Journal of Network Communications and Emerging Technologies (JNCET) www.jncet.org
Volume 5, Special Issue 2, December (2015)

Network Security Based on Rough Set: A Review

Pritam Paul
Dept of CA, JIS College of Engineering.

Urmimala Dey
Dept of CA, JIS College of Engineering

Payel Roy
Dept of CA, JIS College of Engineering.

Srijan Goswami
Dept of IT, Institute of Computer Engineers (ICEI), Kolkata.

Abstract- As the conventional network security evaluation methods have prejudiced aspects when the weights of an evaluation indexes are recognized, it is not easy to make precise and objective evaluation. However, the Rough set theory has the benefits of not needing apriority information when dealing with tentative problems. Therefore, the appliance of the Rough set theory in the network security evaluation is quite essential. This paper mainly includes the recent network security research based on the rough set theory.

Index Terms– Network Security, Rough Set.

This paper is presented at International Conference on Recent Trends in Computer and information Technology Research on 25th& 26th September (2015) conducted by B. S. Anangpuria Institute of Technology & Management, Village-Alampur, Ballabgarh-Sohna Road, Faridabad.

1. INTRODUCTION

Network Security is a perception to preserve network and data transmission over wireless network. Data Security is the main aspect of protecting data transmission over unreliable network. Data Security is a challenging affair of data communications today that touches many areas including defend communication channel, strong data encryption technique and trusted third party to control the database. The rapid development in information technology, the protected transmission of confidential data herewith gets a great accord of attention. The conventional methods of encryption can only manage the data security. The information could be accessed by the unauthorized user for malicious purpose. Therefore, it is crucial to apply effective encryption/decryption methods to enhance data security. Network security involves the authorization of access to data in a network, which is controlled by the network administrator. User's choice or are an assigned ID and password or other authenticating information that al lows them access to information and programs within their authority. Network security includes a variety of computer networks, both public and private, that are used in everyday jobs conducting transactions and communications among businesses, government agencies and individuals. Networks can be private, such as within a company, and others which might be open to public access. Network security is involved in organizations, enterprises, and other types of institutions. It does as its title explains: It guards the network, as well as defends and oversees operations being done. The most common and simple way of insulating a network resource is by assigning it a unique name and a corresponding password [1]. Network security starts with authenticating the user, commonly with a username and a password. Since this requires just one detail authenticating the user name —i.e. the password, which is something the user 'knows'—this is sometimes termed as one-factor authentication. With two-factor authentication, that the user 'has' is also used (e.g. a security token or 'dongle', an ATM card, or a mobile phone); and with three-factor authentication, that the user 'is' is also used (e.g. a fingerprint or retinal scan). Once authenticated, a firewall enforces access policies such as what services are allowed to be accessed by the network users [2]. Though effective to prevent unauthorized access, this component may fail to check potentially harmful content such as computer worms or Trojans being transmitted over the network. Anti-virus software or an intrusion prevention system (IPS) helps detect and inhibit the action of such malware. An anomaly-based intrusion detection system may also monitor the network and traffic for unexpected (i.e. suspicious) content or behavior and

other anomalies to conserving resources, e.g. from denial of service attacks or an employee accessing files at strange times. Individual events occurring on the network may be logged for audit purposes and for later high-level analysis [3]. Communication between two hosts uses a network that may be encrypted to manage privacy.

Rough set theory (RST) is a major mathematical procedure developed by Pawlak in 1982 (Pawlak, 1982)[4]. This rule has been developed to manage uncertainties from information that presents some inexactitude, incompleteness and noises. When the available information is insufficient to determine the exact value of a given set, lower and upper approximations can be used by rough set for the representation of the concerned set. The approximation synthesis of approach from the acquired data is the main objective of the rough set analysis. For example, if it is difficult to define a theory in a given knowledge base, rough sets can approximate' with respect to that knowledge. In decision making, it has confirmed that rough set methods have a Powerful essence in contracting with uncertainties. The RST has been applied in several fields including image processing, data mining, pattern recognition, medical informatics, knowledge discovery and expert systems. In the current literature, several research works have been combined to the rough set theory with other artificial intelligence methods such as neural networks, fuzzy logic, additionally to other methods resulting in some good results. The use of rough set theory to solve a specific complex problem has attracted world-wide attention of further research and development, extending the original theory and increasingly widening fields of application. Additionally, rough set as a computationally efficient technique it presents a basic significance to many theoretical developments and practical applications of computing and automation, especially in the areas of machine learning and data mining, decision analysis and intelligent control. Among other computational problems, rough set addresses problems such as data significance evaluation, hidden pattern discovery from data, decision rule generation, data reduction and data-driven inference interpretation (Pawlak, 2004)[5].

2. LITERATURE REVIEW

There are some related works in this research area. In 1982 Pawlak Z. [6] proposed the operations on sets, approximate equality os sets, and approximate inclusion of sets. In 2005, Guoyin W. et. al. [7] firstly discussed the complications of network security and intrusion, and then some data –mining-based processes were disinterred to settle these dilemmas. Rough set based new techniques such as data reduction, incremental mining, uncertain data mining, and initiative data mining were recommended for intrusion detection systems. In 2005, Wenquing Zhao et. al. [8] described the effects of spam on network. It proposed a new scheme based on decision theoretic rough sets to classify emails into three categories – spam, no-spam,and suspicious. By comparing the anti-spam filter model reduced the error ratio that a non spam is discriminated to spam and the potential security problems of some email systems. In 2009, Rui L. et. al. [9] proposed a model of network security assessment based fuzzy sets and rough sets. The rough sets and the fuzzy sets were combined to find out the association rules in network security. In this the connection degree in set pair analysis was applied into rough sets. The data were done fuzzy clustering firstly and then the assessment rules in network security were extracted based on fuzzy sets and rough sets. In 2009, Lisha Kong et. al. [10] indentified the principle of assessment indexes system of network security and established the indexes system of network security assessment, established security assessment model and common steps of network security assessment which were both based on the rough set theory and finalized and validated a model by an example. In 2012, Jing L. et. al. [11] proposed a network security events correlation scheme based on rough set, build database of network security events and knowledge base, gives rule generation rule and rule matcher. This technique solved the simplification and correlation of massive security events through combining data discretization, attribute reduction, value reduction and rule generation. In 2012, Hai – Sheng W. et. al. [12], innovated a process to improve the computation accuracy and the efficiency of the classification computation by using Rough set combined with SVM classifier. In 2014, Roy P. [12] proposed that how rough-set theory helped in very fast convergence and in avoiding local minima problem, thereby enhancing the performance of the EM. During rough-set-theoretic rule generation, each band was individualized by using the fuzzy-correlation-based gray-level thresholding. In 2014, Chowdhuri S. et. al. [14] proposed the ad hoc routing protocol's design was used in order to detect the unpredictable and rapid mobility of a node. It was created dynamically without any infrastructure. In ad hoc each node was responsible for routing the information between them. To improve the performance of unused information and to

overcome the overhead in maintaining this information the protocols were designed. MANET (Mobile Ad hoc Network) ,the collection of wireless mobile nodes which could dynamically form a network. The notion of the thresholds and the temporal extensions to Rough Sets was applied in several protocols. The successful routing in MANETs using the Random Waypoint mobility model was based on various rough sets based protocol.

2.1. Pawlak Z. [1982]

This paper proposed the operations on sets, approximate equality on sets, and approximate inclusion of sets.

2.2. Wang G. et. al. [2005]

In this paper the concern of network security and intrusion was discussed at first, and then some data –mining-based methods were ascertained to work out these predicaments. Rough set based new techniques such as data reduction, incremental mining, uncertain data mining, and initiative data mining were proposed for intrusion detection systems.

2.3. Zhao W. et. al. [2005]

This paper described the effects of spam on network. It proposed a new scheme based on decision theoretic rough sets to classify emails into three categories – spam, no-spam, and suspicious .By comparing the anti-spam filter model reduced the error ratio that a non spam is discriminated to spam and the potential security problems of some email systems.

2.4. Li R. et. al. [2009]

This paper proposed a model of network security assesement based fuzzy sets and rough sets. The rough sets and the fuzzy sets were combined to find out the association rules in network security. In this the connection degree in set pair analysis was applied into rough sets. The data were done fuzzy clustering firstly and then the assesement rules in network security were extracted based on fuzzy sets and rough sets.

2.5. Kong L. et. al. [2009]

This paper indentified the principle of assesement indexes system of network security and established the indexes system of network security assesement, established security assesement model and common steps of network security assesement which were both based on the rough set theory and finalized and validated a model by an example.

2.6. Liu J. et. al. [2012]

This paper proposed a network security events correlation scheme based on rough set, build database of network security events and knowledge base, gives rule generation technique and rule matcher. This rule solved the simplification and correlation of massive security events through combining data discretization, attribute reduction, value reduction and rule generation.

2.7. Wang H-S et. al. [2012]

This paper innovated a process to improve the computation accuracy and the efficiency of the classification computation by using Rough set combined with SVM classifier.

2.8. Roy P. et. al. [2014]

This paper proposed that how rough-set theory helped in very fast convergence and in avoiding local minima problem, thereby enhancing the performance of the EM. During rough-set-theoretic rule generation, each band was individualized by using the fuzzy-correlation-based gray-level thresholding.

2.9. Chowdhuri S. et. al. [2014]

In this paper the ad hoc routing protocol's design was used in order to detect the unpredictable and rapid mobility of a node. It was created dynamically without any infrastructure. In ad hoc each node was responsible for routing the information between them. To improve the performance of unused information and to overcome the overhead in maintaining this information the protocols were designed. MANET (Mobile Ad hoc Network) ,the collection of wireless mobile nodes which could dynamically form a network. The notions of the thresholds and the temporal extensions to Rough Sets were applied in several protocols. The successful routing in MANETs using the Random Waypoint mobility model was based on various rough sets based protocol.

3. ROUGHSET BASED METHODS FOR NETWORK SECURITY

1. Junk email detection based on rough set

We are habituated with email. It is very essential to eliminate unsolicited emails or junk emails. Rough set based filters can be utilized to perceive junk emails on the Internet. One major procedure is to construct filters in email reassign route. Numerous junk email filters hadn't finished exercise of the entire security information in an email, which subsisted mostly in the junk email header not in the text and attachment. Below are certain guidelines for email headers which is helpful for judging whether an email is junk or not [15].

Eleven condition attributes and one decision attribute are described as follows.

Condition Attributes:

A. Amount of "Received" fields which is the times of email relaying. One "Received" per relay.
B. Amount of addressees.
C. Amount of email route disruption. For example, it's a route disruption when the IP address in the former "Received" field and receiver's domain name are dissimilar from those in the concluding "Received" field;
D. Amount of divergence between the domain name and its consequent IP address. This attribute is quite significant. E. Amount of no domain name of the sending host after "from" in the "Received" field.
F. Amount of no domain name of receiving host after "by" in the "Received" field.
G. Amount of no IP address of sending host after "from" in the "Received" field.
H. Whether the original sender address in the "From" field is accordant with that in the "Received" field the original sender address is given in the previous "Received" field after the "from" or "by".
I. Whether the target address in "To" field is accordant with that in the "Received" field the final is the actual receiver.
J. If "Delivered-To" field subsists, whether it is accordant with the "To" field it's default value is 1 that is yes.
K. If "Return-Path" field subsists, whether it is accordant with the "From" field it's default value is 1 that is yes.

Decision Attribute:

L. Legitimate emails value is denoted by 1 and junk emails value is denoted by 2 irrespective of its type. There are several processes to mine knowledge from a decision table, such as Preprocessing of data, including dealing with values of missing attributes, discretization of data. Attribute reduction. Value reduction. Some useful knowledge about detection of junk email can be obtained from email headers. Our simulation results demonstrate that when mining on selected baleful email corpus, the filter has high efficiency and high identification rate.

2. IDS architecture based on rough set

A network-based IDS based on rough set theory has been developed recently [16]. The system architecture is shown in Figure 1. In case of high-speed network, the detection efficiency of a single host will be very low, for this reason a distributed framework has been adopted. It is mainly composed of Administrator, Alarm/Response Plug-in, Intrusion Detection Module, Protocol Decoder, Rule Generation Component and Sensor. Their functions are as follows.

i. Administrator is an interface between intrusion detector and users. Users can update the set of rules manually by checking the log file, and defining the strategy of the detection, etc. As shown in Figure 1, the system performs its task in two phases, rule training phase and detection phase. In the rule training phase, audit data labeled with attacks are used as training data for rule generation. The output of this phase is a rule tree. Especially, this phase will be executed offline before intrusion detection. In the detection phase, actual detection is implemented through matching incoming network data with the rule tree. The incoming data will be labeled as normal behavior or a certain attack. Based on the detection results, Alarm/Response Plug-in will take action. The system is capable of extracting a set of detection rules from network packet features. It is effective and suitable for online intrusion detection with low cost and high efficiency.

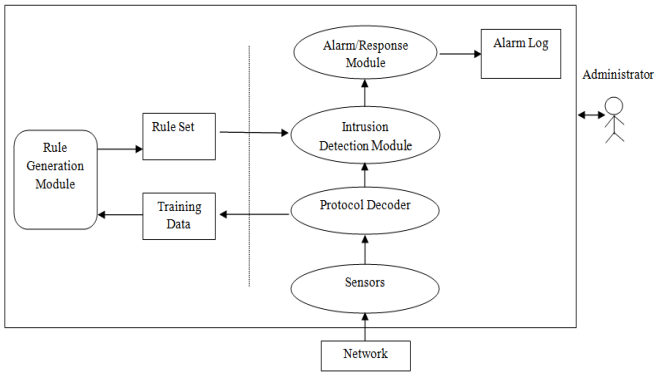

Figure1: Architecture of IDS

ii. Alarm/Response Plug-in is responsible for dealing with results from Intrusion Detection Model. In case of external attacks, it will notify the administrator by means of e-mails, console alerts, log entries, or a visualization tool.
iii. Intrusion Detection Module examines the rule tree generated in Rule Generation Component and matches the incoming raw data with rules. The results of rule matching will be transferred to Alarm/Response Plug-in and the latter will handle them by using the strategy defined in advance.
iv. Protocol Decoder analyzes raw data collected by the Sensor. At the beginning, it reassembles those data according to their protocol. Then, it converts raw packet and connection data into a format so that the rule generation component and intrusion detection module can use.
v. Rule Generation Component integrates rough set theory with rule generation. The whole procedure of rule generation was already discussed in section 3. In this component, a rule

model is generated in the format of rule tree. Each path from the root node of the rule tree to a leaf node represents a rule.

vi. Sensors are responsible for collecting network data. It is actually an interface to capture the information flowing through a network card on a machine or scan port of a switch. Evidently, its location determines the localization of intrusion detection. For example, intrusion detection can be done on a single machine, a network segment, or a gateway.

3. Other rough-set-related methods for network security

Another procedure wasre represented for anomaly intrusion detection with reduced cost and high efficiency [17]. It extracts detection rules using rough set algorithm from the system call sequences generated during the normal execution of a process and considered as the normal behavior model. It is capable of detecting the abnormal operating status of a process and thus reporting a possible intrusion. Compared with other methods, it requires a smaller size of training data set and less effort to collect training data and is more suitable for real-time detection. Empirical results show that this method is promising in terms of detection accuracy, required training data set and efficiency. Not only rule generation based on rough set theory can be used for network security, but other perception may also be useful. For example, rough inclusion is used for matching of normal behaviors and abnormal behaviors [18]. We conclude that rough set method is suitable and promising for network security.

4. CONCLUSION

With the advancement in the field of networking technologies and with the exponential rise in the data stored in various data repositories the security of the data in the vast network is the most serious issue this day. Several strategies, guidelines and algorithms have been implemented for resolving the security issues. Inspite of having all these sophisticated technology achieveing complete network security is still an issue. Implementing Rough set based methods for network security can minimize the gap with its accuracy, efficiency and unique implementation strategies.

REFERENCES

[1] Simmonds, A; Sandilands, P; van Ekert, L (2004) Ontology for Network Security Attacks". Lecture Notes in Computer Science. Lecture Notes in Computer Science 3285, pp.317–323.
[2] A Role-Based Trusted Network Provides Pervasive Security and Compliance - interview with Jayshree Ullal, senior VP of Cisco.
[3] Dave Dittrich, Network monitoring/Intrusion Detection Systems (IDS), University of Washington.
[4] Pawlak. Z (1991). Rough sets, in: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Dordrecht, 1991.
[5] Pawlak Z. (2004). Some Issues on Rough Sets. Transactions on Rough Sets, vol. 1, pp. 1-58.
[6] Pawlak Z. (1982). Rough Sets.International Journal of Computer & Information Sciences,vol.11,pp.341-356.
[7] Wang G., Chen L., Wu Y. (2005). Rough Set Based Solutions for Network Security. Monitoring security and Rescue Techniques in Multiagent Systems Advance in Soft Computing Volume 28, pp 455 - 0465.
[8] Zhao W., Zhu Y. (2005). An Email classification scheme based on decision –Theoretic Rough Set theory and Analysis of Email Security.
[9] Li R., Yang Y. (2009). Network Security Assesement Based on Fuzzy stes and rough sets.Wireless Communications , Networking and Mobile Computing. WiCom '09. 5th International Conference.
[10] Kong L., Ren X., Fan Y. (2009). Study on Assesement method for Computer Network Security based on Rough Set . Intelligent Computing and Intelligent Systems,ICIS 2009, IEEE International Conference on Volume:3.
[11] Liu J., Gu L., Xu G., Niu X. (2012). A Correlation Analysis Method of Network Security events based on rough Set Theory. Network Infrastructure and Digital Content (IC – NIDC), 3rd IEEE International Conference.
[12] Wang H.-S., Gui X.-L. (2012). A Network Security Model Based on Machine Learning. Control Engineering and Communication Technology (ICCECT), International Conference.
[13] Roy P., Goswami S., Chakraborty S., Azar A. T., Dey N. (2014). Image Segmentation Using Rough Set Theory:A Review. International Journal of Rough Sets and Data Analysis, 1(2), 62-74.
[14] Chowdhuri S., Roy P., Goswami S., Azar A. T., Dey N. (2014). Rough Set Based Ad Hoc Network : A Review. 66 International Journal of Service Science, Management, Engineering, and Technology, 5(4), 66-76.
[15] Wu Y., Li Z. J., Luo P., Wang G. Y.(2003). A new anti-Spam filter based on data mining and analysis of email security. Data Mining and Knowledge Discovery: Theory, Tools, and Technology V, pp. 147–154.
[16] Li Z. J., Wu Y., Wang G. Y., Hai Y. J., He. Y. P.(2004). A new framework for intrusion detection based on rough set theory. SPIE Defense and Security Symposium. Orlando, Florida USA. accepted and to appear.
[17] Cai Z. M., Guan X. H., Shao P., Peng Q. K., Sun G. J. (2003) A rough set theory based method for anomaly intrusion detection in computer network systems. Expert Systems Vol.20(5), pp251–259.MATHView Article.
[18] Li X. J., Huang Y, Huang H. K.,(2003). An Computing Immune Model based on Poisson Procedure and Rough Inclusion. Chinese Journal of Computers. Vol.26(1), pp.71–76.
[19] B. H. Sharmistha. (2014). A Study on Bayesian Decision Theoretic Rough Set, IJRSDA, Vol (1), pp. 1-14
[20] V. Renu (2014). Comparing and Contrasting Rough Set with Logistic Regression for a Dataset, IJRSDA, Vol (1), pp. 81-98