Indonesian Journal of Electrical Engineering and Computer Science
Vol. 17, No. 3, March 2020, pp. 1474~1479
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v17.i3.pp1474-1479

New insider threat detection method based on recurrent neural networks

Mohammed Nasser Al-Mhiqani, Rabiah Ahmad, Zaheera Zainal Abidin, Warusia Yassin, Aslinda Hassan, Ameera Natasha Mohammad
Information Security and Networking Research Group (InFORSNET),
Center for Advanced Computing Technology, Faculty of Information Communication Technology,
Universiti Teknikal Malaysia Melaka, Malaysia

Article Info

Article history:

Received Jul 8, 2019
Revised Sep 10, 2019
Accepted Sep 26, 2019

Keywords:

Cyber security
Deep learning
Gated recurrent network
Insider
Insider threat

ABSTRACT

Insider threat is a significant challenge in cybersecurity. In comparison with outside attackers, inside attackers have more privileges and legitimate access to information and facilities that can cause considerable damage to an organization. Most organizations that implement traditional cybersecurity techniques, such as intrusion detection systems, fail to detect insider threats given the lack of extensive knowledge on insider behavior patterns. However, a sophisticated method is necessary for an in-depth understanding of insider activities that the insider performs in the organization. In this study, we propose a new conceptual method for insider threat detection on the basis of the behaviors of an insider. In addition, gated recurrent unit neural network will be explored further to enhance the insider threat detector. This method will identify the optimal behavioral pattern of insider actions.

Corresponding Author:

Mohammed Nasser Al-Mhiqani,
Information Security and Networking,
Research Group (InFORSNET),
Universiti Teknikal Malaysia Melaka,
Melaka, Malaysia.
Email: almohaiqny@gmail.com

1. INTRODUCTION

Insider threat have been a critical threat source for an organization given their increased access and opportunity that can cause considerable damage to the organization. In comparison with outsiders, insiders have more privileged and legitimate access to information and facilities. Moreover, insiders are knowledgeable about an organization and its critical assets. With the additional knowledge of insiders, conducting an attack is easy for these insiders because they can hide their hacking trail/activities [1-3]. Surprisingly, 2018 insider threat reports have shown that 53% of threats come from within an organization in the last 12 months [4]. Moreover, 27% of surveyed organizations have stated that attacks originate from inside [4]. Thus, most organizations that implement cybersecurity techniques, such as intrusion detection, firewall, and electronic access system, aim to protect data not only from outside threats but also from insider threats [5]. In the last decades, many incidents of insider threats have gradually reached the media; for example, well-known cases of data leakage have been conducted by Edward Snowden, Daniel Ellsberg, and Chelsea Manning [6]. In contrast to the threats by outsiders, insider threats are easy to perform with no experience or advanced technical knowledge required given the authorization access that insiders have and the knowledge of the vulnerabilities of business processes and deployed systems. In comparison with outsiders whose hacking trails are hard to hide, malicious insiders are difficult to detect [6, 7].

Recurrent neural network (RNN) considers current value and previous input, thereby making this algorithm different from other neural networks. Therefore, RNN has been extensively used for solving the

problems of input order, such as issues on natural language processing [8, 9]. In the present study, we propose a conceptual method for insider threat detection on the basis of the behaviors of an insider. In addition, gated recurrent unit (GRU) neural network is explored further as a method for enhancing insider threat detection. Similar to long short-term memory (LSTM) approach, GRUs can capture long-term temporal dependencies on the sequence of user actions well considering the hidden units that GRU use to record temporal behavior patterns with simple-structure GRUs, thus saving additional computing resources and training times.

2. RELATED WORK

In recent years, cyber security has become a matter for societal, infrastructures and economic to every country in the world due to the tremendous number of electronic devices that are interconnected via networks communication [10-12, 13-15], one of the cyber security problems is the problem of insider threat that have been increased, thus attracting considerable attention from researchers in the field of insider threats [7]. Considerable studies have been conducted in this field. The works in [16, 17] proposed an approach to insider threat detection by applying the hidden Markov model (HMM). These studies used the HMM in modeling the normal behavior of users to detect any anomalous behaviors that may deviate from the norm. By using the HMM, the number of states has an enhanced impact on the effectiveness of a method. However, an increment in the number of states increases the computational cost of the HMM.

Machine learning techniques have powerful capabilities of detection pereformance improvement [18-20] and self-adaptive abilities to handle the changes in the insider threat environment; these technologies are still affected by the impact of imbalanced data and lack of extensive knowledge on insider behavior patterns [21]. For example, to model the daily log time series, the work in [22] suggested one-class support vector machine (OCSVM), which conceptualizes the detection of an insider threat problem as a stream mining issue and demonstrates higher accuracy and lower false positives than traditional OCSVM.

Recently, deep learning and RNN approaches have been applied in the field of insider threat detection, the proposed work in [23] utilized deep neural networks and RNNs to detect an insider; these neural networks are trained to recognize activities that are characteristic of every single user in the network and simultaneously assess whether the behavior of the user is normal or anomalous in real time.

Similar to our proposed solution with a different technique, the work in [21]utilized the LSTM to model the insider activity log as a sequence of natural language, wherein the model extracts feature and detects anomalies when log patterns deviate from their trained models. The evaluation of the proposed model was based on a limited number of users, wherein eight users are randomly selected as a group from the experimental dataset.

3. PROPOSED METHOD

In this section, details of the proposed insider threat detection method will be discussed. This method utilizes the GRU to model an insider activity as a sequence that is similar to the natural language sequence. GRUs are selected in this study considering their simplicity and rapid training phase over the LSTM.

3.1. Log Files

All the activities performed by employees in an organization which are the events that come from many different sources, such as Logon event, device event, HTTP event, file event, system call, and email event logs.

3.2. Data Processing

During this stage all the operations will be collected and extracted for every user from multiple source files. All the operational data will then be organized into a sequence on the basis of the individual user's daily actions. Similar to the modeling of the natural language, the action corresponds to the word, and the sequence of actions corresponds to a sentence. Thus, every user will have a list of actions performed on each day. Finally, when the log of a user data sequence of actions is inputted into the GRU classifier model, each process must be converted to a one-hot vector. Proposed method as shown in Figure 1.

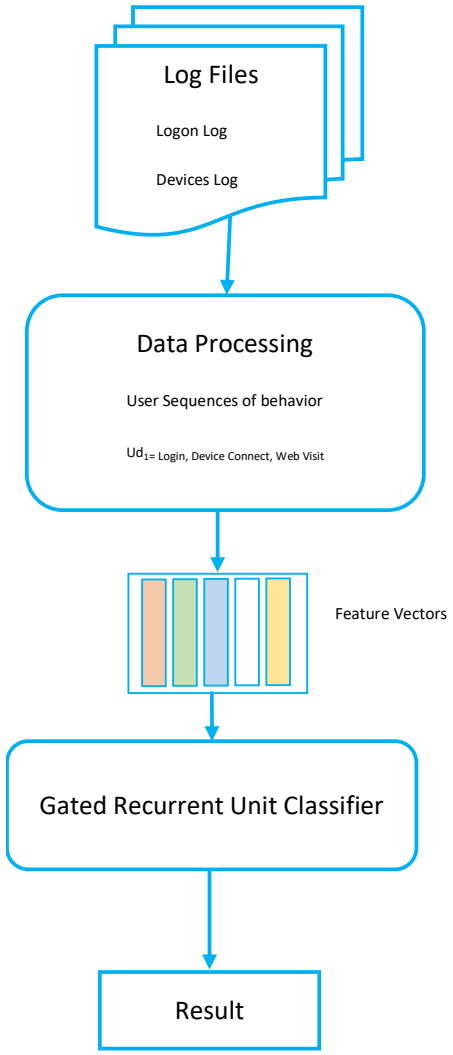

Figure 1. Proposed method

3.3. Anomaly Detection

User activity is difficult to identify as a malicious insider action given the complexity of the insider threat problem. This section discusses an insider anomaly detection technique on the basis of the features of user action sequences for every user. GRU, a variant of the LSTM that was introduced in [24, 25], is the anomaly detection technique used in this study. This technique is similar to the LSTM but without the output gate. Therefore, GRUs fully write the contents from their memory cell to the large net at every time step. The internal structure of the GRU is simple, thereby accelerating training because few computations are required to update the hidden state of GRUs. Gated recurrent unit structure as shown in Figure 2.

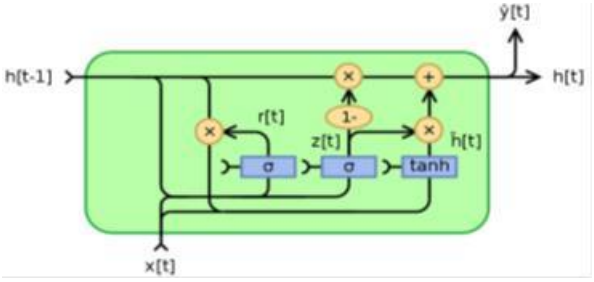

Figure 2. Gated recurrent unit structure

3.4. Experiment

In order to evaluate the performance of the proposed mothod, We run our proposed model on the public insider threat dataset CERT v4.2, the cert v4.2 has been chosen because this version contains more insider threats instances compared to the previous version of cert datasets. Over the period of 17 months 32,770,227 log lines have been generated by 1000 users and among these logs there 7323 anomalous activities which was injected manually by the domain expert to represent the insider threats malicious scenarios that was described by CERT. Detailed data description as shon in Table 1.

Table 1. Detailed Data Description from the Files of the r4.2 Dataset

File	Features
Logon	ID DATE USER PC ACTIVITY
File	ID DATE USER PC FILENAME CONTENT
Device	ID DATE USER PC ACTIVITY
HTTP	ID DATE USER PC URL CONTENT
Email	ID DATE USER PC TO CC BCC FROM SIZE ATTACHMENT_COUNT CONTENT
Psychometric	EMPLOYEE_NAME USER_ID O C E A N
LDAP	EMPLOYEE_NAME USER_ID E-MAIL ROLE PROJECTS BUSINESS UNIT FUNCTIONAL_UNIT DEPARTMENT TEAM SUPERVISOR

3.5. Initial Results

This subsection is the last stage to testing whether an insider conduct can be considered a malicious act or not. The corresponding logs sequences are extracted and sent to the result model, which then outputs

the sequence's classification result on the GRU. The evaluation of the proposed model shows that the model can successfully classify most of the insider with a good accuracy up to 0.92% when it was excuted with 20 epochs and the loss value around 0.29. as shown in Figure 3.

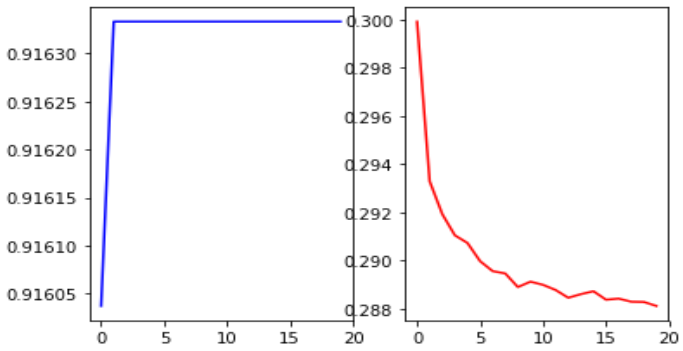

Figure 3. Accurcy and loss result

4. CONCLUSION

Insider threat is a dangerous security threat. The issue of insider threats has become a primary concern for enterprises of all sizes. This study proposed a conceptual method for insider threat detection. The log file stages collect multiple logs from different sources. The preprocessing stage organizes the logs into a sequence on the basis of an individual user's daily actions. Furthermore, GRU is suggested for insider threat detection. Finally, the result stage will output the results of the classification model. Therefore, the proposed method solution will help detect malicious behavior inside an organization. Future works will implement the proposed method on real and public datasets and combined with a modified adaptive synthetic oversampling technique (ADASYN) algorithm to handle imbalanced of the insider threats datasets.

ACKNOWLEDGEMENTS

This project is funded by the Ministry of Higher Education Malaysia under Transdisciplinary Research Grant Scheme (TRGS) with project Number TRGS/1/2016/UTEM/01/3. And this project referred as TRGS/1/2016/FTMK-CACT/01/D00006 at UNIVERSITI TEKNIKAL MALAYSIA MELAKA

REFERENCES

[1] Y. Hashem, H. Takabi, M. GhasemiGol, and R. Dantu, "Towards insider threat detection using psychophysiological signals," in 7th ACM CCS International Workshop on Managing Insider Security Threats (MIST'15), 2015, pp. 71–74.
[2] N. Nguyen, P. Reiher, and G. H. Kuenning, "Detecting insider threats by monitoring system call activity," in *Information Assurance Workshop, 2003. IEEE Systems, Man and Cybernetics Society*, 2003, pp. 45–52.
[3] C. Colwill, "Human factors in information security: The insider threat - Who can you trust these days?," *Inf. Secur. Tech. Rep.*, vol. 14, no. 4, pp. 186–196, 2009.
[4] C. R. P. and C. security Insider, "Insider threat Threats 2018 Report," 2018.
[5] M. L. Collins *et al.*, "Common sense guide to mitigating insider threats 5th edition," CERT, Software Engineering Institute, Carnegie Mellon University, 2016.
[6] I. Homoliak, F. Toffalini, J. Guarnizo, Y. Elovici, and M. Ochoa, "Insight into Insiders: A Survey of Insider Threat Taxonomies, Analysis, Modeling, and Countermeasures," 2018.
[7] M. N. Al-Mhiqani *et al.*, "A new taxonomy of insider threats : an initial step in understanding authorised attack," *Int. J. Inf. Syst. Manag.*, vol. 1, no. 4, 2018.
[8] M. Kim, K. Kim, and H. Lee, "Development trend of insider anomaly detection system," *Int. Conf. Adv. Commun. Technol. ICACT*, vol. 2018-Febru, pp. 373–376, 2018.
[9] P. Sharma, D. Saini, and A. Saxena, "Fault Detection and Classification in Transmission Line Using Wavelet Transform and ANN," *Bull. Electr. Eng. Informatics*, vol. 5, no. 3, p. 284~295, 2016.
[10] M. N. Al-Mhiqani, R. Ahmad, K. H. Abdulkareem, and N. S. Ali, "Investigation study of Cyber-Physical Systems: Characteristics, application domains, and security challenges," *ARPN J. Eng. Appl. Sci.*, vol. 12, no. 22, pp. 6557–6567, 2017.
[11] A. Ahmed, K. A. Bakar, and M. I. Channa, "Countering Node Misbehavior Attacks Using Trust Based Secure Routing Protocol," *TELKOMNIKA*, vol. 13, no. 1, pp. 260–268, 2015.

[12] M. N. Al-mhiqani *et al*., "Cyber-Security Incidents : A Review Cases in Cyber-Physical Systems," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 1, 2018.
[13] M. N. Dazahra, F. Elmariami, A. Belfqih, and J. Boukherouaa, "A Defense-in-depth Cybersecurity for Smart Substations," *Int. J. Electr. Comput. Eng*., vol. 8, no. 6, pp. 4423–4431, 2018.
[14] A. M. Riyad, M. S. I. Ahmed, and R. L. R. Khan, "An adaptive distributed intrusion detection system architecture using multi agents," *Int. J. Electr. Comput. Eng*., vol. 9, no. 6, pp. 4951–4960, 2019.
[15] S. Norussaadah, M. Salleh, R. Din, N. H. Zakaria, and A. Mustapha, "A Review on Structured Scheme Representation on Data Security Application," *Indones. J. Electr. Eng. Comput. Sci*., vol. 11, no. 2, pp. 733–739, 2018.
[16] T. Rashid, I. Agrafiotis, and J. R. C. Nurse, "A new take on detecting insider threats: Exploring the use of Hidden Markov Models," in *MIST '16 Proceedings of the 8th ACM CCS International Workshop on Managing Insider Security Threats Pages 47-56*, 2016, pp. 47–56.
[17] C. Wang, G. Zhang, and L. Liu, "A Detection Method for the Resource Misuses in Information Systems," in *Affective Computing And Intelligent Interaction*, 2012, vol. 137, p. 545+.
[18] Z. Alfikri and A. Purwarianti, "Detailed Analysis of Extrinsic Plagiarism Detection System Using Machine Learning Approach (Naive Bayes and SVM) Detailed Analysis of Extrinsic Plagiarism Detection System Using Machine Learning Approach (Naive Bayes and SVM)," *TELKOMNIKA*, vol. 12, no. July 2016, pp. 7794 ~ 7804, 2014.
[19] H. Wang and R. Ma, "Design of Neural Networks for Intrusion Detection 1," *TELKOMNIKA*, vol. 14, no. 3, pp. 321–325, 2016.
[20] A. Jishan, K. R. Mahmud, A. Kalam, and A. Azad, "Natural language description of images using hybrid recurrent neural network," *Int. J. Electr. Comput. Eng*., vol. 9, no. 4, pp. 2932–2940, 2019.
[21] D. Zhang *et al.*, "Role-based Log Analysis Applying Deep Learning for Insider Threat Detection," *Proc. 1st Work. Secur. Des. Comput. Archit. Process.*, pp. 18–20, 2018.
[22] P. Parveen, Z. R. Weger, B. Thuraisingham, K. Hamlen, and L. Khan, "*Supervised learning for insider threat detection using stream mining*," in 23rd IEEE International Conference on Tools with Artificial Intelligence Supervised, 2011, pp. 1032–1039.
[23] A. Tuor, S. Kaplan, B. Hutchinson, N. Nichols, and S. Robinson, "Deep Learning for Unsupervised Insider Threat Detection in Structured Cybersecurity Data Streams," *arXiv Prepr.*, 2017.
[24] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," *CoRR*, vol. abs/1412.3, pp. 119–124, 2014.
[25] K. Cho, B. van Merrienboer, C. Gulc¸ehre, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *CoRR*, vol. CoRR, pp. 1724–1734, 2014.