Journal of Cases on Information Technology
Volume 26 • Issue 1

Predicting and Visualizing Lateral Movements Based on ATT&CK and Quantification Theory Type 3

Satoshi Okada, Toyo University, Japan*
https://orcid.org/0000-0002-6601-5710
Yosuke Katano, Toyo University, Japan
Yukihiro Kozai, Toyo University, Japan
Takuho Mitsunaga, Toyo University, Japan
https://orcid.org/0009-0003-0089-7997

ABSTRACT

When a cyber incident occurs, organizations need to identify the attack's impacts. They have to investigate potentially infected devices as well as certainly infected devices. However, as an organization's network expands, it is difficult to investigate all devices. In addition, the cybersecurity workforce shortage has risen, so organizations need to respond to incidents efficiently with limited human resources. To solve this problem, this paper proposes a tool to assist an incident response team. It can visualize ATT&CK techniques attacker used and, furthermore, detect lateral movements efficiently. The tool consists of two parts: a web application that extracts ATT&CK techniques from logs and a lateral movement detection system. The web application was implemented and could map the collected logs obtained from an actual Windows device to the ATT&CK matrix. Furthermore, actual lateral movements were performed in an experimental environment that imitated an organizational network, and the proposed detection system could detect them.

KEYWORDS

Automated System, Cyber Attack, Incident Response, MITRE ATT&CK, Quantification Theory Type 3, Security

INTRODUCTION

The[1*] number of cyberattacks continues to increase in the United States. The Internet Crime Complaint Center (IC3), managed by the FBI, serves as a central hub for reporting cybercrime and welcomes reports from anyone who believes they have been a victim of internet crime, including individuals, businesses, and other organizations. It reported that 847,376 complaints about cyberattacks were reported by members of the American public in 2021, regardless of their organizational affiliation and the type of cyberattack. The number of complaints has increased approximately 2.8 times

DOI: 10.4018/JCIT.340722
*Corresponding Author

compared to five years ago (Internet Crime Complaint Center, 2022). In addition, cyberattack techniques are also becoming more sophisticated. Therefore, it has become challenging to prevent all cyberattacks completely. Given this trend, it is very important to detect cyberattacks quickly and take countermeasures to minimize the damage. (Prompt detection and countermeasures are termed "incident response").

When it is clear that cyber-incidents have happened in organizations, the incident response team has to conduct an initial analysis to confirm the extent of the incident. This includes determining which networks, systems, or applications are affected, what is the source of the incident, and how the incident is being carried out (e.g., the attack techniques and tools being used and the vulnerabilities being exploited; Scarfone et al., 2008). However, it is inefficient and even impossible for the team to analyze all devices, systems, and services in the organization, because organizations' internal networks are getting larger and more complex. Furthermore, it is also pointed out that the cybersecurity industry now faces a critical shortage of skilled workers. This means that incident response teams are forced to conduct efficient incident responses with limited human resources.

Our Contribution

In order to solve the above problems, this paper proposes an automation tool to help organizations' incident response teams conduct more efficient incident responses. The proposed tool consists of two parts. The first is a web application to extract ATT&CK techniques from Sysmon log data. It can also visualize the ATT&CK techniques the attacker used by mapping the techniques to the ATT&CK matrix. The second part is an automatic lateral movement detection system based on the similarity scores between the initially compromised devices and other devices. The scores are calculated by using the techniques extracted by quantification theory type 3.

We implemented a web application to realize our proposed method. We also prepared an experimental environment simulating an organizational network, simulated actual attacks, and confirmed that mapping Sysmon logs obtained from Windows terminals to ATT&CK enabled us to visualize attackers' movements. In addition, we confirmed the usefulness of a method to find undetected infected terminals by quantifying the similarity of these ATT&CK techniques. In the following discussion, the main contributions of this paper are summarized:

- Proposal of a method to automatically extract ATT&CK Techniques from collected Sysmon logs.
- Proposal of a method to efficiently find which devices are infected, by lateral movement based on similarity to initially infected devices using quantification theory type 3.
- Development of a web application to realize the proposed methods and confirm their effectiveness.

Roadmap

This paper is organized as follows: Section 2 describes the techniques related to our research and lists previous works to clarify our contributions compared to them. Then, the authors' proposals are explained in Section 3. The results are described and discussed in Section 4. Finally, conclusions are shown in Section 5.

LITERATURE REVIEW

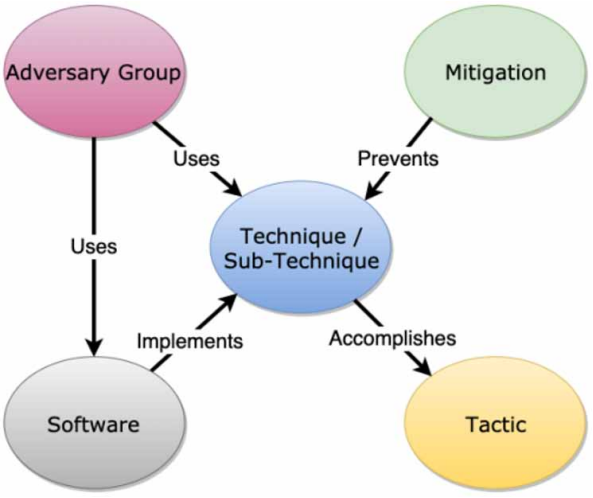
In our proposed method, to detect attackers' lateral movement, we utilize MITRE ATT&CK, a multivariable analysis method, Atomic Red Team, and Windows Sysmon logs. This section explains each of these techniques, lists previous research related to them, and reveals our research position.

MITRE ATT&CK

ATT&CK is a knowledge base on cyberattack tactics and techniques based on actual attacks that have occurred in the past, created by MITRE, a nonprofit organization in the US (the MITRE Corporation, 2023). ATT&CK consists of five components: tactics, techniques, adversary groups, software, and mitigations, which respectively address the tactical goals of attackers, the technical methods used to achieve the tactics, the attackers who conduct cyberattacks, the tools used in attacks, and the mitigation measures used to prevent the attacks. These elements are illustrated in Figure 1. ATT&CK is divided into enterprise, mobile, and industrial control systems. As of August 2022, 14 tactics, 191 techniques, and 133 adversary groups have been published for enterprise ATT&CK.

There are various types of cyberattack prediction or attacker behavior visualization methods utilizing ATT&CK. Al-Shaer et al. (2020) proposed a system to predict late-stage attacks using hierarchical clustering of the APT and software attacks reported in MITRE ATT&CK. There were 98 attacks associated with some techniques, of which 78% showed significant mutual information content. The system proved to be highly predictive. Elitzur et al. (2019) created the "Attack Hypothesis Generator" system from the knowledge graph using ATT&CK. It could provide attack hypotheses from their five recommendation algorithms and preliminary analysis by security analysts. Kuppa et al. (2021) introduced a multi-label text classification (MLTC) task that maps CVE texts to ATT&CK techniques using natural language processing (NLP) techniques and the MLTC model, and a multi-head coupled embedded neural network architecture was proposed. This labeling technique allowed us to map 17 techniques. Andrew et al. (2021) evaluated natural language NLP techniques that map Linux bash commands to MITRE ATT&CK tactics and techniques. They collected Linux bash commands and their associated descriptions and MITRE ATT&CK descriptions, performed some preprocessing to remove stop words and symbols, and projected the resulting sentences into vectors using various methods: Bag of Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF), and pre-trained NLP techniques such as word embeddings to map Linux commands to ATT&CK techniques, sub-techniques, and ATT&CK tactics. Sadlek et al. (2022) proposed the kill-chain attack graph, combining Cyber kill-chain, ATT&CK, and STRIDE as a threat analysis model. Their approach can predict the flow of attacker behavior. It enabled administrators to check the attackers' kill-chain phase and take more appropriate countermeasures to mitigate possible cyber threats. Cho et al. (2018) proposed a new cyber kill-chain model and developed Cyber Common Operational Picture (CyCOP). It can visualize the current situation in cyberspace by utilizing MITRE CAPEC and

Figure 1. ATT&CK Model Relationships (The MITRE Corporation, 2020)



MITRE ATT&CK. Cyber threats are classified into ATT&CK tactics and techniques at each phase in the cyber kill-chain. Their proposed method can help people to predict cyberattacks. Kuwano et al. (2022, 2023) proposed a method to predict additionally compromised devices by lateral movement from an initially infected one using quantification theory type 3 and the ATT&CK technique. This method maps the logs of each device to ATT&CK techniques and predicts the infected device by calculating the similarity score of different devices' logs.

Quantification Theory Type 3

Multivariate analysis is a general term for statistical techniques for analyzing and summarizing multiple data sets to discover hidden associations. Multivariate analysis includes cluster analysis, principal component analysis, etc., which are commonly used in marketing and medicine (Katz, 2011). Quantification theory type 3, which is used in this study, is also a type of multivariate analysis.

These methods can be categorized on the basis of whether they involve an objective variable. In scenarios without an objective variable, explanatory variables are further divided into quantitative or categorical data. Quantification Theory Type 3 is a method that operates without an objective variable and utilizes categorical data as the explanatory variable. This technique is particularly effective in discerning similarities among samples. In this research, we quantify each device's behavior using the ATT&CK technique and analyze it with Quantification Theory Type 3 to identify similarities between devices.

Because of the large number of samples and categories, we use a matrix to perform Quantification Theory Type 3. The algorithm is described using data with four samples and two categories as examples. Let a be the matrix of sample data, b be the matrix of categorical data, D be the data matrix, A be the matrix with the number of each sample diagonal, and B be the matrix with the number of each category diagonal.

$$a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix},\ b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},\ D = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},\ A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},\ B = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$$

The goal is to find a_j and b_j that yield the largest correlation coefficient between the sample data and categorical data. Let V_{aa} be the variance of the sample data, V_{bb} the variance of the categorical data, and V_{ab} the covariance of the sample and categorical data and express them as a matrix.

$$V_{aa} = a^T \cdot A \cdot a$$
$$V_{bb} = b^T \cdot B \cdot b$$
$$V_{ab} = a^T \cdot D \cdot b$$
$$a^T \cdot A \cdot a = b^T \cdot B \cdot b = 1$$

Correlation coefficient R is calculated as follows:

$$R = \frac{V_{ab}}{\sqrt{V_{aa}}\sqrt{V_{bb}}} = \frac{a^T \cdot D \cdot b}{\left(a^T \cdot AA \cdot a\right)^{1/2}\left(b^T \cdot B \cdot b\right)^{1/2}}$$

By using the Lagrange undetermined multiplier, we can solve the equation and put it in the form of eigenequations.

$$G = a^T \cdot D \cdot b - \frac{k}{2}\left(a^T \cdot A \cdot a - 1\right) - \frac{m}{2}\left(b^T \cdot B \cdot b - 1\right)$$

$$B^{-1/2} \cdot D^T \cdot A^{-1} \cdot D \cdot B^{-1/2} \cdot x = k^2 \cdot x$$

$$B^{1/2} \cdot b = x$$

$$a_1 = 0.478, a_2 = -0.120, a_3 = 0.478, a_4 = -0.718, b_1 = 0.365, b_2 = -0.548$$

The value of a is the sample score, the value of b is the category score, and the closeness of these values is the similarity.

Some works proposed a method to identify C&C servers, which are the root of botnets, by using multivariable analysis (Okayasu & Sasaki, 2015; Mihara & Sasaki, 2010). They used a deny-list of C&C servers and analysis of CCCDATAset, a set of bot observation data that contains malware samples. They proposed applying the results to Quantification Theory 2, a type of multivariate analysis. In the initial experiment, the detection rate was 86%. In the subsequent experiment, the detection rate was 96.799%, considering the time of day.

Atomic Red Team

Atomic Red Team is an open-source library that can test ATT&CK techniques on various platforms (Red Canary, 2023). It enables adversarial activities simulation in each user's environment. This library is thought to be helpful in various scenarios, such as visibility verification, detection coverage testing, emulation of enemy behavior, and red team exercises. This research uses it as a reference to correlate the ATT&CK technique with the device logs when generating data to perform Quantification Theory Type 3.

Sysmon and Lateral Movement Detection

Sysmon, a Windows system service, monitors Windows system activities and logs them in the Windows Event Log. Once installed, it provides detailed information on process creation, network connections, and modifications to file creation timestamps. By collecting events generated via Windows Event Collection or SIEM agents and analyzing them, it's possible to identify malicious or anomalous activities, including lateral movement.

Matsuda et al. (2019, 2020) proposed a real-time detection system for targeted attacks based on Window Sysmon logs (Microsoft, 2022) to obtain DLL information. Their later research achieved a high detection rate of 90% even if the DLL was changed by about 10% using deep learning analysis. Bohara et al. (2017) proposed lateral movement detection based on the correlation of various indicators of abnormal host behavior and graph-based modeling of the target system's security state. The method focused on each host's communications especially and used Principal Component Analysis, k-means clustering, and Median-Absolute Deviation. The detection accuracy was approximately 88.7%.

From the above literature review, we can see that there are previous studies that use ATT&CK technology to visualize and predict attacker behavior, Sysmon logs to detect attacks, and host communications and behavior to detect lateral deployments. However, there are no existing studies that use Quantification Type 3 to calculate similarities in behavior within hosts and to detect lateral movement of attackers using ATT&CK techniques.

PROPOSED METHOD

This research has two proposals. The first is to develop a method to extract ATT&CK techniques from Sysmon logs automatically. The second is an efficient prediction system of lateral movements combining the quantification triad and the ATT&CK technique. By combining these two techniques, it is possible to automatically extract ATT&CK techniques from the collected terminal logs and use the results to predict lateral movement.

Automatic Mapping of Log Data to ATT&CK Techniques

The flow of the proposed system is as follows: First, a database is created that records the relationship between ATT&CK technique and actual attack commands on the basis of the examples by Atomic Red Team. Next, input log data is compared with the database, and techniques contained in the log are output. In addition, an Excel sheet mapping the extracted techniques to the Enterprise matrix (The MITRE Corporation, 2023) is also output. Visualization of the ATT&CK techniques the attacker used allows people to more easily understand the attack stages and current status. In the following subsection, the details of how the authors designed and implemented the system is described.

ATT&CK classifies attackers' behavior according to 14 tactics. However, attackers do not necessarily go through all tactics. If their goal is achieved halfway through, they may stop the attack procedure. In addition, it is difficult to distinguish techniques in the reconnaissance, initial access, and impact phases from general IT operation. Therefore, this paper focuses on TA0006, TA0007, and TA0008 when creating the database. Next, frequently (20 or more times) used techniques are picked up from among the techniques. As a result, 14 techniques are selected: T1003, T1555, T1056, T1082, T1083, T1057, T1016, T1033, T1018, T1518, T1049, T1087, T1046 and T1021. Specific attack logs and commands corresponding to each of these techniques are extracted from Atomic Red Team, and a database is created.

Automatic Detection System for Lateral Movements

This research also aims to find devices that are secondarily infected through lateral movement. It is challenging to find laterally moved devices by standard log analysis. Therefore, our approach is to discover secondarily infected devices by the similarity between the initially infected device and others. On the basis of each device's logs, the device's behavior is converted into the ATT&CK technique, and they are analyzed using Quantification Theory Type 3. It is thought that device behavior that is similar to that of initially infected devices is likely to be secondarily infected due to lateral movement.

The system proceeds as follows:

1) ***Collect device activity logs.*** Collect activity logs of the devices to be analyzed using Sysmon.
2) ***Correlate the logs with the ATT&CK Techniques.*** The information in the Atomic Red Team is compared to the logs obtained, and the logs are then correlated to the ATT&CK approach. For example, if the authors refer to the Atomic Red Team's Atomic Test T1550.002, they can confirm the presence of the whoami.exe command. Since the activity log of the device showed "Image: C:\Windows\System32\whoami.exe," they think that T1550.002 in the technique may have been performed on this device. It can be concluded that T1550.002 in the technique may have been performed on this device.
3) ***Pre-process data for analysis.*** A truth table, as shown in Table 1, is created. It has devices in the index and techniques in the columns. The techniques in the columns are those found in one or more of the behaviors of all devices. If a technique is found in each device's log, it is set to 1. Otherwise, it is set to 0.

4) ***Analyze using Quantification Theory Type 3.*** As the data input is created, the author's original Python code based on Quantification Theory Type 3 algorithm is used. The similarity of each device can be visualized by plotting the numbers on a graph on the basis of the sample scores. If the sample score of secondary infected devices is close to the initial one, the authors consider being able to predict the secondary infected device by similarity using Quantification Theory Type 3.

EVALUATION

We implemented our proposed method to map log data to ATT&CK techniques as a web application. Next, we confirmed that it could correctly extract and visualize ATT&CK techniques used by attackers. Third, we also implemented an automatic lateral movements detection system and found it could detect them precisely.

Automatically Mapping Log Data to ATT&CK Techniques

To confirm that our proposed method could work correctly, we implemented it using Python Django and evaluated its performance. The application's user interface is similar to Figure 2. Each line of the input log was matched against a database, and the technique contained in the log was extracted and output.

As an evaluation, we input 5 Windows commands (Figure 3), each of which corresponded to T1003.001, T1056.001, T1518, T1087.002, and T1021.006. A part of the output of the application is shown in Figure 4. An Excel file mapping the technique to the ATT&CK matrix was also output (Figure 5). From these results, we could see that the mapping results by our proposed application worked correctly. Thus, by using this system, users can visualize the technique used and the attack phase simply by inputting the logs into the application.

Automatic Detection System for Lateral Movements

An experiment following the scenario was conducted to verify whether our proposed method could detect the secondarily infected devices by lateral movement using Quantification Theory Type 3.

Scenario

Assuming an office environment, 10 devices were prepared in the same network. All devices were connected to Windows server and managed by active directory. Let the name of each device be Device1, 2, 3, ..., 10. Five of them, Devices1, 2, 3, 4, and 5, are assumed for office workers, and Devices 6, 7, 8, 9, and 10 are for developers. During the 30 minutes of operation at each device, the attacker laterally moved the network from Device 1, the initially infected device, to two devices, Device 2 and Device 10, and performed unauthorized operations on them. The attacker conducted network discovery and pass the hash and executed ransomware.

Device 1 was assumed to be the device that had already been recognized as the initially infected device. The similarity between this device and the other devices was determined using the logs of

Table 1. Sample of Pre-Analysis Data

Device	Txxxx	Txxxx	Txxxx	...
Sample1	1	1	1	...
Sample2	1	0	1	...
Sample3	0	1	0	...

Figure 2. The Proposed Application's User Interface

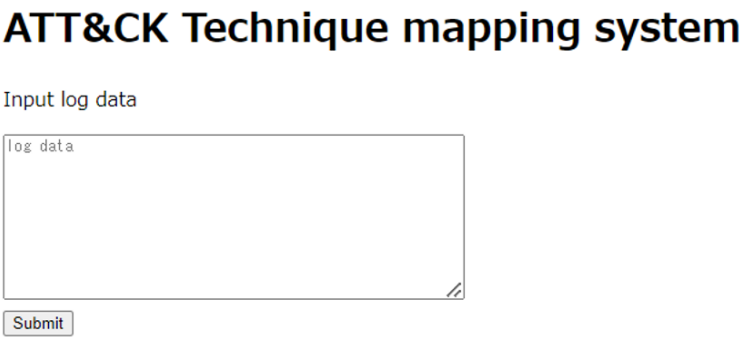


Figure 3. Input Windows Commands

```
#{mimikatz_exe} "sekurlsa::minidump #{input_file}" "sekurlsa::logonpasswords full" exit

trap 'echo "$(date +"%d/%m/%y %H:%M:%S.%s") $USER $BASH_COMMAND" >> #{
    output_file}' DEBUG
echo "Hello World!"
cat #{output_file}

reg query "HKEY_LOCAL_MACHINE\Software\Microsoft\Internet Explorer" /v svcVersion

Invoke-Expression #{adrecon_path}

$SecPassword = ConvertTo-SecureString "#{password}" -AsPlainText -Force
$Cred = New-Object System.Management.Automation.PSCredential("#{username}",
    $SecPassword)
Invoke-Command -ComputerName "#{remotehost}" -Credential $Cred -ScriptBlock {
    whoami}
```

Figure 4. Example of Technique Outputs

```
technique_id: T1021.006

main_technique_name: Remote Services

technique_name: Remote Services: Windows Remote Management

tactic_id: TA0008

technique_id: T1003.001

main_technique_name: OS Credential Dumping

technique_name: OS Credential Dumping: LSASS Memory

tactic_id: TA0006

technique_id: T1056.001

main_technique_name: Input Capture

technique_name: Input Capture: Keylogging

tactic_id: TA0006
```

Figure 5. Example of Mapping to ATT&CK Matrix

Initial Access	Execution	Persistence	Privilege Escalation	Defense Evasion	Credential Access	Discovery	Lateral Movement	Collection	Commar
Drive-by Compromise	Command and Scripting Interpreter	Account Manipulation	Abuse Elevation Control Mechanis	Abuse Elevation Control Mechanis	Adversary-in-the-Middle	Account Discovery	Exploitation of Remote Services	Adversary-in-the-Middle	Application La
Exploit Public-Facing Application	Container Administration Comman	BITS Jobs	Access Token Manipulation	Access Token Manipulation	Brute Force	Application Window Discovery	Internal Spearphishing	Archive Collected Data	Communicatio
External Remote Services	Deploy Container	Boot or Logon Autostart Executio	Boot or Logon Autostart Executio	BITS Jobs	Credentials from Password Stores	Browser Bookmark Discovery	Lateral Tool Transfer	Audio Capture	Data Encoding
Hardware Additions	Exploitation for Client Execution	Boot or Logon Initialization Scripts	Boot or Logon Initialization Scripts	Build Image on Host	Exploitation for Credential Access	Cloud Infrastructure Discovery	Remote Service Session Hijacking	Automated Collection	Data Obfuscat
Phishing	Inter-Process Communication	Browser Extensions	Create or Modify System Process	Debugger Evasion	Forced Authentication	Cloud Service Dashboard	Remote Services	Browser Session Hijacking	Dynamic Reso
Replication Through Removable Me	Native API	Compromise Client Software Binar	Domain Policy Modification	Deobfuscate/Decode Files or Infor	Forge Web Credentials	Cloud Service Discovery	Replication Through Removable Me	Clipboard Data	Encrypted Cha
Supply Chain Compromise	Scheduled Task/Job	Create Account	Escape to Host	Deploy Container	Input Capture	Cloud Storage Object Discovery	Software Deployment Tools	Data from Cloud Storage	Fallback Chan
Trusted Relationship	Serverless Execution	Create or Modify System Process	Event Triggered Execution	Direct Volume Access	Modify Authentication Process	Container and Resource Discovery	Taint Shared Content	Data from Configuration Repositor	Ingress Tool T
Valid Accounts	Shared Modules	Event Triggered Execution	Exploitation for Privilege Escalation	Domain Policy Modification	Multi-Factor Authentication Interc	Debugger Evasion	Use Alternate Authentication Mate	Data from Information Repositories	Multi-Stage C
	Software Deployment Tools	External Remote Services	Hijack Execution Flow	Execution Guardrails	Multi-Factor Authentication Requ	Domain Trust Discovery		Data from Local System	Non-Applicatio
	System Services	Hijack Execution Flow	Process Injection	Exploitation for Defense Evasion	Network Sniffing	File and Directory Discovery		Data from Network Shared Drive	Non-Standard
	User Execution	Implant Internal Image	Scheduled Task/Job	File and Directory Permissions Mo	OS Credential Dumping	Group Policy Discovery		Data from Removable Media	Protocol Tunn
	Windows Management Instrumenta	Modify Authentication Process	Valid Accounts	Hide Artifacts	Steal Application Access Token	Network Service Discovery		Data Staged	Proxy
		Office Application Startup		Hijack Execution Flow	Steal or Forge Authentication Cer	Network Share Discovery		Email Collection	Remote Acces
		Pre-OS Boot		Impair Defenses	Steal or Forge Kerberos Tickets	Network Sniffing		Input Capture	Traffic Signali
		Scheduled Task/Job		Indicator Removal	Steal Web Session Cookie	Password Policy Discovery		Screen Capture	Web Service
		Server Software Component		Indirect Command Execution	Unsecured Credentials	Peripheral Device Discovery		Video Capture	
		Traffic Signaling		Masquerading		Permission Groups Discovery			
		Valid Accounts		Modify Authentication Process		Process Discovery			
				Modify Cloud Compute Infrastructure		Query Registry			
				Modify Registry		Remote System Discovery			
				Modify System Image		Software Discovery			
				Network Boundary Bridging		System Information Discovery			
				Obfuscated Files or Information		System Location Discovery			
				Plist File Modification		System Network Configuration Discovery			
				Pre-OS Boot		System Network Connections Discovery			
				Process Injection		System Owner/User Discovery			
				Reflective Code Loading		System Service Discovery			
				Rogue Domain Controller		System Time Discovery			
				Rootkit		Virtualization/Sandbox Evasion			
				Subvert Trust Controls					
				System Binary Proxy Execution					
				System Script Proxy Execution					
				Template Injection					
				Traffic Signaling					
				Trusted Developer Utilities Proxy Execution					
				Unused/Unsupported Cloud Regions					
				Use Alternate Authentication Material					
				Valid Accounts					

each device obtained from Sysmon. The goal was to detect the remaining two infected devices, Device 2 and Device 10, on the basis of the similarity obtained.

Results and Discussion

1) ***Collect device activity log.*** A total of 2,760 activity logs were output from all Sysmon devices in this experiment. The number of logs output by each device is shown in Table 2.
2) ***Correlate the logs with the ATT&CK technique.*** The output logs were then associated with the techniques. In this experiment, five techniques were associated with T1059.001, T1059.003, T1204.002, T1003, and T1550.002. The rationale for the association with each technique is shown in Table 3.
3) ***Create the data for analysis.*** From the behavior of each device, data was created for analysis in the quantification theory type 3. Table 4 shows this.
4) ***Analyze using Quantification Theory Type 3.*** Quantification Theory Type 3 analyzed the created analysis data, and the sample scores were output, as shown in Table 5. A graphical plot of these sample scores is shown in Figure 6. Figure 6 shows that Device 2 and Device 10, the secondarily infected devices by lateral movement, are close to Device 1, the initially infected device. It can also be seen that other non-infected devices are also close in similarity to each group.

We collected a sufficient number of output logs from each device (Table 2). Thus, we could extract correct ATT&CK techniques from each device's logs (Table 3 and Table 4). From Table 5 and Figure 6, it can be seen that the sample scores of the additionally infected devices (devices 2 and 10) are more similar to the initially infected device (device 1) than the other, uninfected devices. This result quantitatively shows that the proposed method, which combines ATT&CK and Quantification Theory Type 3, can accurately detect lateral movement by setting an appropriate threshold (e.g., the absolute difference in similarity scores, compared to the infected device, is less than 1.00E-08). In addition, the time taken for the proposed method to detect lateral expansion was very short. It is difficult to detect cyberattacks and find lateral movements manually from a large number of logs. The proposed method seems to be very useful in solving this problem.

CONCLUSION

In this paper, we proposed an automatic incident response assist tool. The proposed tool is divided into two main functions. The first is a web application that automatically extracts MITRE ATT&CK techniques from Sysmon logs. The second function automatically detects the attacker's lateral

Table 2. Number of Output Sysmon Logs

Device	Number of Output Logs
Device1	247
Device2	192
Device3	195
Device4	320
Device5	384
Device6	325
Device7	282
Device8	313
Device9	193
Device10	309
Sum	2760

Table 3. List of Techniques Found in the Input Log

Technique ID	Technique Name	Mapping With Atomic Red Team
T1059.001	Command and Scripting Interpreter	Powershell.exe in Create Process.
T1059.003	Command and Scripting Interpreter	Check cmd.exe
T1204.002	User Execution	Check Office software
T1003	OS Credential Dumping	Check mimikatz.exe
T1550.002	Use Alternate Authentication Material	Check whoami.exe

Table 4. The Data Used in the Evaluation

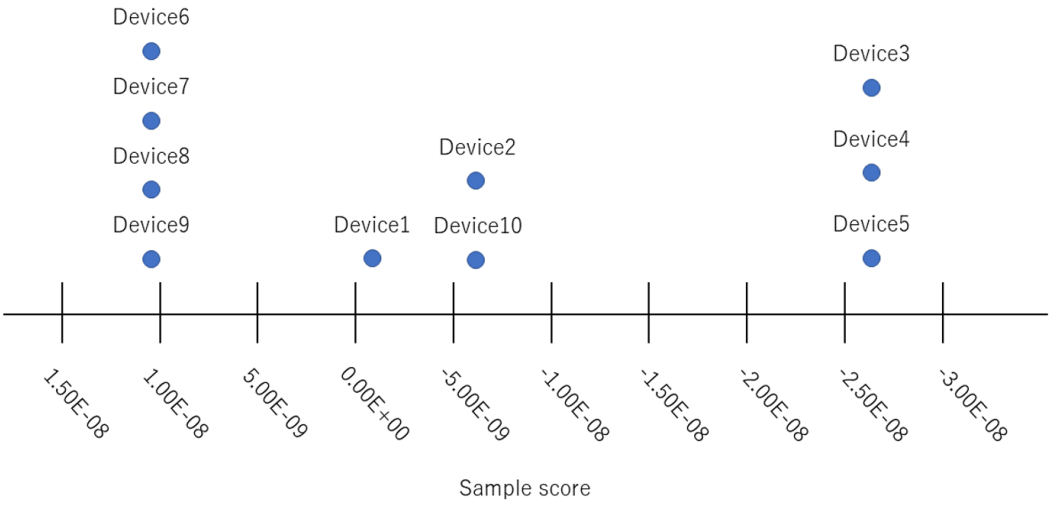
Technique	T1059	T1059	T1204	T1003	T1550
Sub-Technique	.001	.003	.002		.002
Device 1	1	1	1	1	1
Device 2	1	1	1	0	1
Device 3	0	0	1	0	0
Device 4	0	0	1	0	0
Device 5	0	0	1	0	0
Device 6	1	1	1	0	0
Device 7	1	1	1	0	0
Device 8	1	1	1	0	0
Device 9	1	1	1	0	0
Device 10	1	1	1	0	1

movement on the basis of the extracted techniques by calculating scores using Quantification Theory Type 3. We also implemented our proposals. Through the implementation and evaluations, our

Table 5. Sample Score of Each Device

Device	Sample Score
Device 1	-1.67797213e-09
Device 2	-6.79957994e-09
Device 3	-2.73406166e-08
Device 4	-2.73406166e-08
Device 5	-2.73406166e-08
Device 6	1.07476909e-08
Device 7	1.07476909e-08
Device 8	1.07476909e-08
Device 9	1.07476909e-08
Device 10	-6.79957994e-09

Figure 6. Graph Plotting Sample Scores of Each Device



proposed web application extracted ATT&CK techniques from log data and mapped them correctly to the ATT&CK matrix. Furthermore, we prepared a real environment simulating an organization's internal network and conducted lateral movements to check the performance of our detection method. As a result, it became clear that our proposed system accurately predicted the destinations of an attacker's lateral movements by setting appropriate thresholds.

However, there remains some room for improvement. Especially in the lateral movement detection method, ATT&CK techniques were extracted from log data based on the Atomic Red Team. The number of techniques extracted was not very large; so our evaluation and analyses were limited. Thus, we can improve our research in the following ways

1. Collect logs over a longer period
2. Use more functional logging tools than Sysmon
3. Introduce additional libraries of tests mapped to the MITRE ATT&CK framework

COMPETING INTERESTS

The authors of this publication declare there are no competing interests.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. Funding for this research was covered by the authors of the article.

REFERENCES

Al-Shaer, R., Spring, J. M., & Christou, E. (2020). Learning the associations of MITRE ATT & CK adversarial techniques. In *2020 IEEE Conference on Communications and Network Security (CNS)* (pp. 1–9). IEEE. doi:10.1109/CNS48642.2020.9162207

Andrew, Y., Lim, C., & Budiarto, E. (2022). Mapping Linux Shell Commands to MITRE ATT&CK using NLP-Based Approach. In *2022 International Conference on Electrical Engineering and Informatics (ICELTICs)* (pp. 37–42). IEEE. doi:10.1109/ICELTICs56128.2022.9932097

Bohara, A., Noureddine, M. A., Fawaz, A., & Sanders, W. H. (2017). An unsupervised multi-detector approach for identifying malicious lateral movement. In *2017 IEEE 36th Symposium on Reliable Distributed Systems (SRDS)* (pp. 224–233). IEEE. doi:10.1109/SRDS.2017.31

Cho, S., Han, I., Jeong, H., Kim, J., Koo, S., Oh, H., & Park, M. (2018). Cyber kill chain based threat taxonomy and its application on cyber common operational picture. In *2018 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (*Cyber SA*)* (pp. 1-8). IEEE. doi:10.1109/CyberSA.2018.8551383

Elitzur, A., Puzis, R., & Zilberman, P. (2019). Attack hypothesis generation. In *2019 European Intelligence and Security Informatics Conference (EISIC)* (pp. 40–47). IEEE. doi:10.1109/EISIC49498.2019.9108886

Internet Crime Complaint Center (IC3). (2022). *Internet crime report 2021*. https://www.ic3.gov/Media/PDF/AnnualReport/2021_IC3Report.pdf

Katano, Y., Kozai, Y., Okada, S., & Mitsunaga, T. (2022, November). Prediction of infected devices using the Quantification Theory Type 3 based on MITRE ATT&CK technique. In *2022 IEEE International Conference on Computing (ICOCO)*, (pp. 198–203). IEEE. doi:10.1109/ICOCO56118.2022.10031822

Katz, M. H. (2011). *Multivariable analysis: A practical guide for clinicians and public health researchers*. Cambridge University Press. doi:10.1017/CBO9780511974175

Kuppa, A., Aouad, L., & Le-Khac, N. A. (2021). Linking CVE's to MITRE ATT&CK techniques. In *16th International Conference on Availability, Reliability and Security* (pp. 1-12). Association for Computing Machinery. doi:10.1145/3465481.3465758

Kuwano, M., Okuma, M., Okada, S., & Mitsunaga, T. (2022, November). ATT&CK Behavior forecasting based on collaborative filtering and graph databases. In *2022 IEEE International Conference on Computing (ICOCO)* (pp. 191–197). IEEE. doi:10.1109/ICOCO56118.2022.10032036

Kuwano, M., Okuma, M., Okada, S., & Mitsunaga, T. (2023). The attacker might also do next: ATT&CK behavior forecasting by attacker-based collaborative filtering and graph databases. *Journal of Information Processing*, *31*(0), 802–811. doi:10.2197/ipsjjip.31.802

Matsuda, W., Fujimoto, M., & Mitsunaga, T. (2019). Real-time detection system against malicious tools by monitoring DLL on client computers. In *2019 IEEE Conference on Application, Information and Network Security (AINS)* (pp. 36–41). IEEE. doi:10.1109/AINS47559.2019.8968697

Matsuda, W., Fujimoto, M., & Mitsunaga, T. (2020). Detection of malicious tools by monitoring DLL using deep learning. *Journal of Information Processing*, *28*(0), 1052–1064. doi:10.2197/ipsjjip.28.1052

Microsoft. (2022). *Ltd. Sysmon - windows sysinternals — microsoft learn*. Retrieved September 30, 2022, from https://learn.microsoft.com/ja-jp/sysinternals/downloads/sysmon

Mihara, H., & Sasaki, R. (2010). Proposal and evaluation of technique to detect C&C server on botnet using CCC dataset 2009 and quantification methods. *IPSJ Journal*, *51*(9), 1579–1590.

MITRE Corporation. (2020). *MITRE ATT&CK®: Design and Philosophy*. Retrieved September 30, 2022, from https://attack.mitre.org/

MITRE Corporation. (2023). *MITRE ATT&CK®*. Retrieved September 30, 2022, from https://attack.mitre.org/

Okayasu, S., & Sasaki, R. (2015, July). Proposal and evaluation of methods using the quantification theory and machine learning for detecting C&C server used in a botnet. In *2015 IEEE 39th Annual Computer Software and Applications Conference: Vol. 3* (pp. 24-29). IEEE. doi:10.1109/COMPSAC.2015.165

Red Canary. (2023). *Explore atomic red team*. Retrieved September 30, 2022, from https://atomicredteam.io/

Sadlek, L., Čeleda, P., & Tovarňák, D. (2022). Identification of attack paths using kill chain and attack graphs. In *NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium* (pp. 1-6). IEEE. doi:10.1109/NOMS54207.2022.9789803

Scarfone, K., Grance, T., & Masone, K. (2008). *Computer security incident handling guide* (Special Publication 800-61). National Institute of Standards and Technology. https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp.800-61r2.pdf

ENDNOTE

* The preliminary version of this work (Katano et al., 2022) appeared in IEEE International Conference on Computing (ICOCO). This paper extends the contribution by adding the proposal of an automated ATT&CK Technique extraction system (web application).