ISSN:2321-1156
International Journal of Innovative Research in Technology & Science(IJIRTS)

CYBER THREAT ANALYSIS AND QUANTIFICATION BASED ON INDICATOR OF COMPROMISE

Hyeisun Cho, KISA; Seulgi Lee, KISA; Nakhyun Kim, KISA; Byung-ik Kim, KISA; Jun-hyung Park, KISA

Abstract

As a large quantity of new and varied attacks occur in Korea, it is difficult to analyze and respond to them with limited security experts and existing equipment. This paper proposes a method of analyzing the threat of Indicator of Compromise (IoC) used for cyber incidents and calculating it as a quantitative value Threat Level of IoC (TL_IoC) in order to check the analysis priority of cyber incidents that occur in large quantities. Using this method, a large quantity of cyber incidents can be efficiently responded to by checking the TL_IoC objectively to quickly determine the response level of the cyber incident and actively analyze cyber incidents with high threat levels.

Ⅰ. Introduction

As a large quantity of new and varied attacks occur in Korea, a paradigm shift is required in terms of response to cyber incidents [1]. As a matter of fact, the APT attack exposure rate for Korea is 38%, which is more than double the world average (15%) and almost three times higher than the U.S. (13%) [2].

As described, even though attackers create new and varied attacks in large quantities that bypass existing security equipment to achieve their attack goals on a particular target [3], it is difficult to analyze and respond to these massive cyber incidents due to the limits of existing equipment and the limited number of analysis personnel. In light of limited resources, to efficiently defend against a large quantity of intelligent cyber- attacks that occur these days, a new strategy is needed that actively analyzes and responds to cyber incidents with a higher threat level [4][5].

This paper proposes a method of analyzing the threat of each IoC used for each cyber incident and calculating it as a quantitative value in order to efficiently analyze cyber incidents occurring in large quantities.

Ⅱ. Cyber threat analysis based on the IoC

The Indicator of Compromise (IoC) is the data that refers to the evidence trail of a particular cyber incident. Originally, the IoC was used in the digital forensic and incident response area, but its use is gradually expanding now and security equipment and solutions also use IoC to check the history of cyber incidents. IoC refers to attacks that include the infrastructure used to attack the cyber incident (servers and domains used as distribution sites, waypoints, C&C, etc.), date and time of attacks, attacks including malicious code, and related information [6].

IoC is used in this study by collecting the analysis result of cyber incidents occurring in the past and attack information that can be obtained from open channels. <Table 1> shows IoC collected for this study.

Table 1. Collection of cyber threat information

Item		Collection details	Total collection quantity (‘15.10.~‘16.12)
Indicator of Compromise	Threat Indicator	C&CIP	96
		Zombie IP	146,122
		Waypoint	5,131
		Distribution site	4,042
	Reputation data	RBL_IP	31,622,940
		RBL_Domain	69,690
	malware	malware (New)	372,117
		malware (Variant)	
Total			**32,220,138**

To analyze the threat of IoC, factors are needed that can determine the “level of threat”. For this, threat analysis requirements can be extracted. The following is shows threat analysis requirements that consider the detection method, detection time, reputation information, and behavior information of the IoC from the viewpoint of attack detection/response.

1) Viewpoint of IoC detection path (source): IoC threats can be understood by checking the source (C&C, zombie,

waypoint, distribution site, etc.) of the detection system that determines the characteristics of the IoC.

2) Viewpoint of the detection time: As IoC detected recently is more likely to be close to recent or current time, threats can be understood according to the detection time.

3) Viewpoint of blacklist registration: As the threat information registered in the blacklist reflects a good deal of reputation information, the threat of IoC can be understood according to the status of blacklist registration.

4) Viewpoint of the DNS change history: If there are many DNS changes, it can be regarded as a strategy of an attacker to avoid attacker tracking. Therefore, threats can be understood according to the DNS change history.

5) Viewpoint of detected malicious URL: The threat of IoC can be understood by checking the number of malicious URLs associated with the attack information.

6) Viewpoint of variant malicious code: As the scale and spread level of the attack group can be guessed using the number of associated variant malicious codes, the threat of IoC (malicious code) can be understood according to the number of variant malicious codes.

7) Viewpoint of malicious code spread: As the number of malicious codes distributed by the IoC (distribution site) is high, it means that attacks are frequently made, which enables us to understand the threat of the IoC.

8) Viewpoint of using distribution sites/waypoints: If a particular domain is frequently used as a distribution site/waypoint, it can be regarded as an infrastructure frequently exploited by attackers. Therefore, the threat of IoC can be understood by the history of using distribution sites/waypoints.

9) Viewpoint of web alteration history: As we can check web page hacking according to the web alteration status, the threat of IoC can be understood by the web alteration status.

10) Viewpoint of dropped malicious code: As we can understand attack complexity by checking the existence of dropped malicious code, the threat of IoC (malicious code) can be understood.

Ⅲ. A Proposal of threat analysis and quantification

A. Weighted value setting according to the threat level analysis standard

We have analyzed the IoC threat analysis requirements mentioned in Section Ⅱ. The aforementioned requirements enable us to conduct a semantic cyber incident analysis based on IoC analysis, but it takes a long time to understand the meaning by analyzing each cyber incident. Therefore, this section presents a method of analyzing the threat and describing it quantitatively according to the IoC threat analysis requirement.

To figure out IoC threats quantitatively, compose the <IoC threat quantification matrix> (Section B) [7]. Set a "Weighted value" to reflect importance by the IoC threat analysis factor in the threat level before composing the matrix. The "weighted value" is used as a tool to control the impact of each factor on threat analysis using the heuristic technique. <Table 2>

Table 2. Correlation according to the IoC analysis standard and weighted value setting

IoC analysis factor	Threat level correlation	Weighted value*
1) Viewpoint of detection path (source)	Threat level ∝ Detection route 1 (C2, distribution site, malicious code) > Detection route 2 (attack utilization, waypoint) > Detection route 3 (RBL) > Detection route 4 (host infected with malicious code)	10
2) Viewpoint of the detection time	Threat level ∝ detected time (recent)	1
3) Viewpoint of blacklist registration	Threat level ∝ Blacklist registration status	5
4) Viewpoint of the DNS change history	Threat level ∝ DNS history (count)	5
5) Viewpoint of the detected malicious URL	Threat level ∝ History using the malicious URL (count)	10
6) Viewpoint of variant malicious code	Threat level ∝ Variant malicious code (count)	15
7) Viewpoint of malicious code spread	Threat level ∝ Distributed malicious code (count)	15
8) Viewpoint of using the distribution site/waypoint	Threat level ∝ History of using the distribution site/waypoint (count)	9
9) Viewpoint of web alteration history	Threat level ∝ Web alteration status	15
10) Viewpoint of dropped malicious code	Threat level ∝ Dropped malicious code status	15

Total	100

B. Composing the matrix determining the IoC threat level

The <IoC threat level determination matrix> is composed to convert the analysis result of each IoC quantitatively. <Table 3>

The <IoC threat level determination matrix> sets "threat attributes" that can reflect the analysis result by the analysis factor of the IoC, and gives the threat level that quantitatively indicates how serious the threat is. The "threat level" is classified into levels 0~5 according to the characteristics of the threat attributes. The bigger the number, the higher the threat level. 0 indicates "no information" [8][9].

Table 3. IoC threat level determination matrix

IoC analysis factor	Attribute	Threat level	Threat index*	Maximum threat index**
Detection path (source)	Detection route 1 (C2, distribution site)	5	50	50
	Detection route 2 (attack utilization, waypoint)	4	40	50
	Detection route 3 (RBL)	3	30	50
	Detection route 4 (host infected with malicious code)	1	10	50
Detection time	~1 month	5	5	5
	1~3 months	4	4	5
	3~6 months	3	3	5
	6~12 months	2	2	5
	12 months ~	1	1	5
RBL registration status	Existence	3	15	15
	Non-existence	0	0	15
DNS change history	41~	5	25	25
	31 ~ 40	4	20	25
	21 ~ 30	3	15	25
	11 ~ 20	2	10	25
	1 ~ 10	1	5	25
	0	0	0	25
Detected malicious URL	41~	5	50	50
	31 ~ 40	4	40	50
	21 ~ 30	3	30	50
	11 ~ 20	2	20	50
	1 ~ 10	1	10	50
	0	0	0	50
Malicious code variants	41~	5	75	75
	31 ~ 40	4	60	75
	21 ~ 30	3	45	75
	11 ~ 20	2	30	75
	1 ~ 10	1	15	75
	0	0	0	75
Malicious code spread	41~	5	75	75
	31 ~ 40	4	60	75
	21 ~ 30	3	45	75
	11 ~ 20	2	30	75
	1 ~ 10	1	15	75
	0	0	0	75
Using the distribution site/waypoint	41~	5	45	45
	31 ~ 40	4	36	45
	21 ~ 30	3	27	45
	11 ~ 20	2	18	45
	1 ~ 10	1	9	45
	0	0	0	45
Web alteration history	Existence	5	75	75
	Non-existence	0	0	75
Dropped malicious code	Existence	5	75	75
	Non-existence	0	0	75

* (Threat index) = (threat level) x (weighted value)

** (Maximum threat index) = (maximum threat level by index) x (weighted value)

C. Quantification of IoC threats

Quantification by IoC threat analysis can be performed based on the <IoC threat level determination matrix> defined in Section B. The value that analyzes and quantifies the threat of the IoC is expressed as a percentage, which is defined as Threat Level of IoC (TL_IoC)[10][11].

If we assume the number of factors used to figure out TL_IoC is n among 10 factors defined with IoC threat analysis, the TL_IoC can be calculated using the following expression (1) :

$$TL\_IoC(x) = \sum_{i=1}^{n}(t_i \times w_i) / \sum_{i=1}^{m}(m_i \times w_i) * 100 \qquad (1)$$

Here, T is a set of threat indices that indicates the threat attribute by IoC threat analysis factor, M is a set of maximum values that T can have, and W is a set of weighted values that correspond to each factor of T. (Figure 1) The flowchart of the IoC threat analysis system quantification system. The IoC threat level is calculated once a month using batch processing and saved in the IoC management database [12].

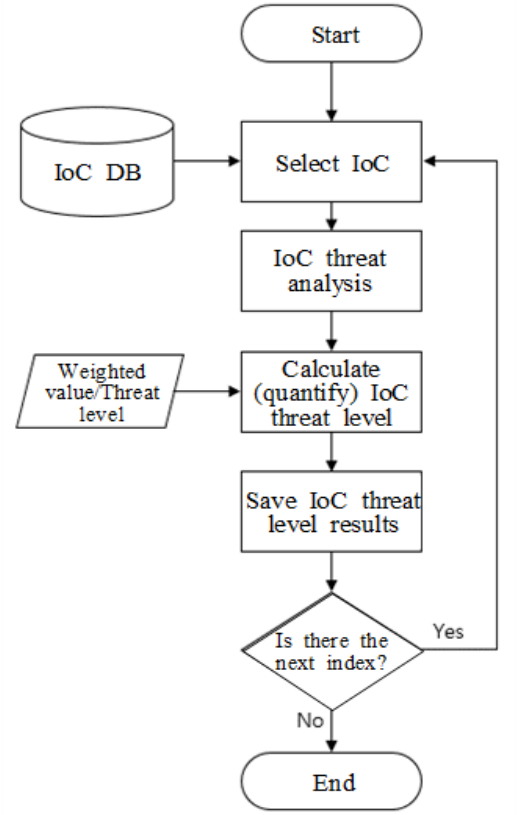

Figure 1. The flowchart of the IoC threat analysis system

Ⅳ. Result of calculating the IoC threat level

A module was implemented to calculate the IoC threat level which refers to the actual IoC management DB using the expression and logic of calculating the IoC threat level. The module is executed by periodical batch processing as the threat level is changed when collecting information about the detection time and additional factors.

When the IoC threat is analyzed regarding the malicious domain exploited for actual cyber incidents, the threat level could be calculated using five factors. (Figure 2)

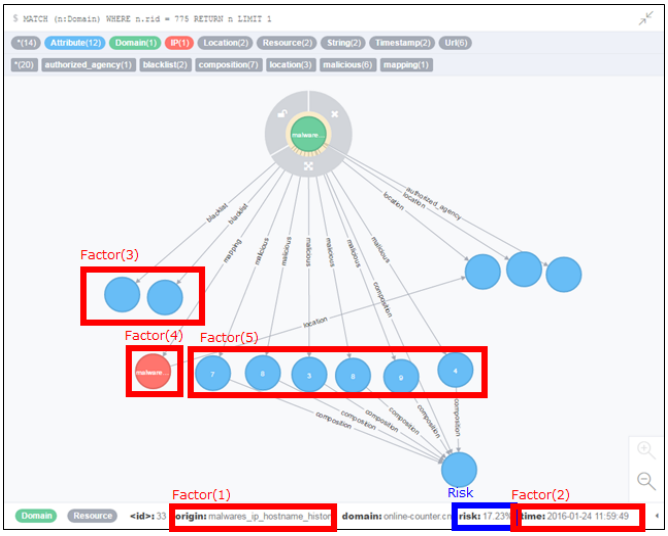

Figure 2. Threat level analysis results regarding the IoC

The figure above is the result of analysis of the cyber threat level against the domain used in the cyber incident. A graph database was used to effectively store and manage relevant information on indicators. As a result of the cyber threat analysis for the phishing domain named 'online-co ****. Cn', the history of the DNS change history in order to avoid detection has not been changed but the history of utilizing the malicious URL was generally high. However, since the time spent as a phishing domain lasted more than a year, the threat level was 17.23%.

As validation of the proposal system, we tried to compare with result of threat analysis. We confirmed the threat level according to the result of maliciousness check of open source and identified the error rate of this study [13][14]. As a result, average of error rate is 8.23%.<Table4>

Table 4. Comparison of Open Source Analysis Results

Item		Detection rate	TL_IoC (%)	Error rate (%)
Proposal system		-	17.23	-
Open source	VirusTotal	4/64	6.25	10.98
	URLvoid	4/34	11.76	5.47

Average of error rate	8.23

V. Conclusion

This paper presents a method of analyzing threats based on the IoC used for cyber incidents, and changing this result to a quantitative value. By applying the IoC threat level that was found, the response level can be determined by understanding the threat level intuitively when cyber incidents actually occur. Likewise, large-scale intelligent cyber-attacks can be responded to efficiently by responding to highly threatening cyber incidents in advance.

However, there is still an issue of securing objectivity in selecting the threat factor to judge the threat level of the indicator. It is a fact that it is difficult to secure objectivity in determining the threat because all the threat evaluation criteria are not 'measurable' because the subjective view of the system operation and analysis subject is likely to be reflected in the process of determining the threat.

However, there is no current standard for evaluating threats and threat levels separately for indicators of compromise that can be resolved objectivity issues. In this sense, this study is a meaningful precedent study for the future researches and works for securing objectivity in that it presents the evaluation criteria of the threat of indicators and presents a measurement method for the threat level.

In future work, we will propose a consistent and measurable threat assessment standard to show the threat level of indicators that are reliable, and to support more effective response to cyber incidents. In addition, in this paper, it is expected that the information used in the attack and the related information are used to analyze the threat level of the cyber indicator, but it can be used variously in the analysis of the cyber incidents as well as the threat level analysis. It is possible to estimate attack group by extracting similar or same characteristics among indicators extracted from different cyber-attacks, and it is also possible to analyze the time series based attack trend through history management of indicators.

Acknowledgments

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT)(2017-0-00158, Development of Cyber Threat Intelligence(CTI) analysis and information sharing technology for national cyber incident response.)

References

[1] "Prospect of 7 cyber attacks in 2017 (report), Korea Internet & Security Agency, 2016.12.
[2] FireEye, "FireEye announced the domestic cyber attack status and ransomware trends", 2016.04.
[3] Kaspersky, "What is known about the Lazarus Group: Sony hack, military espionage, attacks on Korean banks and other crimes", 2016.
[4] "Urgently needed to secure reliability of cyber threat information", The Electronic Times, 2015.11.
[5] Chanil Park , "Vulnerability Risk Computation method for Attack Graph Generation" the Korean Institute of Communications and Information Sciences (2017): 123-124.
[6] ″Indicator of compromise″, 『Wikipedia』 (2011).
[7] "Basic matrix to respond to cyber attacks", Ryu Hakyong, Ryu Dongyeong, Korea Internet & Security Agency, 2014.06
[8] Sallam, Hany. "Cyber Security Risk Assessment Using Multi Fuzzy Inference System." IJEIT 4.8 (2015): 13-19.
[9] "A study on the malicious code detection method", Kim Taekyeong, Security engineering research papers 9.5 (2012): 387-400.
[10] CVE, Available: http://www.cve.mitre.org/
[11] CVSS, Available: https://www.first.org/cvss/
[12] Hyeisun Cho, "Analysis of Cyber Threat Level Based on Indicator of Compromise", Spring KIPS Conference (2017).
[13] VirusTotal, Available: https://virustotal.com/
[14] URLvoid, Available: http://www.urlvoid.com/

Biographies

HYEISUN CHO received the B.S. degree in Computer Science from the University of Sejong, Korea, in 2013, the M.S. degree in Information Security from the University of SungKyunKwan, Korea, in 2017, respectively. Currently, She is a Researcher of Security Technology R&D Team 1 at Korea Internet & Security Agency. Her research areas include cyber threat analysis, cyber attack related data correlation and cyber threat reputation analysis. Hyeisun Cho may be reached at hscho@kisa.or.kr

SEULGI LEE received the B.S. degree in Computer Science from the University of Chungnam, Korea, in 2013. Currently, He is a Researcher of Security Technology R&D Team 1 at Korea Internet & Security Agency. His research areas include cyber threat analysis, cyber attack related data correlation and machine learning algorithm. Seulgi Lee may be reached at sglee@kisa.or.kr

NAKHYUN KIM received the B.S. degree in Computer Science from the University of Seoil, Korea, in 2008, the M.S. degree in Network Security from the University of SoongSil, Korea, in 2011, respectively. Currently, He is a Deputy General Researcher of Security Technology R&D Team 1 at Korea Internet & Security Agency. His research areas include cyber threat analysis, cyber attack related data correlation, and Sensor Network Security. Nakhyun Kim may be reached at knh@kisa.or.kr

BYUNG-IK KIM received the B.S. degree in Computer Science from the University of Ajou, Korea, in 2010. Currently, He is a Deputy General Researcher of Security Technology R&D Team 1 at Korea Internet & Security Agency. His research areas include cyber threat analysis, cyber attack related data correlation, and Sensor. Byung-Ik Kim may be reached at kbi1983@kisa.or.kr

JUN-HYUNG PARK received the B.S. degree in Computer Science from the University of Ajou, Korea, in 1999, the M.S. degree in Multimedia from the University of Chonnam, Korea, in 2002, the PhD degree in Information Security from University of Chonnam, Korea, in 2004, respectively. Currently, He is a Manager of Security Technology R&D Team 1 at Korea Internet & Security Agency. His research areas include cyber threat analysis, malware analysis and mobile billing fraud detection. Junhyung Park may be reached at junpark@kisa.or.kr