports based on their evaluation by a designated committee, i.e., the WT, after the conclusion of the exercise. The WT can assign 30, 70, or 100 points if each filled report is evaluated as fair, good, or excellent, respectively.

The final ranking is determined by combining each BT's remaining points at the conclusion of the exercise, along with the points earned from the reports.

Scenario. Figure 1 shows the scenario used for the competition. It represents the ICT infrastructure of an enterprise, namely *Potenzio*, that manufactures audio speakers for musical events. It consists of a segmented network that includes a *Demilitarized Zone* (DMZ) subnet for hosting public services, a *Server* subnet dedicated to corporate servers, and a *Client* subnet for corporate users. The DMZ network connects two Linux servers, *srv-www* and *srv-ns*, which are used for hosting the corporate website and the name server for the public domain name, i.e., *potenzio.com*.

Meanwhile, the Server subnet includes the *srv-intranet* node dedicated to the intranet website. Access to the intranet website requires authentication, and it is partially published on the Internet via a reverse proxy configured on the corporate website server.

The BT must also defend a user workstation connected to the Client subnet, namely *client*. The system is a Linux Desktop that runs an automatic agent to simulate a legitimate user, specifically performing tasks related to the Yellow Team (YT). This includes activities such as word processing and browsing the simulated

Internet (see below). Other connected workstations are dedicated to automated YT activities and are not controlled by the BT.

Policy enforcement between the subnets is provided by a firewall called *FW*, implemented with the OPNsense project [7]. It also connects to a router, namely *Provider*, which mimics the equipment used by Internet Service Providers to give customers access to the public Internet. In this scenario, the router connects the enterprise to a *Simulated Internet* (SI).

The SI consists of a single large public network (58.16.0.0/15) that hosts a public DNS server, i.e., *provider-ns*, automated YT agents, and the RT.

The public DNS server resolves any domain name with the IP addresses of the SI and provides each of them with a fake public website. The above functionality allows for simulating the web browsing activity of the company's clients.

Instead, the automated YT agents access the company's public services using random IP addresses. The traffic that is generated makes it difficult to distinguish RT attacks that originate from the SI from legitimate traffic.

One of the above YT nodes, a Linux desktop named *client-yellow*, can be managed by the BT to access the infrastructure from the SI. For example, the BT can use this node to check if an update to the firewall rules has resulted in the unavailability of public services.

Lastly, all the nodes under the control of the BT are connected to a special network, i.e., the *Management* network. The RT cannot exploit this network and ensures access to the BT's assets even after firewall misconfigurations or malfunctions.

The Management network hosts two nodes, namely *Kali Blue* and *Monitoring*. The first node provides the desktop client that each BT can use to access the scenario and includes the facilities to execute the exercise. This includes features such as shortcuts to access via secure shell the nodes to be defended and bookmarks to scoring and company websites. Instead, the role of the Monitoring node is twofold: (i) checking for the availability and integrity of BT services for the scoring system, i.e., *health checkers*, and (ii) providing the web interface through which BTs can visualize the status of services they are defending. Briefly, we configured 14 health checkers that monitor the reachability from the SI of public services, the content integrity and functionality of the login system on websites, and the SI access along with the availability of the web browser and office automation software on the client.

Table 1: Attacks summary.

Id	Assets	Vulnerabilities	Impact
C1.1	Intranet server (<i>srv-intranet</i>)	<ul style="list-style-type: none"> – a <i>SQL injection</i> on the custom PHP website – a <i>command injection</i> on the custom PHP website – a <i>privilege escalation</i> through the Linux <code>sudoedit</code> command due to CVE-2023-22809 [24] 	ransomware encrypts files on the intranet website
C2.1	Web server (<i>srv-www</i>)	– a <i>remote code execution</i> through the Bolt [1] Content Management System due to CVE-2022-36532 [22]	website defacement by replacing the main index file
C2.2	Client (<i>client</i>)	– a <i>macro virus</i> in a document that can open a reverse shell to a remote system	attackers install a malware that kills and does not allow the user to restart the browser and office automation software
C3.1	DNS server (<i>srv-ns</i>)	– a <i>command injection</i> through the Webmin software due to CVE-2019-15107 [21]	website defacement by replacing the DNS record with the address of a SI rogue site
C3.2	Firewall (<i>FW</i>)	– a <i>backdoor</i> on the firewall through an API key with admin privileges	denied access to SI by changing the routing configuration on the firewall

Attacks. In Table 1, we summarize the attacks that the RT executed during the three campaigns, namely C1, C2, and C3. For each attack, we specify the involved asset, i.e., the target host, the vulnerabilities that the RT exploits to compromise the asset, and details about the impact.

Vulnerabilities are injected during the scenario design, including realistic flaws enabling common attack patterns [23], publicly disclosed security issues from the Common Vulnerabilities and Exposures (CVE) database, and hidden backdoors on the company assets. BTs can prevent RT attacks by identifying and fixing vulnerabilities and can earn points by reporting them and describing the remediation they have implemented.

Attacks are automated through scripts. Each script mimics a realistic attacker's behavior by executing a reconnaissance activity, e.g., brute forcing the website password from a dictionary or gathering information about an asset through shell commands before exploiting vulnerabilities. The impact of the attacks is always noticeable and causes some health checkers to fail. Moreover, each attack leaves cybertrails and does not cause irreversible damage. As a result, students are encouraged to find evidence of the attacks to restore services and fill in a more comprehensive report.

For example, in Figure 2, we show the impact of the attack C1.1. It represents a screenshot of the Intranet website being defaced and encrypted with ransomware. The new page hints at the command to be used for decrypting files but omitting the required password. As stated above, students are challenged to understand the attackers' actions and find evidence they left in log files containing the secret to restore the website.

Incidents reports. In addition to dealing with the attacks, students were also asked to fill in a report for each attack. The aim is to mimic the activity performed during the incident response procedure within an organization. For this task, we used our university learning management system (based on Moodle [16]) that students use daily for their official courses.

**Figure 2: Intranet defacement.**

To assist students in the process of reporting, we structured it in the form of a questionnaire. Due to time constraints, the questionnaires were short, identical for each attack, and could be filled either at the end of each attack campaign or at the end of the entire competition.

For each attack, we asked to specify which asset was compromised, evidence and possible remediation, and whether they could identify the IP address of the attacker or recognize a specific CVE. The questions are shown in Table 2.

Students feedback. Before the exercise started and after its execution, BTs were surveyed with a *pre* and a *post-exercise* questionnaire, respectively. The survey methodology and questions are inspired by [3]. The pre-exercise questionnaire was used to identify the participant's profile and understand expectations. Instead, the post-exercise questionnaire asked students to specify how they used their prior knowledge, learning outcomes, and any suggestions for improving the experience.

Table 3 shows an excerpt of the questions used in the pre (B) and post-exercise (E) questionnaires.

Table 2: Questions asked to BTs after each attack.

Id	Question after attack	Type
Q1	Which asset was compromised with this attack? Assets: 1) Client, 2) DNS server , 3) Firewall, 4) Intranet server, 5) Web server	Multiple choice
Q2	Which is/are the IP addresses of the attacker?	Short text
Q3	For each asset (see Q1), detail the evidences of the attack, if you detect any.	Long text
Q4	Did you find a CVE used by the attacker? If yes, write the ID here (e.g., CVE-2021-44228).	Short text
Q5	Which activities did you perform to recover from the attack?	Long text
Q6	Can you describe possible remediation?	Long text

Table 3: Feedback asked to BTs at pre- and post-exercise.

Id	Question pre-exercise	Type
B1	Have you ever participated in similar CDX? Please, provide some details about the exercises, your role, your lessons learned.	Short text
B2	How would you rate your skills, knowledge, and abilities in cybersecurity? (from 1=beginner, to 5=expert)	Single choice
B3	What are your expectations for the event? 1) Show off yourself, 2) Have a good time, 3) Establish contacts, 4) Get a certificate	Multiple choice
B4	What are your strengths? 1) Firewall management, 2) Network analysis, 3) Forensics, 4) MS Windows OS, 5) Linux OS, 6) Database administration, 7) Soft skills	Multiple choice
Id	Question post-exercise	Type
E1	Was it worth participating in the CDX?	Yes/No
E2	How much the CDX met your expectations? (from 1=definitely no, to 5=definitely yes)	Single choice
E3	What did you learn during the CDX?	Short text
E4	What did you like in the CDX (atmosphere, particular incidents, etc.)? Why?	Short text
E5	What did you not like in the CDX? Why?	Short text
E6	Did you find any attack, task, or disruption challenging? If yes — which and why?	Short text

3.5 Level 5: Cyber mentor




Some students who attended all the program phases became passionate and asked to take part in the organization of the following year edition. In a sort of *give back*, they offered to coach next generation of cyber defenders sharing the knowledge and experience acquired during the program.

4 CDX RESULTS AND FEEDBACK

This section contains the results of the CDX exercise and a summary of the pre and post-exercise feedback, detailing the most noteworthy students' responses.

4.1 CDX Results

To present the CDX results, we report the performance of each BT on attacks and reports. Then, we provide the overall score used to determine the final ranking.

Attacks. In Table 1, we summarize how students were affected by each attack. In particular, we use  and  to denote whether the attack was successful or blocked, respectively. We use  to distinguish when the attack was successful but the student was able to recover from it.

Although BT1 managed to block attack C2.1, it was later discovered that this was due to a firewall configuration change that blocked all access from the SI to external services. As a result, BT1

was successful in blocking the attack but getting penalized by the unavailability of monitored services.

BT5 and BT6 were able to block attacks C3.1. BT5 also succeeded in preventing attack C3.2. This outcome was a result of having properly filtered with firewall policies the access to unnecessary services and to the firewall itself from the SI.

As expected, given the limited time available for the familiarization, none of the BTs were able to prevent the attacks by identifying and fixing the vulnerabilities. However, a significant proportion of BTs succeeded in understanding and analyzing the events of at least one attack and restoring the affected service.

Incidents reports. All the responses to the questionnaires have been collected into a document and evaluated by the WT.

In general, the quality of the reports was not notably high. In most cases, short answers were provided, in a few cases the reports were even submitted empty.

For the BTs that submitted reports, almost all were able to indicate the affected asset (Q1) and the IP address of the attacker (Q2).

Responses to Q3 were often limited to describing the effects of the attacks without indicating the causes. Only BT4 and BT5 were able to partially highlight what happened by including extracts from some log files. No one was capable of providing a concise and

Table 4: Results of the attacks campaigns.

Attack ID	BT1	BT2	BT3	BT4	BT5	BT6	BT7	BT8
C1.1								
C2.1								
C2.2								
C3.1								
C3.2								

comprehensive explanation of the attackers' steps and what they exploited.

Again, no BTs answered to Q4 with the indication of the CVEs used by the attacker. BTs who correctly restored services from attacks were able to provide the description of the necessary operations in Q5. Finally, Q6 has received less attention or insufficiently detailed responses.

Final scoring. Table 5 shows the final scoreboard of the CDX, combining the results of the competition. As detailed in Section 3.4, each BT had an initial score of 5000 points to protect. The first row, namely *Exercise*, indicates the remaining score that decreased as a result of services unavailability and attacks received. The second row, namely *Reports*, shows the score gained by producing the reports. The *Total* row is the sum of *Exercise* and *Reports* scores. The last row indicates the position in the final ranking.

Analyzing the *Exercise* score, BT1 was severely penalized by availability for having enforced a firewall policy that was too restrictive. Moreover, with the exception of the highest-scoring student, i.e., BT5, who only lost less than 60 points, the scores of the other BTs were very well balanced.

The challenge of generating high-quality reports is apparent in the score. Although BT6 was proficient in defending against attacks and maintaining service availability, the lack of emphasis on producing quality reports resulted in a penalty that prevented the student from securing the second position.

4.2 Students feedback

According to the answers to the pre-exercise feedback (Table 3, questions B1-B4), for all eight students this was the first experience in a CDX, and their expectation was to "Have a good time". In the self-assessment of their competencies, they used only the values from 1 to 3 with an average value of 2.375, which is rather low, but expected since this was their first experience in a CDX. We report here also the following comment "Regardless of how it goes, I think it is an excellent opportunity to experience a simulation of an attack on the own infrastructure."

In the post-exercise feedback (Table 3, questions E1-E6) all students declared that it was worth participating and that the CDX positively met their expectations, with an average value of 3.67. We report here a few comments on the open questions, which somehow summarize how the participants perceived this experience.

- E3: What did you learn during the CDX?
 - "I learned how far an attack on an infrastructure can go."
 - "That I still have too much to learn."
- E4: What did you like in the CDX (atmosphere, particular incidents, etc.)?

Table 5: Results of the cyber defense exercise.

	BT1	BT2	BT3	BT4	BT5	BT6	BT7	BT8
Exercise	3636	3803	3719	3846	4641	4005	3918	3832
Reports	0	150	160	300	470	90	120	150
Total	3636	3953	3879	4146	5111	4095	4038	3982
Position	8	6	7	2	1	3	4	5

- "Working under pressure and that the attacks were above my current level."
- "The atmosphere. Even though the exercise was difficult they made us feeling little pressure."
- E5: What did you not like in the CDX?
 - "Requesting CVEs in reports."
 - "Not enough lectures."
- E6: Did you find any attack, task, or disruption challenging? If yes — which and why?
 - "The attacks were all difficult on average, precisely because they were based on aspects that were still unknown or relatively poorly explained within the academic world. Such experience helps to increase knowledge in the cyber field."

5 DISCUSSION

Based on the results outlined in the previous section, we draw the following considerations related to CDXs and our experience with students within the academic setting.

An essential point to note is that our experience confirms important distinctions between CDXs and CTF competitions. Students who made it to the final phase are frequently active in CTFs, possess strong technical skills, and are knowledgeable about Linux and cybersecurity issues. However, they find themselves struggling in CDX. One of the main difficulties is that CDX requires a much broader range of competencies than just technical proficiency, including risk management, incident response, and reporting.

Some students emphasized the pressure they felt during this type of competition. This feeling also arises from the need to understand how to organize tasks across different stages and manage time efficiently. Having the option to participate as a team could have helped them, but the constraints of this scholarship prevented us from offering students this opportunity (see Section 5.1).

Additionally, it is worth noting that none of the students managed to identify vulnerabilities and correlate attacks with CVEs. This observation underscores a weakness in their hands-on experiences, which includes CTFs. These experiences often lack a focus on real-world vulnerabilities and incidents, thus limiting their readiness for practical cybersecurity scenarios.

Moreover, as pointed out in one of the answers, CDXs can offer the opportunity to learn "how far an attack on an infrastructure can go". Again, managing a complex network with heterogeneous components, analyzing and correlating different data sources, and experiencing the lateral movements of attackers are unique features of CDXs that can enhance their level of preparedness, as mentioned above.

Another crucial element that students can learn from CDXs is reporting. Writing a report after an attack is paramount in a company. However, it was not perceived as such during the competition: it was seen as a task to be completed because technical (hard) skills were considered more important and took precedence. On our side, we need to put more emphasis on this aspect and spend more time during the meetings teaching how to fill them out.

Finally, an answer underscores that attacks of CDXs revealed “aspects that were still unknown or relatively poorly explained within the academic world”. This comment should encourage academic institutions that have already integrated their cybersecurity curriculum with CTF-like experiences to further enhance their offerings and learning outcomes by adopting CDXs.

From our experience as WT, evaluating the BT reports was the most demanding aspect. This challenge primarily arises from the manual and time-consuming nature of the evaluation process, which requires the involvement of cybersecurity experts for a comprehensive assessment. Considering this factor is crucial to encourage the extensive integration of CDXs into academic contexts. Therefore, investigating techniques to support and automate the evaluation of such reports is a noteworthy component of our future work.

Considering the GT perspective, our framework provides all the core facilities to host a CDX and allows us to run a significant scenario requiring few resources. Similar to CTF challenges, the process of designing and implementing the scenario content proved to be a particularly challenging task. Specifically, the recreation and implementation of vulnerabilities and planning attacks emerged as the most crucial aspects. As in the case of CTFs, growing the community associated with CDXs and ensuring that vulnerable scenarios or scenario components can be shared and reused is of considerable importance and a possible answer to this issue. For our part, we are close to releasing our framework and the scenario we have used in this exercise as open source.

5.1 Limitations

While participants enjoyed the CDX, their number is small by design, and this limits the generalization of results and does not guarantee anonymity for the answers. However, we believe that the general feedback has been positive and encourages us to move forward.

One of the most interesting aspects of this type of event is teamwork: participants collaborate in a team to defend a realistic scenario. Due to the constraint of the grants to be assigned to the best-performing students, our BTs were formed by single individuals. However, we should find an alternative configuration in which students can play in teams to get closer to what they will experience in reality. Also, working in a team is less stressful and much more fun.

There is room for improvement, but most importantly, now that we have the expertise in organizing a CDX at the academic level, we should aim to expand it in a follow-up with different scenarios, more students and more time for the exercises.

6 CONCLUSION

The job market increasingly demands IT professionals, placing a strong focus on cybersecurity defenders while valuing that other professionals are aware of cybersecurity practices, i.e., they evolve in cybersecurity stakeholders.

As educators, we are responsible for preparing university students for their future endeavors to the best of our ability. While theoretical courses have been a longstanding tradition in universities, practical skills are equally essential in preparing them for real-world challenges.

In this paper, we have focused on cybersecurity defenders and presented our experience designing and organizing a CDX for university students competing for a grant offered by ACME company. We already presented our experience organizing more “traditional” Jeopardy CTFs, but this is the first time we report on a different type of exercise run on a CR.

Gamification plays a critical role in the above hands-on learning experiences as it allows students to learn through practical application in a simulated environment that is safe, legal, and live. Additionally, it encourages motivation, competition, and recognition of achievements.

Some tuning is necessary, and some new configurations should be defined to scale up. Creating a CDX demands significant resources, specialized personnel, and considerable time investment. Starting from our “pilot” experience, now we should try to put into practice the phase (v) *repetition* of the CDX life cycle [30], i.e., reusing the work done and involving more and more participants.

To this aim, we need also to involve cybersecurity stakeholders and test the outcomes of this type of exercise with them.

ACKNOWLEDGMENTS

This work was partially funded by the NextGenerationEU project “Security and Rights in Cyberspace” (SERICS) and “EMPHAsis Evolution” project (CUP J89J22003110005), under “Azione 1.1.5 POR Calabria FESR-FSE 2014-2020”. It was carried out while Giacomo Longo was enrolled in the Italian National Doctorate on Artificial Intelligence run by the Sapienza University of Rome in collaboration with the University of Genoa.

REFERENCES

- [1] 2023. Boltcms. <https://boltcms.io/>.
- [2] M. Beltrán, M. Calvo, and S. González. 2018. Experiences Using Capture The Flag Competitions to Introduce Gamification in Undergraduate Computer Security Labs. In *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*. IEEE, 574–579.
- [3] Agnė Brilingaitė, Linas Bukauskas, and Aušrius Juozapavičius. 2020. A framework for competence development and assessment in hybrid cybersecurity exercises. *Computers & Security* 88 (Jan. 2020), 101607. <https://doi.org/10.1016/j.cose.2019.101607>
- [4] Nestoras Chouliaras, George Kittes, Ioanna Kantzavelou, Leandros Maglaras, Grammati Pantziou, and Mohamed Amine Ferrag. 2021. Cyber Ranges and TestBeds for Education, Training, and Research. *Applied Sciences* 11, 4 (Feb. 2021), 1809. <https://doi.org/10.3390/app11041809>
- [5] Gabriele Costa, Martina Lualdi, Marina Ribaud, and Andrea Valenza. 2020. A NERD DOGMA: Introducing CTF to Non-Expert Audience. In *Proceedings of the 21st Annual Conference on Information Technology Education (Virtual Event, USA) (SIGITE '20)*. Association for Computing Machinery, New York, NY, USA, 413–418. <https://doi.org/10.1145/3368308.3415405>
- [6] CTFd LLC. 2023. CTFd: The Easiest Capture The Flag Framework. <https://ctfd.io/>.
- [7] Deciso B.V. 2021. OPNsense firewall. <https://opnsense.org/>.
- [8] Luca Demetrio, Giovanni Lagorio, Marina Ribaud, Enrico Russo, and Andrea Valenza. 2019. ZenHackAdemy: Ethical Hacking @ DIBRIS. In *Proceedings of*

- the 11th International Conference on Computer Supported Education, CSEDU 2019, Heraklion, Crete, Greece, May 2-4, 2019, Volume 1. SciTePress, 405–413.
- [9] Magdalena Glas, Manfred Vielberth, and Guenther Pernul. 2023. Train as You Fight: Evaluating Authentic Cybersecurity Training in Cyber Ranges. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 622, 19 pages. <https://doi.org/10.1145/3544548.3581046>
 - [10] Tommy Gustafsson and Jonas Almroth. 2021. Cyber Range Automation Overview with a Case Study of CRATE. In *Secure IT Systems*. Springer International Publishing, 192–209. https://doi.org/10.1007/978-3-030-70852-8_12
 - [11] Marcus Knüpfer, Tore Bierwirth, Lars Stiemert, Matthias Schopp, Sebastian Seeber, Daniela Pöhn, and Peter Hillmann. 2020. Cyber Taxi: A Taxonomy of Interactive Cyber Training and Education Systems. In *Model-driven Simulation and Training Environments for Cybersecurity*. Springer International Publishing, 3–21. https://doi.org/10.1007/978-3-030-62433-0_1
 - [12] Maria Leitner, Maximilian Frank, Gregor Langner, Max Landauer, Florian Skopik, Paul Smith, Benjamin Akhras, Wolfgang Hotwagner, Stela Kucek, Timea Pahi, Lenhard Reuter, and Manuel Warum. 2021. Enabling exercises, education and research with a comprehensive cyber range. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)* 4 (December 2021).
 - [13] Kees Leune and Salvatore J. Pettrilli. 2017. Using Capture-the-Flag to Enhance the Effectiveness of Cybersecurity Education. In *Proceedings of the 18th Annual Conference on Information Technology Education*. ACM. <https://doi.org/10.1145/3125659.3125686>
 - [14] Giacomo Longo, Alessandro Orlich, Stefano Musante, Alessio Merlo, and Enrico Russo. 2023. MaCySTe: A virtual testbed for maritime cybersecurity. *SoftwareX* 23 (July 2023), 101426. <https://doi.org/10.1016/j.softx.2023.101426>
 - [15] Jelena Mirkovic and Peter A. H. Peterson. 2014. Class Capture-the-Flag Exercises. In *2014 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 14)*. USENIX Association, San Diego, CA. <https://www.usenix.org/conference/3gse14/summit-program/presentation/mirkovic>
 - [16] Moodle Pty Ltd. 2023. Moodle - Open-source learning platform. <https://moodle.com/>.
 - [17] Djedjiga Mouheb, Sohail Abbas, and Madjid Merabti. 2019. Cybersecurity Curriculum Design: A Survey. In *Transactions on Edutainment XV*. Springer Berlin Heidelberg, 93–107. https://doi.org/10.1007/978-3-662-59351-6_9
 - [18] Enrico Russo, Gabriele Costa, and Alessandro Armando. 2020. Building next generation Cyber Ranges with CRACK. *Computers & Security* 95 (Aug. 2020), 101837. <https://doi.org/10.1016/j.cose.2020.101837>
 - [19] Sam Scholefield and Lynsay A Shepherd. 2019. Gamification techniques for raising cyber security awareness. In *HCI for Cybersecurity, Privacy and Trust: First International Conference, HCI-CPT 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings* 21. Springer, 191–203.
 - [20] Max Smeets. 2022. The Role of Military Cyber Exercises: A Case Study of Locked Shields. In *2022 14th International Conference on Cyber Conflict: Keep Moving! (CyCon)*, Vol. 700. 9–25. <https://doi.org/10.23919/CyCon55549.2022.9811018>
 - [21] The MITRE Corporation. 2019. CVE-2019-15107. <https://www.cve.org/CVERecord?id=CVE-2019-15107>
 - [22] The MITRE Corporation. 2022. CVE-2022-36532. <https://www.cve.org/CVERecord?id=CVE-2022-36532> Accessed on 02/05/2023.
 - [23] The MITRE Corporation. 2023. Common Attack Pattern Enumeration and Classification. <https://capec.mitre.org/index.html> Accessed on 02/05/2023.
 - [24] The MITRE Corporation. 2023. CVE-2023-22809. <https://www.cve.org/CVERecord?id=CVE-2023-22809> Accessed on 02/05/2023.
 - [25] Manfred Vielberth, Magdalena Glas, Marietheres Dietz, Stylianos Karagiannis, Emmanouil Magkos, and Günther Pernul. 2021. A Digital Twin-Based Cyber Range for SOC Analysts. In *Data and Applications Security and Privacy XXXV*. Springer International Publishing, 293–311. https://doi.org/10.1007/978-3-030-81242-3_17
 - [26] G. Vigna. 2011. The 2010 International Capture the Flag Competition. *IEEE Security Privacy* 9, 1 (2011), 12–14.
 - [27] Jan Vykopal and Miloš Barták. 2016. On the Design of Security Games: From Frustrating to Engaging Learning. In *2016 USENIX Workshop on Advances in Security Education (ASE 16)*. USENIX Association, Austin, TX. <https://www.usenix.org/conference/ase16/workshop-program/presentation/vykopal>
 - [28] Jan Vykopal, Radek Oslejsek, Pavel Celeda, Martin Vizvary, and Daniel Tovarnak. 2017. KYPO Cyber Range: Design and Use Cases. In *Proceedings of the 12th International Conference on Software Technologies*. SCITEPRESS - Science and Technology Publications. <https://doi.org/10.5220/0006428203100321>
 - [29] Jan Vykopal, Valdemar Švábenský, and Ee-Chien Chang. 2020. Benefits and Pitfalls of Using Capture the Flag Games in University Courses. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. ACM. <https://doi.org/10.1145/3328778.3366893>
 - [30] Jan Vykopal, Martin Vizvary, Radek Oslejsek, Pavel Celeda, and Daniel Tovarnak. 2017. Lessons learned from complex hands-on defence exercises in a cyber range. In *2017 IEEE Frontiers in Education Conference (FIE)*. IEEE. <https://doi.org/10.1109/fie.2017.8190713>
 - [31] Brad Wolfenden. 2019. Gamification as a winning cyber security strategy. *Computer Fraud & Security* 2019, 5 (2019), 9–12.
 - [32] Muhammad Mudassar Yamin and Basel Katt. 2022. Modeling and executing cyber security exercise scenarios in cyber ranges. *Computers & Security* 116 (May 2022), 102635. <https://doi.org/10.1016/j.cose.2022.102635>

Received 2023-08-03; accepted 2023-08-17