Deployment of Honeypot and SIEM Tools for Cyber Security Education Model in UITM

https://doi.org/10.3991/ijet.v17i20.32901

Muhammad Azizi Mohd Ariffin[1(✉)], Mohamed Yusof Darus[1], Haryani Haron[1], Aditya Kurniawan[2], Yohan Muliono[2], Chrisando Ryan Pardomuan[2]
[1]Universiti Teknologi MARA, Shah Alam, Malaysia
[2]Bina Nusantara University, Jakarta, Indonesia
mazizi@fskm.uitm.edu.my

Abstract—Nowadays the threat of cyber-attacks is increasing as more organizations undergo digital transformation. Therefore, organizations need to take proactive measures to mitigate the cyber threat to avoid a further loss to their business. To mitigate cyber risk effectively, organizations need to employ competent people in the IT security team to implement effective security controls. But there is a shortage of cyber security talent or professionals in the job market and to produce talents in the cyber security field requires extensive effort in education and training. A good cyber security education program should have to date curriculum and provide practical experience. To achieve this, the program must be supported by a cyber security lab equipped with various software, equipment, and tools used by a real professional in the industry. Therefore, this paper proposed a model of a cyber security lab equipped with honeypot and SIEM systems to enhance the quality of cyber security education. The cyber security lab based on the model was deployed at Universiti Teknologi MARA (UiTM) and used for teaching and learning activities. The honeypot will provide student experience analyzing the behavior of hackers while the SIEM system will aggregate the logs data of the Campus Network Firewall in real-time. To evaluate the effectiveness of the proposed lab model, a functional test and a survey was conducted. The survey result shows that majority of the respondent agreed that the cyber security lab improve their teaching and learning experience while taking the cyber security subject.

Keywords—cyber security, education, honeypot, SIEM, network security

1 Introduction

As more organization undergo digital transformation, the threat of cyber-attacks also increases as the attack surface become bigger. Cyber-attacks on an organization's IT infrastructure will lead to a data breach, and financial and reputation loss. A survey conducted by IBM [1] states that the cost of a data breach in an organization increase from $3.86 million in 2020 to $4.24 million in 2021. The survey also states that the financial loss is higher for an organization with a less mature security posture. Meanwhile in Malaysia, the reported cyber incidents such as Denial-of-service (DoS) attacks, system

intrusion and malware infection are rising from 2,429 in 2016 to 3,787 in 2019 [2]. Therefore, organizations need to take proactive measures to mitigate the cyber threat to avoid a further loss to their business.

For the organization to have effective cyber security control, it must employ talents in cyber security as part of the IT security team. A competent security team will implement effective security controls, formulate the right security policy and be able to respond to incidents correctly. This will result in more prudent cyber security measures for the organization. But nowadays there is a shortage in IT talent [3] and specifically a cyber security talent which will impact the recruitment of competent staff. A report published in the year 2020 highlighted that 64% of the organization reported cyber security staff shortages while 22% reported significant cyber security staff shortages [4]. To fill up the shortage in the workforce, cyber security education and training need to be given to the student or reskilling the existing workforce.

To produce talents in the cyber security field requires extensive effort in education, training, and certification to develop and nurture the required skillset. Besides, the cyber security education and training given must adhere to certain qualities so that prospective student skillset is in line with industry and job market requirements. Cyber security education which did not adhere to certain qualities will produce incompetent graduates in the workforce. A study [5] conducted in a developing country such as Ecuador regarding cyber security education shows that most universities do not have labs or equipment to deliver cyber security education, only 11% of respondents have a lab. This may impact the quality of students.

Besides a competent instructor, a good cyber security education program should have to date curriculum and practical exercises such as the "Capture the Flag" event and simulation of real-world security threats and breaches [6]. To achieve this, the program must be supported by a cyber security lab equipped with various software, equipment, and tools used by real professionals in the industry just as vocational school providing student with a hands-on maker space to nurture practical skills [7]. This will provide valuable experience to the students and instill confidence in the organization to recruit the student into their workforce.

Several works have been done in using software, hardware, or tools to aid cyber security education. The work of [8] proposed the use of augmented reality (AR) during cyber security training activities. The AR system visualized the implementation of a cyber-physical system and focused on the implementation of data acquisition, storage and processing platform for new sensor networks and instruments. But the contents are only limited to a cyber-physical system and did not provide the student with practical experience such as analyzing traffic data or analyzing appliance logs. The work of [9] introduces a game called HackLearn for cyber security education. The game was developed based on the COFELET framework and the student will be presented with interactive content related to information security via a GUI interface. This paper [10] also proposed a game for cyber security education called Sherlocked, it uses a puzzle approach to enhance student understanding of a concept in information security. Although game-based learning helps the student to understand the fundamental theory and hands-on skills with common tools such as Nmap or Metasploit, it lacks the experience in analyzing the threat on a real system.

Besides that, the cyber security educational experience can be enhanced using cyber range has been proposed by [11] [12]. By having a cyber range, students can sharpen their penetration skills in a virtual or simulated environment. But cyber range only provides practical experience from the Red-Team perspective. There is a need for the student to also sharpen their Blue-Team skillset if the organization needs to hire the student as an analyst for their Security Operation Centre (SOC). Meanwhile, the work of [13] acknowledges the need for the student to sharpen their Blue-Team skillset as security analysts in SOC. The work proposed a concept of a digital twin-based cyber range for SOC analysts equipped with a security information and event management (SIEM) system which displays real-time log data. However, the concept lacks the component to sharpen the Red-Team skillset such as penetration testing.

Based on the review of the previous work. Lack of work has been done to provide cyber security students with tools which able to improve both Blue-Team (Threat Analysis) and Red-Team (Penetration Test) skillset with real system experience. Therefore, to address the gap which has been identified, this paper proposed a model of a cyber security lab equipped with Honeypot and SIEM systems to enhance the quality of cyber security in education. The honeypot will simulate a real vulnerable system such as a web server and student can sharpen their penetration skill on the honeypot and can also analyze the behavior of external hackers attempting to penetrate the honeypot. Meanwhile, the SIEM will aggregate the logs data of the Campus Network Firewall and display it in real-time to the student. Using the correlated threat data from the SIEM, students were able to sharpen their Blue-Team skills by analyzing the threat data. The cyber security lab based on the model was deployed at Universiti Teknologi MARA (UiTM) which is the largest public university in Malaysia. The deployment was evaluated by surveying students and lecturers. The lab has benefited many students taking information security subjects for their undergraduate or postgraduate studies.

2 Literature review

This section will review topics related to cyber security education, honeypot deployment and SIEM technology. All the topics are fundamental for the proposed cyber security lab model.

2.1 Cyber security education

As society is transforming into a digital society, many human daily tasks depend on a digital system. This will increase the individual digital footprint and vulnerability to cyber-attack. Therefore, cyber security education is becoming more important not only to create awareness in society but also to produce potential cyber security professionals in the future. Due to the lack of cyber security education, we can see the rise of cyber-crime, online fraud, and cyber-bully in society [14]. Many recent works and research have been done to improve the quality of delivering cyber security education. Some proposed a gamification [15] [16] [17] of cyber security content for education, multimedia tools [18] [19] and simulation such as cyber range [20] [21]. The work of [15] proposes a framework for information security awareness and training programs

called as Cyber-Hero, but the framework only focuses on raising awareness of human weakness and social engineering and left out other technical skills in cyber security. There are also authors [16] who proposed gamification of cyber security education via Facebook messenger to increase cyber security knowledge, but the proposal only provides awareness and did not provide any practical knowledge and skill. There is also a security awareness program [17] being proposed to improve company employee awareness of social engineering attacks, but the delivery of content is via embedded video and games and did not provide the employee with real-world experience.

Meanwhile, the work of [18] develops a multimedia curriculum to teach about online privacy but it lacks other cyber security topics and more focus on public education. The work of [19] introduces ADA which is an open-source robot tool that provides cyber security information interactively, it can read the Twitter feed, RSS feed and show articles related to cyber security, but it does provide any mechanism to help students in sharpening their practical skill. For the simulated environment, the work of [20] proposes CyExec*, a cyber range setup which able to randomize the scenario using containerized environment. Although it is a good platform for student to develop their pen testing skill, it lacks the mechanism to sharpen their blue team skillset. The work of [21] presents a model for implementation of a hybrid cyber-range based on the model of a real water supply system, but the work is conceptual and did not demonstrate workable implementation of the cyber range. It will be interesting to see the development of augmented reality learning such as [22] being applied for cyber security education to make it more immersive to the students.

A finding from a study [23] indicates that the best outcome for cyber security education is when the structured simulated environment is paired with live competitive activities. Therefore, the model proposed by this research used live traffic from the campus network firewall to provide the student with live real threat data to enhance their experience.

2.2 Honeypot deployment

Honeypot is a vulnerable system which often used by the administrator to detect, deflect, and counteract unauthorized access to the IT infrastructure. The honeypot system is purposely made to be vulnerable, and it will emulate various real services. The honeypot can be deployed on a bare-metal machine, in a virtual machine such as in [24] or in a more recent trend deployed in a container [25]. Besides that, a honeypot can be divided into three low interaction levels, medium, and high interaction.

A low interaction refers to a honeypot that provides limited interaction to the hacker and only emulates a certain part of the operating system or the protocol. An example of the implementation of low interaction honeypot is Honeyd [26]. A high interaction refers to a real system (non-production) that is used as a honeypot to provide hackers interaction with a real system making it harder for them to guess, whether they are being observed or diverted. The work of [27] implements a high interaction honeypot

for SQL injection analysis. Lastly, medium interaction refers to honeypot which can provide a more advanced response to hackers compared to low interaction honeypot. It can emulate certain aspect of the system and provide some depth in the interaction; example of a medium interaction honeypot is [28] which implement a honeypot for the internet of things target. This research project proposed to use a medium to high interaction honeypot to provide the student with a platform to improve their red team skills.

2.3 SIEM technology

Security Information and Event Management System or often abbreviated as SIEM is often deployed by an organization at their SOC, Figure 1 shows an example of SIEM being deployed at the company security operation center (SOC).

Fig. 1. SIEM in Security Operation Center (SOC) [29]

SIEM system is responsible for performing real-time analysis, correlation and visualization of security alerts generated by the network or security appliances. Many SIEM systems have been developed either commercial or open source. An example commercial SIEM solution is QRadar developed by IBM, Splunk and Data dog. Meanwhile, an example of open source SIEM is AlienVault OSSIM and ELK stack. In the recent development of SIEM technology, there is an emerging trend to apply big data technologies to enhance SIEM data processing such as [30] which integrates Apache Spark to the SIEM. Besides that, there is also a recent trend to apply artificial intelligence elements to the SIEM to improve the accuracy in detecting threats such as the AI-SIEM system developed by [31].

3 Methodology

This section will discuss the methodology used in this research to achieve its aim which is to propose and deploy the model of a cyber security lab for cyber security education and training. Figure 2 shows the methodology of the research.

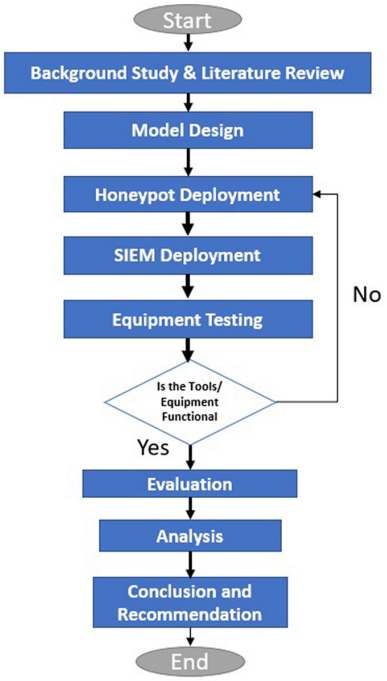


Fig. 2. Research flowchart

Based on Figure 2, the research activity starts with a background study and literature review. This is to obtain the current state-of-art of cyber security education especially on teaching aid tools and identify the gap in current practice. Besides, at this stage, feedback from lecturers, instructors and students of information security subjects was also gathered. Based on information gathered during background study and literature review, this research project designs the proposed model for the cyber security lab. After that, the research project begins the deployment of honeypot and SIEM devices. The honeypot will be deployed on a Virtual Machine (VM) and will be placed inside the DMZ subnet, meanwhile, the SIEM will be developed to process real-time logs from the UiTM gateway firewall and from the Honeypot. After the deployment phase, the functional test was conducted to determine whether the honeypot and SIEM are functioning as been designed in the model. If the deployment did not satisfy the requirement, it will be undergoing fine-tuning and re-deployed.

After the deployment phase is done, the model which includes the honeypot and SIEM was evaluated quantitatively via a survey. The cyber security lab which is set up based on the proposed model will be used for teaching and learning activities at UiTM. After the activities has concluded, a survey was conducted on the lecturer and the student. After that, the survey result was analyzed and after that, the research project provides a conclusion and recommendation. The next section will discuss in more detail the proposed cyber security lab model.

3.1 Proposed cyber security lab model

The model proposed by the research project was designed based on information gathered from literature and feedback from lecturers and students taking information security subjects in UiTM. Figure 3 shows the proposed cyber security lab model. Based on the model, the SIEM and honeypot devices were deployed inside a cyber lab subnet which also resides inside the UiTM campus network. The UiTM's campus network provides connectivity and internet access to staff workstations, Wi-Fi network and the production server of the data center. All inbound or outbound traffic of the campus network will pass through the gateway firewall at the network border. The gateway firewall was equipped with Fortinet Unified Threat Management (UTM), therefore most of the threats passing through the firewall will be filtered and logged based on the up-to-date signature. The logs which are generated by the firewall will be sent to the SIEM in real-time for further analysis via the Syslog protocol. This will provide the student with a wealth of threat data for their learning activities.

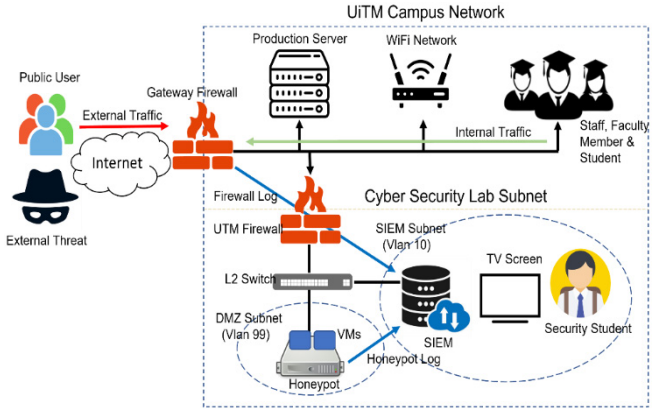


Fig. 3. Proposed cyber security lab model

Meanwhile, the cyber security lab subnet will be connected to the campus network via an internal firewall. The purpose of the internal firewall is to filter traffic and perform layer 3 routing to route traffic from different VLANs and port forwarding. The VLAN was used to separate the traffic of the DMZ subnet and the SIEM subnet. The DMZ subnet will host the physical server which hosts the honeypot software. The honeypot software will be hosted in the virtual machine to ensure privileges separation if hackers able to break out of the honeypot environment. The DMZ subnet will be assigned with public IPs to expose the honeypot services to external hackers. The student can also access the DMZ subnet to practice their system penetration skills. The server log generated by the honeypot during operation will be forwarded to the SIEM via Syslog for correlation and visualization of threat data.

The SIEM subnet in a separate VLAN will host the SIEM device, Smart TV, and computer lab workstation. The SIEM device will be assigned with private IP which is reachable by the gateway firewall and the honeypot in the DMZ. The SIEM devices will process and correlate the raw log and visualized the threat data on the dashboard. The dashboard will then be displayed on the smart TV screen to provide real-time threat data for teaching and learning activities. Students can also further inspect the threat data by

accessing the SIEM system using the workstation available at the lab. The next section will discuss the SIEM tools and the process they undertake to process the threat data.

3.2 SIEM tools

The purpose of the SIEM tool is to process, correlate and display the threat data received from the gateway firewall and honeypot. The processed threat data is then displayed on the dashboard to enhance the learning experience of the student. The SIEM tools were developed in partnership with the Techforte MSSguard platform [32]. The platform was further refined and customized based on the proposed model and UiTM training requirements. Figure 4 shows the SIEM data processing flowchart.

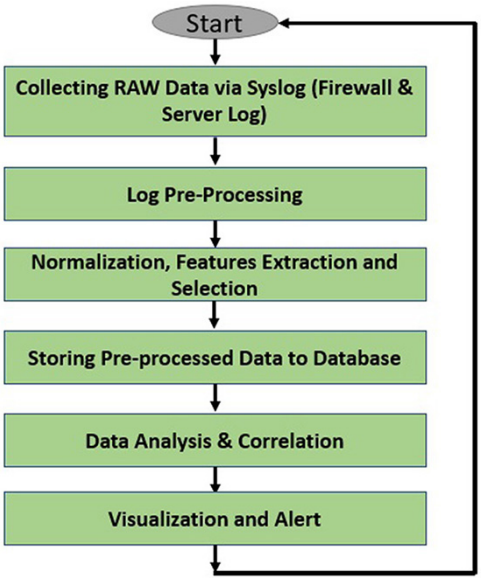


Fig. 4. SIEM data processing flowchart

Based on Figure 4, the SIEM process starts with collecting raw log data via the Syslog protocol (UDP 514). The raw log received at the network interface is first put in a queue and segregated into different directories based on source device type and IP. After the raw log has been stored in its respective directory, it will undergo a log pre-processing process. The purpose of pre-processing is to filter out any error in the format and prepare it for primary processing. During pre-processing the system will check whether the received log follows the format defined in RFC 5424 [33], if the format is malformed, it may indicate an error during log transmission. The pre-processing also helps the system identify the log timestamp, device ID, facility code, severity level, message number and message text. If required, the timestamp will be converted from the UTC zone to Malaysia (GMT+8) time zone.

After pre-processing the system will conduct data normalization, features extraction and selection. The normalization will process the log into the readable and structured format of the SIEM system. After that, the system will select and extract features in

the log message. The message will contain information such as protocol type, port number, attack type, source IP, destination IP, URL, signature information, and other information regarding the attack. For example, if the firewall detects malicious traffic from Mirai Botnet, it will block the traffic and generate logs about the event. The logs will contain information about Mirai Botnet traffic such as using the HTTP protocol to relay malicious information, using either port 80 or 8080, the IP of the slave and the command-and-control center. The message will also contain the hash value of the malicious payload. After the data has been processed, normalized, and undergone features selection and extraction, the information will be stored in a database as an event. The information will then be used for further analysis and correlation for attack pattern detection.

The next process is about data analysis and correlation. This is the phase where the SIEM connects the dots and correlates events from different sources to identify attack patterns. The analysis and correlation will be made based on rules being defined for a specific sequence of events that could be indicative of a breach in security. For example, the rules may state that if the number of requests sent from certain IPs and ports exceeds the threshold value, it may indicate an ongoing denial of service attack. After analyzing and correlating the data, the SIEM will visualize the threat or ongoing attack on the dashboard in the form of a chart and graph. This will improve situational awareness and allows a security analyst to easily view the data. This will provide the student with real-life experience of working in a SOC environment. If needed, the SIEM can also send an alert to the administrator and generate a report regarding the security threat event. The next section will discuss the honeypot deployment setup.

3.3 Honeypot deployment

The honeypot was deployed inside virtual machines hosted by a physical server in the DMZ subnet. This allows a different type of honeypot service and environment to be hosted on a physical host. Moreover, running a honeypot service inside a VM increase privilege separation as any breakout from any of the honeypot environment will only be confined inside a VM. Table 1 shows the list of the simulated protocol in the honeypot deployment.

Table 1. Honeypot simulated protocol

No	Simulated Protocol	Software	Description
1	Secure Shell (SSH)	Cowrie [34]	A Medium to High interaction honeypot which simulates a remote Linux Shell.
2	Telnet		
3	HTTP	Wordpot [35]	A Medium interaction honeypot that simulates a WordPress-based website.

Based on Table 1, the honeypot deployment simulates three common protocols which are SSH, Telnet and HTTP Protocol. Both SSH and Telnet are simulated using Cowrie which is a medium to high interaction honeypot developed using Python. It was designed to log brute force attacks and the shell interaction performed by the attacker. Cowrie can emulate the UNIX shell and act as SSH and telnet proxy. Meanwhile, the HTTP protocol is simulated using word pot software which is medium interaction.

It emulates WordPress based site and detects probes for plugins, themes and other common files used to fingerprint a WordPress installation. The logs generated by both cowrie and wordpot are sent to SIEM. The next section will discuss how the deployment was tested and evaluated.

3.4 Testing and evaluation

After the proposed model has been deployed in the cyber security lab, it will be tested and evaluated. After deployment, a functional test was carried out to ensure the SIEM tool and the honeypot met the proposed model requirement. During the functional test, the SIEM will be tested to whether able to correlate and visualized the threat correctly. Meanwhile, the honeypot was tested to see whether it's able to attract hacking attempts from the internet and generate the necessary log. The functional test result will be presented in the next section.

Once the deployment of SIEM and honeypot passed the functional test, the cyber security lab was used to conduct an information security class for one semester at UiTM. At the end of the semester, a survey was conducted on the students and lecturers to quantitatively evaluate the experience using the cyber security lab. The questions used during the survey are listed in Table 2 and the survey was conducted using the Google form platform. The result of the survey will be discussed further in the next section.

Table 2. Evaluation survey question

No	Question	Answer Option
1	Please enter your role in University Teknologi MARA	1. Lecturer 2. Student 3. Staff 4. Others
2	What is your education level?	1. High school 2. Undergraduate 3. Postgraduate 4. No Formal Education
3	Did you teach or enroll in a subject related to cyber or network security in UiTM?	1. Yes 2. No 3. Unsure
4	Did the FSKM SOC setup help you when studying Cyber or Network Security related subjects?	1. Yes 2. No 3. Unsure
5	Did FSKM SOC provide exposure to real-life Cyber Threat Data?	1. Yes 2. No 3. Unsure
6	Did FSKM SOC able to visualize cyber threat data?	1. Yes 2. No 3. Unsure

4 Result and discussion

This section discusses the result collected after the deployment of the proposed model in a cyber security lab. The section will discuss further the Honeypot and SIEM functional test and the evaluation survey result.

4.1 Honeypot functional test

This section will present the functional test result for the Cowrie and Wordpot honeypot.

Cowrie data. As been stated earlier in the methodology section, the Cowrie honeypot simulates SSH and Telnet protocol to provide remote access to the UNIX shell. Every attempt to connect to the SSH and Telnet will be logged, Figure 5 shows an example of Cowrie's raw log. The Figure 5 shows that the raw log contains much useful information such as the hacker's IP, username, and password they used during the attempt and the command that they used when they were able to access the shell. To determine whether the cowrie honeypot is functional, we observed and visualized the log generated by the honeypot.

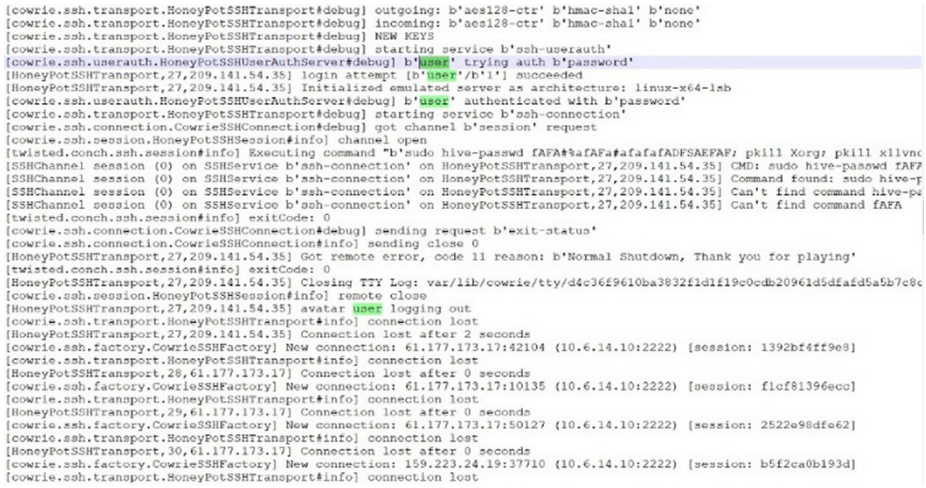


Fig. 5. Example of Cowrie Honeypot raw log

When the hacker's IP is visualized using SIEM in the form of a graph as shown in Figure 6, it shows there are multiple hacking attempts from a different external network. The graph shows that the highest attempts are from Petersburg Internet Network Ltd ASN, and the second-highest attempts are from Skynet Ltd ASN. This shows that the honeypot setup is working and able to attract hacking attempts from various networks on the Internet. Besides that, this project also visualized the common password and command used by the hackers into a word cloud as shown in Figures 7 and 8. The figure shows that the most common password used during the attempts are root and admin while the shell command is echo and uname command. This information provides lecturers and students which involve in information security subjects with valuable experience in a real-world scenario.

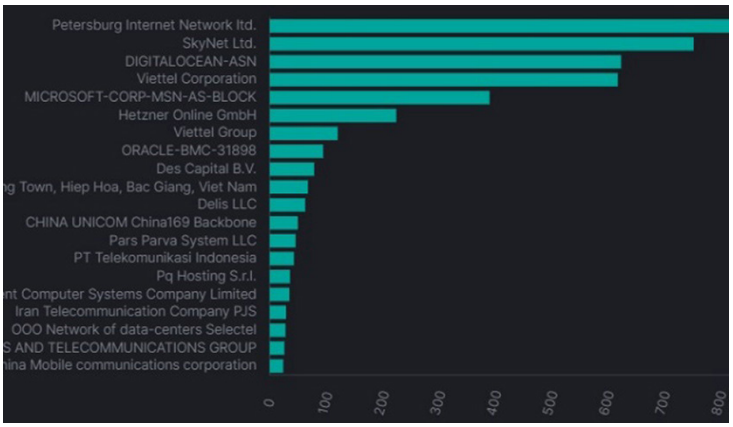


Fig. 6. Top honeypot attacker IP ASN

Fig. 7. Common password used by honeypot attacker

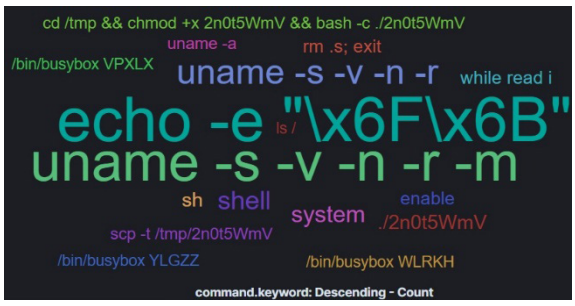


Fig. 8. Common shell command executed by honeypot attacker

Wordpot data. To simulate HTTP protocol or web service, this project uses wordpot honeypot to emulate a real WordPress site. Figure 9 shows the functional website complete with an admin panel which is emulated by the honeypot. Every attempt by hackers to conduct an SQL injection attack or brute force attack on a guest the admin password will be recorded to a log. Figure 10 shows the raw log generated by the wordpot honeypot. The raw log contains valuable information such as the hacker's IP address, URL path, username, and password that hackers attempt to login to the admin panel. This indicates that the honeypot is running as intended and able to gather valuable threat data.

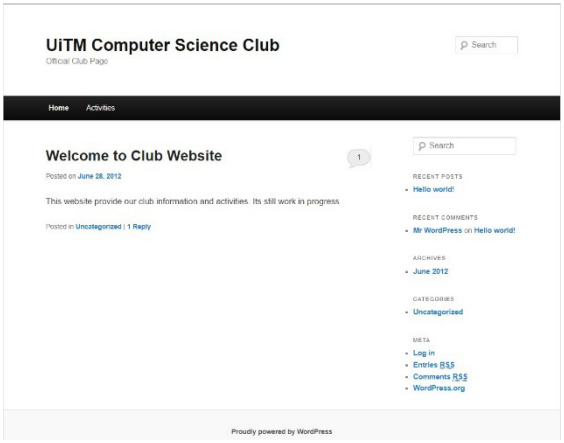


Fig. 9. WordPress website emulated by wordpot

```
2021-08-28 14:30:30,587 - Honeypot started on 10.5.14.10:8080
2021-08-30 10:30:31,679 - 10.5.14.1 probed for the admin panel with path: /
2021-08-30 10:30:31,691 - 10.5.14.1 probed for the login page
2021-08-30 10:30:35,952 - 10.5.14.1 tried to login with username test and password test1234
2021-08-30 10:30:36,637 - 10.5.14.1 tried to login with username  and password
2021-08-30 10:30:37,778 - 10.5.14.1 tried to login with username asdf and password
2021-08-30 10:38:23,557 - 10.5.14.1 tried to login with username asdf and password
2021-08-30 10:38:25,703 - 10.5.14.1 tried to login with username asf and password
2021-08-30 10:38:28,147 - 10.5.14.1 tried to login with username sdf and password dfasdf
2021-08-30 10:45:44,526 - 10.5.14.1 tried to login with username sdf and password dfasdf
2021-08-30 10:45:48,049 - 10.5.14.1 tried to login with username test123 and password test123
2021-08-30 10:48:12,421 - 10.5.14.1 tried to login with username test and password
2021-08-30 10:48:14,728 - 10.5.14.1 tried to login with username test and password 123
2021-08-30 11:05:44,508 - 10.5.14.1 probed for the admin panel with path: /
2021-08-30 11:05:44,516 - 10.5.14.1 probed for the login page
2021-08-30 11:05:47,877 - 10.5.14.1 tried to login with username uitm and password uitm123
2021-08-30 11:07:03,200 - 10.5.14.1 tried to login with username efwegwteg and password werwerwewer
2021-08-30 11:07:06,008 - 10.5.14.1 tried to login with username qewrwqerqwe and password ewqrerqewr
2021-08-30 11:07:08,384 - 10.5.14.1 tried to login with username qwerq and password qewr
2021-08-30 11:07:10,754 - 10.5.14.1 tried to login with username qwerqw and password qwerewq
```

Fig. 10. Example of wordpot honeypot raw log

4.2 SIEM functional test

After the SIEM has been deployed at the lab in UiTM, it needs to undergo functional testing to ensure it can process and correlate threat data received from the UiTM gateway firewall and the honeypot. This section will discuss the ability of the SIEM system to visualize the outbound and inbound attacks.

The first thing that the SIEM system will display on its dashboard is the Top Ten 10 attack event. Figure 11 shows SIEM Dashboard visualizing the Top 10 attacks. The Figure 11 shows that the highest type of attack is the EternalBlue Downloader Botnet, followed by Andromeda Botnet and Sinkhole botnet. This may indicate that a lot of hosts inside the campus network have been compromised with malware.

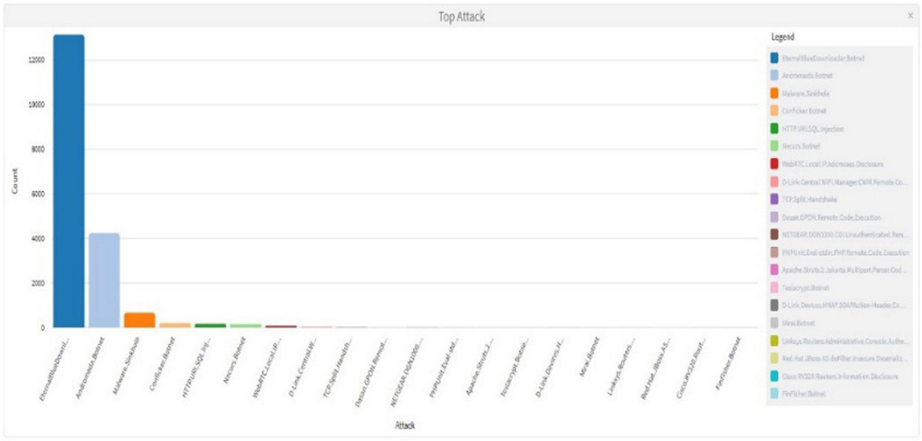


Fig. 11. SIEM dashboard visualizing top 10 attack

Outbound attack. Outbound attack refers to an attack that originates from the host inside the campus network and targets the host on the Internet. In the scenario of this project, most of the outbound attacks were originated from the host inside the campus network which had been compromised with botnet malware. This indicates that the compromised host has become a botnet slave and is either communicating with the command and control (C&C) host on the external network or performing an attack on behalf of the botnet master. Figure 12 shows the IP address of the Internal host which the attack or malicious traffic originated from.

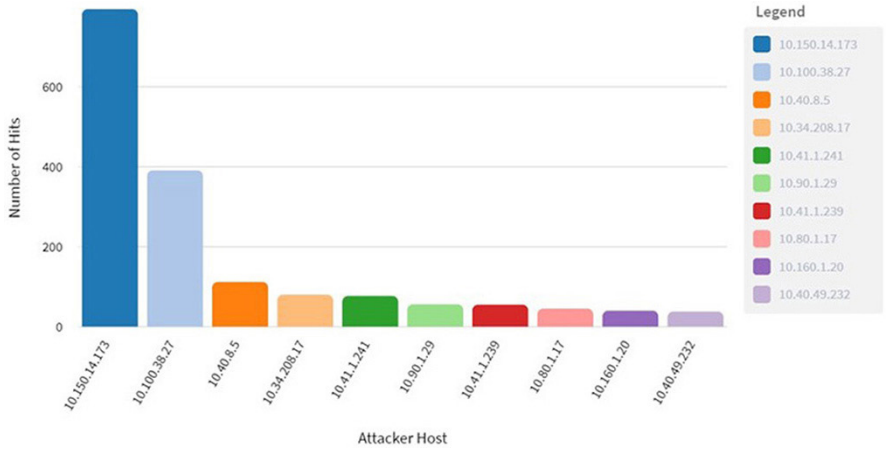
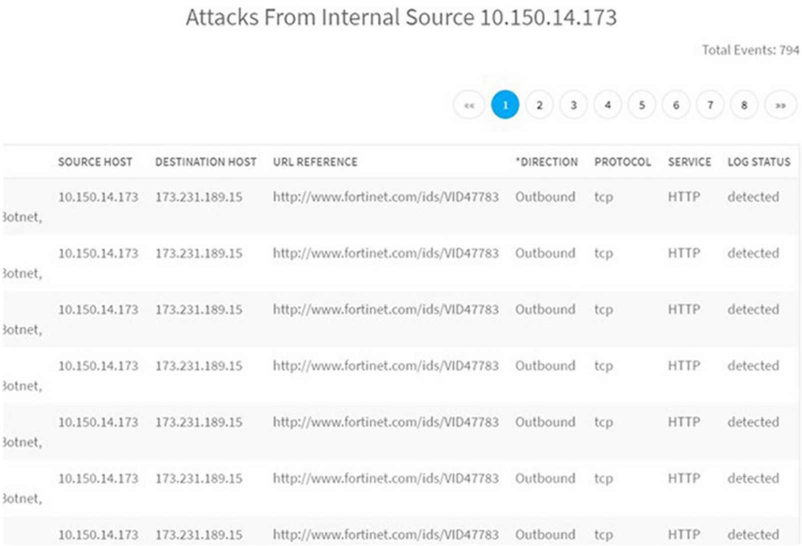


Fig. 12. Top internal attacker IP

Based on Figure 12, the highest number of malicious traffic came from 10.150.14.173 IP addresses and the second highest came from 10.100.38.27. The SIEM system was able to further show the type of attack of malicious traffic coming from the top internal attacker IP. Figure 13 shows the Attack Type from the top Internal Attacker. Based on the figure shows that the malicious traffic is HTTP based and the destination IP is 173.231.189.15 and the traffic is using TCP as the transport protocol.


Attacks From Internal Source 10.150.14.173

Total Events: 794

«« 1 2 3 4 5 6 7 8 »»

	SOURCE HOST	DESTINATION HOST	URL REFERENCE	*DIRECTION	PROTOCOL	SERVICE	LOG STATUS
3otnet,	10.150.14.173	173.231.189.15	http://www.fortinet.com/ids/VID47783	Outbound	tcp	HTTP	detected
3otnet,	10.150.14.173	173.231.189.15	http://www.fortinet.com/ids/VID47783	Outbound	tcp	HTTP	detected
3otnet,	10.150.14.173	173.231.189.15	http://www.fortinet.com/ids/VID47783	Outbound	tcp	HTTP	detected
3otnet,	10.150.14.173	173.231.189.15	http://www.fortinet.com/ids/VID47783	Outbound	tcp	HTTP	detected
3otnet,	10.150.14.173	173.231.189.15	http://www.fortinet.com/ids/VID47783	Outbound	tcp	HTTP	detected
3otnet,	10.150.14.173	173.231.189.15	http://www.fortinet.com/ids/VID47783	Outbound	tcp	HTTP	detected
	10.150.14.173	173.231.189.15	http://www.fortinet.com/ids/VID47783	Outbound	tcp	HTTP	detected

Fig. 13. Attack type from internal attacker

Moreover, the SIEM can perform a detailed inspection of threat events. Figure 14 shows further inspection of the threat event from an internal attacker. The Figure 14 shows that the traffic matches the Fortinet signature of Mirai Botnet, and the destination port is 80. This shows that the SIEM system can successfully process and correlate the event of an outbound attack. This will provide the student with experience in inspecting real botnet traffic.

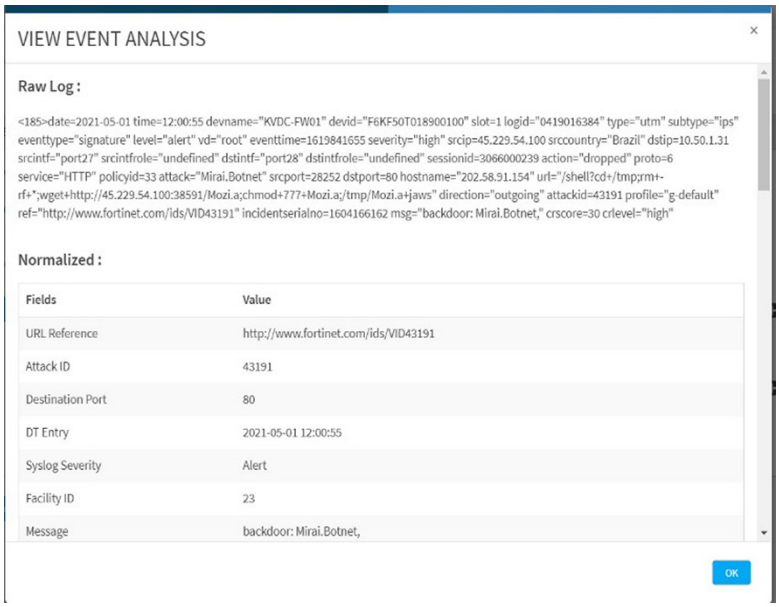


Fig. 14. Further inspection of Mirai Botnet log from internal attacker

Inbound attack. Inbound attack refers to an attack that originates from an external host or Internet and targets the host inside the campus network. The external hackers often targeted hosts inside the campus network which are assigned with public IPs. Figure 15 shows the top external attacker IP being correlated by the SIEM system.

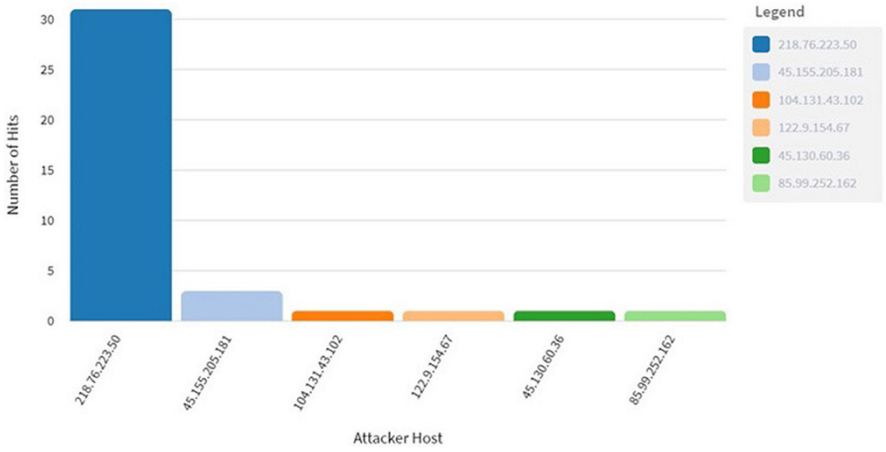


Fig. 15. Top external attacker IP

Based on Figure 15, the inbound attack comes from multiple public IPs on the Internet and the highest number of attacks came from 218.76.223.50 and follows by 45.155.205.181. Meanwhile, Figure 16 shows the top internal host which became the top target from the external host.

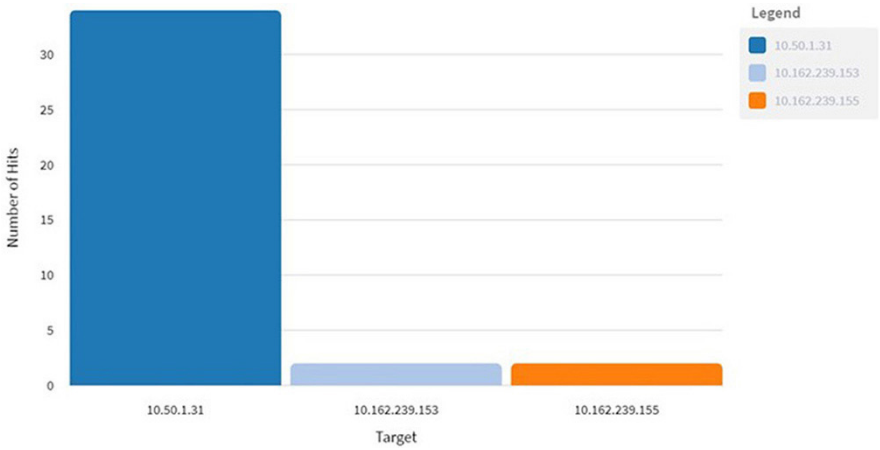


Fig. 16. Top internal target IP

Based on Figure 16 shows that 10.50.1.31 become the most targeted host inside the campus network. Upon further inspection, the host is a University production server serving a web service. Figure 17 shows different attack types from the external attacker to the top internal target. Figure 17 shows that different type of web-based attacks such as PHP remote code execution and SQL Injection has been attempted on the host. The results shows that the SIEM system can successfully process and correlate the event of an inbound attack.

Attacks To Internal Target 10.50.1.31

Total Events: 34

« 1 2 »

*PROCESSED TIME	ATTACK ID	MESSAGE	SOURCE HOST	DESTINATION HOST
2021-05-17 15:27:25	32416	applications3: Generic.XXE.Detection,	45.155.205.181	10.50.1.31
2021-05-17 15:27:21	45765	web_server: PHPUnit.Eval-stdin.PHP.Remote.Code.Execution,	45.155.205.181	10.50.1.31
2021-05-17 12:53:34	49427	applications3: vBulletin.tabbedcontainer.Template.Remote.PHP.Code.Execution,	218.76.223.50	10.50.1.31
2021-05-17 12:53:33	41512	web_misc: HTTP.Header.SQL.Injection,	218.76.223.50	10.50.1.31
2021-05-17 12:53:33	41851	web_app3: Joomla!.Core.Session.Remote.Code.Execution,	218.76.223.50	10.50.1.31
2021-05-17 12:53:32	48872	applications3: Tongda.Office.Anywhere.Unauthorized.File.Upload,	218.76.223.50	10.50.1.31
2021-05-17 12:53:32	47291	web_app3: ThinkPHP.Controller.Parameter.Remote.Code.Execution,	218.76.223.50	10.50.1.31

Fig. 17. Attack type from external attacker to internal target

4.3 Evaluation survey result

After the functional test, a survey has been conducted to evaluate the effectiveness of the proposed model during teaching and learning activities. The survey was conducted at the end of a semester at UiTM after the lab has been used for teaching and learning activities in an information security class. A survey has been sent to a group of 50 individuals and 27 provide their responses.

The first three questions that have been asked are about knowing the respondent's background. Figure 18 shows the result of question 1 which ask about the respondent's role, Figure 19 shows the result of question 2 which asks respondent's education level and Figure 20 shows the result of question 3 which ask whether the respondent teaches or enrolls in cyber security subject at UiTM.

Based on Figure 18, shows that most of the respondents (92.6%) are a student and minority are the lecturer (7.4%). Figure 19 shows the majority (59.35%) of the respondents are undergraduate students while the rest are postgraduate (40.7%). Meanwhile, Figure 20 shows that all respondents are either teach or enroll in cyber security subjects. Based on the result of questions 1, 2 and 3 shows that the survey is targeting the right respondent and the feedback they gave regarding the model and lab in the next set of questions are relevant.

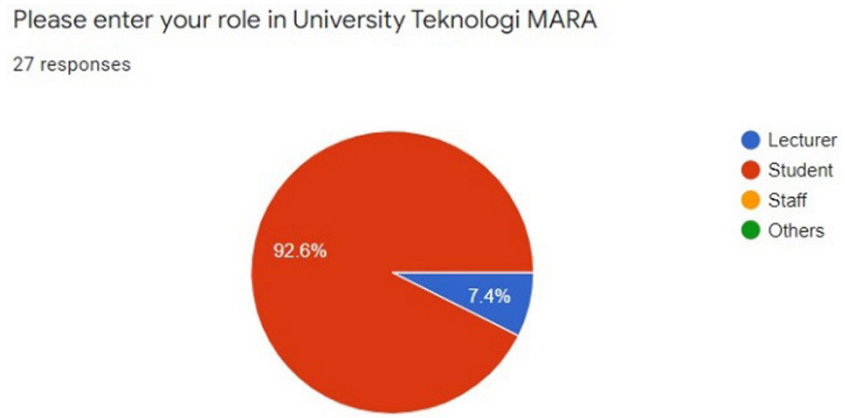


Fig. 18. Question 1 survey result

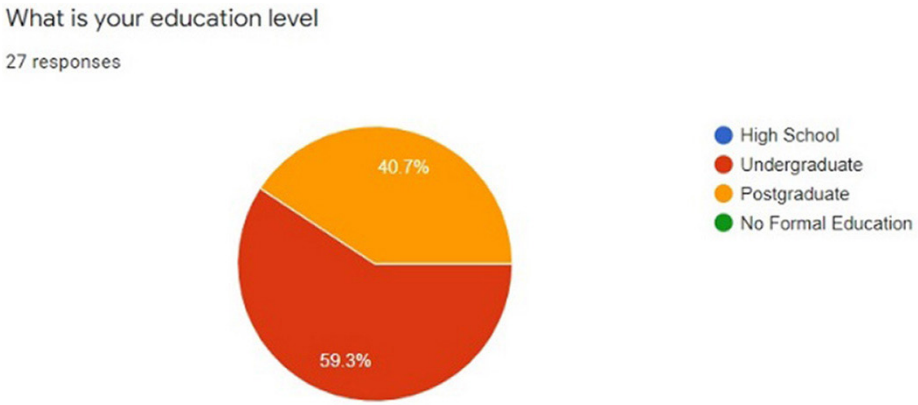


Fig. 19. Question 2 survey result

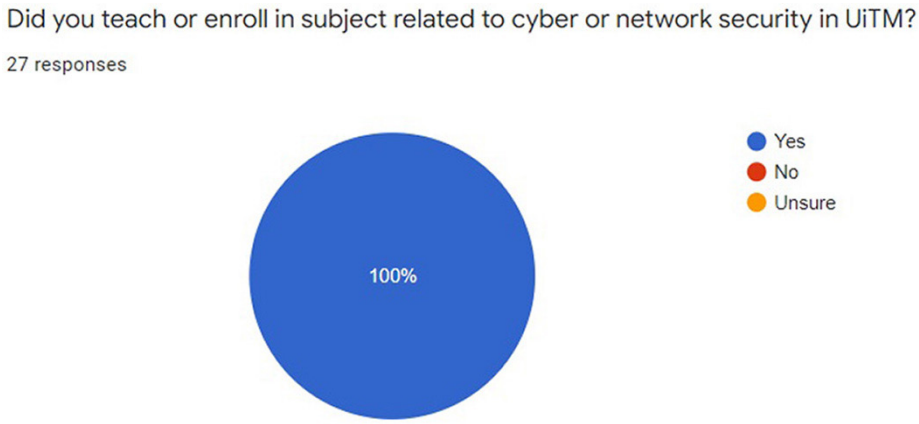


Fig. 20. Question 3 survey result

Meanwhile, the last three questions are asking about the respondent's experiences and opinions using the cyber security lab. Figure 21 shows the result of question 4 which ask whether the SOC lab helped with their study, Figure 22 shows the result of question 5 which ask the respondent whether they agree the SOC lab provide exposure to real-life threat data and Figure 23 shows the result of question 6 which ask whether respondent agree that the SOC lab was able to visualize the threat data.

Based on Figure 21, shows that most of the respondents (77.8%) agreed that the lab setup helps them when studying cyber security subjects, and only 22.2% of respondents are unsure. Figure 22 shows that the majority (81.5%) of the respondents agree that the lab able to provide exposure to threat data in a real environment. Meanwhile, Figure 23 shows that the majority (81.5%) of the respondent agree that the lab can be visualized cyber threat data. Based on the result of questions 4,5 and 6, most of the respondents

feel that the deployed cyber security lab can improve their teaching and learning experience on cyber security subjects. It shows that the proposed cyber security lab model is effective in enhancing the experience in cyber security education.

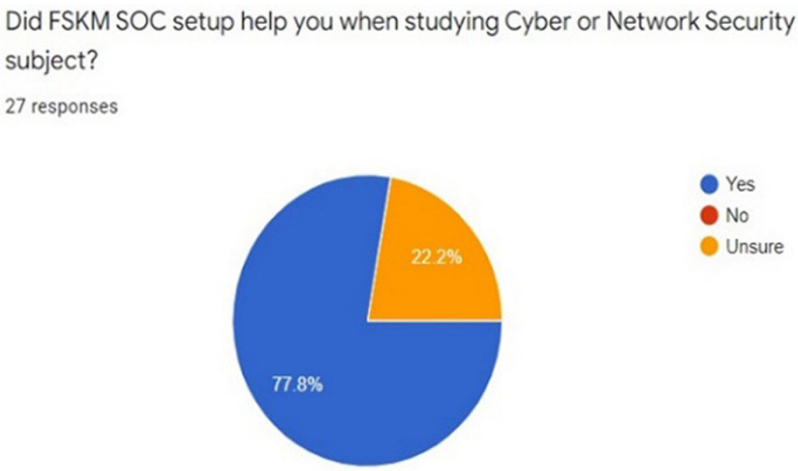


Fig. 21. Question 4 survey result

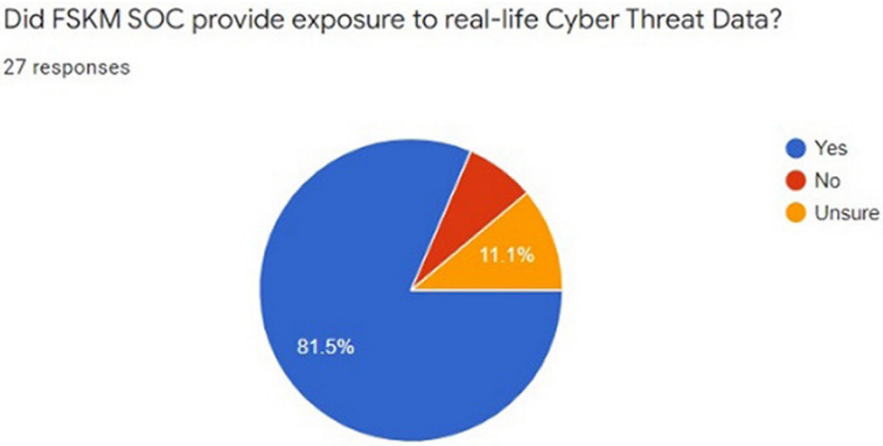


Fig. 22. Question 5 survey result

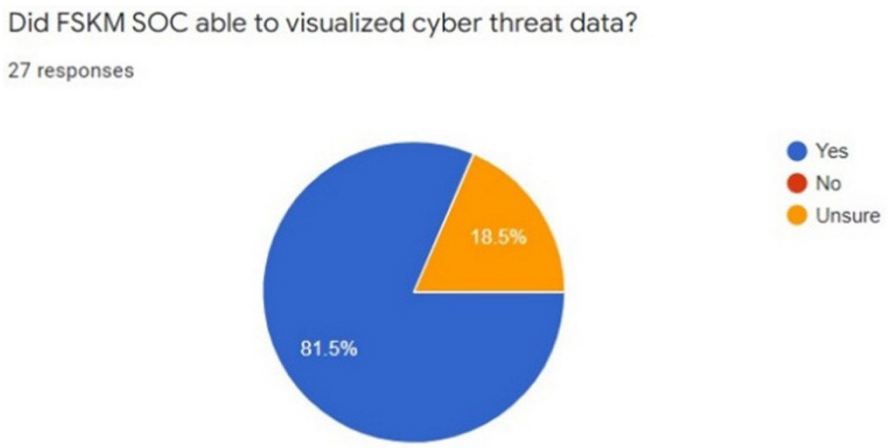


Fig. 23. Question 6 survey result

5 Conclusion

In conclusion, to address the gap in cyber security education which has been identified, this research project proposed a model of cyber security lab which incorporates the SIEM tools and honeypot to enhance the teaching and learning experience in cyber security. The lab was deployed at Universiti Teknologi MARA (UiTM) and used for teaching and learning activities on information security subjects. To evaluate the effectiveness of the proposed lab model, a functional test and a survey was conducted. The functional test results show that the honeypot can attract hacking attempts and can gather useful threat data. Meanwhile, the functional test for the SIEM system shows that it was able to effectively process and correlate the threat data coming from outbound and inbound attacks. Moreover, the survey result shows that majority of the respondent agreed that the cyber security lab improve their teaching and learning experience while taking the cyber security subject. For future work, the research project can be expanded by including threat data sources from a variety of security appliances such as web application firewalls and email gateways and different types of servers.

6 Acknowledgment

The authors would like to extend appreciation to Universiti Teknologi MARA for supporting this project by providing a grant (100-TNCPI/INT 16/6/2 (063/2021)) and the facilities needed to complete this research.

7 References

[1] IBM Security, "Cost of a Data Breach Report 2021," New York, Jun. 2021. Accessed: Mar. 08, 2022. [Online]. Available: https://www.ibm.com/downloads/cas/OJDVQGRY; https://doi.org/10.1016/S1361-3723(21)00082-8

[2] National Security Council, "Malaysia Cyber Security Strategy 2020–2024," Putrajaya, 2020. Accessed: Mar. 08, 2022. [Online]. Available: https://asset.mkn.gov.my/wp-content/uploads/2020/10/MalaysiaCyberSecurityStrategy2020-2024.pdf

[3] L. Li, K. Zhang, and T. Li, "A Performance Analysis Model for the Training and Education of Information Security Talents," *Int. J. Emerg. Technol. Learn.*, vol. 15, no. 5, pp. 140–155, Mar. 2020, https://doi.org/10.3991/ijet.v15i05.13329

[4] (ISC)², "Cybersecurity Workforce Study 2020," 2020. Accessed: Mar. 09, 2022. [Online]. Available: https://www.isc2.org/-/media/ISC2/Research/2020/Workforce-Study/ISC2ResearchDrivenWhitepaperFINAL.ashx?la=en&hash=2879EE167ACBA7100C330429C7EBC623BAF4E07B

[5] F. E. Catota, M. Granger Morgan, and D. C. Sicker, "Cybersecurity Education in a Developing Nation: The Ecuadorian Environment," *J. Cybersecurity*, vol. 5, no. 1, Jan. 2019, https://doi.org/10.1093/cybsec/tyz001

[6] B. J. Blažič, "Changing the Landscape of Cybersecurity Education in the EU: Will the New Approach Produce the Required Cybersecurity Skills?" *Educ. Inf. Technol.*, pp. 1–26, Sep. 2021, https://doi.org/10.1007/s10639-021-10704-y

[7] Z. Hu and X. Gong, "The Practice of a New Maker Teaching Model in Vocational and Technical Education," *Int. J. Emerg. Technol. Learn.*, vol. 17, no. 9, pp. 241–256, May 2022, https://doi.org/10.3991/ijet.v17i09.30935

[8] L. Chongrui *et al.*, "Use of Augmented Reality-Enabled Prototyping of Cyber-Physical Systems for Improving Cyber-Security Education," *J. Phys. Conf. Ser.*, vol. 1840, no. 1, p. 012026, Mar. 2021, https://doi.org/10.1088/1742-6596/1840/1/012026

[9] M. N. Katsantonis and I. Mavridis, "Evaluation of HackLearn COFELET Game User Experience for Cybersecurity Education," *Int. J. Serious Games*, vol. 8, no. 3, pp. 3–24, Sep. 2021, https://doi.org/10.17083/ijsg.v8i3.437

[10] A. Jaffray, C. Finn, and J. R. C. Nurse, "SherLOCKED: A Detective-Themed Serious Game for Cyber Security Education," *IFIP Adv. Inf. Commun. Technol.*, vol. 613, pp. 35–45, Jul. 2021, https://doi.org/10.1007/978-3-030-81111-2_4

[11] K. B. Vekaria, P. Calyam, S. Wang, R. Payyavula, M. Rockey, and N. Ahmed, "Cyber Range for Research-Inspired Learning of 'Attack Defense by Pretense' Principle and Practice," *IEEE Trans. Learn. Technol.*, vol. 14, no. 3, pp. 322–337, Jun. 2021, https://doi.org/10.1109/TLT.2021.3091904

[12] M. Leitner *et al.*, "Enabling Exercises, Education and Research with a Comprehensive Cyber Range," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)*, vol. 12, no. 4, pp. 37–61, Dec. 2021, https://doi.org/10.22667/JOWUA.2021.12.31.037

[13] M. Vielberth, M. Glas, M. Dietz, S. Karagiannis, E. Magkos, and G. Pernul, "A Digital Twin-Based Cyber Range for SOC Analysts," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12840 LNCS, pp. 293–311, Jul. 2021, https://doi.org/10.1007/978-3-030-81242-3_17

[14] N. A. A. Rahman, I. H. Sairi, N. A. M. Zizi, and F. Khalid, "The Importance of Cybersecurity Education in School," *Int. J. Inf. Educ. Technol.*, vol. 10, no. 5, pp. 378–382, May 2020, Accessed: Mar. 10, 2022. [Online]. Available: https://doi.org/10.18178/ijiet.2020.10.5.1393

[15] H. Qusa and J. Tarazi, "Cyber-Hero: A Gamification Framework for Cyber Security Awareness for High Schools Students," *2021 IEEE 11th Annu. Comput. Commun. Work. Conf. CCWC 2021*, pp. 677–682, Jan. 2021, https://doi.org/10.1109/CCWC51732.2021.9375847

[16] B. Fatokun Faith, Z. Awang Long, S. Hamid, O. Fatokun Johnson, C. Ifeanyi Eke, and A. Norman, "An Intelligent Gamification Tool to Boost Young Kids Cybersecurity Knowledge on FB Messenger," *2022 16th Int. Conf. Ubiquitous Inf. Manag. Commun.*, pp. 1–8, Jan. 2022, https://doi.org/10.1109/IMCOM53663.2022.9721733

[17] F. Abu-Amara, R. Almansoori, S. Alharbi, M. Alharbi, and A. Alshehhi, "A Novel SETA-Based Gamification Framework to Raise Cybersecurity Awareness," *Int. J. Inf. Technol.* vol. 13, no. 6, pp. 2371–2380, Aug. 2021, https://doi.org/10.1007/s41870-021-00760-5

[18] S. Egelman, J. Bernd, G. Friedland, and D. Garcia, "The Teaching Privacy Curriculum," *SIGCSE 2016 – Proc. 47th ACM Tech. Symp. Comput. Sci. Educ.*, pp. 591–596, Feb. 2016, https://doi.org/10.1145/2839509.2844619

[19] "Ada The Cyber Security Education Robot." https://www.unhcfreg.com/single-post/2015/10/13/Ada-The-Cyber-Security-Education-Robot (accessed Apr. 05, 2022).

[20] R. Nakata and A. Otsuka, "CyExec*: Automatic Generation of Randomized Cyber Range Scenarios," *Proceedings of the 7th International Conference on Information Systems Security and Privacy – ICISSP*, pp. 226–236, 2021, https://doi.org/10.5220/0010324502260236

[21] S. Ahmad, N. Maunero, and P. Prinetto, "EVA: A Hybrid Cyber Range*," Accessed: Apr. 05, 2022. [Online]. Available: https://www.cbsnews.com/news/cyber-war-sabotaging-the-system-06-11-2009/

[22] N. Nordin, N. Rasyidah, M. Nordin, and W. Omar, "The Efficacy of REV-OPOLY Augmented Reality Board Game in Higher Education," *Int. J. Emerg. Technol. Learn.*, vol. 17, no. 7, pp. 22–37, Apr. 2022, https://doi.org/10.3991/ijet.v17i07.26317

[23] M. D. Workman, J. Anthony Luevanos, and B. Mai, "A Study of Cybersecurity Education Using a Present-Test-Practice-Assess Model," *IEEE Trans. Educ.*, vol. 65, no. 1, pp. 40–45, Feb. 2022, https://doi.org/10.1109/TE.2021.3086025

[24] S. Sentanoe, B. Taubmann, and H. P. Reiser, "Virtual Machine Introspection Based SSH Honeypot," *ACM Int. Conf. Proceeding Ser.*, vol. Part F131370, pp. 13–18, Jun. 2017, https://doi.org/10.1145/3099012.3099016

[25] A. Kyriakou and N. Sklavos, "Container-Based Honeypot Deployment for the Analysis of Malicious Activity," *2018 Glob. Inf. Infrastruct. Netw. Symp. GIIS 2018*, Feb. 2019, https://doi.org/10.1109/GIIS.2018.8635778

[26] A. M. Nasution, M. Zarlis, and S. Suherman, "Analysis and Implementation of Honeyd as a Low-Interaction Honeypot in Enhancing Security Systems," *Randwick Int. Soc. Sci. J.*, vol. 2, no. 1, pp. 124–135, Jan. 2021, https://doi.org/10.47175/rissj.v2i1.209

[27] J. Ma, K. Chai, Y. Xiao, T. Lan, and W. Huang, "High-Interaction Honeypot System for SQL Injection Analysis," *Proc. – 2011 Int. Conf. Inf. Technol. Comput. Eng. Manag. Sci. ICM 2011*, vol. 3, pp. 274–277, 2011, https://doi.org/10.1109/ICM.2011.287

[28] E. D. Saputro, Y. Purwanto, and M. F. Ruriawan, "Medium Interaction Honeypot Infrastructure on the Internet of Things," *IoTaIS 2020 – Proc. 2020 IEEE Int. Conf. Internet Things Intell. Syst.*, pp. 98–102, Jan. 2021, https://doi.org/10.1109/IoTaIS50849.2021.9359711

[29] Balaji N, "How to Build and Run a Security Operations Center," *GB Hackers on Security*, Jan. 2022. https://gbhackers.com/how-to-build-and-run-a-security-operations-center/ (accessed Mar. 19, 2022).

[30] O. Azeroual and A. Nikiforova, "Apache Spark and MLlib-Based Intrusion Detection System or How the Big Data Technologies Can Secure the Data," *Information*, vol. 13, no. 2, p. 58, Jan. 2022, https://doi.org/10.3390/info13020058

[31] J. Lee, J. Kim, I. Kim, and K. Han, "Cyber Threat Detection Based on Artificial Neural Networks Using Event Profiles," *IEEE Access*, vol. 7, pp. 165607–165626, 2019, https://doi.org/10.1109/ACCESS.2019.2953095

[32] "MSSGard & SectorGard – Tecforte." http://tecforte.com/sectorgard/ (accessed Apr. 03, 2022).

[33] "RFC 5424 – The Syslog Protocol." https://datatracker.ietf.org/doc/rfc5424/ (accessed Apr. 03, 2022).

[34] Michel Oosterhof, "GitHub – Cowrie/Cowrie: Cowrie SSH/Telnet Honeypot https://cowrie.readthedocs.io," *Github*, 2022. https://github.com/cowrie/cowrie (accessed Mar. 19, 2022).

[35] B. Gianluca, "GitHub – gbrindisi/wordpot: A Wordpress Honeypot," *Github*, 2018. https://github.com/gbrindisi/wordpot (accessed Mar. 19, 2022).

8 Authors

Muhammad Azizi Mohd Ariffin is currently a lecturer at Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam Malaysia. His research interest is in Cybersecurity, Cloud Computing, and Internet of Things (IOT) (email: mazizi@fskm.uitm.edu.my).

Mohamed Yusof Darus is a head of department for Computer Technology and Networking department at Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam Malaysia. His research interest is in Wireless sensor network, Information Security and Cloud Computing (email: yusof@fskm.uitm.edu.my).

Haryani Haron is a dean for Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam Malaysia. His research interest is Knowledge Management, Industrial Informatics, Information Technology, Information System Methodology, and Citizen Science (email: haryani@fskm.uitm.edu.my).

Aditya Kurniawan is a senior lecturer at School of Computer Science, Bina Nusantara University, Jl. Raya Kb. Jeruk No.27, Kota Jakarta Barat, Daerah Khusus Ibukota Jakarta, 11530, Indonesia. His research interest is Cybersecurity and Internet of Things (IOT) (email: adkurniawan@binus.edu).

Yohan Muliono is a senior lecturer at School of Computer Science, Bina Nusantara University, Jl. Raya Kb. Jeruk No.27, Kota Jakarta Barat, Daerah Khusus Ibukota Jakarta, 11530, Indonesia. His research interest is Cybersecurity, privacy, and information security (email: ymuliono@binus.edu).

Chrisando Ryan Pardomuan is a senior lecturer at School of Computer Science, Bina Nusantara University, Jl. Raya Kb. Jeruk No.27, Kota Jakarta Barat, Daerah Khusus Ibukota Jakarta, 11530, Indonesia. His research interest is Cybersecurity, privacy, and information security (email: chrisando.pardomuan@binus.edu).

Article submitted 2022-05-31. Resubmitted 2022-07-13. Final acceptance 2022-07-13. Final version published as submitted by the authors.