An Effective Cybersecurity Exercises Platform CyExec and its Training Contents

Nobuaki Maki, Ryotaro Nakata, Shinichi Toyoda, Yosuke Kasai, Sanggyu Shin, and Yoichi Seto

Abstract—Recently the threats of cyberattacks, especially of targeted attacks are increasing rapidly and a large number of cybersecurity incidents are occurring frequently. On the other hand, capable personnel are greatly lacking, and strengthen the systematic human resource development cultivating capabilities for cybersecurity activities is becoming an urgent issue. However, only a few parts of academia and private sector in Japan can carry out the cybersecurity exercises because of high cost and inflexibility of commercial or existing training software. On this account, in order to enforce cybersecurity practical exercises cost-effectively and flexibly, we developed a virtual environment Cybersecurity Exercises (CyExec) system utilizing VirtualBox and Docker. We also implemented an open source vulnerability scanner tool WebGoat and our original cyberattack and defense training contents on CyExec.

Index Terms—Cyberattack and defense exercise, cyber range exercise, ecosystem, human resource development on cybersecurity sector, threads, virtualization, vulnerability, WebGoat.

I. INTRODUCTION

In this paper, we propose a cybersecurity exercises system in a virtual computer environment. This exercises system enables effective human resource development and contributes cybersecurity level of society. Backgrounds, characteristics, constitution and training contents of the exercises system are described below.

Cyberattacks are bringing serious social influences, causing vast cybersecurity incidents and even affecting business continuity. In January 2018, $350 million cryptocurrency was stolen in Japan, and in February 2018, organizations associated with the Pyeongchang Winter Olympics were targeted by cyberattack. These matters directly link to people’s lives and cybersecurity is becoming a matter of deep social concern [1].

In the cybersecurity strategy of the Government of Japan, human resource development is cited as a serious issue. Human resource with skill insufficiency on cybersecurity is estimated at 190,000 by 2020 in Japan. The lack of technical knowledge and skill is worried even in personnel engaged in cybersecurity operations [2], [3].

As efforts towered human resource development and training for knowledge and skill regarding cybersecurity, some universities and public organizations are carrying out vulnerability learning exercises using dedicated software, and cyberattack and defense exercises using Cyber Range [4], [5].

Participants of the Cyber Range exercises learn practical defense technology against assumed cyberattack on the network in virtual environment. Participants also learn systematic correspondence method depending on roles in organization by using possible practical scenarios such as real malware infection. Therefore, high training effects can be expected [6].

However, universities have not enough exercises infrastructure to bring up cybersecurity human resources because of the high cost to introduce the practical exercises system and the lack of personnel to maintain the practice environment.

Therefore, a cybersecurity exercises platform which can promote joint development and common use is strongly required in the universities. This is the reason why we developed a cybersecurity exercises platform “Cybersecurity Exercises” (hereinafter referred to as CyExec) using a virtual computer environment of VirtualBox and Docker [7], [8].

Training contents implemented on CyExec is composed of a basic part and an applied part. Regarding the basic part, we implemented an open source vulnerability scanner tool WebGoat on CyExec and we developed a curriculum and a training guidance for the WebGoat exercises. Regarding the applied part, we developed and implemented our original cyberattack and defense training contents on CyExec.

In this paper, we describe the constitution of the cybersecurity exercises platform CyExec and training contents we implemented on it. We explain the outline of the cybersecurity exercises platform CyExec in Chapter II.; the problems and measures of training using open source vulnerability scanner tool WebGoat in Chapter III.; and the constitution of the training contents implemented on CyExec including WebGoat and our original cyberattack and defense training contents in Chapter IV.

II. OUTLINE OF THE CYBERSECURITY EXERCISES PLATFORM CYEXEC

A. Subjects of Existing Cyberattack and Defense Exercises

There are two kinds of existing cybersecurity exercises; using open source vulnerability scanner tool, and using commercial Cyber Range software.

1) Vulnerability scanning exercises

Vulnerability Scanning Exercises are aimed at learning the outline of the vulnerability, the detective method, and the
countermeasures using open and free training software, such as WebGoat provided by OWASP (Open Web Application Security Project) [9].

Participants attending the lecture create the practice environment by installing the training software on their own PC and acquire diagnosis method and countermeasures against web application vulnerabilities systematically utilizing the software and the environment inside the PC.

However, training of correspondence method in organization is out of the scope from the vulnerability scanning exercises. In addition, the exercises are lacking interactive cyberattack and defense training and viewpoint of the exercises is limited to vulnerability detection and countermeasure on the static environment.

WebGoat is constantly revised in line with rapid technological changes, but only program materials are released. Therefore, frequent maintenance of the curriculum and the renewal of training guidance are necessary to correspond to the newest practice.

2) Cyber range exercises

Cyber Range exercises are aimed at upbringing of personnel available for responding to cybersecurity incidents in organization. The practice environment is constructed on a virtual environment imitating the real-world including clients, servers and network [10].

Trainees can learn attack techniques and knowledge of various types of malware, and train on confirmation of damaged situation and response method, assuming all stages of cybersecurity incidents recovery process from the beginning of the detection to the end of the response. The exercises are applicable to train personnel of Computer Security Incident Response Team (CSIRT) and Security Operation Center (SOC) [7], [8]. However, the introduction and operation of the Cyber Range exercises takes very high cost. In addition, the Cyber Range exercises are lacking the flexibility to change the curriculum in accordance with the intention of universities.

Universities need an exercises system to train the basics of vulnerability measures and response method in organization using the existing computer environment without adding anything. The vulnerability detecting exercises are suitable for learning the basics but are lacking the cyberattack and defense interactive training. On the other hand, the Cyber Range exercises are difficult to introduce in universities because of their limited budget and staff. For this reason, we developed the exercises system CyExec, described in the next section [7], [8], [11].

B. Characteristics of CyExec

CyExec is a cost-effective and flexible exercises system in a virtual environment to learn the basic technology of cyberattack and defense practically. It is expected to be introduced in universities and small and medium-sized enterprises. Characteristics of CyExec are shown below [7], [8].

1) Low cost and highly portable exercises environment

Most of the costs for installing and maintaining the cybersecurity exercises system are the costs of equipment and software licensing. To update the exercises system, personnel having specialized skills and high labor cost are required.

In order to reduce these costs, we developed an exercises environment using virtualization technology that can easily implement the training program in existing client and server computer environment. We utilized VirtualBox, which can operate a guest OS (virtual OS) on a host OS (Windows, macOS, etc.). On the virtual environment, we implemented the operating environment for the exercises program.

2) Practice environment for easy joint development and utilization

A high level of specialty and a long period of time are necessary for developing the cybersecurity training program. On the other hand, in the field of cybersecurity, technological progress is rapid. Therefore, it is difficult for a single university or private enterprise to develop a new cybersecurity training program and several organizations need to work together for the development. For this reason, we adopted the concept of ecosystem which will realize joint development and common utilization of the training programs between some organizations.

The word “ecosystem” means that whole associated organizations develop not on each single organization’s own but through the collaboration of related organizations. CyExec also enriches the training program not only by a single organization but also by joint development and utilize of related organizations [7], [8]. We have realized the joint development and utilization between multiple organizations by container technology using Docker.

We implemented Docker on the virtual environment constructed in VirtualBox, then we installed a container on Docker. It is easy for universities and private enterprises to build the training environment according to each purpose by implementing and operating various training programs on the container such as vulnerability diagnosis training or cyberattack and defense training. By making and releasing image files of the containers that operate the developed training programs, associated organizations can utilize them jointly.

The architecture of the CyExec system is shown in Fig. 1.

<table>
<thead>
<tr>
<th>Exercise</th>
<th>Process</th>
<th>Process</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>Container</td>
<td>Container</td>
<td>Docker Engine</td>
<td></td>
</tr>
<tr>
<td>Guest OS (Linux)</td>
<td>VirtualBox</td>
<td>Host OS (Windows / macOS)</td>
<td></td>
</tr>
<tr>
<td>Infrastructure</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1. The architecture of the CyExec system.

The architecture of the exercises system we developed has two-layer structure using Docker container. Docker is installed on the guest OS that operates on VirtualBox on the host OS. The processes, on which WebGoat and the cyberattack and defense training program run, is implemented on Docker container. VirtualBox has superior portability and Docker container has high extensibility for
any existing computer environment. They enable joint development and utilization of the training program.

III. OUTLINE OF WEBGOAT

A. Configuration of WebGoat Exercises

WebGoat is an open source software to teach web application security lessons designed and maintained by OWASP experts [9]. The detection method and countermeasures of vulnerability can be learned through the exercises.

As shown in Table I, WebGoat contains totaled 12 lesson plans and each lesson plan is consists of one or more subtopic.

<table>
<thead>
<tr>
<th>No.</th>
<th>Category</th>
<th>Lesson Plan</th>
<th>Subtopic</th>
<th>Number of Assignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>Introduction</td>
<td>WebGoat</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Basic Knowledge</td>
<td>General</td>
<td>RUTF, Python</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CIA tool</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Injection Flaws</td>
<td>SQL Injection</td>
<td>SQL Injection (introduction)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SQL Injection (advanced)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SQL Injection (mitigation)</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Authentication Flows</td>
<td>Cross Site Scripting (XSS)</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cross Site Scripting (cross)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cross Site Scripting (rules)</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Vulnerabilities Assumptions</td>
<td>Cross Site Scripting (XSS)</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Access Control</td>
<td>Access Control</td>
<td>Access Control</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>Secure Communication</td>
<td>Secure Communication</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Secure Decentralization</td>
<td>Secure Decentralization</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Request Forgeries</td>
<td>Request Forgeries</td>
<td>Request Forgeries</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>Vulnerable Components</td>
<td>Vulnerable Components</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Client side</td>
<td>Client side</td>
<td>Client side</td>
<td>2</td>
</tr>
</tbody>
</table>

For example, summary of one lesson plan “Injection Flaws” is shown as bellow.

1) Subtopic

The lesson plan “Injection Flows” contains 4 subtopics: SQL Injection (introduction), SQL Injection (advanced), SQL Injection (mitigation), and XXE (XML eXternal Entity).

2) Contents

Each subtopic consists of some detailed contents: Explain the vulnerability; Assignments to learn about how to exploit the vulnerability; Describe the possible mitigation scenarios. For example:

- SQL Injection (introduction): What is SQL, What is SQL Injection
- SQL Injection (advanced): Combining SQL Injection Techniques, Blind SQL Injection
- SQL Injection (mitigation): Defense against SQL
- XXE (XML eXternal Entity): XML External Entity attack

3) Configuration

Each subtopic begins with Concept describing the explanatory policy, the Goal describing the achievement of the lesson, followed by the explanation of the vulnerability and some assignments that confirms the understanding.

B. Problems and Measures of Exercises Using WebGoat

1) Curriculum

The exercises theme of WebGoat consists of the latest technical issue selected by OWASP experts, but the learning level is unclear. Proper level setting is necessary for the exercises in the curriculum of universities in accordance with participants’ practical skills and purpose of the training. See reference for details [11].

We adopted HMM (Hunting Maturity Model) proposed by Sqrrl, and SecBok (security knowledge field) human resource skill map published by JNSA (Japan Network Security Association) for the level setting [12], [13]. We matched the contents of WebGoat exercises with HMM level definitions and SecBok skill items.

The outline of the level setting is shown in Fig. 2. After clarifying the level setting, we developed the customizable curriculum using WebGoat.

IV. DEVELOPMENT OF TRAINING CONTENTS

A. Basic Concept

The training contents implemented on CyExec consist of a basic part and an applied part. Fig. 3 shows learning configuration of CyExec.

Regarding the basic exercises, we utilized the latest
version of WebGoat, v8.0.0.M25 [11]. We expect trainees to use inspection tools together such as OWASP ZAP (OWASP Zed Attack Proxy) to detect vulnerabilities [14].

WebGoat exercises are developed based on OWASP experts’ technical regular research on high risk vulnerabilities, therefore the basic exercises have high training effects.

Regarding the applied exercises, we constructed interactive environment of attacker and defender to realize practical exercises. Fig. 4 shows an example of the applied exercises environment.

The environment for the attack defense training is constructed using Docker on a virtual guest OS in a closed network separated from the outside. Both of the attacker’s and the defender’s practice environments are built on the Docker container. The trainees playing the role of the attacker exploit vulnerabilities from the attacker’s environment and the trainees playing the role of defender monitor the network traffic and analyze the log regarding the cyberattack in the defender’s environment.

High expertise and a long periods of time are required to develop the new cyberattack training contents by a single organization. CyExec enables joint development of the training contents in short time by cooperation of multiple universities and private enterprises [7], [8].

In addition to the basic exercises and the applied exercises, we also focused on law and ethics practice before the exercises to prevent participants from illegal and injustice use of acquired skill by intention or fault.

We expect active learning where participants are engaged in solving problems, learn lessons at home using a training guidance in advance and exercises with a lecturer’s help after learning the necessary skills.

For example, summary of “SQL Injection” exercises is shown as bellow.

1) Purpose of the exercises

Purpose of the basic exercises is to understand basic knowledge of SQL, outline of SQL injection and detection method and to acquire basic skills on cyberattack and defense through assignments.

2) Capable skills of being acquired

The following are example of learnable skills. These items are selected from the SecBok skill table described in Section III. B.

- Basic knowledge of vulnerability assessments
- Knowledge of system and application security threats and vulnerabilities
- Skill in recognizing and categorizing types of vulnerabilities and associated attacks

3) Basic knowledge of SQL

SQL (Structured Query Language) is a language for data definition, data control, and data manipulation. It enables accessing and updating records of a database. SQL consists of three types of statements:

Data Manipulation Language (DML): SELECT, INSERT, UPDATE, DELETE

Data Definition Language (DDL): CREATE, ALTER, DROP, TRUNCATE

Data Control Language (DCL): GRANT, REVOKE

4) Outline of SQL injection

SQL injection is code injection technique using vulnerability which allows an application to execute an unintended malicious SQL statements inserted into request of an entry field to manipulate the database improperly. Exploiting this vulnerability causes falsification and leakage of data in the database.

5) Harmful effects of SQL injection

SQL injection induces disclosure or destruction of the confidential data, improper program execution and file reference, and theft of database server administrator authority.

6) Attack example

An overflow of a literal (a constant in the SQL statement) causes the SQL injection. The following is an attack example using the vulnerability.

"select * from users where name="'+username+'"; (1)
The variable userName in Statement (1) stores the input value received from the request. For example, when the attacker supplies unexpected string "Smith' or '1'='1" in the variable userName, the range of the SQL literal becomes to be "name='Smith'" and the part of "or '1'='1" is pushed out and executed. Since "or '1'='1" is always true, information that does not match the condition leaks.

7) Assignments

Fig. 5 shows an example of SQL Injection assignments.

Try It! String SQL Injection

The query in the code builds a dynamic query as seen in the previous example. The query in the code builds a dynamic query by concatenating strings making it susceptible to String SQL injection:

```
"select * from users where LAST_NAME = '?" and userName = """
```

Using the form below try to retrieve all the users from the users table. You should now need to know any specific user name to get the complete list, however you can use 'Smith' to see the data for one user.

```
Account Name: Admin
You have successfully:
USERID, FIRST_NAME, LAST_NAME, CC_NUMBER, CC_TYPE, COOKIE, LOGIN, COUNT, 
101, Joe, Snow, 967654321, VISA, 0, 0, 
101, Joe, Snow, 9234500005411, MC, 0, 
102, John, Smith, 2435400000022, MC, 0, 
102, John, Smith, 4332200020222, AMEX, 0, 
103, Jane, Plane, 135456789, MC, 0, 
103, Jane, Plane, 33546750933, AMEX, 0, 
103, Jolly, Hornby, 175960769, MC, 0, 
103, Jolly, Hornby, 333300000333, AMEX, 0, 
1032, Grumpy, yoshiotheworkskeith, ET3564468, MC, 0, 
10332, Grumpy, yoshiotheworkskeith, ET3564468, AMEX, 0, 
10003, Peter, Sand, 023506970, MC, 0, 
10003, Peter, Sand, 20080945032, AMEX, 0, 
10013, Joseph, Something, 2345634533, AMEX, 0, 
10037, Chico, Monkey, 3245989653, CM, 0, 
19004, Mr. Goat, 3359035533, VISA, 0, 
```

Fig. 5. Example of assignments about string type SQL injection.

Trainees try to acquire and display the user information from the database exploiting the vulnerability of the SQL injection without access permission.

C. Applied Exercises Using Original Cyberattack and Defense Program

After learning the basics of vulnerability and countermeasures in the basic exercise CyExec provides the applied exercises to offer more practical cyberattack and defense techniques. Trainees can improve the response ability in organization against various kind of cyberattacks through the applied exercises simulating the different roles and viewpoints such as attacker and defender, manager and general user.

The exercises are carried out separately in the attacker’s side and the defender’s side. The outline of the applied exercises is described below.

1) Purpose of the exercises

The purpose of the exercises is to acquire cyberattack and defense skill comprehensively from the following viewpoints. Purpose of training attack skill is limited only to deep understandings of defense technology.

- To understand cyberattack methods exploiting vulnerabilities: Vulnerability detection using tools such as OWASP ZAP, attacks exploiting vulnerabilities of software or server
- To understand defending methods against cyberattacks: detection and analysis of cyberattacks using access log file, countermeasures against cyberattacks

2) Capable skills of being acquired

Examples of specific learnable skills are as following.

- Ability to identify systemic security issues based on the analysis of vulnerability and configuration data
- Knowledge of penetration testing principles, tools, and techniques
- Skill in conducting vulnerability scans and recognizing vulnerabilities in security systems
- Skill in using network analysis tools to identify vulnerabilities

3) Configuration of exercises system

Fig. 6 shows an outline of configuration of the exercises system.

The trainees exercise after logging in either attacking terminal or defending terminal operating on the Docker container on CyExec. Two terminals communicate each other through a virtual network constructed on the Docker container.

The attacker operating the attacking terminal exploits vulnerability, logs in the defending web server via virtual network without proper authorization, and attempts to infect the server with an attack script. The main goal for the attacker is to steal confidential information using the script.

The defender tries to find the sign of the attack by monitoring the network traffic logs from the attacker. The main goal for the defender is to consider the attacks, to implement countermeasures and to ensure that the attacks can be prevented.

4) Program specification

The specifications of the exercises program are shown below.

a) Guest OS
- OS: Ubuntu 18.04
- Memory: 2GB
- Storage: 20GB
- Container platform: Docker 18.09

b) Exercises program
- Programming language: Ruby 2.5.1, PHP 7.2
- Database management system: MySQL 8.0
- Web server software: Apache 2.4

c) Software development process model

We adopted the spiral development model. The development period was divided into several phases to develop and improve the exercises program at the proper time.
We promoted the development efficiently by using Docker Hub to share the created containers on the cloud, and GitHub to manage source code versions.

d) Development man-month
Approximately 6 man-month

5) Exercises scenario
Fig. 7 shows the image of the exercises scenario.

![Fig. 7. Image of the exercises scenario.](image)

a) Scenario of the attacker
The contents of the exercises of the attacker are shown below.
1) Test the vulnerability of the web server using OWASP ZAP and make a report of inspection results.
2) Exploit an SQL injection attack on the vulnerable web application and attempt unauthorized login.
3) Upload an attack script using the file upload function after login.
4) Access to and execute the attack script from the browser on the attacker’s terminal and execute the script to stop the firewall using the attack script.
5) Unauthorized access to the confidential information in the web server using SSH command.

b) Scenario of the defender
The contents of the exercises of the defender are shown below.
1) Monitor the network log from the attacker using tools such as Apache Log Analyzer to detect the SQL injection attack and the attack script.
2) Modify source code causing SQL injection vulnerability and confirm the improvement.
3) Implement a Web Application Firewall (WAF) and confirm that WAF can prevent the unauthorized access to the confidential information on the Web server.

V. CONCLUSION
Cyberattacks including targeted attacks are increasing and becoming serious issues of digital society. Enforcement of the human resource development for personnel having cyberattack and defense skills is an urgent priority, but the environment for growing up the cyber security personnel is still poor because of the high cost of the exercises system and the shortage of the personnel to maintain and manage the exercises environment.

Therefore, we developed a cybersecurity exercises system CyExec consisting of virtual environments using VirtualBox and Docker container based on ecosystem.

The basic contents on CyExec are using open source vulnerability scanning tool WebGoat. The applied contents on CyExec are our original cyberattack and defense exercises programs.

In this paper, we introduced the contents of the vulnerability diagnosis exercise using WebGoat implemented in CyExec and the cyberattack and defense exercises program we developed.

In the future, we plan to develop and utilize CyExec with other universities and small and medium-sized enterprises jointly.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS
RN, SS and YS conceived of the presented idea; ST, YK, SS and YS developed the theory and implemented the exercises contents; NM, SS and YS wrote the paper; YS supervises the findings of this work; all authors had approved the final version.

ACKNOWLEDGMENT
This work was supported by JSPS KAKENHI Grant number JP 19K03006. This study was supported in part by Research and Study Project of Tokai University Educational System General Research Organization.

REFERENCES

Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).

Nobuaki Maki was born in Tokyo, Japan in 1972. He received the B.Sc. in physics from Kyoto University, Kyoto, Japan, in 1996. Since 2018, he has been enrolled in the master course of Advanced Institute of Industrial Technology, Tokyo, Japan. In 1998, he joined the Ministry of Health Labour and Welfare, Government of Japan. In 2012, he became First Secretary, Embassy of Japan in Indonesia. In 2015, he became Deputy Director, Overseas Cooperation Division, Human Resource Development Bureau, Ministry of Health Labour and Welfare. Since 2016, he has been Cyber Security Officer, Compensation Operation Division, Labour Standards Bureau, Ministry of Health Labour and Welfare.