International Science Index, Computer and Information Engineering Vol:6, No:12, 2012 waset.org/Publication/13529

A Review on Soft Computing Technique in Intrusion Detection System

Noor Suhana Sulaiman, Rohani Abu Bakar, and Norrozila Sulaiman

Abstract**—**Intrusion Detection System is significant in network security. It detects and identifies intrusion behavior or intrusion attempts in a computer system by monitoring and analyzing the network packets in real time. In the recent year, intelligent algorithms applied in the intrusion detection system (IDS) have been an increasing concern with the rapid growth of the network security. IDS data deals with a huge amount of data which contains irrelevant and redundant features causing slow training and testing process, higher resource consumption as well as poor detection rate. Since the amount of audit data that an IDS needs to examine is very large even for a small network, classification by hand is impossible. Hence, the primary objective of this review is to review the techniques prior to classification process suit to IDS data.

Keywords—Intrusion Detection System, security, soft computing, classification.

I. Introduction

ELECTRIC commerce and the recent online consumer boom have forced a change in the basic computer security design for systems on a shared network. Systems are now designed with more flexibility and less barrier security. Furthermore, as computers become more financially available to the masses, they also become increasingly consumer-oriented. The combination of user friendliness and public accessibility, although advantageous for the average person, inevitably renders any exchanged information vulnerable to criminals. Consumer information, employee data or intellectual property stored in internal data warehouses are all at risk, from external attackers and disgruntled employees who might abuse their access privileges for personal gain. Security policies or firewalls have difficulty in preventing such attacks because of the hidden weaknesses and bugs contained in software applications [1].

Moreover, hackers constantly invent new attacks and disseminate them over the Internet. Disgruntled employees, bribery and coercion also make networks vulnerable to attacks from the inside. Mere dependence on the stringent rules set by security personnel is not sufficient. Intrusion detection systems, which can detect, identify and respond to unauthorized or abnormal activities, have the potential to mitigate or prevent such attacks [2].

Intrusion detection systems (IDS) were proposed to complement prevention based security measures. An intrusion is defined to be a violation of the security policy of the system. Intrusion detection thus refers to the mechanisms that are developed to detect violations of system security policy. Intrusion detection is based on the assumption that intrusive activities are noticeably different from normal system activities and thus detectable. The system is not introduced to replace prevention-based techniques such as authentication and access control. Instead, it is intended to complement existing security measures and detect actions that bypass the security monitoring and control component of the system. Intrusion detection is therefore considered as a second line of defense for computer and network systems. Generally, an intrusion would cause loss of integrity, confidentiality, denial of resources, or unauthorized use of resources [3].

II. Data Representation of Intrusion Detection System

The large amount of audit data that an IDS needs to examine and analysis is difficult because extraneous features can make it harder to detect suspicious behavior patterns [4]. Audit data capture various features of the connections. For example, the audit data would show the source and destination bytes of a TCP connection or the number of failed login attempts or duration of a connection. Complex relationships exist between the features, which are difficult for humans to discover. The intensive application of the intrusion detection system (IDS) in the network, it appears some significant problems, such as slow detect speed, big load, and mass data which cannot deal with in time. Even for a small network, the quantities of network traffic data that an IDS needs to examine are very large. Detect speed already becomes an important index of real time requirement in intrusion detection. How to reduce network feature attributes without decreasing detect rate already becomes current research hotspot [5].

In the network intrusion detection, the system needs to handle massive amounts of network data in real-time manner, typically, the network data contains a large number of features [6], which significantly increases the load of IDS, but at the same time, there are many irrelevant and redundant features that will decline detection accuracy during the intrusion

Noor Suhana Sulaiman is with the Faculty of Computer, Media and Management Technology (FKMPT), Kolej Universiti TATi, 24000 Kemaman, Terengganu, Malaysia (e-mail: suhana@tatiuc.edu.my).

Rohani Abu Bakar is with Faculty of Computer Systems and Software Engineering, University Malaysia Pahang, Lebuhraya Tun Razak, 26300 Kuantan, Pahang, Malaysia (e-mail: rohani@ump.edu.my).

Norrozila Sulaiman is with Faculty of Computer Systems and Software Engineering, University Malaysia Pahang, Lebuhraya Tun Razak, 26300 Kuantan, Pahang, Malaysia (e-mail: norrozila@ump.edu.my).

detection process based on machine learning mechanism and bring additional of the complexity of learning algorithms. All of these require IDS must be able to select the right subset of the most important features to improve the detection accuracy and efficiency.

IDS must therefore reduce the amount of data to be processed. This is very important if real time detection is desired. Some data may not be useful to the IDS and thus can be eliminated before processing. In complex classification domains, features may contain false correlations, which hinder the process of detecting intrusions. Further, some features may be redundant since the information they add is contained in other features. Extra features can increase computation time, and can have an impact on the accuracy of the IDS. Feature selection improves classification by searching for the subset of features, which best classifies the training data. The clustering method could produce high quality dataset with far less instances that sufficiently represent all of the instances in the original dataset. The reduced feature set should yield the best detection rate based on the set of important features [7]. It indicates that elimination of the insignificant and/or useless inputs leads to a simplified problem and possibly faster and more accurate detection, so feature selection is very important in intrusion detection [8]. Feature selection will delete unimportant features according to certain rules in order to low the dimension of the feature space. Also, it can find the most effective subset of original feature set to improve prediction accuracy rate for classification and prediction models or to lower the complexity of the model structure in the guarantee of forecasting accuracy [9].

The question in the large number of features that can be monitored for IDS, which are truly useful, which are less significant, and which may be useless are relevant because the elimination of useless features (or audit trail reduction) enhances the accuracy of detection while speeding up the computation, thus improving the overall performance of the detection mechanism [10]. In cases where there are no useless features, concentrating on the most important ones may well improve the time performance of the detection mechanism, without affecting the accuracy of detection in statistically significant ways. Intrusion detection is a kind of technology of finding intrusion through collecting and analyzing protected system information. Its main function is to monitor computer system and network, and to find the intrusion activities in the system. Then intrusion alarm is presented by the intrusion detection system. At first, the intrusion detection model presented is to improve the security of computer system through audit data. But it also has bugs to use audit data because so much information is presented that security administrators can not manage it in effect.

The main field in intrusion detection system (IDS) research, it can detect unknown intrusions stably without too much knowledge on system flaw, but it has the shortcoming of high false alarm rate. The key of anomaly intrusion is to establish the system or user's normal behavior pattern (storehouse) and use the pattern (storehouse) to carry on the comparison and judgment to the current behavior [11]. There are a lot of technologies being used as anomaly detection methods, such as the neural network, the data mining, the support vector machine, and the hidden Markov model [12-20], each kind of these technologies has shown its advantages to detect novel intrusions, but it still has some shortcomings. Anomaly detection is generally based on machine learning. The normal system behavior is modeled by a systematic method and the intrusion detection is performed by tracing and the model. It has the advantage of being able to detect unknown intrusions and therefore is a very active research area. However, anomaly detection may have a higher rate of false positives. Data collection and model training is the first step in anomaly detection. In a UNIX or a UNIX compatible operating system, a series of system calls will be invoked during the execution of a process. Forrest [21] discovered that by analyzing the system call sequences, it is possible to detect abnormal behaviors of a process. Forrest also suggested that the normal behavior of a process could be characterized by using a database of fixed length short sequences of system calls. The above researchers have made various contributions to intrusion detection techniques based on analyzing system call sequences, but their methods generally require high quality training data to obtain the normal behavior model. However, collecting training data is difficult in implementing an IDS and in most circumstances only the data of normal behavior are available. Therefore the approach of Asaka [19] that requires a properly labeled datasets for both normal and abnormal behaviors in order to search the optimal classification surface is difficult to implement in a practical system. If a complete normal behavior database is not available, using the methods of Forrest may cause a significant number of false alarms.

III. A Review on Soft Computing Technique in Intrusion Detection System

Of late, a variety of published algorithms have been applied to assemble a computer-aided analysis system in medical field. The common used algorithms are Neural Network [22-24], Genetic Algorithm [25-27], Support Vector Machine [28,29] [30-32], [33,34], Particle Swarm Optimization [35,36], Artificial Immune System [37-40] , Wrapper And Filter [41,42], Dempster-Shafer Theory Of Evidence [43,44] and Rough Set Theory [45]. These techniques are described in the section.

A. Neural Network

In last years, several neural network (NN) techniques such as multilayer perceptron network (MLPN)/Back Propagation network (BPN) have been applied in many IDS models and have obtained the corresponding detection performance [22-24]. However, the experiments also revealed many problems such as slow convergence, overfitiing, local minimum, difficult to determine the number of hidden layers and hidden nodes, as well as the quantity and quality of samples have deep impact on generalization ability and accuracy of neural network. The current improvements of MLPN/BPN include using simulated annealing and genetic algorithms to overcome the local minima; using Levenberg-Marquardt algorithm and conjugate gradient algorithm to speed up the training. Nevertheless, to a large number of high-dimensional data, all the above improvements will inevitably result in the structure of the network too large to train, and increase contradictory in samples so that the network has poor generalization capability and lower accuracy. All these limitations seriously affected the alert reliability and the performance of the NN improvement. In addition, the key of the anomaly detection is to form a user or system profile of the normal activities. In order to achieve data storage, analysis and sharing of components in the cooperative intrusion detection system, the data must be reduced as much as possible so as to reduce the storage and computational cost [46].

B. *Genetic Algorithm*

In feature selection of Intrusion Detection, the SNFS [35] algorithm used Neural network and Support Vector Machine. CFSSGA [36] proposed a hybrid algorithm with correlation-based feature selection (CFS), and employed the SVM and genetic algorithm to achieve the optimization of intrusion detection. While, the FSRGA [37] algorithm is based on rough sets and improved genetic algorithms to improve feature selection. Both SNFS and CFSSGA algorithm need data classification for their each iteration. This produces much more time complexity and do not take care of the combination of characteristics as well as the balance of the number and classification accuracy. However, FSRGA algorithm did not optimize the genetic operation, which will easily to make the algorithm be trapped into a local optimal solution [47].

C. *Support Vector Machine*

Support vector machine [28, 29] is a new kind of machine learning algorithm proposed recently which is based on structural risk minimization of statistical learning theory. Many researchers verified that SVM performed well in intrusion detection classification [30-32]. However, when applying standard SVM on high dimension and large-scale dataset, such as network connection dataset, it often suffers memory storage and time consuming problem because a standard SVM solver will solve a dual quadratic optimization problem [38, 39]. Decreasing the dimension of training samples by feature selection or attribution reduction method is benefit to help alleviate this problem.

Comparing with traditional ANN, SVR possesses prominent advantages such as excellent properties in learning limited samples, good generalization ability, etc. SVR is originally developed for solving classification problems and later extended to solve regression problems, and exhibits good learning performance [50]. But there exists a problem in the practical application of SVR. This problem is how to select some of SVR parameters so that the performance of SVR can be brought into the best.

In [31] Mukkamala and Sung compare performance of ANN and SVMs in intrusion detection. SVMs outperform ANN in speed and scalability, And SVMs are relatively insensitive to the number of data points and the classification complexity does not depend on the dimensionality of the feature space. High dimensions of attribute space and parameters to be optimized are often suffered from by SVM for intrusion detection [51].

Tarun Ambwani reports his multi-class SVM to intrusion detection in [33]. The standard method for N-class SVM is one-versus-rest which constructs N SVMs. Ambwani uses one-versus-one method which constructs all possible n*(n-1)/2 two-class SVMs. The detection rate of Ambwani's one-versus-one multi-class SVM seems good, but he did not present the time cost of his methods, as well as the problem of confusions of multi-hyper-planes. For one-versus-rest or one-versus-one SVMs, many hyper-planes are constructed, but in these methods, the hyper-planes do not promise a perfect separation. Therefore, Takahashi and Abe propose decision-tree-based SVMs [52]. In their method, in training, a decision tree in which each node presents a decision hyper-plane which separates one or some classes from others are constructed by top-down ways. By this method, when training is finished, the feature space is divided by N-1 hyper-planes and there are no unclassifiable regions. DT SVMs are also efficient to deal with confusions. Unfortunately, it is hard to control the classification error of decision tree, which is performance of one node will influence the whole sub-tree below it [52].

D. *Particle Swarm Optimization*

PSO is robust and has been proven theoretically and empirically to be able to search the optimum solution or near-optimal solution to a complex problem. However, if current optimal position is discovered by certain particle, the other particle will draw close to the optimal position rapidly in the process of running. If this optimal position is local optimal point, particle population will not research in the solution space. Thus, PSO is easy to sink into local optimized solution that is to say, premature phenomena is appeared. Many scholars resolve problems above by combination with particle swarm optimization and genetic algorithms [35] or selection inertia weight. Berhart and Shi [36] put forward the strategy of linear decreasing inertia weight, which is applied in particle swarm optimization algorithm. Although this strategy improves the performance of particle swarm optimization, it is related with iteration times of PSO and cannot really reflect the algorithms’ characteristics of complex and nonlinear in the process of running [53].

E. *Artificial Immune System*

Artificial immune method was firstly used to protect computer by Forrest et al in [37].They viewed the problems of protecting computer system as abnormal process and normal process and an intrusion detection system based on immune model was proposed by them. Some research in intrusion detection system using data mining has been done by W. Lee [38].In recent surveys, representative designs are Mukkamala et al.[39] and W. Lee and Stolfo[40]. A good representative of the IDSs using data mining tools is NADAM ID[40], which is used as experiment platform by many IDS researchers. However, the structure of NADAM ID is a little complicated and has low detection rate of DOS attacks [54].

F. Wrapper and Filter

Several literatures have tried to figure out important intrusion features through wrapper and filter approaches; two broad types of feature selection. Wrapper method exploits a machine learning algorithm to evaluate the goodness of features. It provides better performance of selecting suitable features since it employs performance of learning algorithm as an evaluation criterion. On the other hand, Filter method which is faster than wrapper utilizes the underlying characteristics of the training data to evaluate the relevance of the features by some independent measures such as distance measure, correlation measure, consistency measure [41, 42]. Wrapper method demands heavy computational resource for training and cross validation while filter method lacks the capability of minimization of generalization error.

G. Dempster-Shafer Theory of Evidence

Dempster-Shafer theory of Evidence (DST) [43], also known as the theory of belief functions, is a tool for representing and combining evidence. Being a generalization of Bayesian reasoning, it does not require probabilities for each question of interest, but the belief in a hypothesis can be based on the probabilities of related questions. Contributing to its success is the fact that the belief and the ignorance or uncertainty concerning a question can be modeled independently. However, some problems may be solved in its applications including fault diagnosis, target identification, decision making and so on. Firstly, it is hard to acquire the objective basic probability assignment (BPA). Secondly, it is difficult to make all evidences independent. Then, when the reasoning chain is too long, its application becomes inconvenient [44].

H. Rough Set Theory

Rough Set Theory is the most widely used baseline technique of single classifier approach on intrusion detection. In addition, it has also been considered recently for model comparisons. Before training, the step of feature or variable selection may be considered. The process of feature selection identifies which features are more discriminative than the others. This has the benefit of generally improving system performance by eliminating irrelevant and redundant features [55].

Many analyses have been done to come out with significant rules. Since the generated rules possible to have in large number of rules, it is important to know whether all rules played a role in the classification process. Indranil Bose, [45] has suggested that, to find the most significant rules for each sample, the rules are sorted according to the value of their support. The generated rules do not differ much in terms of length and thus support is used as the criterion for ranking the rules. Subsequent analysis is the evaluation of the rules length to obtain testing accuracy. Typically, rules of less length ascend to a higher overall testing accuracy. This indicates that the dataset led to the formation of a smaller number of rules, can correctly recognize the problem. The overall testing accuracy is highest when the training sample is reduced to 10% from the original sample. Then, the experiment of changing the parameters associated with the testing procedure is implemented. The experiment is conducted using four factors; balance of sample, a ratio of training to testing sample size, testing sample size and training sample. The experiment resultant that there was no significant effect of changing the balance of the sample, training to testing sample size, training sample size and testing sample size on testing accuracy across all samples. The best classification result is obtained and the comparisons are made with two other statistical approaches; logistic regression and discriminant analysis. The reported results reveal that Rough Set Theory method generally performed better than the others in terms of classification accuracy on training and testing samples.

IV. Conclusion

An attempt has been made in this study to explore the essential of Intrusion Detection System. Consequently the more accurate detection is needed in real time IDS. The representation of Intrusion Detection System data understanding is important to probe the high resultant classification. This study has presented a review of Intrusion Detection System data representation and a review of other soft computing techniques applied in Intrusion Detection System prior to select the best technique suit to Intrusion Detection System classification.

Acknowledgment

Authors would like to thank to *Faculty Computer, Media and Management Technology (FKMPT)*, Kolej Universiti TATi and *Faculty Computer Systems & Software Engineering (FSKKP)*, Universiti Malaysia Pahang for the support in making this study a success.

References

[1] Dasarathy, B.V. (2003). Intrusion detection. Information Fusion, 4, pp.243-245.

[2] Ye, Q., Wu X.P. a*, Liu Y.Q. b*, Huang, G.F. c*. (2010). A Hybrid Model of RST and DST with Its Application in Intrusion Detection. Depart. of Information security, Naval University of Engineering a,c*, Naval Institute of Compute Technology, Beijing, China b*.

[3] Bishop, M. (2003). Computer security e art and science: Addison Wesley

[4] Sung, A.H., Mukkamala, S. (2003). Identifying important features for intrusion detection using support vector machines and neural networks. Proceedings of International Symposium on Applications and the Internet (SAINT 2003), p. 209e17.

[5] WenJie, T., JiCheng, L. (2009). Intrusion Detection Quantitative Analysis with Support Vector Regression and Particle Swarm Optimization Algorithm. Beijing Union University, Beijing, China

[6] Chen, L., Shi, L., Jiang, Q., & Wang, S. (1983-1992). Supervised Feature Selection for Dos Detection Problems Using a New Clustering Criterion. Journal of Computational Information Systems, 3(5)

[7] Shi-Jinn, H., a,b,*, Ming-Yang, S., c*, Yuan-Hsin, C., b*, Tzong-Wann, K., d*, Rong-Jian, C., b*, Jui-Lin, L., b*, Citra Dwi P., a*, (2011). A novel intrusion detection system based on hierarchical clustering and support vector machines.
a Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan b Department of Electronic Engineering, National United University, Miaoli, Taiwan, c Department of Computer Science and Information Engineering, Ming Chuan University, Taoyuan, Taiwan d Department of

Electronic Engineering, Northern Taiwan Institute of Science and Technology, Taipei, Taiwan.

[8] Ling, Y.,Bo, C., Junmo, X. (2007). An Integrated System of Intrusion Detection Based on Rough Set and Wavelet Neural Network. Nanjing, China

[9] Koller, D., & Sahami, M. (1996). Toward optimal feature selection. Proceedings of the International Conference on Machine Learning.

[10] Xiang, C., Bing-Xiang, L., & Yi-Lai, Z. (2010). Attribute Reduction Method Applied to IDS. Information engineering Institute, Jingdezhen Ceramic Institute.

[11] Fanping, Z., Kaitao, Y., Minghui, C., & Xufa, W. (2009). A New Anomaly Detection Method Based on Rough Set Reduction and HMM. University of Science and Technology of China, Anhui, China

[12] Mukherjee, B., Heberlein, L.T., & Levitt, K.N. (1994, May). Network intrusion detection. IEEE Network, (3):pp 26-41

[13] Susan, C.L., & David V.H. (2001). Training a neural–network based intrusion detector to recognize novel attacks. IEEE Transactions on systems, man and cybernetics-part a: System and Humans, Vol 31 No4, pp294-299

[14] Debar, H., Becke, M., & Siboni, D. (1992). A neural network component for an intrusion detection system. Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy.

[15] Manganaris. (2000). A data mining analysis of RTID alarms. Computer Networks, Vol 34, No 4 pp571-577.

[16] Tran, Q., Zhang, Q.L., & Li, X. (2002). SVM classification-based intrusion detection system. Journal of China Institute of Communications, Vol 23, No.5.

[17] Warrender, C., Forrest, S., & Pealmutt, B. (1999). Detecting intrusion using system calls: alternative data mode. IEEE Symposium on Security and Privacy.

[18] Hofmeyr, S.A., Forrest, S., & Somayaji, A. (1998). Intrusion detection using sequences of system calls. Journal of Computer Security, Vol. 6, pp 151–180.

[19] Asaka, M., Onabuta, T., Inoue, T., Okazawa, S., & Goto, S. (2001). A New Intrusion Detection Method Based on Discriminant Analysis. IEICE Transactions on Information and Systems, pp 570-577, 5.

[20] Beynon, M. J., Curry, & B., Morgan, P.H. (2000). Classification and Rule Induction Using Rough Set Theory. Expert Systems, Vol 17, NO 3, pp 136-148

[21] Forrest, S., Ofmeyr, S.A., Somayaji, A. (1996). A Sense of Self for Unix Processes. IEEE Computer Society, In Proceedings of 1996 IEEE Symposium on ComputerSecurity and Privacy, New York, pp120–128

[22] Jian, L., Zhang G., Gu G. (2004). The research and implementation of intelligent intrusion detection system based on artificial neural network. *The 3rd International Conference on Machine Learning and Cybernetics*, Shanghai.

[23] Hofmann, A., Schmitz, C., & Sick, B. (2003). Rule extraction from neural networks for intrusion detection in computer networks systems. *IEEE Transactions on system, Man and Cybernetics*, IEEE Inc,CA, pp.1259-1265.

[24] Golovko, & V., Kochurko, P. (2005, September). Intrusion Recognition Using Neural Networks. *IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, Sofia, Bulgaria, pp.108-111.

[25] Sung, A.H., & Mukkamala, S. (2003). Identifying Important Features for Intrusion Detection using Vector Machines and Neural Networks. *Proceedings of International Symposium on applications and the Internet Technology*, pp. 209-216

[26] Shazzad, K.M., & Jong S. P. (2005, Dec). Optimization of Intrusion Detection through Fast Hybrid Feature Selection. *Parallel and Distributed Computing, Applications and Technologies, PDCAT 2005. Sixth International Conferenc,* pp. 264-267.

[27] Luyin, C., Qingshan J., Lifei, C. (2008). A Feature Selection Method for Network Intrusion Detection. *Computer Research and Development Supplement, 45*(10):156-160.

[28] Vapnik, V. (1995). The Nature of Statistical Learning Theory. *Springer-Verlag Press*, New York , American

[29] Cortes, C., & Vapnik, V. (1995). Support vector networks. *Machine Learning, Vol.20*, No.3, 273-297

[30] Hansung, L., Jiyoung, S., & Daihee, Park. (2005). Intrusion Detection System Based on Multi-class SVM. *Lecture Notes in Computer Science, vol.3642,* Springer Berlin, 9, pp.511-519.

[31] Mukkamala, S., Janoski, G., Sung, A.H. (2002). Intrusion Detection Using Neural Networks and Support Vector Machines. *Proceedings of IEEE International Joint Conference on Neural Networks, Vol 2,* Honolulu, 5, pp. 1702-1707.

[32] Dong, S. K., Ha, N.N., Jong, S.P. (2005). Genetic algorithm to improve SVM based network intrusion detection system. *19th International Conference on Advanced Information Networking and Applications, Vol.2,* Taiwan, 3, pp.155–158.

[33] Ambwani, T. (2003). Multi class Support Vector Machine Implementation to Intrusion Detection. *Proc. IEEE International Joint Conference on Neural Networks,* pp.2300-2305

[34] Mukkamala, S., Sung A.H., & Abraham, A. (2005). Intrusion detection using an ensemble of intelligent paradigms. Journal of Network and Computer Applications, vol.28, pp. 167-182

[35] Shi, Y., & Eberhart, R. (1998). A Modified Particle Swarm Optimizer. *IEEE World Congress on Computation Intelligence*, pp.69-73.

[36] De Castro, L., & Von Zuben, F. (2002). Learning and Optimization Using the Clonal selection principle. *IEEE Transactions on Evolutionary Computation, Vol. 6(3),* pp. 239–251.

[37] Forrest, S., Perrelason, A.S., Allen, L., & Cherukur, R. (1994). Self-Nonself discrimination in a computer. *Proceedings of the 1994 IEEE Symposium on Research in Security and Privacy,* Oakland, CA:IEEE Computer Society Press, pp.202-212.

[38] Lee, W., & Stolfo, S.J. (1999). A data mining framework for building intrusion detection model. *Proceedings of the 1999 IEEE Symposium on Security and Privacy.Oakland,* CA:IEEE Computer Society Press, pp.120-132.

[39] Mukkamala, R.K., Gagnon, J., & Jajodia, S. (2000). Integrated data mining techniques with intrusion detection. *Research Advances in Database and Information Systems Security.* Kluwer Publisher, pp. 33–46.

[40] Lee, W., & Stolfo, S.J. A framework for constructing features and models for intrusion detection systems. *ACM Trans .Inform. Syst. Security , vol.3,* pp. 227-261.

[41] Dash, M., Liu, H., & Motoda, H. (2000). Consistency based feature selection. *Proc. of the Fourth PAKDD 2000,* Kyoto, Japan, pp. 98–109.

[42] Almuallim, H., & Dietterich, T.G. (1994). Learning Boolean Concepts in the Presence of Many Irrelevant Features. *Artificial Intelligence,* vol. 69, nos. 1-2, pp. 279-305.

[43] Shafer, G. (1976). A mathematical theory of evidence. *Princeton, NJ: Princeton University Press*

[44] Ye, Q., Wu, X.P., Liu, Y.Q., Huang, G.F. ( ). A Hybrid Model of RST and DST with Its Application in Intrusion Detection. Naval University of Engineering, Wuhan, China

[45] Bose, I. (2006). Deciding The Financial Health Of Dot-Coms Using Rough Sets. *School of Business, University of Hong Kong.*

[46] Ling, Y., Bo, C., Junmo, X. (2007). An Integrated System of Intrusion Detection Based on Rough Set and Wavelet Neural Network. Nanjing China.

[47] Liang, S.Y., Yuteng, G., Beizhan, W., Xinxing, Z., Xiaobiao, X., Lida, L., & Qingda, Z. (2010). Feature Selection Based on Rough Set and Modified Genetic Algorithm for Intrusion Detection. *The Research of Complex-Intrusion-oriented Alert Information Aggregation and Association Analysis Technology (NO.2008F3101)"* A.P. Software School of Xiamen University, Xiamen, China

[48] Cortes, C., & Vapnik, V. (1995). Support Vector Networks. *Machine learning, vol.20, no.3*, Springer Berlin, pp.273-297.

[49] Burges, C.J.C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery, vol 2(2),* Springer US, pp.121-167.

[50] Steve, R.G. (1998). Support Vector Machines for Classification and Regression. *Technical Report,* University of Southampton Press, Southampton, UK

[51] Mukkamala, S., Janoski, G., Sung, A.H. (2002). Intrusion Detection Using Neural Networks and Support Vector Machines. *Proceedings of IEEE International Joint Conference on Neural Networks, Vol 2,* Honolulu, 5, pp. 1702-1707.

[51] Huaping, L., Yin, J., & Sijia, L. (2010). A New Intelligent Intrusion Detection Method Based on Attribute Reduction and Parameters Optimization of SVM. Xihua University, Chengdu, China.

[52] Takahashi, F., & Abe, S. (2003). Decision-Tree-Based Multi class Support Vector Machines. *Proc. International Conference on Neural Information Processing, Vol.3*, pp.1418-1422

[53] Zhenying, M., Lei, Z., Xiaofeng, L. (2005). On the Efficiency of Support Vector Classifiers for intrusion detection. Chongqing, China

[54] Qingxiang, W., Jianmei, S. (2009). Fusion of Rough Set Theory and Linear SVM for Intrusion Detection System. University of Science and Technology of China, Hefei, China.

[55] Chih-Fong, T. a*, Yu-Feng, H. b*, Chia-Ying, L. c*, & Wei-Yang, L. d,*. (2009). Intrusion Detection by Machine Learning: A review, a Department of Information Management, National Central University, Taiwan b Department of Information Management, National Sun Yat-Sen University, Taiwan c Department of Accounting and Information Technology, National Chung Cheng University, Taiwan d Department of Computer Science and Information Engineering, National Chung Cheng University, Taiwan.