A study on Classification of Insider threat using Markov Chain Model

Dong-Wook Kim[1], Sung-Sam Hong[1] and Myung-Mook Han[1]
[1] Department of Computer Engineering, University of Gachon
Seongnam-Si, Korea
[e-mail: kog7306@naver.com, sungsamhong0@gachon.ac.kr, mmhan@gachon.ac.kr]
*Corresponding author: Myung-Mook Han

Received October 4, 2017; revised January 26, 2017; accepted March 5, 2017; published April 30, 2018

Abstract

In this paper, a method to classify insider threat activity is introduced. The internal threats help detecting anomalous activity in the procedure performed by the user in an organization. When an anomalous value deviating from the overall behavior is displayed, we consider it as an inside threat for classification as an inside intimidator. To solve the situation, Markov Chain Model is employed. The Markov Chain Model shows the next state value through an arbitrary variable affected by the previous event. Similarly, the current activity can also be predicted based on the previous activity for the insider threat activity. A method was studied where the change items for such state are defined by a transition probability, and classified as detection of anomaly of the inside threat through values for a probability variable. We use the properties of the Markov chains to list the behavior of the user over time and to classify which state they belong to. Sequential data sets were generated according to the influence of n occurrences of Markov attribute and classified by machine learning algorithm. In the experiment, only 15% of the Cert: insider threat dataset was applied, and the result was 97% accuracy except for NaiveBayes. As a result of our research, it was confirmed that the Markov Chain Model can classify insider threats and can be fully utilized for user behavior classification.

Keywords: Insider threat, Markov Chain, Classification,

A preliminary version of this paper was presented at APIC-IST 2017, and was selected as an outstanding paper. This research was supported by a research grant from NRF(National Research Foundation) NRF-2015R1D1A1A01060874

http://doi.org/10.3837/tiis.2018.04.027 **ISSN : 1976-7277**

1. Introduction

For insider's activity, employee's behavioral tasks can be identified by using orgsnization system resources. Inside employees show an activity pattern and behavior list according to their own task assignment. In such behavior, the inside employee intentionally leaks information within the limited system inside the organization, the information leakage is realized by an error without their awareness. For the subject for such activitys, inside threat types within the organization can be classified. The types can be classified as workers and retirees, or can also be divided into the case of stealing organization information for turnover and the case of resisting against the organization due to dissatisfation with the organization in terms of detail. In such typem, the inside information leakage may pertain to the partner subjects related to the organization rather than employees. In addition, to prevent indiscreet leakage of inside information, protection of sensitive data and documents is being realized through authority asignment and use policies of the inhouse employees within the organization.

The inside threats are achieving the objectives through the attribute of being well aware of the inside system within the organization, while the inside subjects are doing so through the social engineering methods or the methods of byassing the system[1][10]. Socially, the insider threats are not limited within the organization such as business in terms of social engineering threats. Intelligent Advanced Persistent Threat(APT) as one of the cyber attacks on a national level also pertains to this case.As in APT attacks, the inside threats continue to attempt to achieve the objecttives in the form of intelligence by ceaseless reconnaissance and finding vulnerabilities.

In the present paper, an analytical study has been implemented on insiders' activitys to prevent information leakage under inside threats. For the activity analysis, changes in the operation processes for employees's task states were analyzed by application of the Markov chain model. In the Markov chain model, state transition is realized as a function of probability variables, and the probabilities for the state transition can be obtained by identification of the corresponding state frequencies. Through such transition probabilities, threatening activitys and normal users are classified for the insider activitys, and the processes for insider activitys are understood. For the data set employed for the present study, the data provided by Cert insider threat Center of Carnegie Mellon's Software Engineering was used, and was checked through the classification algorithms of SVM, NaiveBayes, Multilayerperceptron, and RandomForest to perform clasification. According to the experimental results. Generally, more than 90% of the clasification was affirmed to be realized for the normal activitys and the threatening activitys depending on the configured data set.

2. Related Work

For the studies on insiders' attacking activitys or invading activitys, studies have been performedby such methods as signature-based filtering, blockage and detection that is fundamentally based on rules. However, new attack patterns of diversified forms are occurring today due to the complexity of an orgnization and the changes of systems. Such changes are being made to a study of prediction from the standpoint of defenders or detection, for the studies are being carried out based on the intelligent methods by way of predicting the data convergence technology for the system and the system changes as well as understnding on insiders. In the present chapter, related studies to revent such insider threats are introduced.

2.1 Understanding the Insider Threat[1]

These are the paper for understanding of insider threat beginning with the fundamental studies on the insider threat. The US information institution of Advanced Research and Development Activity (ARDA) is conducting system studies for reliance on the sensitive information such as insider threat. Under the theme of "Understanding the Insider Threat", observable activitys for the vulnerabilities for insider threat have been classified as in **Fig. 1**[1].

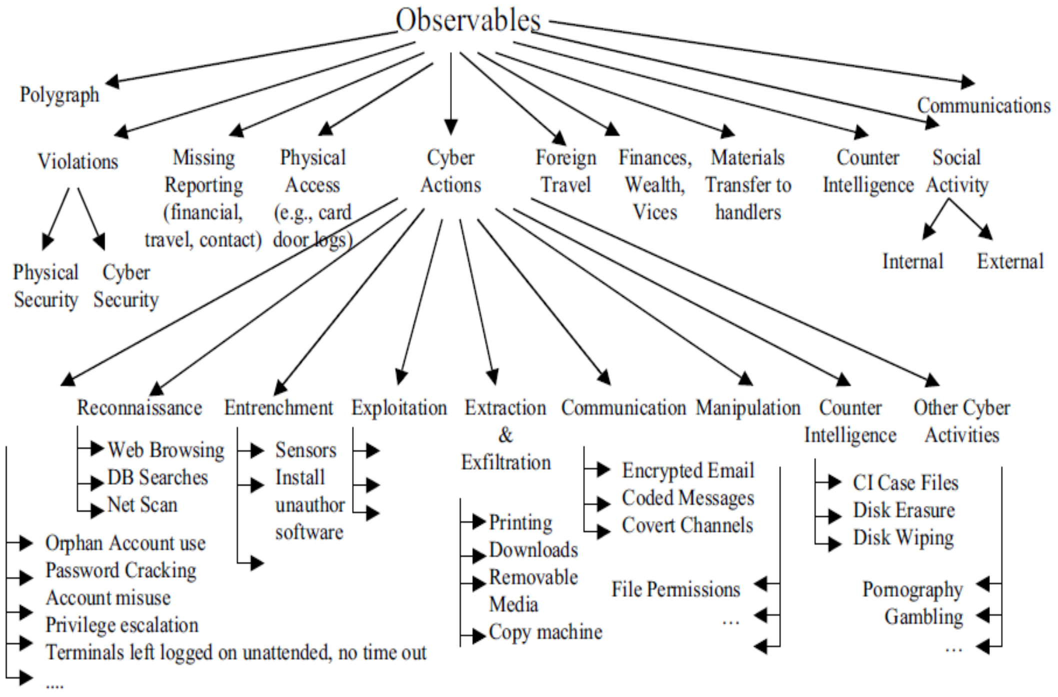

Fig. 1. Taxonomy of Observables

This classification helps determining anomalous behavior by tracking anomalous activitys of the insiders in the list. Using this classification, a discussion was made on the problems including diversified vulnerabilities of the IC system model deriving scenarios concerning insiders' threat and the characteristics of the events related to insider attacks in the attacker model. This shows attacking methods from internal threats, scenarios, precondtions, and observable elements.

2.2 A Comparison of System Call Feature Representations for Insider Threat Detection[4]

Detection methods for insider threat concerning inappropriate accessing activity, distributing activity of sensitive information, and user activity damaging the information system have been studied with the detection technology being focused on the insider threat[4]. Since such detection is realized through anomaly detection on the system calling level, the functions based on system calling have been expressed from three kinds of viewpoints. These three kinds of expression methods include N-gram system calling name, histogram system calling name, and individual system calling. To analyze the ability capable of detecting malicious insider activites, the data set used has generated a data set including normalness and malignancy for one user. Employed was the data of user activites monitored by using the tool of SNARE business for the study team of system operation recognition. In the normal activities, activities of interaction that use word processing, web browsing, and command prompt were included.

In the malignant activities,, acquisition of administrator (root) authority, file duplication and mail transmission of the protected files by mobile drive change of file extensions, encoding release, searching for system files, etc. are included. For the anomaly detection algorithm, abnormalities were detected by using the outlier detection algorithm based on KNN (K-nearest neighbor), which was supervised. According to the experimental results, the detection algorithm showed performance of a reliable level through ROC. This enabled affirmation that there was a clear difference between normal operations and malignant operations.

2.3 Detection of Anomalous Insiders in Collaborative Environments via Relational Analysis of Access Logs[5]

The study on the autonomous learning framework of the CADS (Community based Anomaly Detection System) was performed to detect insider threats based on the information recorded in the access log of CIS(Collaborative information systems)[5]. CIS refers to a collaboration system through general operation in user group communication and virtual environments, while CADS is a system for detection of potential unlawful activities of those having low affinity based on the community's affinity for the normal users. In the present paper, an anomaly detection model is proposed through extraction of Relational Patterns of CIS and CADS.

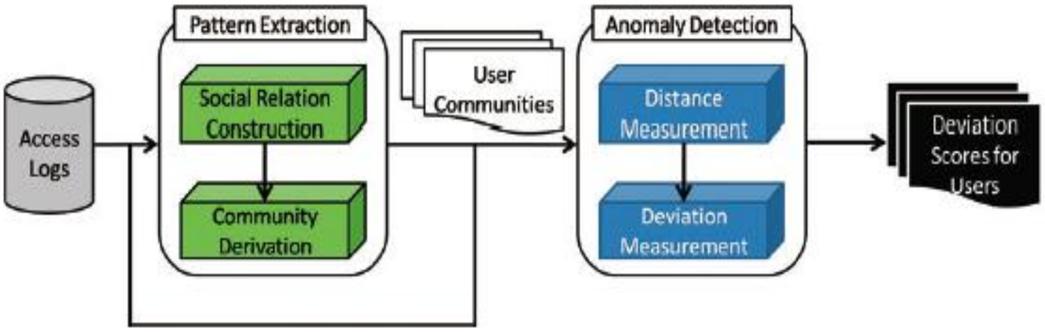

Fig. 2. Community based anomaly detection system overview

In general, the framework of CADS consists of 2 components including pattern extraction (CADS-PE) and anomaly dedetion (CADS-AD) as shown in **Fig. 2**. In the CADS-PE component, Access log is collected in the user's community CIS. Since the access has the character of tranjection, this needs conversion of data structure for inference of community relationships. CADS-PE is composed of a series of stages of the community pattern set, and the tranjection first captures the relationship between the user and the subject for mapping into the data structure where the structure is converted into the network relationship. Decomposition of the network is realized for the user community into the pattern spectrum by a probability model. For the user activity, a comparison between the community pattern and the Access log is realized in the CADS-AD component. Such comparative analysis algorithm is performed by using K-Nearest Neighbors algorithm and Principal component analysis (PCA) technique, and the framework which predicts the user with severe deviation from the normal pattern as the anomalous user is propsed as the detection model in the present paper. According to the experimental results, the discovery of anomalous users through deviation of CADS was found to be difficult when the number of user Access subjects simulated according to the data set was small.

3. Methodology for classification of user activity

In the present chapter, the user activitys are classified by using the data set provided by the insider threat response team of CERT(computer emergency response team)[7]. In the Markov Chain for classification, the user activitys as a function of time elapse can be measured. By reflecting the user's operation on a specified date, the user operations can be reflected for classification as to what state the relevant operation belongs to, allowing sensing of anomalous activitys[8]. Inside behavior change occurs by being affected by the previous events that occur for 1, 2, and 3 times through stochastic processes according to Markov attributes[9]. We analyze a series of the user's sequential defining the insider's activities as "n" states. In this methodology, the contents of the user 's behavior were studied in the related paper[1], and the information leakage path of the insider was grasped. The contents of the state transition of the Markov Chain Model are analyzed in [4] [5] and analyzed for insider activity to refer to data set generation for behavior classification.

3.1 Markov Chain Model

Contemporary probability theory studies the processes of events affecting prediction of the future. When a series of random experiments are observed, all results in the past can affect the next event. For instance, when a student's usual behavior is observed, there are examples allowing prediction of the student's temperament. However, proving what allows so many generalities is very difficult[5].

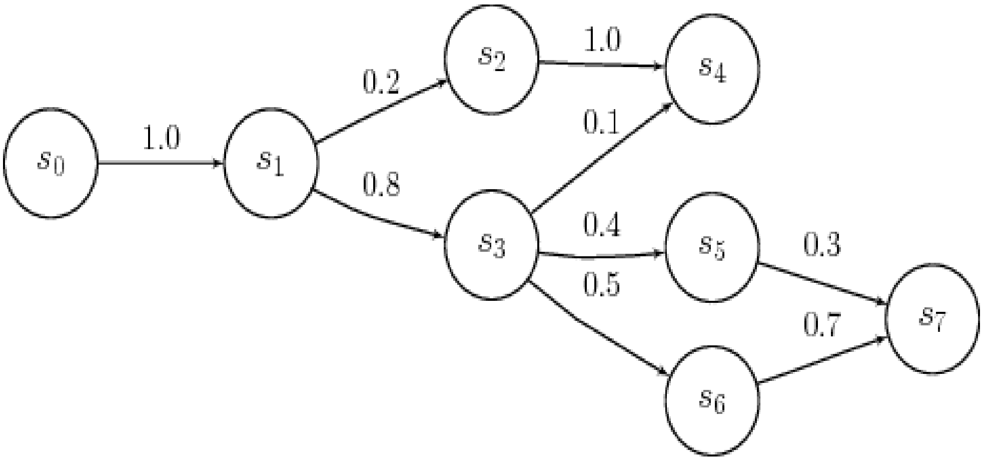

Fig. 3. State transition probability diagram

In 1907, Markov started a study on the chance processes for a new type, in which process he demonstrated that the results of a given experiment in this process could affect the results of the next experiment. The process of such type is referred to as a Markov Chain[6][8].
Markov Chain is a discrete stochastic process. The stochastic process means modelling that a rando variable having a probability distribution generates values at a given time interval. Among such models, the stochastic process where the current state is affected only by the

previous state is called MP, and transition of each state is realized at a discrete time. The time of dwelling in an arbitrary state belonging to a state set X is defined as a step. If the current step is called " n". then the next step can be described as " n+1".

Here, the value of state transition probability at this time can be shown as follows.

pij means the value of probability of transition from the state i to the state j. Xn means the state dwelling in the step "n", and more accurately, a random variable (r,v) for the relevant state.

$$p_{ij} = p(X_{n+1} = j | X_n = i)$$

$$\forall \mathrm{i}, \mathrm{j} \in \mathrm{X}$$

Importantly, the value of trnsition probability from the state i to the next state j is always the same irrespective of what state the previous visit was made in , when the state i is visited. Putting this as a formula, it becomes as follows.

$$p_{ij} = p(X_{n+1} | X_n = i_n, X_{n-1} = i_{n-1}, \ldots X_0 = i_0) = p(X_{n+1} | X_n) = i_n$$

Such characteristic (the charateristic that the probability of the next state is affected only by the previous state) is called Markovian Property. In addition, the condition for the transition probability is as follows.

$$p_{ij} \geq 0$$

$$\sum_{j \in X} p_{ij} = 1$$

$$\forall \mathrm{i} \in \mathrm{X}$$

3.2 Data Set

The research office for insider threat of CERT(computer emergency response team) is conducting a study, modelling, analysis programs to respond to the evoution of insider threat. The present data set has the database for insider threat based on the scenario that affects negative secret for malignant insider threat of the organiztion, or grant the access authority to the system and the data to intentionally provide to other businesses on the foundation of more than 700 cases of insider threat. Configuration of the data set consists of the folllowing as shown in **Table 1**[7].

As a description for the table, the use status of the mobile device can be checked in Device file, while the login status at work hours and in holiday, etc, can be checked in Logon file, and PC state can be seen by using Logoff depending on variations in operation time and Device file relationships. The http file allows checking for URL of the accessed domain, through which the visit status to malignant Web can be determined. In addition, the words included in URL are assumed to be the words highly associated with Web page, and the text anlaysis is presented to be implementable. In the email file, the amount of the mails sent in a day is recorded, also containing the record of mail recipients. In addition, employees and non-employees are divided for specification. In the File, the header including extensionof the file is configured, and normalness/anomaly concerning file duplication can be analyzed through data figures.

Table 1. Cert:insider Dataset Configuration

Domain	Features
Logon	id, date, user, pc, activityivity
Device	id, date, user, pc, activityivity
Email	id, date, user, pc ,to, cc, bcc, from, size ,attachments, content
File	id, date, user, pc, filename, content
http	id, date, user, pc, url, content

Description of data set.

- The device file allows checking for the use of mobile drive.
- Login status such as working time, holiday, etc. can be checked in the logon file. State of the PC can be seen through logoff of the PC depending on the working time and the device file relationship.
- The http file allows checking for URL of the connected domain, through which the malignant web can be visited. Words of the URL are associated with the web page.
- In the email file, the user records the amount of mails sent in a day, and the record for mail recipients is also included. Division of employees and non-employees is specified.
- The file consists of header including extension of the file, and the figure value of the file can be used for the analysis of normal values for file duplication.

Insider behavior data set has such charactivityeristics. The user records implementation situation from logon to logoff. In this operation, operations of each user are analyzed.

3.2 Insider behavior classification

In the present chapter, the methodology for classificatio of the inside threat activitys using Markov Chain Model is introduced. For the classification of inside behavior, the data set of 3.2 has been used. First, to extract the user’s activitys, attribute value activityivity of the dat set has been utilized. In the activityivity, the classification for the user’s activitys according to each domain is shown. For application of the Markov Chain, a series of the user’s activitys need to be listed. For this purpose, the operation of extracting the activitys of data set was first attempted.

Table 2. Status Definition

State	Description
S1	Start and end of operation
S2	State of performing Email operation
S3	State of performing http
S4	State of connecting device and instrument in Device operation
S5	State of performing File operation
S6	State of Disconnecting device and instrument in Device operation

Each user sequence has been generated with state definitions according to **Table 2** for the value extracting the activity, and the state transition probability was obtained subsequently. The sequence of one user is expressed as shown below by **Table 3**. In **Table 3**, "User" refers to the user name, and "Data" refers to the user during the day. "Sequence" is made up of a list of behaviors of **Table 2** during the day.

Table 3. User state sequence

Features	**Values**
User	ABC0174
Date	2010-01-02 07:37:00 ~ 2010-01-02 15:56:00
Sequence	s1,s3,s3,s3, ··· s4,s3,s5,s5,s3,s3,s3,s3,s3,s3,s3,s3,s3,s3,s3,s3,s3,s3,s6,s1

The transition frequency is obtained from the state listed as in **Table 3**, based on which the state transition determinant is calculated as follows.

Fig. 4 is the transition matrix of one dat according to ‘ABC0174’, and the case of state transition shown in **Table 2** is defined as 6 states, In S1, Login and Logoff are recorded as a pair, while S2 is the state of performing the activity of Email. The state of S3 represents the operation of performing the activity of http, while the staes of S4 and S6 define Connect and Disconnect that divided the activity of Device. Lastly, the state of S5defines the state of performing the operation of File. When marked as the state transition diagram for the user defined in the data set, an expression as shown below in **Fig. 5** is possible.

In the probability distribution of the transition matrix geneated as shown in **Fig. 5**, classification for the inside activitys is performed. To perform the classification and verify the classification, 10-fold cross-validation was applied to the training set, and checking was made through the classification algorithms of SVM, NaiveBayes, Multilayerperceptron, and RandomForest. An overall Flow Chart for propsed method is as shown below in **Fig. 6**.

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0.7333 & 0.2666 & 0 & 0 & 0 \\ 0.006 & 0.024 & 0.938 & 0.123 & 0.006 & 0.012 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 & 0 \end{bmatrix}$$

Fig. 4. State transition matrix

In summary, when one user data set is configured, sequnce is generated through the activity feature composed of **Table 3**. The generated Sequence constructs the probability distribution by applying the state transition probability to apply to the Markov chain model. Then, a process of classifying through the machine learning algorithm is performed to detect the abnormal behavior of the insider.

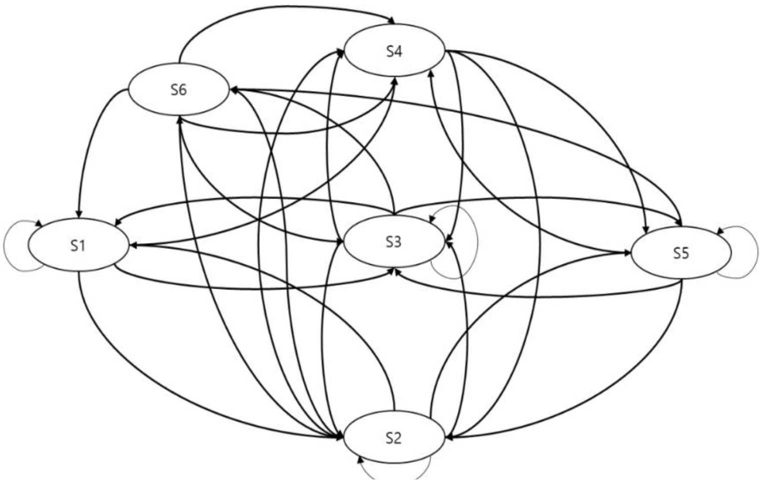

Fig. 5. User State Transition Diagram

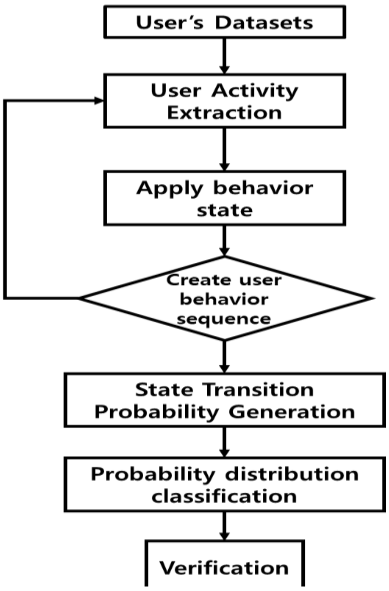
User's Datasets
User Activity Extraction
Apply behavior state
Create user behavior sequence
State Transition Probability Generation
Probability distribution classification
Verification

Fig. 6. A flowchart of the proposed method

4. Experiment

In the present experiment, classifiation for the insider activitys is performed. The classification was realized into normal user and threat user, the transition matrix was constructed by using CERT:insider threat data set, and a total of 4 classification algorithms of SVM, NaiveBayes, Multilayerperceptron, and RandomForest were excuted. For Tranining Data and Test Data, experiments were implemented by application of 10-fold cross-validation[12][13]. The whole users of the data set consisted of 1,000 people in total, of whom 30 people were defined as threats. In the makeup of the data set for the present experiment, 170 users were made as the subject by randomly selecting only about 15% of the whole users. Here, the transition probability for the 164 normal people and the 6 threats selected was obtained to conduct the experiment.

In the experimental resuts shown in **Table 4**, an accuracy higher than 97% was indicated excluding NaiveBayes classifier. The best result among them shown by RandomForest classification algorithm had SVM higher by 0.14%. Howeve, the lowest FP Rate was evaluated by NaivieBayes, after which RandomForest was satisfactory. Even when considered in terms of overall Precison and Recall ratio, the RandomForest classification algorithm shows the highest results.

Table 4. Experiment results

Classification	Naïve Bayes	SVM	Multilayer perceptron	Random Forest
Accuracy(%)	83.33%	97.68%	97.58%	97.82%
FP Rate	0.306	0.977	0.936	0.826
Precision	0.97	0.954	0.963	0.972
Recall	0.833	0.977	0.976	0.978
F-Measure	0.89	0.965	0.967	0.972

When the normal users ans the insider threats were classified by each classification algorithm, RandomForest algorithm can be seen to show a good accuracy for classification of the normal users, while NaiveBayes classification algorithm can be affirmed to show good results when the inside threat are classified. The reason for producing such result is affirmed when the state frequencies of normal users and inside threat shown in the data set are compared. Considering **Fig. 7**, it may be identified that the inside threateners ewere realizing more activitys excluding the state of S1. In the case of S1, transition to a different state can be predicted as the frequencies of Logon and Logoff were high, and the inside threat are identified to conduct more activitys than the normal users for the remaining states. Namely, as a result, the number of data for than activitys of the inside threateners in the data set can be induced to be small. However, in terms of classification accuracy, classification can also be seen sufficiently possible through the Markov Chain Model.

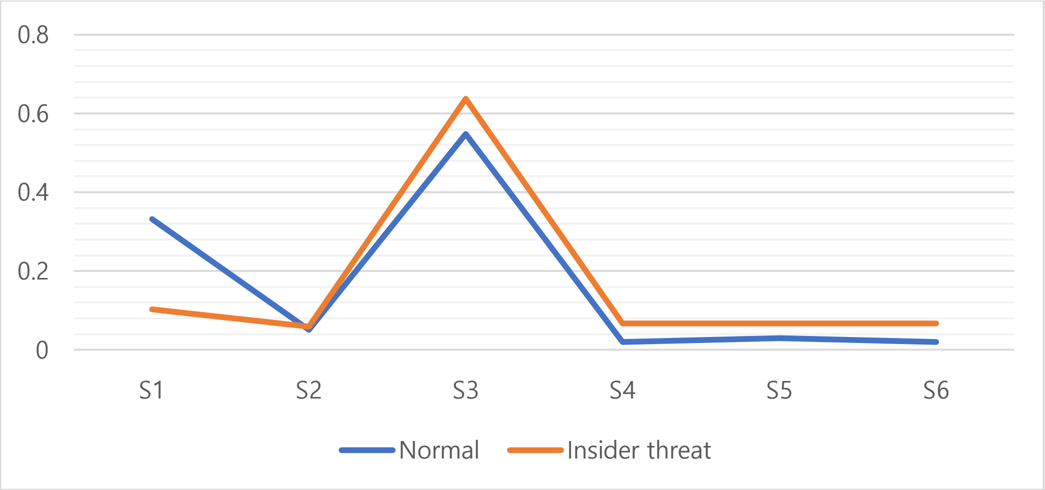

Fig. 7. Probability distribution according to frequency

5. Conclusion

We have carried out a study to classify insider threats through the insider activitys. To detect anomalous activitys for the inside threateners and classify the same, Markov Chain Model has been applied in the present this paper. What could be seen through the present paper is[11][14]; first, classification of inside threatening activitys through state transition probabilities of Markov Chain was possible, and secondly, a direction of improving the characteristics for RandomForest algorithm and NaiveBayes algorithm was needed in the classification algorithms. Thirdly, the data ratio for the inside threateners should be expanded for more accurate classification in the current data set. Based on the consideration of this point , the accuracy for the insider activitys will be improved further, and necessary expansion will be made in the future by the application of related skills to utilize the features that could not be used due to configuration in a text format in the data set.

References

[1] Anderson, Robert H., and Richard Brackney, "Understanding the insider threat," in *Proc. of a March 2004 Workshop,* 2004. https://www.rand.org/pubs/conf_proceedings/CF196.html

[2] Eldardiry, Hoda, et al., "Multi-domain information fusion for insider threat detection," *Security and Privacy Workshops (SPW)*, 2013 IEEE. IEEE, p. 45-51. 2013. Article (CrossRef Link)

[3] Malek Ben Salem, Shlomo Hershkop, Salvatore J. Stolfo, "A Survey of Insider Attack Detection Research," *Insider Attack and Cyber Security Advances in Information Security*, 2008 Article (CrossRef Link)

[4] Liu, A., et al. "A comparison of system call feature for insider threat detection," in *Proc. of the 6th Annual IEEE Systems, Man & Cybernetics, Information Assurance Workshop*. p. 341-347. 2005. Article (CrossRef Link)

[5] Chen, You, and Bradley Malin, "Detection of anomalous insiC in collaborative environments via relational analysis of access logs," in *Proc. of the first ACM conference on Data and application security and privacy*. ACM, p. 63-74. 2011. Article (CrossRef Link)

[6] Grinstead, Charles Miller, and James Laurie Snell. "Introduction to probability." *American Mathematical Soc*., p.405-469. 2012.

[7] http://www.cert.org/insider-threat/tools/index.cfm

[8] Eberle, William, Jeffrey Graves, and Lawrence Holder, "Insider threat detection using a graph-based approach." *Journal of Applied Security Research* 6.1 p32-81. 2010. Article (CrossRef Link)

[9] Wen-Hua Ju and Yehuda Vardi, "A hybrid high-order markov chain model for computer intrusion detection," *Journal of Computational and Graphical Statistics*, June, p 277-295, 2001. Article (CrossRef Link)

[10] Dawn M. Cappelli, Andrew P. Moore, Randall F. Trzeciak, "The CERT Guide to Insider Threats: How to Prevent, Detect, and Respond to Information Technology Crimes (Theft, Sabotage, Fraud)," *Addison-Wesley Professional*, 2012. https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=30310

[11] Y. Liao and V. R. Vemuri, "Using Text Categorization Techniques for Intrusion Detection," *11 USENIX Security Symposium*, 2002. https://dl.acm.org/citation.cfm?id=720290

[12] Cortes, C., Vapnik, V., "Support-vector networks," *Machine Learning,* 20 (3): 273, 1995. Article (CrossRef Link)

[13] Press, William H., Teukolsky, Saul A., Vetterling, William T., Flannery, B. P. Section 16.5. Support Vector Machines. Numerical Recipes: The Art of Scientific Computing *3 Edition. New York: Cambridge University Press*. 2007.

[14] STOLFO, Salvatore J., et al., "A comparative evaluation of two algorithms for windows registry anomaly detection," *Journal of Computer Security*, 13.4: 659-693. 2005.
Article (CrossRef Link)

Dong-Wok Kim received the Bachelor degree in Computer Software from Gachon University, Korea in 2015 and Master degree Computer Engineering from Gachon University, Korea in 2017. He is currently a Ph.D. candidate in the Department of Computer Engineering, Gachon University, Korea. His research interests include Insider Threat, Information Security, Data Mining, Machine Learning.

Sung-Sam Hong was born in Seoul, Korea, in 1983. He received the Master degree Computer Science from Gachon University, Korea in 2011 and the Ph.D degree Computer Science from Gachon University, Korea in 2016. He is a researcher professor in the Department of Computer Engineering in Gachon University, Korea. His research interests include Multimedia Security, Intelligent Information Security, Machine Learning, Cryptology, Data Mining, Big Data.

Myung-Mook Han received MS degree in computer science from New York Institute of Technology in 1987 and Ph.D. degree in information engineering from Osaka City University in 1997, respectively. From 2004 to 2005, he was a visiting professor at Georgia Tech Information Security Center(GTISC), Georgia Institute of Technology. Currently, He is a professor in the Department of Computer Engineering, Gachon University, Korea. His research interests include Information Security, Intelligent System, Data Mining, Big Data.