A Study on Analysis of Malicious Code Behavior Information for Predicting Security Threats in New Environments

Seul-Ki Choi[1], Taejin Lee[2] and Jin Kwak[3*]
[1] ISAA Lab., Department of Computer Engineering, Ajou University, Republic of Korea
[e-mail: skchoi.isaa@gmail.com]
[2] Department of Computer Engineering, Hoseo University, Republic of Korea
[e-mail: kinjecs0@gmail.com]
[3] Department of Cyber Security, Ajou University, Republic of Korea
[e-mail: security@ajou.ac.kr]
*Corresponding author: Jin Kwak

*Received October 2, 2018; revised January 8, 2019; accepted February 28, 2019;
published March 31 2019*

Abstract

The emergence of new technologies and devices brings a new environment in the field of cyber security. It is not easy to predict possible security threats about new environment every time without special criteria. In other words, most malicious codes often reuse malicious code that has occurred in the past, such as bypassing detection from anti-virus or including additional functions. Therefore, we are predicting the security threats that can arise in a new environment based on the history of repeated malicious code. In this paper, we classify and define not only the internal information obtained from malicious code analysis but also the features that occur during infection and attack. We propose a method to predict and manage security threats in new environment by continuously managing and extending.

Keywords: Cyber security, Malicious code, Malware, Security threat

A preliminary version of this paper was presented at APIC-IST 2018, and was selected by the conference review process.

http://doi.org/10.3837/tiis.2019.03.028 **ISSN : 1976-727**

1. Introduction

The development of ICT technology is changing our lives. Especially, due to the advanced technology of IoT (Internet of Things), which is a technology in which all the objects around are connected with each other through networks, various services and various types of ICT environments are emerging [1][2]. ICT technology has come to coexist near us in residential, work and living spaces such as Smart home, smart factory, and smart city environments [3][4]. In addition, ICT technology is being applied to transportation that has a big impact on our lives, such as the next-generation intelligent transportation system called C-ITS (Cooperative-Intelligent Transport System) [5].

However, the development of ICT does not always bring us the benefit. Malicious code is evolving as fast as the development of ICT technologies. According to the report released in McAfee lab in September 2018, the total number of malicious codes is about 80 million in Q2 of 2018[6]. In particular, mobile malware that infects and spreads through mobile devices accounts for about 27 million cases [6].

As such, the development of ICT technology and the evolution of malicious code are closely related. In particular, devices and ICT technologies that are popular around the world are being used as a good infection route and attack target for malicious code. Especially, the rapid improvements of smartphone technology have resulted in the evolution of mobile botnets [7]. The most representative case is the Mirai Botnet incident in October 2016 [8]. In case of ransomware, it is designed to encrypt system user's files and documents, but it can do more than that depending on which family of ransomwares it belongs to [9]. Recently, ransomware which locks the screen of a smartphone instead of file encryption is also emerging.

Through this, we have found that malicious code is scalable to adapt quickly to various environments and it is using success cases in existing environment. At the time when various malicious apps appeared for smartphones, there were some malicious apps with DDoS (Distributed Denial of Service) attack function, but they did not achieve great results. However, IoT devices have emerged to replace smartphones and IoT devices have received worldwide attention. Like the Mirai Botnet described above, DDoS attacks using IoT devices reappeared and achieved great results. We analyze the evolution of these malicious codes and try to predict how malicious codes will appear in the new environment.

Therefore, in this paper, we propose a method to predict security threats that can occur in new environment based on malicious behavior information that can be acquired through malicious code analysis.

The contents of this paper are as follows. Chapter 2 presents threat statistical related to malicious code. Chapter 3 presents structure of malicious code behavior information and its sections and subsections. Chapter 4 presents predicting security threats in new environments method. The last chapter presents the conclusion and describes future research.

2. Related Work

2.1 Threats Statistics

As ICT technology evolves, malicious code is also becoming more intelligent and automated, posing a significant threat to users of ICT technologies. Recently, malicious codes are getting out of the level of taking information of users. It also implemented a large number of DDoS attacks using a large number of zombie devices by inserting automation functions into malicious code. In addition, a variety of new malicious codes are being generated, including Ransomware, which encrypts important data for individuals and corporate storage and devices and requires money. As a result, malicious codes are changed into various forms and malicious codes of variants are being generated.

In September of 2018, McAfee Labs provides McAfee Labs Threats Reports to provide insights into recent security threats [6]. **Fig. 1** shows the statistics of total malware that occurred until 2018 Q2.

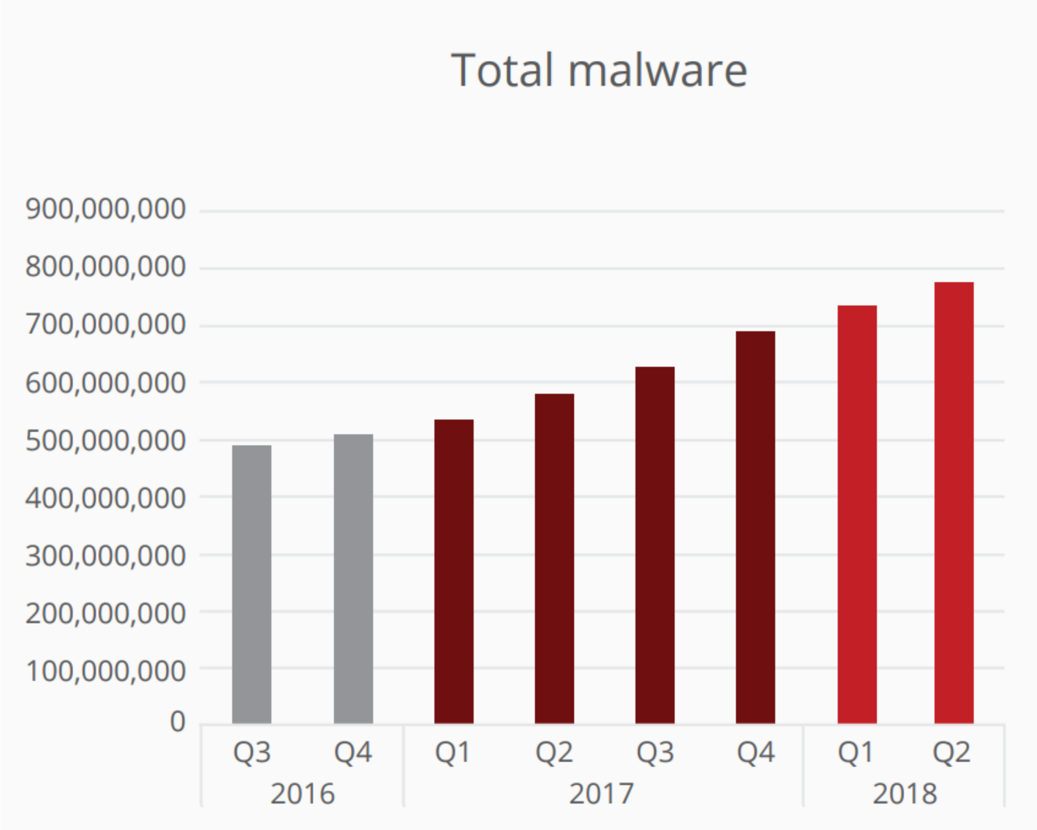


Fig. 1. Overall development of total malware

The total malware shown in the above statistics includes new malware and variant malware. The statistics of new malware are shown in **Fig. 2** below.

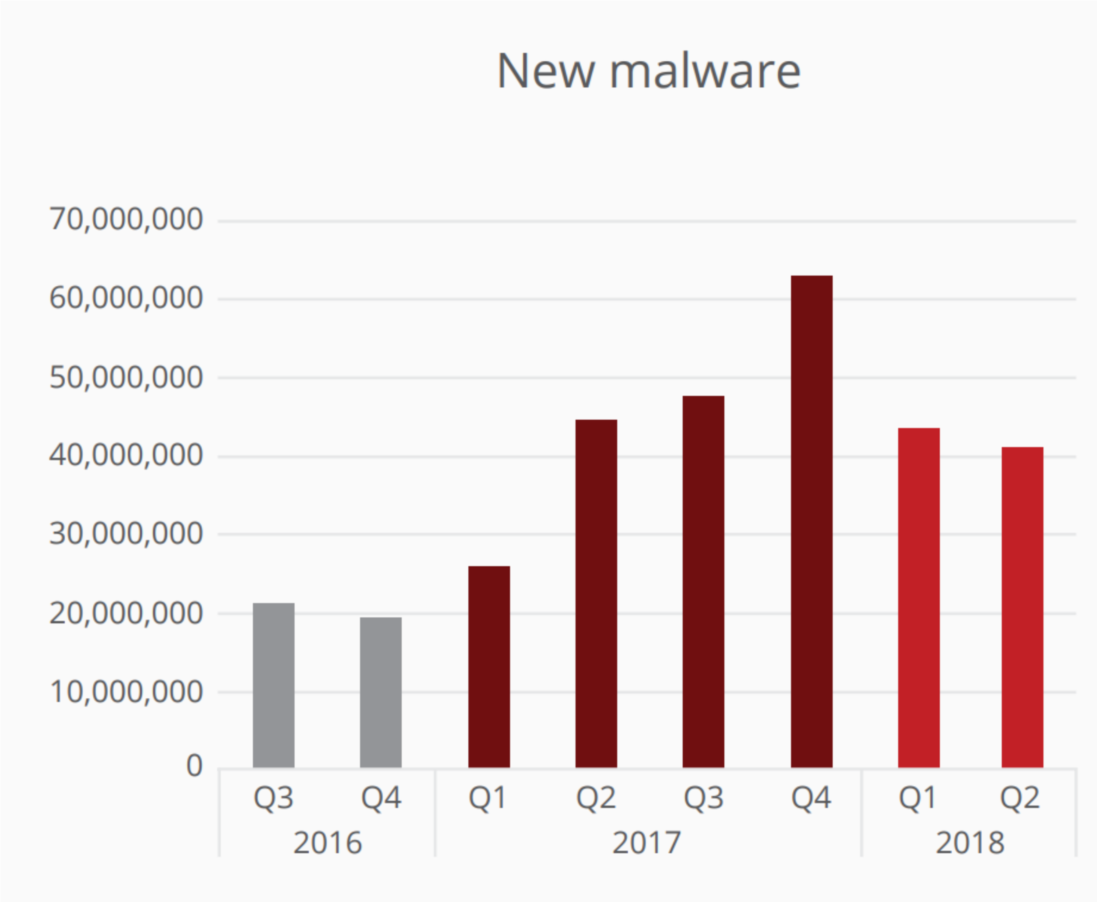


Fig. 2. Overall development of new malware

Fig. 1 and **Fig. 2** show that the ratio of new malware among total malware does not exceed 10%. As a result, we can see that the ratio of variant malware that reuse known malicious codes is higher than new malware.

In addition, the report also shows that mobile malware targeting mobile devices is steadily increasing. Also, Android lockscreen malware with Ransomware has also been growing rapidly since 2017[6].

Therefore, we analyze the fact that most malicious codes are reusing already generated malicious codes, and that there are many malicious codes in the environment (mobile, IoT, etc.) that can be called the latest trend.

3. Malicious Code Behavior Information

3.1 Definition

In this paper, Malicious Code Behavior Information refers to information such as infection route, attack target and attack behavior that can be obtained by analyzing malicious code. That is, it refers to all information related to malicious activity as well as internal information contained in malicious code. It is used to predict security threats that can occur in new environment by classifying them in detail. In this paper, MCBI refers to Malicious Code Behavior Information.

3.2 Structure of MCBI

The MCBI can be divided into an Infection & Propagation section and an attack section. First, the Infection and propagation section consists of the information that appears when malicious code infects an attacking target. The Attack section is also made up of information related to the attack target, executor and attack behavior of malicious code. **Table 1** below shows the subsection of each main section of MCBI.

Table 1. Subsection of MCBI

Section	Subsection	Description
Infection & Propagation	Infection & propagation route	- Media or service used as a route of infection and propagation of malicious code
	Infection & propagation type	- Features classified in malware infection and propagation process (e.g., the method of transmitting malicious code itself directly to an attack target)
	User dependency	- User dependency in malicious code infections and propagation.
Attack	Attack target	- The target for malicious code to perform attack behavior
	Executor	- An entity that execute malicious code that reaches the target system
	Attack behavior	- Actions caused by malicious code to attack targets

3.3 Infection & Propagation Section

All malicious codes use various paths to infect attack targets. There are also various types of malicious codes, such as downloading itself directly when the malicious code reaches an attack target or exploiting vulnerabilities of the system. User involvement may or may not be needed in the process of infecting and propagating malicious code. Therefore, this section classifies and defines the features that occur when malicious code propagates and infects. The section consists of infection & propagation route, infection type and user dependency subsection.

3.3.1 Infection & Propagation Route

Infection & propagation route classifies devices or services used as route of malicious code infection. This part can be added to new devices or services in consideration of the rapidly changing and evolving ICT environment, and can be deleted or included in other parts.

Table 2. Contents of infection & propagation route

Subsection	Contents	Description
Infection & Propagation route	Removable data storage	- A device for storing and moving data, such as USB, CD, HDD, flash memory etc. - It is a malicious code infecting and propagating means that occurs even in an environment where security is maintained by disconnecting from the external network.
	Web/Cloud storage	- Storage that exists in a networked space such as web and cloud environment. - It is used to use web service or to store data.
	Local network	- path that infects or propagates malicious code using internal network protocols such as ARP spoofing.
	P2P	- Data sharing technique used to share data between users without going through the server.
	Mobile device	- Mobile devices are data terminal with mobility and computing capabilities, such as Smartphones, IoT devices and wearable devices. - These mobile devices have been recently used in malicious code infections and propagation routes because they have appropriate computing ability and various functions to infect or propagate malicious codes.
	Message service	- Specific message service available for each device such as SMS / MMS of smart phone. - It is a malicious code infecting and propagating method that is still used today, such as providing a path to access malicious code using social engineering attack technique called smishing.
	E-mail	- A method of exchanging messages between people using electronic devices without periodic access - Although it is not widely used due to the development of application services provided by message service and mobile device, it is still used as a malicious code infecting and propagating path because of its advantage of sending messages without mutual periodic access.
	Web board	- Community space where users can upload / download information between users. - Malicious code can be spread to infected and unspecified users, or attacked to administrator by using vulnerability of web board.
	Web application	- Applications that utilize Web services, such as Web browsers - Current web applications are used in a mixture of different kinds of applications to give users a visible effect. An attacker can attempt to propagate and infect malicious code to users of the web application through some of these vulnerable applications.
	TCP/UDP port service	- TCP / UDP-based services - Various services are provided through ports as well as well-known ports.

3.3.2 Infection & Propagation Type

Infection & propagation type subsection classifies the features used by malicious code to succeed in infecting and propagating to attack target.

Table 3. Contents of infection & propagation type

Subsection	Contents	Description
Infection & Propagation Type	Direct	- Infection and propagation through direct access such as downloading malicious code directly through network service or copying directly.
	Indirect	- Indirectly infecting and propagating by suggesting a method of accessing malicious code (Access link included in SMS/MMS etc.) through a social engineering method
	Vulnerabilities	- There are various types such as unauthorized access or elevation of privilege by infecting and spreading using security vulnerabilities.
	Weak setting	- Security related settings are not properly set up such as using default password, activating guest user etc.

3.3.3 User Dependency

User dependency subsection defines whether or not the system user intervention is necessary for malicious code to infect and propagate the attack target.

Table 4. Contents of user dependency

Subsection	Contents	Description
User Dependency	Need	- System user involvement is essential to successful infection and propagation - Since system user intervention is necessary, social engineering attack is often used. (e.g., malicious code download & execute, permission agreement etc.)
	Unnecessary	- Malicious code can be infected and propagated without the involvement of the system user. - Malicious code infections and propagation methods using security vulnerabilities and weak settings, such as unauthorized access through privilege elevation

3.4 Attack Section

Malicious code has a specific or unspecified attack target, and after it reaches the attack target, it executes itself through any entities. In addition, when the malicious code is executed, malicious code performs ultimate aim. There are a variety of types that can be used to accomplish the end goal themselves, or to be used as a method for secondary infection and propagation, etc.

Therefore, the Attack section classifies and defines malicious code attack targets, malicious code executors, and attacking behavior. The section consists of attack target, executor and attack behavior.

3.4.1 Attack Target

The attack target subsection classifies and defines the target that malicious code wants to cause an attack.

Table 5. Contents of attack target

Subsection	Contents	Description
Attack Target	User	- System user infected with malicious code is the victim of malicious code. - The user may be attacked to perform a secondary security threat using the user's identity. (e.g., disguised user, false information leakage, access to other user etc.)
	System	- Malicious code infected system itself is the target of malicious code. - If the infected system is an attack target, the attacker targets the service running on the system and the configuration information of the system. - In order to ensure the continuous operation of malicious code, the boot process can be attacked or the network service can be attacked for the second intrusion. - In order to utilize the resources of the infected system, the operation policy of the system can be targeted. (e.g., crypto mining, DDoS botnet etc.)
	Information	- Information of the user/system stored and managed inside the system is target of malicious code. - In order to sniff information such as OTP and user password, input information can be targeted for attack. - Malicious code can target the stored data to collect data related to the user or system. - System information can be attacked for continuous intrusion and attack.

3.4.2 Executor

The Executor subsection classifies and defines entities that actually execute malicious code.

Table 6. Contents of Executor

Subsection	Contents	Description
Executor	User	- Users execute malicious code that reaches the target system directly. - If the executor of the malicious code is a user, it is related to the user dependency of the infection & propagation section.
	OS	- OS execute malicious code. - There is a case where malicious code is inserted into the scheduler of the operating system. - Malicious code can be executed by modifying processes that run automatically in the operating system.
	Application	- Malicious code execution through applications with macros. - Execute malicious code using vulnerable applications. - Execute malicious code through other malicious code.
	BIOS	- The BIOS that manages booting the system is directly related to the execution of the malicious code. - These malicious code are executed before the operating system is loaded. - By inserting malicious code into the MBR(Master Boot Record), the malicious code can be run before the operating system is booted.

3.4.3 Attack behavior

Malicious code performs malicious behavior as it can be understood by its name. In attack behavior subsection, malicious code classifies and defines entities that are mainly used to perform malicious actions.

Table 7. Contents of attack behavior

Subsection	Contents	Description
Attack Behavior	System	- The attacking behavior of the malicious code is related to the operation of the infected system. - Malicious code exploits system privileges to perform unwanted actions. - It deliberately depletes system resources and causes system failure. - Network and security-related configuration information is forcibly changed so that malicious code can perform a attack.
	Process	- Malicious code attacks the process of infected system. - An attack that can control the execution and termination of processes. - Checking the status of the process and system to check whether the anti-virus application is running or not.

	Filesystem	- Malicious code attacks files stored and managed by an infected system. (e.g., file deletion, modulation, generation etc.)
	Network	- Malicious code performs network-based attack. - To leak information about system and user or stored data. - When a remote attack is performed by server-client communication between an attacker and a victim, such as RAT(Remote Access Trojan). - When it is related to an attack that uses network traffic such as DDoS attack.
	Device	- Malicious code attacks device managed by an infected system. - It causes malfunction of input/output device in PC environment. - In the case of a recent IoT environment, when attacking a server or a central control unit managing a plurality of IoT devices.

3.5 MCBI Example

In this section, MCBI for representative malicious code is shown to help understanding of MCBI structure generation. **Table 8** below shows the MCBI for the Mirai botnet code. Through the following results, contents information corresponding to Mirai botnet code can be utilized as tag information for type and grouping of malicious codes.

Table 8. MCBI example for Mirai botnet code

Section	Subsection	Contents	Description
Infection & Propagation	Infection & Propagation Route	TCP/UDP port service	- Mirai botnet code uses the Telnet port(23) as the initial infection and penetration route.
	Infection & Propagation Type	Weak setting	- Mirai botnet code attempts to infect IoT devices using the administrator default password.
	User Dependency	Unnecessary	- Mirai botnet code can be infected and propagated without the involvement of the system user.
Attack	Attack Target	System	- The Mirai botnet code targets the operating policy of the system in order to utilize the resources of the infected system.
	Executor	Application	- The Mirai botnet code performs a second malicious action by launching a malicious application that is installed after successful penetration into the IoT device.

	Attack Behavior	System	- The Mirai botnet code infects the system with additional command code to perform DDoS attacks on infected systems. - In order to keep the Mirai botnet code running, code is inserted to disable the reboot function.
		Network	- The Mirai botnet code utilizes the network system resources of the infected IoT devices to perform DDoS attacks.

4. Predicting Security Threats in New Environments

4.1 MCBI Management

The MCBI proposed in this paper is a structure that can be continuously expanded and developed. In order to predict security threats in a new environment, such as the emergence of new ICT devices or new ICT technology, MCBI structure should be continuously developed through MCBI analysis for various malicious codes. Therefore, this section describes the process for continuous management of MCBI.

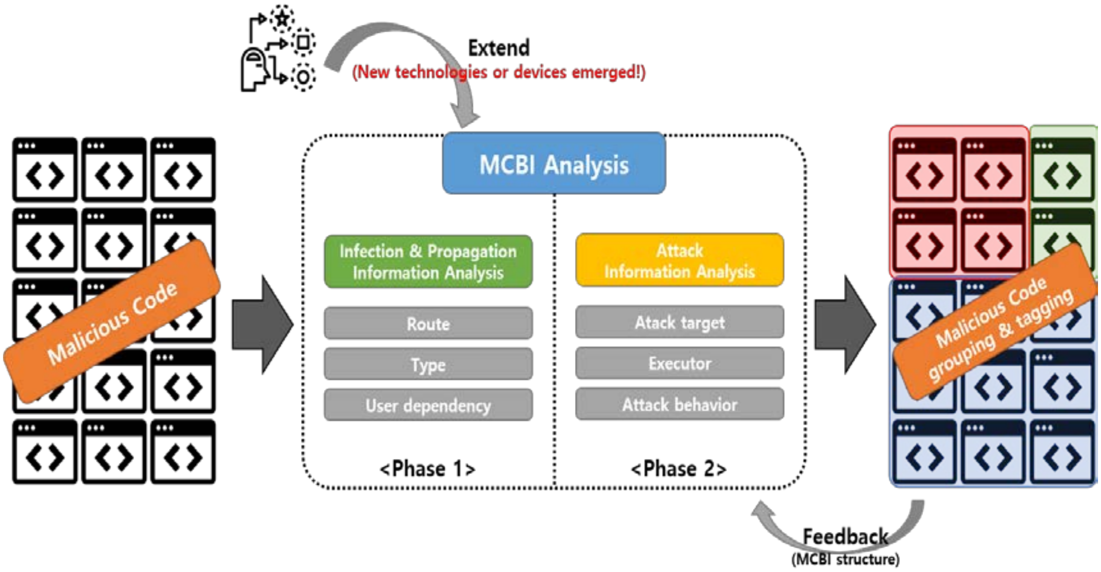

Extend
(New technologies or devices emerged!)
Malicious Code
MCBI Analysis
Infection & Propagation Information Analysis
Route
Type
User dependency
<Phase 1>
Attack Information Analysis
Atack target
Executor
Attack behavior
<Phase 2>
Malicious Code grouping & tagging
Feedback
(MCBI structure)

Fig. 3. MCBI management process

The process shown in **Fig. 3** shows a mixture of regular operation process and event conditions with feedback and extension.

- Regular operation process
  Step 1. Collect malicious code from various media and services.
  Step 2. Perform MCBI analysis. (Phase 1 and Phase 2 are performed in sequence)
  Step 3. Perform malicious code grouping and tagging based on the analysis result.

Regular operation process is aimed at gathering data on the recent behavior of malicious code by continuing to collect malicious code as possible and perform MCBI analysis. After analyzing the malicious code for MCBI, group similar malicious codes using MBCI contents.

- Feedback process

The feedback process is performed when it is difficult to define and classify a specific malicious code with the current MCBI structure. It is performed when a completely different type of malicious activity occurs although the ICT environment has not changed.

- Extension process

The process of extending the MCBI structure when a new technology or device emerged is called an extension process. New contents can be added to the infection & propagation route subsection when a new technology or device appears. Therefore, it is necessary to consider the process of expanding the MCBI structure.

4.2 Predicting Security Threats Process

Most of the security threats in the new environment are being reused from malicious code that has already occurred. Therefore, it is possible to predict the security threats that may occur in the new environment by performing the extension process of the MCBI structure proposed in this paper. When a new technology or device emerged with a new environment in the security ecosystem, the extension process of MCBI structure is performed through the following process and we predict security threats.

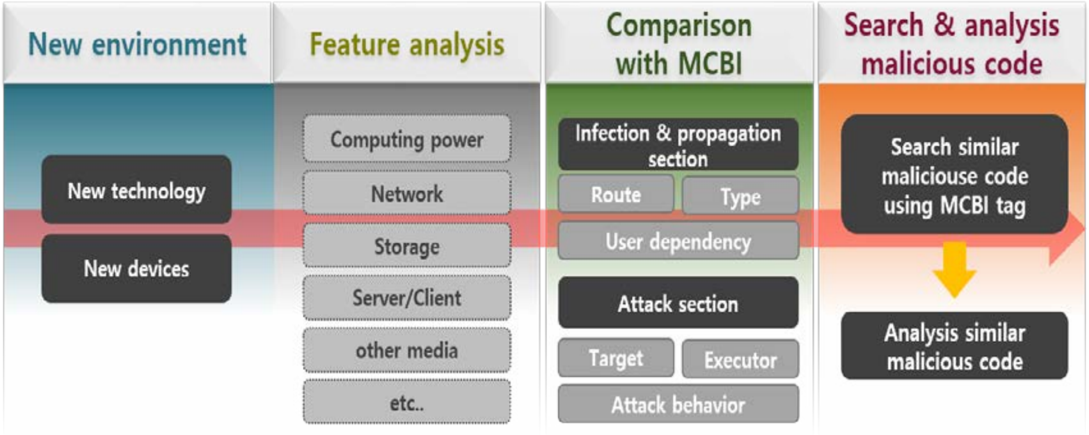


Fig. 4. Predicting security threats process

As shown in **Fig. 4**, the process of predicting possible security threats in a new environment consists of 4 steps.

- New environment

Initiate the MCBI's extension process to realize the emergence of new technologies or devices and to predict potential security threats in the upcoming new environment.

- Feature analysis

As a step of analyzing features for a new technology or device it analyzes not only computing ability, network function, storage function, but also connectivity with other media and data type that can be newly created and managed.

- Comparison with MCBI

This is the stage of comparing and analyzing with MCBI based on the analysis result of feature analysis step. If an entirely new technology or device emerges, new content can be added to the MCBI during this process. This process expands the MCBI to accommodate new environments.

- Search and analysis malicious code

The analysis of MCBI for new technologies and devices and malicious codes that have related MCBI tags are searched. Based on MCBI information of searched malicious codes, security threats that can occur in new environment are derived.

5. Conclusion & Future Research

Malicious code is evolving in line with the pace of ICT technology development. There may be some difficulties in predicting all possible security threats in a new environment that will be met by the emergence of new technologies and devices without special criteria. Malicious code that causes most security threats is evolving into a form that applies security threats that already existed in existing environments to new environments rather than entirely new forms. Therefore, in this paper, we classify and define not only internal information of malicious code that can be acquired from malicious code samples but also malicious behavior information such as infection process and attack process of malicious code. We refer to this information as Malicious Code Behavior Information (MCBI) and we propose a method to predict security threats in a new environment by continuously managing it and defining an extension process to apply it to new environment. We believe that it can help to predict and respond to repeated malicious code attacks in a rapidly changing ICT environment.

In the future, we will continue to manage and extend the MCBI information to predict the security threats to new technologies and devices. We also plan to provide effective and continuous operation of these processes.

Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. NRF-2017R1E1A1A01075110).

References

[1] Somia Sahraoui and Azeddine Bilami, “Asymmetric End-to-End Security for Human-to-Thing Communications in the Internet of Things,” in *Proc. of IoT’16 Proceedings of the 6th International Conference on the Internet of Things*, pp.131-139, November 07-09, 2016. Article (CrossRef Link).

[2] Meesun Kim, Hyun Ahn and Kwanghoon Pio Kim, “Process-Aware Internet of Things: A Conceptual Extension of the Internet of Things Framework and Architecture,” *KSII Transactions on Internet and Information Systems*, vol. 10, no. 8, August 31, 2016. Article (CrossRef Link).

[3] Vu-Anh-Quang Nguyen, “Study on realtime control system in IoT based smart factory: Interference awareness, architectural elements, and its application,” in *Proc. of Information Science and Technology (ICIST), 2017 Seventh International Conference on*, April 16-19, 2017. Article (CrossRef Link).

[4] H. Arasteh, V. Hosseinnezhad, V. Loia, A. Tommasetti, O. Troisi, M. Shafie-khah and P. Siano, “Iot-based Smart Cities: A Survey,” in *Proc. of Environment and Electrical Engineering (EEEIC), 2016 IEEE 16th International Conference on*, June 7-10, 2016. Article (CrossRef Link).

[5] Jorge Alfonso, Nuria Sánchez, José Manuel Menéndez and Emilio Cacheiro, “Cooperative ITS communications architecture: the FOTsis project approach and beyond,” *IET Intelligent Transport System*, vol. 9, issue. 6, pp.591–598, August 06, 2015. Article (CrossRef Link).

[6] McAfee Labs, McAfee Labs Threats Report September 2018, September, 2018. Article (CrossRef Link).

[7] Ahmad Karim, Syed Adeel Ali Shah, Rosli Bin Salleh, Muhammad Arif, Rafidah Md Noor and Shahaboddin Shamshirband, “Mobile Botnet Attacks – an Emerging Threat: Classification, Review and Open Issues,” *KSII Transactions on Internet and Information Systems*, vol. 9, no.4, April 30, 2015. Article (CrossRef Link).

[8] James A. Jerkins, “Motivating a market or regulatory solution to IoT insecurity with the Mirai botnet code,” in *Proc. of Computing and Communication Workshop and Conference (CCWC), 2017 IEEE 7th Annual*, January 09-11, 2017. Article (CrossRef Link).

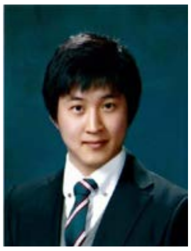
[9] Ahmed El-Kosairy and Marianne A. Azer, “Intrusion and ransomware detection system,” in *Proc. of 2018 1st International Conference on Computer Applications & Information Security (ICCAIS)*, September 27, 2018. Article (CrossRef Link).

[10] Taejin Lee and Jin Kwak, “Effective and Reliable Malware Group Classification for a Massive Malware Environment,” *International Journal of Distributed Sensor Networks, Hindawi Publishing Corporation*, Volume 2016, 2016. Article (CrossRef Link).

[11] Taejin Lee, Bomin Choi, Youngsang Shin and Jin Kwak, “Automatic malware mutant detection and group classification based on the n-gram and clustering coefficient,” *The Journal of Supercomputing, Springer*, 18 December, 2015. Article (CrossRef Link).

[12] Zhang Fuyong and Zhao Tiezhu, “Malware Detection and Classification Based on ngrams Attribute Similarity,” in *Proc. of 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, 21 July, 2017. Article (CrossRef Link).

[13] Arzu Gorgulu Kakisim, Mert Nar, Necmettin Carkaci and Ibrahim Sogukpinar, “Analysis and Evaluation of Dynamic Feature-Based Malware Detection Methods,” *Innovative Security Solutions for Information Technology and Communications (SECITC 2018)*, pp 247-258, Feb, 2019. Article (CrossRef Link).



Seul-Ki Choi received Korea B.S. and M.S degrees in Department of Information Security Engineering from Soonchunhyang University. He is currently pursuing the Ph.D. degree in Department of Computer Engineering with Ajou University, Korea. His research interests include IoT Security, Vulnerability & Malware analysis and Cryptographic protocols.

Tae-Jin Lee is a professor at Dept. Of Information Security in Hoseo University, Korea. He received the Ph.D. degree from Ajou University, Korea. Professor Lee’s current research interests focus on System security, Malware Analysis.

Jin Kwak is a professor at Dept. Of Cyber Security in Ajou University, Korea. He received the Ph.D. degree from SKKU, Korea. His research interests include Cryptographic protocols, Applied security mechanisms for Cloud and Big Data system and so on.