ABSTRACT

Intrusion Detection Systems (IDS) are developed to be the defense against these security threats. Current signature based IDS like firewalls and anti viruses, which rely on labeled training data, generally can not detect novel attacks. A method that offers a promise to solve this problem is the anomaly based IDS. Literature has shown that direction towards reducing false positive rate and thus enhancing the detection rate and speed have shifted from accurate machine learning classifiers to the adaptive models like bio-inspired models. Consequently, this study has been introduced to enhance the detection rate and speed up the detection process by reducing the network traffic features. Moreover, it aimed to investigate the implementation of the bio-inspired Immune Network approach for clustering different kinds of attacks. This approach aimed at enhancing the detection rate of novel attacks and thus decreasing the high false positive rate in IDS. Rough Set method was applied to reduce the dimension of KDD CUP '99 dataset which used by this study and select only the features that best represent all kinds of attacks. Immune Network clustering was then applied using aiNet algorithm in order to cluster normal data from attacks in the testing dataset. The results revealed that detection rate and speed were enhanced by using only the most significant features. Furthermore, it was found that Immune Network clustering method is robust in detecting novel attacks in the test dataset. The principal conclusion was that IDS is enhanced by the use of significant network traffic features besides the implementation of the Immune Network clustering to detect novel attacks.

ABSTRAK

Sistem Pengesanan Pencerobohan (IDS) dibangunkan untuk menangani ancaman keselamatan ini. Sistem pengesanan berteraskan tandatangan seperti dinding api dan anti-virus kebiasaannya tidak dapat mengecam serangan-serangan baru manakala sistem pengesanan berbentuk 'anomaly' berupaya menyelesaikan masalah sebegini. Kajian menunjukkan tumpuan telah beralih kepada pengurangan *false alarm* serta pembaikan terhadap tahap pengesanan dan kelajuan pengesanan dengan aplikasi seperti pengesanan berbentuk 'machine-learning' kepada kaedah berdasarkan 'bio-inspired'. Oleh itu, kajian projek ini dibangunkan untuk menambahbaik kadar pengesanan serta kelajuan proses pengesanan ini dengan mengurangkan atribut pada paket rangkaian trafik. Secara spesifiknya, fokus kajian projek ini tertumpu kepada pengaplikasian pendekatan 'bio-inspired Immune Network' pada sistem pengesanan pencerobohan dengan mengumpukkan (cluster) kepada kelas-kelas serangan. Tujuannya adalah untuk menambahbaik kadar pengesanan terhadap serangan-serangan baru dan menurunkan kadar 'false positive'. Kaedah 'Rough Set' digunakan untuk mengurangkan dimensi atribut pada paket-paket rangkaian set data KDD CUP '99 serta memilih atribut-atribut yang terbaik bagi mewakili semua jenis serangan. Algoritma aiNet digunakan bagi mengasingkan kan data normal dari data serangan pada set data pengujian. Hasil kajian menunjukkan kadar pengesanan serta kelajuan proses pengesanan boleh dicapai dengan menggunakan atribut-atribut rangkaian yang penting. Disamping itu, Immune Network berupaya mengesan jenis-jenis serangan yang baru. Kesimpulannya, penambahbaikan pada sistem pengesanan pencerobohan ini boleh diperolehi dengan menggunakan atribut-atribut rangkaian yang penting sahaja mewakili semua jenis serangan dan Immune Network berupaya mengklusterkan serangan-serangan baru.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Due to the increase use of computer networks in many aspects of life, the number of vulnerabilities also is increasing causing the network resources unavailable and violates the system confidentiality, integrity and availability. Intrusions pose a serious security threat for the stability and the security of information in the network environment. A network intrusion attack encompasses a wide range of activities. It includes attempting to destabilize the network, gaining unauthorized access to files with privileges, or mishandling and misusing of software (Jiang *et al.*, 2006).

Intrusion Detection Systems (IDS's) are security tools that, like other measures such as antivirus software, firewalls, and access control schemes, are intended to strengthen the security of information and communication systems (Teodoro, 2009). An Intrusion Detection System is an important component of the computer and information security framework. Its main goal is to differentiate between normal activities of the system and behaviors that can be classified as intrusive.

The purpose of IDS is to detect unauthorized use or access to the computer system or network from the outside environment by those who do not have the authority or access rights to such systems. The main goal of intrusion detection is to build a system that could automatically scan the network activity and detect such intrusion attacks. An IDS is used to detect several types of malicious behaviors that can compromise the security and trust of a computer system or network. This includes network attacks against vulnerable services, data driven attacks on applications, host based attacks such as privilege escalation, unauthorized logins and access to sensitive files, and malware (viruses, Trojan horses, and worms).

There are two main intrusion detection approaches: anomaly intrusion detection system and misuse intrusion detection system .The anomaly detection focuses on the unusal activities of patterns and uses the normal behavior patterns to identify an intrusion. The misuse detection recognizes known attack patterns and uses well-defined patterns of the attack. On the other hand, IDS's may be categorized according to the host system into two types:

i. Host-based IDS (HIDS)
ii. Network-based IDS(NIDS)

The first operates at the host level and monitors a single host machine using the audit trails of the host operating system, whereas the other operates at the network level and monitors any number of hosts on the network.

According to Zainal *et al*., (2006), IDS can be treated as pattern recognition problem or rather classified as learning system. They stated that, an appropriate representation space for learning by selecting relevant attributes to the problem domain is an important issue for learning systems. Bello *et al., (*2005) suggested that feature reduction is necessary to reduce the dimensionality of training dataset. They claimed that feature reduction also enhances the speed of data manipulation and improves the classification rate by reducing the influence of noise.

Furthermore, Kim *et al.,* (2005) stated that the goal of feature reduction is to find a feature subset in order to maximize some performance criterion, such as classification accuracy. They claimed that selecting important features is an important issue in intrusion detection.

In literature, numbers of anomaly detection systems were developed based on many different machine learning techniques. For example, some studies apply single learning techniques, such as neural networks, genetic algorithms, support vector machines, bio-inspired algorithms , etc. On the other hand, some systems are based on combining different learning techniques, such as hybrid or ensemble techniques. In particular, these techniques were developed as classifiers, which were used to classify or recognize whether the incoming network access is normal access or an attack.

Computing models inspired by biology are a way to make use of concepts, principles and mechanisms underlying biological systems. This type of computing includes among others, fields as evolutionary algorithms, neural networks, molecular computing quantum computing and immunological computation. The trend now is going towards the bio-inspired systems because of the ability of those systems to adapt naturally with the environment in which they applied. The human immune system provides inspiration for solving a wide range of innovative problems.

The Immune Network theory as originally proposed by Jerne, (1974a) hypothesized a novel viewpoint of lymphocyte activities, natural antibody production, pre-immune repertoire selection, tolerance and self/non-self discrimination, memory and the evolution of the immune system. It was suggested that the immune system is composed of a regulated network of cells and molecules that recognize one another even in the absence of antigens. The immune system was formally defined as an enormous and complex network of paratopes that recognize sets of idiotopes, and of idiotopes that are recognized by sets of paratopes, thus it could recognize as well as be recognized (de Castro *et al.,* 2001).

By dealing with very large number of data over networks, it is very difficult to classify them manually to detect possible intrusions. Labeled data can be obtained by simulating intrusions, but this is limited to the set of known attacks and will fail to address the new types of attacks that may occur in the future. As a result of that limited ability in detecting unknown attacks, the detection system will not be able to play its role in securing the network data. So a technique for detecting intrusions when the data is unlabeled is needed, as well as detecting new and unknown types of attacks .

1.2 Problem Background

Kim and Park, (2003) proposed a network-based Support Vector Machine SVM IDS, and demonstrated it through results of 3 kinds of experiments. They have shown that SVM IDS can be an effective choice of implementing IDS. They discovered that there are some miss-classified input vectors and those degrade the performance of SVM IDS. They claimed that the performance of SVM IDS can be improved by applying Genetic Algorithm (GA) based feature extraction. They further suggested that Decision Tree (DT) method can also be used to extract features instead of GA.

Shon *et al.*, (2005) employed a GA for selecting proper TCP/IP packet fields to be applied to support vector learning in order to distinguish anomaly attacks from normal packets. Their results showed that their proposed GA and the time delay preprocessing were reasonable for feature reduction. Moreover, they claimed that the two approaches using supervised and unsupervised SVM provide a high correction rate, but high false positive alarms.

A multi-level hybrid classification model combining DT and Bayesian Network (BN) clustering has been proposed by Xiang, (2008). Their results on KDD CUP'99 dataset, a benchmark dataset used to evaluate IDS, were compared with other popular approaches such as multi-level tree classifier and winners of KDD CUP'99. It was shown that this new approach is very efficient in detecting intrusions with an extremely low false-negative rate of 3.23%, while keeping an accepTable level of false-positive rate of 3.2%. Although the false positive rate was reported to be as low as 0.5% for the KDDCUP'99 winner (Pfahringer, 2000), the corresponding false negative rate was merely 9.1%, which is much higher than their result (3.23%).

Zanero and Savaresi, (2004) stated that the problem of IDS does not lie only in the sheer number of vulnerabilities that are discovered every day. They claimed that there are also an unknown number of unexposed vulnerabilities that may not be immediately available to the experts for analysis and inclusion in the knowledge base. In order to overcome this problem, they introduced an unsupervised anomaly detection based on clustering. They satated that their approach increase the detection rate of different kinds of unknown attacks

In most circumstances, labeled data or purely normal data is not readily available since it is time consuming and expensive to manually classify it. Purely normal data is also very hard to obtain in practice, since it is very hard to guarantee that there are non intrusions when they were collecting network traffic (Leung and

Leckie, 2005). They stated that to address these problems, they used a new type of intrusion detection algorithm called unsupervised anomaly detection.

Data reduction can be achieved by filtering, data clustering and feature selection (Chebrolu *et al*., 2004). Generally, the capability of anomaly intrusion detection is often hindered by the inability to accurately classify a variation of normal behavior as an intrusion. Additionally, network traffic data is huge, and it causes a prohibitively high overhead and often becomes a major problem in IDS (Sung and Mukkamala, 2004).

According to Chakraborty, (2005), the existence of these irrelevant and redundant features generally affects the performance of machine learning or pattern classification algorithms. Hassan, *et al*., (2003) proved that proper selection of feature set has resulted in better classification performance. (Sung and Mukkamala, 2004) have demonstrated that the elimination of these unimportant and irrelevant features did not significantly lowering the performance of IDS.

Chebrolu *et al*., (2004) tackled the issue of effectiveness of an IDS in terms of real-time and detection accuracy from the feature reduction perspective. In their work, features were reduced using two techniques, Bayesian Network (BN) and Classification and Regression Trees (CART). They have experimented using four sets of feature subset which are 12, 17, 19 and all the variables (41) from one network connection. Dataset used was KDD CUP 99. They suggested that using the full 41 features do drop the detection rate of the IDS.

1.3 Problem Statement

In most circumstances, labeled data is not available since it is time consuming and very expensive to classify it manually. Meanwhile there is always a dynamic change in traffic patterns as well as an emerging of novel attacks. All these problems make the supervised approach unpractical solution for IDS. To solve these problems, researchers move to work on unsupervised approach, namely clustering as this unsupervised approach does not depend on the labeled data. Various unsupervised methods have been proposed but the detection rate of IDS is rather inadequate in comparison to supervised approaches.

1.4 Research Question

To cater the problems stated in secion 1.3 this study was carried out to answer the following questions:

1. How to reduce the dimension of input data or how to choose the features that best represent all types of attacks?
2. How Immune Network clustering method perform in detecting new attack types?
3. How to evaluate the performance of the Immune Network method for clustering?

1.5 Project Hypothesis

The Feature Reduction using Rough Set method and the Immune Network clustering for network-based IDS would yield a good performance in terms of detection accuracy, speed and identifying new attack types.

1.6 Project Aim

The aim of this project was to reduce the data features and classify /cluster the data such that it gives a lower false positive rate, improves accuracy, speed, and improves the ability of classifying new attack types of unknown dataset.

1.7 Objectives

To accomplish the aim of this project, a few objectives have been identified:

1. To improve the accuracy and speed of detecting attacks by selecting significant data features that best represent all attack types.
2. To investigate the ability of Immune Network method for clustering normal traffic from attacks thus detection rate for unknown attacks can be further improved.
3. To compare the performance of the Immune Network clustering with a traditional clustering method, namely K-Means on the test data.

1.8 Project Scope

The scope of this project is defined as follows:

1. The dataset that was used in this study is the KDD Cup 1999, a common benchmark for evaluation of intrusion detection techniques.
2. The feature reduction model was developed using Rough Set Theory to get the most significant features that best represent all attack types.
3. The Immune Network model was adopted using the original Immune Network algorithm (aiNet).
4. Evaluation / Comparison was done with a traditional clustering method, namely K-Means.

1.9 Significance of the Project

This study evaluates the performance of feature reduction and Immune Network clustering for the network-based IDS in terms of accuracy, speed of detection and the ability to classify new attack types in the test dataset. The results were compared with K-Means clustering method to see whether this approach can give better performance. The result of the study was contributed to the identification of an improved approach for IDS. This new approach could be used to develop a methodology that will be valuable in future studies of anomaly based IDS improvements.

REFERENCES

Allen, J., Christie, A., Fithen, W., McHugh, J., Pickel, J., and Stoner, E. (2000). State of the Practice of Intrusion Detection Technologies. Tech. rep., *Carnegie Mellon University January.*

Ayara, M., Timmis, J., de Lemos, R., de Castro, L.N., Duncan, R. (2003).Negative Selection: How to Generate Detectors. *Technical Report, Computing Laboratory.University of Kent at Canterury*. U.K.

Bivens, A., Palagiri, C ., Smith, R., Szymanski, B ., and Embrechts, M. (2002). Network-Based Intrusion Detection using Neural Networks. *Rensselaer Polytechnic Institute Troy,* New York 12180-3590.

Burbeck, K. and Nadjm-Tehrani,S. (2005). ADWICE – Anomaly Detection with Real-Time Incremental Clustering. LNCS 3506, pp. 407–424, Springer. *In proceedings of the 7th International Conference on Information Security and Cryptology,* Seoul, Korea.

Bello, R. , Nowe, Y. Caballero, Y. Gomex, and P. Vrancx (*2005*), A Model Based on Ant Colony System and Rough Set Theory to Feature Selection. *GECCO'05, June 25-29,* Washington DC, United States. Pp. 275-276.

Chandola,V., Banerjee, A., and Kumar V. (2007). Anomaly Detection: A Survey. *Technical Report, TR 07-017*.

Chandola,V., Banerjee, A., and Kumar V. (2009). Anomaly Detection: A Survey. *ACM Computing Surveys, Vol. 41(3).*

Chakraborty, B. (2005). Feature Subset Selection by Neuro-rough Hybridization. *LNCS, Springer Hiedelberg,* pp. 519-526.

Chebrolu, S ., Abraham, A. and Thomas, J.P. (2004). Feature Deduction and Ensemble Design of Intrusion Detection Systems. *International Journal of Computers and Security* .Vol 24, Issue pp. 295-307.

Chimphlee, W., Abdul Hanan Abdullah, Mohd Noor Md Sap and Chimphlee, S (2005). Unsupervised Anomaly Detection with Unlabeled Data using Clustering, *Proc. Int. Conf. on ICT- Mercu Buana ICT2005*. pp. 42-49.

Dasgupta, D., Ji, Z., and González , F. (2003). Artificial Artificial Immune System (AIS) Research in the Last Five Years. *Proceedings of The 2003 Congress on Evolutionary Computing.*

De Castro, L.N., Zuben ,F. (1999). Artificial Immune System: part I -Basic theort and aoolications. *Technical Report TR – DCA 01/99.*

De Castro, L.N and Zuben, F.J. (2000). The Clonal Selection Algorithm with Engineering Applications. *In Workshop Proceedings of GECCO, pp. 36-37, Workshop on Artificial Immune Systems and Their Applications,* Las Vegas , USA.

De Castro, L.N., and Zuben, F., (2001). aiNet: An Artificial Immune Network for Data Analysis . *Idea Group Publishing*, USA.

De Castro, L.N. and Timmis, J. (2002). An Artificial Immune Network for Multimodal Function. *Proceedings of the 2002 Congress on, Vol.1, pp. 699-704.*

De Castro,L.N., and Timmis, J. (2002). Artificial Immune Systems: A Novel Paradigm to Pattern Recognition. *In Artificial Neural Networks in Pattern Recognition,*University of Paisley, UK, pp. 67-84.

De Castro L.N., Timmis, J. (2003). Artificial immune systems as a novel soft computing paradigm. *Soft Computing 7 526–544 Springer-Verlag.*

Denning, DE. (1987). An intrusion-detection model. *IEEE Transactions on software engineering.*

Depren, O., Topallar, M., Anarim, E. and Ciliz, M.K. An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks, *Expert Systems with Applications*, Volume 29, Issue 4 , pp. 713-722, 2005.

Duda, R., Hart, E. ,.Tork, D.G. (2001): Pattern classification. *(2nd Edition)* (9780471056690): Richard O. Duda, Peter E. Hart, David G. Stork: Books.

Eick,C., and Zeidat N., (2005). Using Supervised Clustering to Enhance ClassifierS. *The15th International Symposium on Methodologies, Springer.*

Eskin, E., W. N. Grundy, and Y. Singer (2000). Protein family classification using sparse Markov transducers. *In Proc. of the Eighth Intl. Conf. on Intelligent Systems for Molecular Biology,* Menlo Park, CA, pp. 134–145. AAAI Press.

Eskin, E., Arnold, A., Prerau, M., Portnoy, L., and Stolfo, S.(2002). A geometric framework for unsupervised anomaly detection: detecting intrusions in unlabeled data. *In Applications of Data Mining in Computer Security.* Kluwer.

Fearon, D. T. and Locksley, R. M. (1996). The Instructive Role of Innate Immunity in the Acquired Immune Response, Science 272: 50-53.

Finley,T , Joachims, T. (2005). Supervised Clustering with Support Vector Machines*.* *Proceedings of the 22 nd International Conference on Machine Learning,* Bonn, Germany.

Forrest, S., Perelson, A.S., Allen, L. and Cherukuri, R. (1994) Self-nonself discriminatioin a computer. *Proceedings of IEEE Symposium on Research in Security and Privacy*. 16(18):202-212

Forest. S., Hofmeyr. S. A., Somayaji. A. (1997). Computer Immunology.

Fred, A.L.N., and Jain, A.K. (2005): Combining multiple clusterings using evidence accumulation. *IEEE Transactions on Pattern Analysis and Machine Intelligence 27(6).*

Fawcett, Tom (2006). An introduction to ROC analysis. *Pattern Recognition Letters 27, 861–874.* Elseiver.

Gao, J., Chengy, H., Tan, P.N. (2006). A Novel Framework for Incorporating Labeled Examples into Anomaly Detection. *In Proc of SDM'06: SIAM Int'l Conf. on Data Mining, Bethesda, MD, Apr 20-22.*

Gerhard Münz, Sa Li and Georg Carle. (2007) Traffic Anomaly Detection Using k-means Clustering. *In GI/ITG Workshop MMBnet*.

Estevez-Tapiador, J. M. , Teodoro,GP.,and Diaz-Verdejo J.E (2004). Anomaly detection methods in wired networks: a survey and taxonomy, *Computer Communications vol. 27, pp. 1569-1584*, 15 October 2004.

Ghali, N. (2009). Feature Selection for Effective Anomaly-Based Intrusion Detection. *IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.3.*

Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. *Addison-Wesley, Reading, M.A.*

González ,F.A., and Dasgupta,D. (2003). Anomaly detection using real-valued negative selection. *Genetic Programming and Evolvable Machines, Vol. 4 , No. 4, pp. 383-403.*

Grira, N., Crucianu, M., Boujemaa, N. (2005). Unsupervised and Semi-supervised Clustering: a Brief Survey. *in A Review of Machine Learning Techniques for Processing Multimedia Content, Report of the MUSCLE European Network of Excellence (6th Framework Programme).*

Hassan, A., Nabi Baksh, M.S. Shaharoun, A.M. and Jamaluddin, H. (2003).Improved SPC Chart Pattern Recognition Using Statistical Feature. *International Journal of Production Research 41(7), , pp. 1587-1603.*

Huang,W. and Jiao ,L. (2008). Artificial immune kernel clustering network for unsupervised image segmentation. *Elseiver Progress in Natural Science 18 455–461*

Hunt, J.E and Cooke, D.E.. (1995). Recognizing promoter sequences using an Artificial Immune System. *In Proceedings of Intelligent Systems in Macular Biology, IEEE Press*. 81-86.

Ingham K. (2003). Protecting network servers.Technical Report, *Department of Computer Science,* University of New Mexico.

Janeway, A. (1993). How the Immune System Recognizes Invaders. *Scientific American. 41-47.*

Jerne, N. K. 1974. Network theory of the immune system. 1974. Ann. Immunol. (Paris). 125c: 373..

Jiang, S., Song, X. , Wang, H. , Han, J. And Li , Q. (2006). *A* clustering-based method for unsupervised intrusion detections. *Elseiver Pattern Recognition Letters 27 802–810.*

Kabiri, P. , and Ghorbani, A. A (2005). Research on intrusion detection and response: A survey. *In International Journal of Network Security, volume 1.*

Kate,R.J., and Mooney, R.J. (2007). Semi-Supervised Learning for Semantic Parsing using Support Vector Machines. *In Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Companion Volume, Short Papers. Rochester,* New York, pp. 81-84. Association for Computational Linguistics.

Kim, B. J. and Kim, I. K. ,*(2005)*. Machine Learning Approach to Realtime Intrusion Detection System. *In Proceedings of 18 th .Australian Join Conference on Artificial Intelligence*, Sydney, Australia. Vol. 3809. Pp. 153-163.

Kim ,J. and Bentley,P. (1999). An Evaluation of Negative Selection in an Artificial Immune System for Network Intrusion Detection. *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO).*

Kim, DS., Nguyen, HN., Park, JS. (2005). Genetic algorithm to improve SVM based network intrusion detection *.Information Networking and Applications*, 2005.

Kim ,JW. (2002). Integrating Artificial Immune Algorithms for Intrusion Detection. *PhD thesis, University College London*.

Kim, D. and Park ,J. (2003). Network-Based Intrusion Detection with Support Vector Machines. *Lecture notes in computer science, 747-756, Springer.*

Kim, J. and Bentley, P. (2002). Towards an artificial immune system for network intrusion detection: an investigation of dynamic clonal selection. *In: Proceeding of the Congress on Evolutionary Computation (CEC-2002),* Honolulu, Hawaii.

Kim, J. and Bentley, P. (2004). Immune memory and gene library evolution in the dynamical clonal selection algorithm. *Journal of Genetic Programming and Evolvable Machines 5(4): 361–391.*

Knidel ,H., de Castro, L.N., and Zuben, F.J. (2005): a realvalued antibody network for data clustering. *In: GECCO'05, Washington, DC, USA, p. 371–2.*

Laskov, P., SchÄafer, C., and Kotenko, I.(2004). Intrusion detection in unlabeled data with quarter-sphere support vector machines. *In: Proc. DIMVA. 71{82}.*

Laskov, P., SchÄafer, C., Kotenko, I., and MÄuller, K.R.(2004). Intrusion detection in unlabeled data with quarter-sphere support vector machines (extended version).

Laskov,P., Däussel, P., SchÄafer, C. and Rieck, K. (2005). Learnin intrusion detection:supervised or unsupervised?. *Lecture Notes in Computer Science, 2005 – Springer.*

Lazarevic, A. (2005). Data Mining for Intrusion Detection. *Encyclopedia of Data Warehousing and Mining, Idea Group.*

Leung, K. and Leckie, C. (2005). Unsupervised Anomaly Detection in Network

Intrusion Detection Using Clusters. *Appeared at the 28th Australasian Computer Science Conference,* The University of Newcastle, Australia.

Lichodzijewski P. (2002). Network based anomaly detection using self organizing maps, *Technical Report, Nova Scotia*: Dalhousie University Halifax.

Li, Z. and H.-Z. Tan (2006), A Combinational Clustering Method Based on Artificial Immune System and Support Vector Machine, In KES 2006, Part I, LNAI 4251, Gabrys, B., Howlett, R.J. and L.C. Jain (Eds.), *Springer-Verlag Berlin Heidelberg, 153-162.*

Li Ming(2004). An approach for reliably identifying signs of DDoS flood attacks based on LRD traffic pattern recognition. *Computers Security;23(7):549e58.*

Lu, W. and Traore,I. (2005). A New Unsupervised Anomaly Detection Framework for Detecting Network Attacks in Real-Time. *The 4th international conference of Cryptography and Information Security(CANS 2005).*

Mandala, S., Asri Ngadi, Md., Hanan Abdullah, A.,(2008). A Survey on MANET Intrusion Detection. *International Journal of Computer Science and Security, Volume (2), Issue (1).*

Mosqueira, E. , Alonso-Betanzos, A., del R´ıo, B. , and Lago Pi˜neiro,J. (2007). A Misuse Detection Agent for Intrusion Detection in a Multi-agent Architecture. *Lecture Notes in Computer Science, Springer.*

Ohrn, A., and Komorowski, J. (*1997*). A Rough Set Toolkit for Analysis of Data, *In Proceedings of the third Joint conference on Information Sciences, Vol(3), pp.403- 407, USA.*

Panda, M., and Patra, M. (2005).SOME Clustering Agorithm to enhance the performance of network intrusion detection system. *Journal of Theoretical and Applied Information Technology.*

Pawlak, Z. (1991). Rough Sets: Theoretical Aspects of Reasoning about Data, *Kluwer Academic Publishers*.

Pfahringer Bernhard (2000). Winning the KDD99 classification cup:bagged boosting. *ACM SIGKDD Explorations Newsletter*. Volume 1 , Issue 2 (January 2000).

Portnoy, L., Eskin, E., and Stolfo, S.(2001). Intrusion detection with unlabeled data using clustering. *In: Proc. ACM CSS Workshop on Data Mining Applied to Security*.

Powers,S., and He , J. (2008). A hybrid artificial immune system and Self Organising Map for Network Intrusion Detection. *Information Sciences. Elseiver. 178 3024–3042.*

Quinlan, J.(1992). C4.5: Programs for Machine Learning. Morgan Kaufmann.

Ramos,V. , and Abraham, A. (2006). ANTIDS: Self-Organized ANT-Based Clustering model for Intrusion Detevtion System. *Arxiv preprint cs/0412068, arxiv.org.*

Rojas, R.(1996). Neural Networks: A Systematic Approach. *Springer-Verlag,* Berlin, Deutschland.

Shin, M., Mi, E., Seon,K. , and Park, H. (2003). Automatically Finding Good Clusters with Seed *k*-Means. *Bioinformatics Research Team, Electronics and Telecommunications Research Insti- tute, 161 Gajeong-dong, Yuseong-gu,* Daejeon, 305-350, Korea.

Shon,T., Kim, Y., Lee, C., Moon, J. (2005). Machine learning framework for network anomaly detection using svm. *Information Assurance Workshop*, IAW'05

SNORT. URL*: http://www.snort.org. Last accessed 07/05/2006. 2006*

Soleymani Baghshah, M. and Bagheri Shouraki, S. (2008). A Novel Semi-supervised Clustering Algorithm for Finding Clusters of Arbitrary Shapes. *13th International CSI Computer Conference, CSICC 2008 Kish Island, Iran.*

Soleymani Baghshah, M., and Bagheri Shouraki, S, (2008).:A fuzzy clustering algorithm for finding arbitrary shaped clusters. In: Soleymani Baghshah, M., Bagheri Shouraki, S..(eds.) *6th ACS/IEEE International Conference on Computer Systems and Applications.*

Spafford, E. H. and Zamboni, D.(2000). Intrusion detection using autonomous agents. *Computer Networks,34(4):547–570.*

Srinoy, S., Kurutach, W., Chimphlee, W., and Chimphilee, S. (*2005*). Network Anomaly Detection Using Soft Computing, *Proceedings of World Academy of Science, Engineering and Technology, vol. 9, pp.140-144,.*

Stibor, T., and Timmis , J.(2007). An Investigation on the Compression Quality of aiNet *.In: Fogel D (ed) Proc of foundations of computational intelligence.*

Sung, A.H. and Mukkamala, S. (2003). Identifying Important Features for Intrusion

Detection Using Support Vector Machines and Neural Networks. *In Proceedings of the Symposium on Applications and Internet (SAINT'03), pp. 209-216.*

Sung, A.H. and Mukkamala, S. (2004). The Feature Selection and Intrusion Detection Problems. *ASIAN 2004. LNCS, vol. 3321, Springer Hieldelberg, pp. 468-482.*

Tan, L. and Sherwood, T.,(2005) A High Throughput String Matching Architecture for Intrusion Detection and Prevention, *ISCA* 2005.

Tang,W. , Xiong, H., Zhong,S. ,and Wu ,J. (2007). Enhancing Semi-Supervised Clustering: A Feature Projection Perspective. *International Conference on Knowledge Discovery and Data Mining. Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining* San Jose, California, USA.

Taylor, D. W., Corne, D. W. (2003). An investigation of the Negative Selection Algorithm for Fault detection in Refrigeration System. *ICARIS 2003. Springer-Verlag Berlin Heidelberg.* LNCS 2787: 34-45

Teng, H., S., Chen, K. and Lu, S., C. Security Audit Trail Analysis Using Inductively Generated Predictive Rules. *In Proceedings of the 11th National Conference on Artificial Intelligence Applications, pages 24-29, IEEE*, IEEE Service Center, Piscataway, NJ, March 1990.

Teodoro,G, Verdejo, Dı´az. , J., Macia´-Ferna´ndez, G., and Va´zquez, E. ,(2009). Anomaly-based network intrusion detection: Techniques, systems and challenges. *0167-4048 Elsevier Ltd.*

Timmis, J., Neal, M., and Hunt, J. (2000). An Artificial Immune System for Data Analysis. *Journal of the BioSystem.* 55: 143-150.

Timmis, J. I. (2001). Artificial Immune System A Novel Data Analysis Technique Inspired by the Immune Network theory. *PHD Thesis. University of Wales*. Aberystwyth, .

Timmis, J., Neal, M. (2001). A Resource Limited Artificial Immune System for Data Analysis. *Journal of Knowledge Based System*. 14 (3-4): 121-130.

Wang , L. ,Wang ,L. and Nie,Y., (2006). A Novel Artificial Immune Network Model and Analysis on Its Dynamic Behavior and Stabilities. *ICNC 2006, Part II, LNCS 4222, pp. 83 – 9. Springer-Verlag Berlin Heidelberg..*

Wroblewski, J., 1995. Finding Minimal Reducts Using Genetic Algorithms.

Proc. of the second Annual Join Conference on Information Sciences. Wrightsvillle Beachm, NC, p.186-189.

Xiang Cheng, Yonga Png Chin and Swee Menga Lim (2008). Design of multiple-level hybrid classifier for intrusion detection system using Bayesian clustering and decision trees . *Elsevie Pattern Recognition Letters*Volume 29, Issue 7, 1 May 2008, Pages 918-924

Xiao, R. B., Wang, L., Liu, Y. (2002). A Framework on AIS Based Pattern Classification and Matching for Engineering Creative Design. *IEE Proceedings of the First International Conference on Machine Learning and Cybernatics.* 4-5 September:1554-1558

Ye, N. and Li, X. , (2001). A Scalable Clustering Technique for Intrusion Signature Recognition. *Proceedings of the 2001 IEEE Workshop on Information Assurance and Security United States Military Academy, West Point, NY.*

Yasami, Y., Khorsandi,S., Mozaffari,SP., and Jalalian ,A., (2008).An unsupervised network anomaly detection approach by k-Means - *IEEE Symposium on Computers* .

Zainal, A., Maarof, M.A. and Shamduddin, S.,(2006). "Feature selection using rough set in intrusion detection", *in Proc. IEEE TENCON, p.4.*

Zainal,A., Maarof ,M.A. and Shamsuddin,S. ,(2006). Research Issues in Adaptive intrusion detection. *PARS06 FSKSM,UTM,* Skudai.

Zainal, A., Maarof M.A., Shamsuddin S.,(2007), Feature Selection using Rough-DPSO in Anomaly Intrusion Detection*, ICCSA 2007, LNCS* 4705.

Zanero, S. and Savaresi ,S, (2004). Unsupervised learning techniques for an intrusion detection system *.In proceedings of the 2004 ACM symposium on Applied computing,* Nicosia, Cyprus.

Zanero, S.,(2005). Improving Self Organizing Map Performance for Network Intrusion Detection. *In: SDM 2005 Workshop on "Clustering High Dimensional Data and its Applications".*

Zhang, L., Zhang, G. , Yu , L., Zhang , J., and Bai,Y ., (2004). Intrusion Detection Using Rough Set Classification. *Journal of Zheijiang University Science. pp. 1076-1086.*

Zhao, Z., Guo,S., Xu,Q., and Ban,T., (2009).G-Means: A Clustering Algorithm for Intrusion Detection. *M. Köppen et al. (Eds.): ICONIP Part I, LNCS 5506, pp. 563–570. Springer-Verlag Berlin Heidelberg.*