Journal of information and
communication convergence engineering
http://jicce.org

J. Inf. Commun. Converg. Eng. 18(1): 16-21, Mar. 2020 Regular paper

Event Log Validity Analysis for Detecting Threats by Insiders in Control System

Jongmin Kim[1*], **Jiwon Kang**[2], **and DongHwi Lee**[3*], ***Member*, *KIICE***

[1]Department of Convergence Security, Kyonggi University, Suwon 16227, Korea
[2]Department of Information Security, Sejong University, Seoul 05006, Korea
[3]Department of Information Security, Dongshin University, Naju 58245, Korea

Abstract

Owing to the convergence of the communication network with the control system and public network, security threats, such as information leakage and falsification, have become possible through various routes. If we examine closely at the security type of the current control system, the operation of the security system focuses on the threats made from outside to inside, so the study on the detection system of the security threats conducted by insiders is inadequate. Thus, this study, based on "Spotting the Adversary with Windows Event Log Monitoring," published by the National Security Agency, found that event logs can be utilized for the detection and maneuver of threats conducted by insiders, by analyzing the validity of detecting insider threats to the control system with the list of important event logs.

Index Terms: AHP, Control System, Information Security, Event Log

I. INTRODUCTION

Control systems were generally constructed with independent closed networks. However, with increased emphasis on convenience at a workplace and an increase in cooperative work with outside institutions, the usage rate of the control system with a Windows operating system and the universal protocol increased as well [1].

Due to the increase in the usage of the Windows operating system, the working with insider management accounts became frequent, to match the security patch and control system patch of the operating system. This resulted in a new exposure of security threats to the control system managed by insiders.

Although the related environments are changing, security systems, such as defense wall, vaccine system, IPS, and PMS, are focused on the attacks made from outside to inside. Thus, security measures for preventing attacks made from the inside, caused through obligatory processes such as installation of security patches and updates of application programs, are currently inadequate.

Therefore, in this study, to conduct a credible analysis, the important event logs of "Spotting the Adversary with Windows Event Log Monitoring," published by the National Security Agency, were used. We selected event logs that can be utilized to detect security threats conducted by insiders in a control system based on the Windows operating system and utilized an analytic hierarchy process to examine the levels of relative significance by different event logs.

Received 15 November 2019, **Revised** 24 March 2020, **Accepted** 25 March 2020
***Corresponding Author** DongHwi Lee (E-mail: dhclub@dsu.ac.kr, Tel: +82-61-330-3954)
Dept. of Information Security, Dongshin University, Naju 58245, Korea.

Open Access **https://doi.org/10.6109/jicce.2020.18.1.16** print ISSN: 2234-8255 online ISSN: 2234-8883

II. RELATED WORK

A. Precedent Research on Security Threats on Control System

The precedent research related to the threat detection system of the control system is described below.

In “Using Model-based Intrusion Detection for Supervisory Control and Data Acquisition (SCADA) Networks,” published in 2006, three model-based methods for detecting dubious symptoms in the Modbus Transmission Control Protocol(TCP) network are proposed: ① protocol-level model, ② communication pattern model, and ③ learning-based model. These were the experimented methods that were proposed in the SCADA test bed of the Sandia National Laboratories (SNL) in the US, of which the model-based methods were found to be most effective in the SCADA network. First, in relation to the protocol-level model, a set of usable function codes by each Modbus device and the rules containing cross-field relations based on the Modbus protocol specification were defined. Second, regarding the communication pattern model, a connectable communication set based on IP addresses and TCP port numbers was defined. Third, apropos the learning-based model, a Bayesian network was structured with the Modbus function codes and abnormal symptoms were detected using conditional probability [2].

In “Communication Pattern Anomaly Detection in Process Control Systems,” published in 2009, two methods for detecting abnormal symptoms were presented: ① pattern-based anomaly detection and ② flow-based anomaly detection. The pattern-based anomaly detection method was the existing method of pattern anomaly detection, which used a pattern that was created by the sent and received IP addresses and port numbers. The flow-based anomaly detection method measured the mean byte size of packets at intervals of certain time by different flows and the mean inter-arrival time, and compared them with the normal model to detect the anomaly symptoms [3].

In “Bloom Filter Based Intrusion Detection for Smart Grid SCADA,” published in 2012, the host-based intrusion detection system, which applied the bloom filter that was based on Modbus function codes and data sequences, minimized the memory space, and shortened the detection speed, considering the performance issue of the automated substation system, was presented. The authors asserted that different models should be applied according to the operating status of the control system. However, they found that there was a difficulty in detection in cases of the non-existing test samples included in the training set. This disadvantage can be complemented when the whitelist, which includes all normal activities, is defined to guarantee the effectiveness in terms of the memory space and detection speed [4].

In “Network Security for Substation Automation,” published in 2001, four types of attacks were proposed in relation to the vulnerability of automated substations and possible scenarios of threats and attacks: ? attack of message modification, which falsifies the control order message in the process of delivery between the source and destination; ② replay attack, in which an attacker sends a new control order or harrows the control order of the source and uses it for the attack; ③ message injection & replay attack, in which an attacker can steal the packet and deliberately omits them; and ④ an attack type of dropping certain messages with malignant codes by controlling the switches for automated substation, routers, and other network devices for making it impossible to deliver messages. To counter these threats to the electric power facilities, the ensuring message integrity and encryption security policy were proposed in International Electrotechnical Commission (IEC) 62351. Particularly, the Goose message, a type of IEC 61850 communication service, is used for the protection of electric power facilities by delivering the important information related to alarming, status, and control, such as trip, interlocking, and status, among each protection intelligent electrical device (IED). Furthermore, the sampled value (SV) message is used for the measurement-reporting services of current and voltage. However, in the 1st edition standard of IEC 61850, only the communication service, object model on data, and the security alternatives that should be provided in the network and smart IED are proposed. In the 2nd edition of IEC 61850, which was presented recently, the functions of Generic Security Application (GSAL), Generic Log (GLOG), and Logical Node (LN), were added to provide the log information of smart-type IED [5].

According to the “Exploiting the Generic Object Oriented Substation Event (GOOSE) Protocol: A Practical Attack on Cyber-infrastructure,” published in 2012, although the security requirement of Goose recommended by IEC 62351 includes message certification through encryption and digital signature, it cannot fulfill the time (latency, ta+tb+tc) of 4 ms, the standard required for IEC 61850, as it takes at least 8.3 ms for verification with message certification or digital signature processes [6].

In “Cyber-attack detection in SCADA systems using temporal pattern recognition techniques,” published in 2019 [7], we proposed a cyber-attack detection technology based on pattern recognition for network attacks. In addition, in “Machine Learning Based Intrusion Detection in Control System Communication,” published in 2020 [8], the intrusion detection method was proposed by applying machine learning.

Previous papers have introduced the potential of control system intrusion detection algorithms and index-based Function codes, and there is insufficient information on insider threat detection. Therefore, this study, based on “Spotting the Adversary with Windows Event Log Monitoring,” published by the NSA, aims to verify that event logs can be utilized for the detection and maneuvering of threats conducted by insiders, by analyzing the validity of detecting insider threats to the control system with the list of important event logs.

B. Windows Event Log

The Windows operating system records three types of logs in the event list: application log, security log, and system log, where the directory service log, file-replication service log, and DNS server log can be added according to the OS compositions [9]. The main features of each event are listed in Table 1.

III. PROPOSED SCHEME

In this section, we analyze the validity levels for discovering whether event logs can be applied for insider threat detection in a control system, based on “Spotting the Adversary with Windows Event Log Monitoring,” published by the NSA.

A. Structure of the Analytic Hierarchy Research

Setting items is important in the analytic hierarchy process (AHP). Fig. 1 shows a hierarchy diagram for finding the validity levels of event logs on inside threats. The most significant stage of the evaluation element is categorized with 16 NSA events, and each higher stage has sub-items.

Table 2 shows the list of the 16 categories of important events based on “Spotting the Adversary with Windows Event Log Monitoring” of the NSA.

Table 3 shows the organized sub-items subordinate to the higher stages.

Table 1. Types of Windows event logs [9]

Event Log	Explanation
Application	Various events recorded by application programs are stored, and the recorded events are decided by the developer of the product. ex. Anti-virus provides record of the detection of malignant codes and update lists. General application programs record information about the activation status and success status.
Security	Events related to the use of resources, such as logging in trials whether they are valid or invalid and creation/reading/deletion of files. By setting auditing logs, diverse security events can be stored.
System	Recorded events by the composition elements of the Windows system, which records the errors occurring in the composition elements, such as the unloaded driver in system booting.

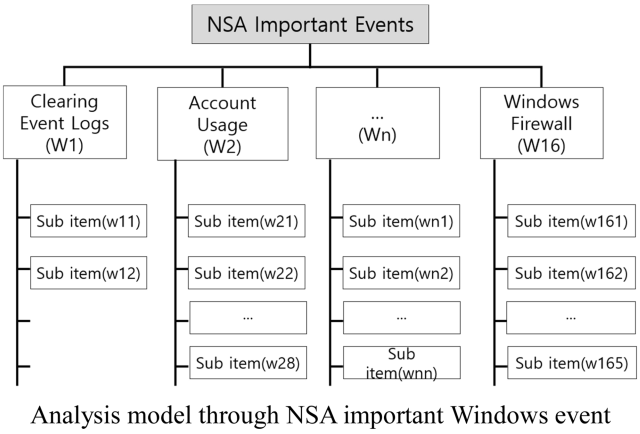

Fig. 1. Analysis model.

Table 2. List of elements of the most significant items [10]

Categories			
Clearing Event Logs	**Account Usage**	**Remote Desktop Logon Detection**	**Windows Defender Activities**
Application Crashes	Software & Service Installation	External Media Detection	Pass the Hash Detection
AppLocker	System or Service Failures	Windows Update Errors	Kernel Driver Signing
Group Policy Errors	Mobile Device Activities	Printing Services	Windows Firewall

Table 3. List of elements of sub-items [10]

General Event Descriptions	General Event IDs
Account and Group Activities	4624, 4625, 4648, 4728, 4732, 4634, 4735, 4740, 4756
Application Crashes and Hangs	1000 and 1002
Windows Error Reporting	1001
Blue Screen of Death (BSOD)	1001
Windows Defender Errors	1005, 1006, 1008, 1010, 2001, 2003, 2004, 3002, 5008
Windows Integrity Errors	3001, 3002, 3003, 3004, 3010 and 3023
EMET Crash Logs	1 and 2
Windows Firewall Logs	2004, 2005, 2006, 2009, 2033
MSI Packages Installed	1022 and 1033
Windows Update Installed	2 and 19
Windows Service Manager Errors	7022, 7023, 7024, 7026, 7031, 7032, 7034
Group Policy Errors	1125, 1127, 1129
AppLocker and SRP Logs	865, 866, 867, 868, 882, 8003, 8004, 8006, 8007
Windows Update Errors	20, 24, 25, 31, 34, 35
Hotpatching Error	1009
Kernel Driver and Kernel Driver Signing Errors	5038, 6281, 219
Log Clearing	104 and 1102
Kernel Filter Driver	6
Windows Service Installed	7045
Program Inventory	800, 903, 904, 905, 906, 907, 908
Wireless Activities	8000, 8001, 8002, 8003, 8011, 10000, 10001, 11000, 11001, 11002, 11004, 11005, 11006, 11010, 12011, 12012, 12013
USB Activities	43, 400, 410
Printing Activities	307

IV. RESEARCH RESULTS

In this section, the consistency rates and validity levels are obtained by calculating the weighted values with relative comparison matrices of NSA important event log elements, and verifying the consistencies of the calculated values.

A. Verifying Credibility Through Consistency Analysis

The advantage of the AHP method is that the respondents' consistency can be verified in the process of calculating the weighted values by using the relative comparison matrices.

Let A be a relative comparison matrix and $\underline{w}$ be a weighted value vector. The order of the matrix, which is defined as the number of evaluation elements, is n. Element a_{ij} of the relative comparison matrix A is the degree value of rationality for the evaluation element j of the reference element i.

$$A = \begin{pmatrix} 1 \; a_{12} \; a_{13} \cdots a_{1n} \\ 1 \;\; 1 \;\; a_{23} \cdots a_{2n} \\ \cdots \\ 1 \; a_{n2} \; a_{n3} \cdots \;\; 1 \end{pmatrix}, \; \underline{w} = [w_1, w_2, \cdots, w_n] \qquad (1)$$

The scalar λ that fulfills $|A - \lambda I| = \underline{0}$ about matrix A is called eigenvalue, and the vector fulfilling $A\underline{w} = \lambda\underline{w}$ about eigenvalue λ is called eigenvector. The eigenvector corresponding to the maximum eigenvalue of the relative comparison matrix A is used for calculating the weighted value of an evaluation element. When the element of the eigenvector is divided by the sum of the elements, and the sum becomes 1, the weighted value can be calculated.

If the relative comparison matrix has complete consistency, $a_{ij}a_{jk} = a_{ik}$ is fulfilled. However, it is almost impossible for an actual respondent to maintain complete consistency. If complete consistency is maintained, $\lambda_{\max} > n$ is fulfilled, and if it breaks, $\lambda_{\max} > n$. Using these principles, Saaty defined the consistency index (CI) as follows [11, 12]:

$$\text{Consistency Index : CI} = \frac{\lambda_{max} - n}{n-1} \qquad (2)$$

Furthermore, consistency indexes were induced to the following concept. If the relative importance a_{ij} of the reference element i for the comparison element j and the inconsistency level is $\delta_{ij} > -1$, a_{ij} can be expressed as $a_{ij} = (1 + \delta_{ij})w_i/w_j$. Then, the following equation is fulfilled.

$$\lambda_{\max} - n = \frac{1}{n} \sum_{1 \le f < f \le n} \frac{\delta_{ij}^2}{1 + \delta_{ij}} \ge 0 \qquad (3)$$

Table 4. Consistency rates of survey respondents

Respondent	Maximum Eigenvalue $\lambda_{\max}$	Consistency Index CI	Consistency Rate CR (%)
1	78.429	4.162	2.793
2	69.297	3.553	2.385
3	51.277	2.352	1.578
4	74.161	3.877	2.602
5	43.139	1.809	1.214
6	68.298	3.487	2.340
7	89.796	4.920	3.302
8	80.447	4.296	2.884
9	50.491	2.299	1.543
10	84.652	4.577	3.072

From the above formula, if an evaluator has complete consistency, in other words, if his/her δ_{ij} becomes 0, $\lambda_{\text{mas}} = n$ is obtained. We used this to induce the concept of CI.

As consistency increases, $\lambda_{\max}$ approaches n. Therefore, the consistency level can be measured using the CR as follows [13].

$$\text{CR(Consistency Rate)} = \frac{\text{CI (Consistency Index)}}{\text{RI (Random Index)}} \qquad (4)$$

Here, as the consistency increases, it approaches 0. RI is the random index, which is the mean of the CIs of a comparison matrix produced with the RIs of 1 to 9.

Table 4 shows the maximum eigenvalues, CIs, and consistency rates (random number index = 1.12) solved from the relative comparison matrix of the 10 respondents who evaluated the relative rationality on important event logs. The consistency rates of all respondents stayed less than 10%, verifying the consistency in their evaluations.

By calculating the single relative comparison matrix using the relative comparison matrix of the respondents' evaluation elements, which maintained their consistency, the result shown in the table could be drawn.

B. Result of Validity Analysis

The analysis result on validity, as shown in Table 5, and the value 0.577 in 1 row 2 columns was calculated through the following calculation process.

Element value in 1 row 2 columns of Respondent 1 = 0.333

Element value in 1 row 2 columns of Respondent 2 = 1

Element value in 1 row 2 columns of Respondent 3 = 1

Element value in 1 row 2 columns of Respondent 4 = 0.333

Element value in 1 row 2 columns of Respondent 5 = 1

Table 5. Analysis result of evaluation element validity

Evaluation Validity	Clearing Event Logs	Account Usage	Omitted	Printing Services	Windows Firewall
Clearing Event Logs	1.000	0.577	…	9.000	3.323
Account Usage	1.732	1.000	…	8.559	5.171
Omitted	…	…	1.000	…	…
Printing Services	0.111	0.117	…	1.000	8.777
Windows Firewall	0.301	0.193	…	0.114	1.000

Element value in 1 row 2 columns of Respondent 6 = 1
Element value in 1 row 2 columns of Respondent 7 = 0.333
Element value in 1 row 2 columns of Respondent 8 = 0.333
Element value in 1 row 2 columns of Respondent 9 = 0.333
Element value in 1 row 2 columns of Respondent 10 = 1.

Therefore, the geometric mean of the four values is

$$\sqrt[10]{0.333 \times 1 \times 1 \times 0.333 \times 1 \times 1 \times 0.333 \times 0.333 \times 0.333 \times 1} = 0.577.$$

The maximum eigenvalue of Table 5 is 60.682. If this value is substituted in formula (2), the CI is 2.991 and the consistency rate when substituted in formula (4) is 2.007% (less than 10%). This means that the group maintains consistency in the evaluation.

V. CONCLUSION

The defensive methods developed so far were mostly security devices, and as the attacking methods of invasion cases have been diversified, new forms of threats and those from insiders have become difficult to counteract. The actions held within the inside system are gathered in event logs as log data, and it is possible to analyze the event logs to judge whether there are security problems in the system.

Therefore, through this study, we could determine the applicability of using the list of important event logs to prevent the security threats conducted by insiders with the validity analysis. Based on this result, the method can contribute to reducing and preventing the security threats by insiders, and we expect another study on the danger detection system by developing further scenarios and analyzing the patterns.

REFERENCES

[1] D. H. Lee and K.H. Choi, "A study of an anomalous event detection using white-list on control networks," *Journal of Convergence Security*, vol. 12. no. 4, pp. 77-84, 2012.

[2] S. Cheung, B. Dutertre, M. Fong, U. Lindqvist, K. Skinner, and A. Valdes, "Using model-based intrusion detection for scada networks," in *Proceedings of the SCADA Security Scientific Symposium*, Miami Beach: FL, pp. 1-7, 2007.

[3] A. Valdes and S. Cheung, "Communication pattern anomaly detection in process control systems technologies for homeland security," in *Proceedings of IEEE International Conference on Technologies for Homeland Security*, Boston: MA, pp. 22-29, 2009. DOI: 10.1109/THS.2009.5168010.

[4] S. Parthasarathy and D. Kundur, "Bloom filter based intrusion detection for smart grid SCADA," in *Proceedings of IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, Montreal: QC, pp. 1-6, 2012. DOI: 10.1109/CCECE.2012.6334816.

[5] M. Naedele, D. Dzung and M. Stanimirov, "Network security for substation automation systems", in *Proceedings of the 20th International Conference on Computer Safety, Reliability and Security*, Berlin: Heidelberg, pp. 25-34, 2001.

[6] J. Hoyos, M. Dehus and T. X. Brown, "Exploiting the GOOSE protocol: A practical attack on Cyber-infrastructure," in *Proceedings of IEEE Globecom Workshops*, Anaheim: CA, pp. 1508-1513, 2012. DOI: 10.1109/GLOCOMW.2012.6477809.

[7] M. Kalech, "Cyber-attack detection in SCADA systems using temporal pattern recognition techniques," *Journal of Computers & Security*, vol. 84, pp. 225-238, 2019. DOI: 10.1016/j.cose.2019.03.007.

[8] T. Onoda, "Machine learning based intrusion detection in control system communication," *Design and Analysis of Distributed Energy Management Systems*, pp. 167-202, 2020, DOI: 10.1007/978-3-030-33672-1_9.

[9] M. Minasi, D. Gibson, A. Finn and B. Henry, *Mastering Windows Server 2008 R2*, Indianapolis, IN, Wiley, p. 921, 2010.

[10] National Security Agency, Spotting the Adversary with Windows Event Log Monitoring, [Internet], Available: https://apps.nsa.gov/iaarchive/library/ia-guidance/security-configuration/applications/spotting-the-adversary-with-windows-event-log-monitoring.cfm/.

[11] S.T. Ung, "The development of safety and security assessment techniques and their application to port operations," Ph.D. Thesis, School of Engineering, Liverpool John Moores University, UK, 2007.

[12] S.W. Kim, A. Wall and J. Wang, "Application of AHP to fire safety based decision making of a passenger ship," *Opsearch*, vol. 45, No. 3, pp. 249–262, 2017. DOI: 10.1007/BF03398817.

[13] T.L. Saaty, "Decision making with the analytic hierarchy process," *International Journal of Services Sciences*, vol. 1. no. 1, pp. 83-98, 2008. DOI: 10.1007/BF03398817.

Jongmin Kim
received the Ph.D. degree in Industrial Security from the Kyonggi University, Korea, in 2015. He is currently a professor of Convergence Security Department, Kyonggi University, Korea. His research interests are Industrial Security, Malware Prediction, Convergence Security, and Information Security.

Jiwon Kang
received M.S. degree in Computer Science (Information Security) from Yonsei University and Ph.D. degree in Information Security from Kyonggi University, Korea. He is currently university-industry collaboration professor of Computer Engineering Department in Sejong University, Korea. His research areas include Cyber-space Operations in Defense and Convergence Security.

DongHwi Lee
received Ph.D degree in Information Security from Kyonggi University, Korea, in 2001 and 2007. He was a research scholar at the University of Colorado Denver in 2010 and 2011, USA. He is currently a Professor at DongShin University. His research areas include Information Security and Control System Security.