Int. J. Advance. Soft Comput. Appl., Vol. 2, No. 3, November 2010
ISSN 2074-8523; Copyright © ICSRS Publication, 2010
www.i-csrs.org

Intrusion Detection System Using Unsupervised Immune Network Clustering with Reduced Features

Murad Abdo Rassam[1,2], Mohd. Aizaini Maarof [2], and Anazida Zainal [2]

[1]Faculty of Engineering and Information Technology
Taiz University
Taiz, Yemen
eng.murad2009@gmail.com

[2]Faculty of Computer Science and Information Systems,
Universiti Teknologi Malaysia,
81310, Skudai, Johor, Malaysia.
aizaini, anazida@utm.my

Abstract

Intrusion Detection Systems (IDS) are developed to be the defense against security threats. Current signature based IDS like firewalls and anti viruses, which rely on labeled training data, generally cannot detect novel attacks. The purpose of this study is to enhance the detection rate by reducing the network traffic features and to investigate the feasibility of bio-inspired Immune Network approach for clustering different kinds of attacks and some novel attacks. Rough Set method was applied to reduce the dimension of features in DARPA KDD Cup 1999 intrusion detection dataset. Immune Network clustering was then applied using aiNet algorithm to cluster the data. Empirical study revealed that detection rate was enhanced when most significant features were used to represent input data. The finding also revealed that Immune Network clustering method is robust in detecting novel attacks in the absence of labels.

Keywords: ***Feature Reduction; Artificial Immune Network; Intrusion Detection System.***

1 Introduction

Due to the increasing use of computer networks in many aspects of life, the number of vulnerabilities also is increasing and causing the network resources unavailable and violates the system confidentiality, integrity and availability. 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 [1]. Its main goal is to differentiate between normal activities of the system and behaviors that can be classified as intrusive.

There are two main intrusion detection approaches: anomaly intrusion detection system and misuse intrusion detection system. The anomaly detection focuses on the unusual 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.

Various researchers [2, 16, 18] have treated IDS as pattern recognition problem or rather classified as learning system. An appropriate representation space for learning by selecting relevant attributes to the problem domain is an important issue for learning systems [2, 17] as irrelevant and redundant features may lead to complex classification model and reduce accuracy [17].

Bello *et al*. in [3] suggested that feature reduction was necessary to reduce the dimensionality of training dataset. They claimed that feature reduction also enhanced the speed of data manipulation and improved the classification rate by reducing the influence of noise.

In literature, numbers of anomaly detection systems were developed based on different machine learning techniques. For example, some studies apply single learning techniques, such as neural networks [19], genetic algorithms [20], support vector machines [21], bio-inspired algorithms [22] and many more.

On the other hand, some systems [9, 24, 25] 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. Many computing models have been introduced to solve complex anomaly detection systems for better solutions such as biological-based computing.

Computing models inspired by biology are a way to make use of concepts, principles and mechanisms underlying biological systems. Some biologically inspired techniques are 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 [23].

This paper describes the anomaly detection system using reduced network traffic features. Rough Set theory was used to reduce the feature dimension and Immune Network was used to detect novel attacks.

The rest of this paper is structured as follows. Section 2 gives a background on the techniques used in this study which are Rough Set Theory and Artificial Immune Network. Section 3 describes the related works in both areas namely, feature reduction and unsupervised Immune Network Clustering. Section 4 describes the experiments using the KDD CUP 99 dataset and the results obtained. It also includes an analysis of the results and performance comparison against k-Means technique. Finally, section 5 concludes the study.

2 Background

2.1 Rough Set Theory

Zhang et al., in [4] defined Rough Set Theory as a mathematical tool for approximate reasoning for decision support and is particularly well suited for classification of objects. They stated that, it can also be used for feature reduction and feature extraction. The most attractive characteristics of Rough Set theory is that it deals with inconsistencies, uncertainty and incompleteness of data instances by determining an upper and a lower approximation to set membership. It has been successfully used in the literature as a selection tool to discover data dependencies, find out all possible feature subsets, and remove redundant information. The following definitions as given in [5] show how to derive reducts.

Definition 1:

An information system is defined as a four-tuple as follows, $S = \prec U, Q, V, f \succ$, where U=$\{x_1, x_2, ..., x_n\}$ is a finite set of objects (n is the number of objects); Q is a finite set of attributes, Q=$\{q_1, q_2, ..., q_n\}$;V= $U_{q \varepsilon Q} V_q$ and V_q is a domain of attribute q ; f: $U \times Q \rightarrow V$ is a total function such that $f(x, q) \in V_q$ for each $q \in Q, x \in U$. If the attributes in S can be divided into condition attribute set C and decision attribute set D, i.e. $Q=C \cup D$ and $C \cap D=\Phi$, the information system S is called a decision system or decision Table.

Definition 2: Let *IND(P), IND(Q)* be indiscernible relations determined by attribute sets *P, Q,* the *P* positive region of *Q*, denoted *POS IND(P) (IND(Q))* is defined as follows: $POS_{IND(P)} (IND (Q)) = U_{X \in U/IND(Q)} / IND (P)- (X)$.

Definition 3: Let *P, Q, R* be an attribute set, we say *R* is a reduct of P relative to *Q* if and only if the following conditions are satisfied:

(1) $POS_{IND(R)} (IND (Q)) = POS_{IND(P)} (IND (Q))$;

(2) For every $r \in R$ follows that

$POS_{IND(R-\{r\})} (IND (Q)) \neq POS_{IND(R)} (IND (Q))$

2.2 Artificial Immune Network

Jerne [6] proposed a network theory for the immune system where it has been widely used in the development of Artificial Immune System (AIS) [7]. This theory suggests that for each antibody molecule, there is a portion of their receptor that can be recognized by other antibody molecules. As the results, a network communication can occur within the immune system, and it is called as Immune Network.

According to de Castro and Timmis [8], the recognition of antigen by an antibody results in network activation, whereas the recognition of an antibody by another antibody results in network suppression. Antibody Ab_2 is said to be the internal image of the antigen Ag, because Ab_1 is capable of recognizing the antigen and also Ab_2. According to the Immune Network theory, the receptor molecules contained in the surface of the immune cells present markers, named idiotopes, which can be recognized by receptors on other immune cells [8], (Fig. 1).

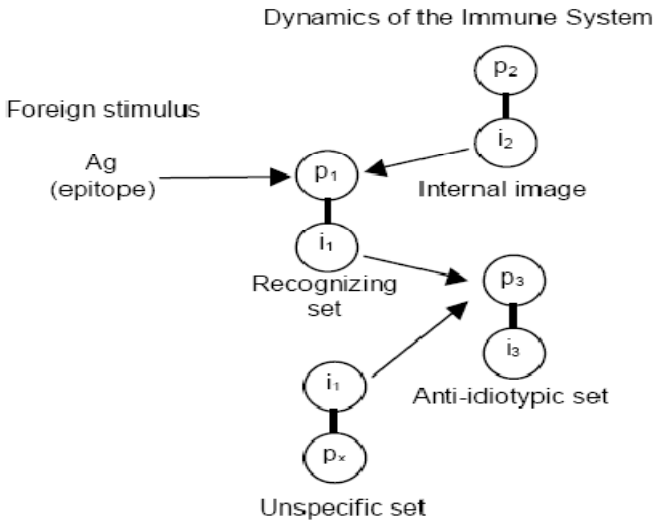

Fig 1: A view on idiotypic Immune Network (de Castro and Zuben, 2001).

Artificial Immune Network is a dynamic unsupervised learning method. The Artificial Immune Network model consists of a set of cells called antibodies interconnected by links with certain strengths. These networked antibodies (idiotypic network) represent the network internal images of pathogens (input patterns) contained in the environment which it is exposed. The algorithm of Immune Network aiNet is given below:

1. Load antigen population.
2. Initialize the Immune Network by randomly selecting an antigen from antigen population as a seed for each cluster.
3. While the termination condition is not true:
    a. For each antigen pattern in the antigen population
        i. Present an antigen to the network
        ii. Determine the affinity of each antibody in each cluster to the antigen
        iii. Select the n highest affinity antibodies from the network
        iv. For each of these highest affinity antibodies:
            - If its affinity is greater, then the affinity threshold is σ
            - Reproduce the antibody proportionally to its affinity
            - Each clone undergoes a mutation inversely proportional to its affinity
            - Increase the fitness of those antibodies
        v. End for

vi. If none of the highest affinity antibodies could bind the antigen then generate a new cluster by using the antigen as a seed.

b. End for
c. Compute the affinity between antibody-antibody within each cluster and do suppression.
d. Calculate affinity between cluster-cluster and do suppression.
e. Delete the antibodies in each cluster whose fitness is less than a threshold σ.

4. End while
5. Output each cluster in the network
6. Output each cluster in the network

3 Related Works

3.1 Feature Reduction in IDS

Most of IDS examine all features of dataset to detect intrusions [9]. Some of the features may be redundant or somehow contribute little to the detection process. The purpose of this phase of the study is to identify important input features in the IDS dataset that contribute to the efficiency and the effectiveness of our proposed model.

Chebrolu *et al*., in [10] have investigated the performance of two feature reduction algorithms involving Bayesian networks (BN) and Classification and Regression Trees (CART) and an ensemble of BN and CART. Their results indicated that input feature reduction is important to design an IDS that is efficient and effective for real world detection systems. Zhang *et al*. in [4] investigated the use of rough set theory and its capability of getting classification rules to determine the category of attack in IDS. In their work, they did not show the features that were implemented in the classification process.

Data reduction can be achieved by filtering, data clustering and feature selection [10]. 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 [11].

According to Chakraborty in [12], the existence of these irrelevant and redundant features generally affects the performance of machine learning or pattern classification algorithms. Hassan, *et al*., in [13] proved that proper selection of feature set has resulted in better classification performance. Sung and Mukkamala [11] have demonstrated that the

elimination of these unimportant and irrelevant features did not significantly lowering the performance of IDS.

3.2 Unsupervised Immune Network Clustering

Zanero and Savaresi [14] 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 stated 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 no intrusions when they were collecting network traffic [15]. To address these problems, an unsupervised anomaly detection approach using artificial immune network was used to show the ability of this bio-inspired algorithm to adapt and cluster normal and attacks data without any prior knowledge.

4 Experiments and Results

This section gives the dataset description, and the experiments done by this study. The experiments have been done in two phases, the first phase; the feature reduction using rough set was implemented on the data samples. The result of this phase was the feature subset which later used as input to the second phase, immune network clustering. The general framework of this study is shown in the following figure.

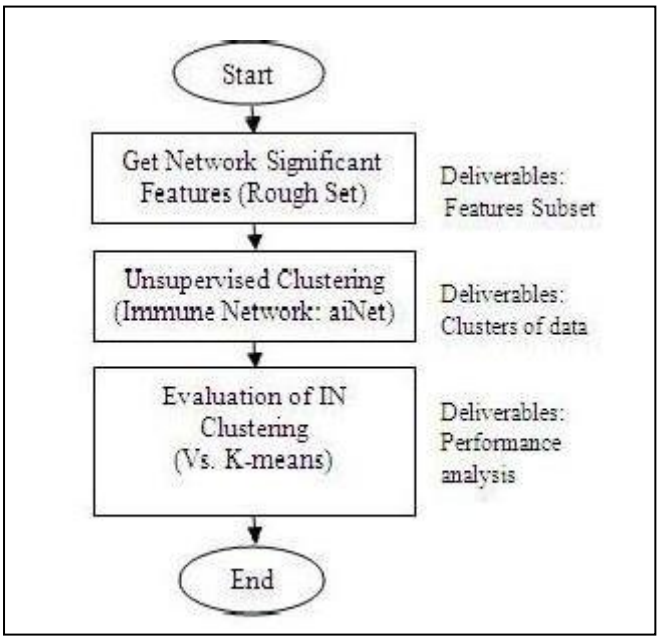

Fig 2: The General Framework of the study

4.1 Dataset

The dataset that was employed in this study, the KDD Cup 1999, is a common benchmark dataset usually used by many researchers for evaluation of intrusion detection techniques.

The original dataset contain 744 MB data with 4,940,000 records. However, most of researchers dealt only with a small part of the dataset (10% percent) which have been chosen for conducting experiments on this dataset. The 10% of the data contains 494021 records**.** The dataset has 41 features for each connection record plus one class label. Some features are derived features, which are useful in distinguishing normal connection from attacks. These features are either nominal or numeric.

There are 4 main categories of attacks in the KDD CUP dataset. A brief description of each class is given in the subsequent sections.

a. **Denial-of-service attack**: is a class of attacks where an attacker makes some computing or memory resource too busy or too full to respond to requests.
b. **Probing:** is a class of attacks where an attacker scans a network to get some information about potential vulnerabilities in the network.
c. **User to Root Attacks**: is a class of attacks where an attacker gets an access to a normal user account on the system to get a root user access to the system later.

d. **Remote to User Attacks** is a class of attacks where an attacker sends some packets to a system over a network remotely, then it gets some information about the potential vulnerabilities in this system.

4.2 Feature Reduction

Three different samples of the dataset were used, each of which contains 10,000 instances. The distribution of data and the number of instances for each class in these samples are shown in Table 1.

Table 1: The distribution of attacks in the data samples

Normal	Probe	DoS	U2R	R2L
2000	684	6907	34	375

After data samples preparation, Rough Set operations were applied on these data samples. Rough Set was implemented using ROSETTA (Rough SET Toolkit for data Analysis) system developed by Ohrn [16].

The experimental steps can be summarized as follows: First, the raw data samples were transformed into Tables recognized by ROSETTA. After the preprocessing of data samples, each data sample was split into two parts: the training dataset and the testing dataset based on the splitting factor determined by the user (i.e. split factor is 0.4 means that 40% of the data sample for training and the remaining 60% for testing). Many algorithms can be used to reduce the data samples i.e. GA, Johnson Holte1R, and Dynamic algorithms. The GA algorithm was used to reduce the data sample features in this study. We are interested in GA, because according to Ohrn [16]; It is used to find minimal hitting sets and it gives less number of reducts as compared to Johnson's algorithm. The set of reducts obtained in the third step was used to generate the rules using the GA built in algorithm in ROSETTA system. These rules were used later to classify the other part of data sample which is the testing part. After a number of experiments, the most 8 significant features were obtained and shown in the following table.

Table 2: The most 8 significant features on three different data samples

Data Sample	**8 most significant features**							
Sample 1	C	E	F	Y	AD	AF	AG	AI
Sample 2	C	E	F	W	AG	AF	AH	AJ
Sample 3	C	E	F	Y	W	AE	AI	AN

Table 2 suggests that all samples shared 3 common features and the rest varies in the number of occurrences in each sample. Features C, E, and F are common in all samples. Features AF, and AG are common between sample1 and sample2. Features Y and AI are common between sample1 and sample3. Feature W is common between sample2 and sample3. According to this commonality we found that the most 8 significant features in the three samples and in the whole dataset are shown in the following table.

Table 3: The most 8 significant features

C	E	F	W	Y	AF	AG	AI

The corresponding network features and the description of each feature is shown in table 4.

Table 4 : The corresponding network features and its description

Feature label	**Corresponding Network Feature**	**Description of feature**
C	Service	Type of service used to connect (e.g. fingure, ftp, Telnet, SSh, etc.).
E	Src_bytes	Number of bytes sent from the host system to the destination system.
F	Dst_bytes	Number of bytes sent from the destination system to the host system.
W	Count	Number of connections made to the same host system in a given interval of time
AF	Dst_host_count	Number of connections from the same host to destination during a specified time window.
AG	Dst_host_srv_ count	Number of connections from the same host with same service to the destination host during a specified time window.
AI	dst_host_diff_ srv_rate	Number of connections to different services from a destination host.

Beside its use in feature reduction, Rough Set was also applied to classify the data to evaluate the performance of the classification for pre-reduct and post-reduct features. The results are shown in Table 5.

Table 5: The classification accuracy on three different samples using all 41 features

Type	Sample 1	Sample 2	Sample 3	Mean	StdDev
Normal	92.82%	95.17%	87.99**%**	91.99%	0.036
Probe	94.34%	100%	99.29%	97.88%	0.030
DoS	99.92%	99.85%	99.96%	99.91%	0.0005
U2R	46.66%	66.66%	26.66%	46.66	0.200
R2L	92.48%	84.25%	93.98%	90.24	0.052

The result of classifying the data samples using the whole 41 features. From this table, we notice how the imbalanced classes U2R and R2L are misclassified. These classes are rare in the main KDD CUP 99 dataset and their ratio in the dataset is very small. Therefore, the data used by this study was grouped into samples to maintain the original distribution as in the main dataset.

We also applied Rough Set classifier on the dataset with the new reduced feature subset for the same data samples to see the effect of feature reduction (Table 6).

Table 6: The classification accuracy on three different samples using only the most 8 significant features

Type	Sample 1	Sample 2	Sample 3	Mean	StdDev
Normal	93.20%	97.52%	88.83%	93.18%	0.643
Probe	95.50%	94.70%	96.40%	95.53%	0.008
DoS	99.36%	99.69%	99.27%	99.44%	0.002
U2R	34.27%	66.66%	80.00%	60.31%	0.235
R2L	85%	84.93%	99.31%	90%	0.082

By looking at the result of classification of the samples using only the most 8 significant features, we notice that there is no great reduction in accuracy for some classes but also there is an increase of accuracy in others. The reason is that the instances of some classes that occupy most of the data space (i.e. Normal, and DoS) have redundant features that do not play any role in detecting these instances.

In addition, the features in these instances were less correlated. As a result, the feature reduction process did not affect the performance of the classifier in these classes. Meanwhile, the instances in other classes (i.e. R2L and U2R) which are called imbalanced classes have noisy and uncorrelated features that affect the classification accuracy. Furthermore, these classes contain attacks that are rare in the data space. The feature reduction process plays a role in eliminating the uncorrelated features and hence increases the accuracy of the classifier.

The features obtained by our model were compared with the features selected by Chebrolu *et al.* in [9] using Bayesian Networks approach (BN). We found that the 8 features obtained by our study were among the 12 features selected by their study and they are: C, E, F, L, W, X, Y, AB, AE, AF, AG, AI (see Fig. 3).

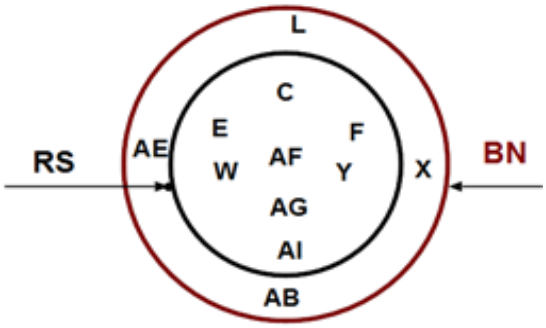

Fig 3: A comparison with BN approach in Chebrolu *et al.*[9].

4.3 Immune Network Clustering

In this phase, the same data samples that have been used for feature reduction using Rough Set were also used here to examine the ability of the aiNet algorithm in clustering different classes of data. In these data samples the distribution of attacks and normal instances is as shown in Table 7.

Table 7: the distribution of the normal and attack instances in data samples

Sample/Class	Normal	Probe	DoS	U2R	R2L
All samples	2000	684	6907	34	375

Before the data samples were fed to the immune network model the normalization process was applied. In the KDD CUP'99 dataset the attributes are either numerical or nominal. By normalization the nominal attributes are converted into linear discrete values (integers). For example, 'ftp' protocol is represented by 1 and 'http' protocol is represented by 2. Then, the attributes fall into two main types: discrete-valued features and continuous-valued. If one of the features has a large range, it can overpower the other features. Many methods can be used for normalization like distance-based method and Mean/Median Scaling method among others.

The parameters of the aiNet algorithm are set up accordingly: N*gen*= 10, σ_d=1, σ_S=0.3, Percentile amount of clones to be re-selected=10, and the learning rate=0.4. The results of using aiNET algorithm are shown in Table 8.

Table 8: Clustering Results obtained by AiNet

Sample/Class	Normal	Probe	DoS	U2R	R2L
Sample 1	3580	1598	4776	9	37
Sample 2	3495	640	5823	6	36
Sample 3	3420	1590	4949	5	36

From Table 7, we see that for each class the actual distribution of data is different from the result clusters. This is common in all clustering methods because it depends on the distances between data instances. In our dataset, there are similarities between normal traffic and attacks and also between the attacks themselves. These similarities make it difficult to differentiate between normal and attack instances.

Table 9: The result of clustering data samples into two categories (Normal and Anomalies)

Sample/Class	Normal	Anomalies (attacks)
Sample 1	3580	6420
Sample 2	3495	6505
Sample 3	3420	6580

The results shown in Table 7 can be presented in binary classification format, segregating between normal and anomalies (attacks) as shown in Table 9. Binary-classification representation is useful especially in

obtaining detection rate (DR) and false positive rate (FPR). Table 10 shows DR and FPR based on our experiments on the three sample sets.

Table 10: Detection rate and false positive rate for the clustering process obtained by AiNet

Sample/Class	Detection Rate	False Positive Rate
Sample 1	80.25%	0.1975
Sample 2	81.31%	0.1868
Sample 3	82.25%	0.1775

The relation between FPR and DR can be expressed using the ROC curve. The following Figures show the ROC curves for sample1, sample2 and sample3 data respectively.

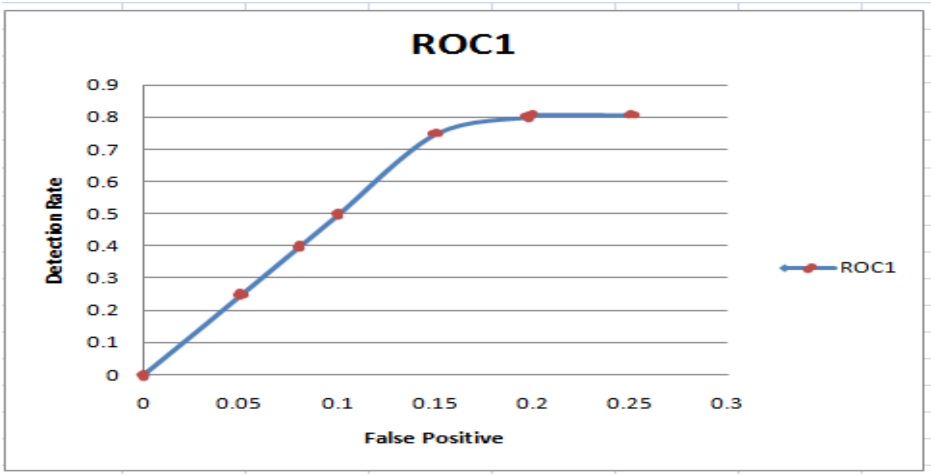

Fig 4: ROC curve for sample1

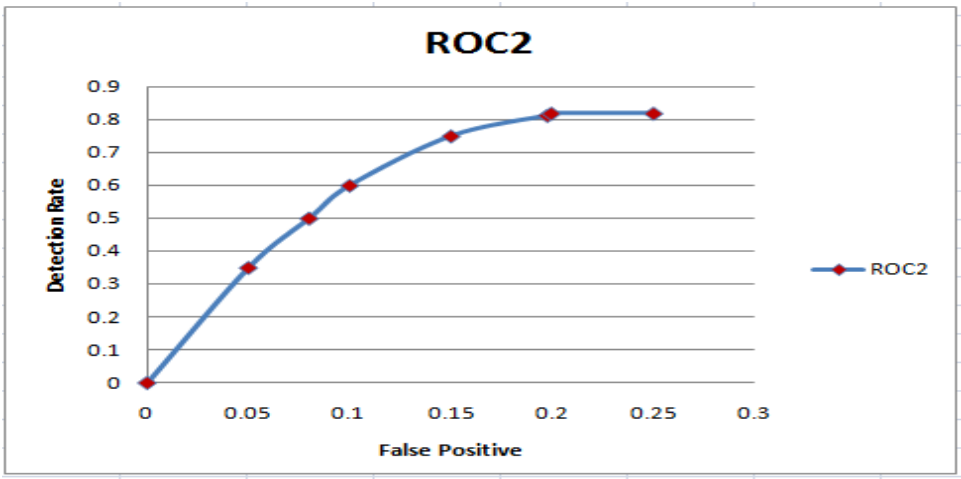

Fig 5: ROC curve for sample2

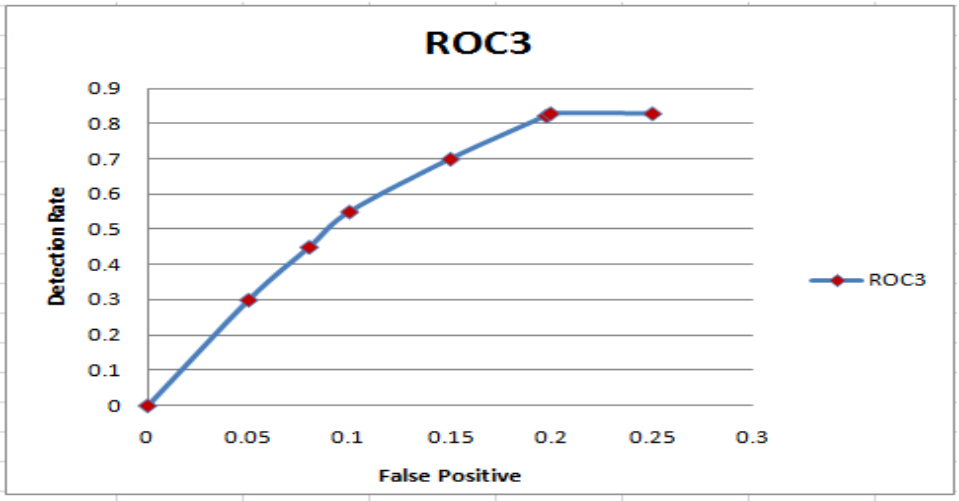

Fig 6: ROC curve for sample3

From Fig. 4 to Fig. 6, we found that the performance of aiNet algorithm is quite consistent on the three sample sets. The FPR is quite low than other approaches used for intrusion detection as we will see later in the analysis. Based on the above results, aiNet seems to be robust enough to distinguish attacks from normal traffic. It is shown that aiNet can cluster attacks in the absence of labels and without any prior knowledge.

To further evaluate the performance of aiNet, a comparison was done with k-Means, a commonly used clustering method in many fields including intrusion detection. We have applied k-Means algorithm on the same data

samples. Before applying k-Means, *k* which denotes the number of clusters has to be set and the seeds for all of *k* clusters were then randomized.

The following figures (Fig. 7 – Fig. 8) show the ROC curve for the performance of K-Means method. It shows the relation between the DR and the FPR.

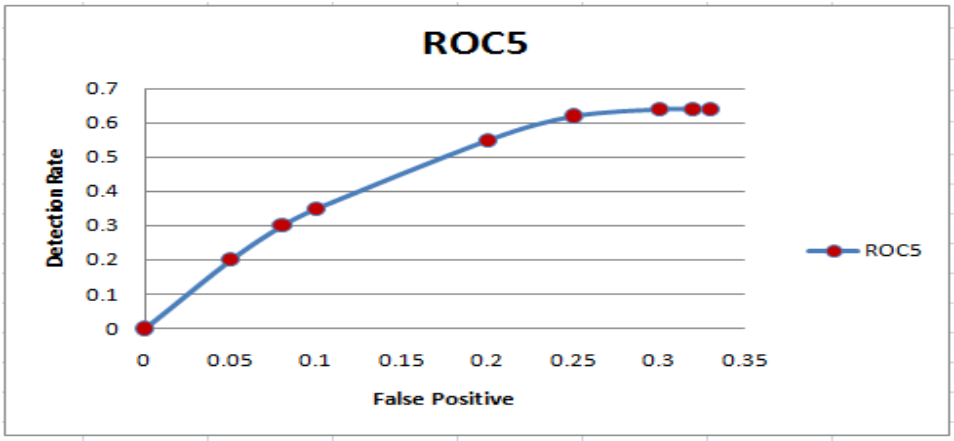

Fig 7: ROC curve for sample1 of group 2 using K-Means.

It suggests that K-Means has a high FPR and relatively low DR. This is due to the nature of intrusion detection data where the distribution of attacks among the different classes is not balanced and there are similarities between instances from different classes. The results also indicate that k-Means which heavily relies on distance measure, could poorly assign the data into their right clusters.

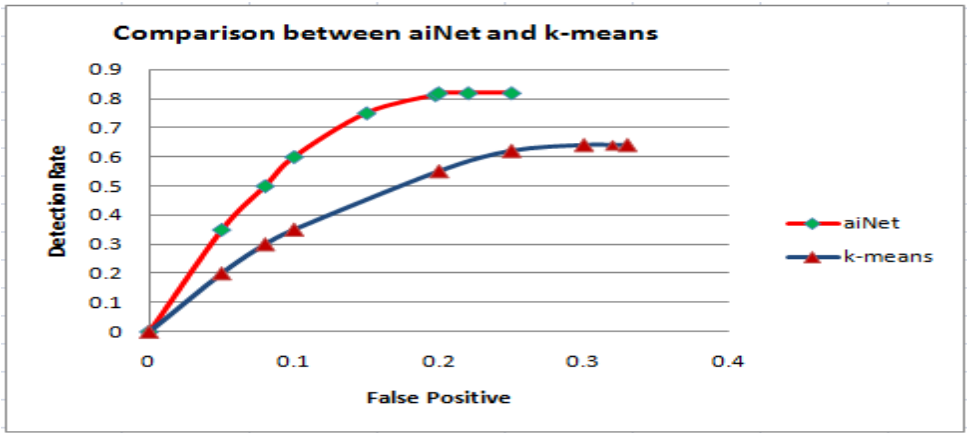

Fig 8: The comparison between the ROC curves of both aiNet and K-Means methods for the same data sample.

In Fig 8, we show the ROC comparison between aiNet algorithm and the K-Means algorithm to indicate the performance of aiNet relative to k-Means. We see that aiNet performs better than K-Means in both performance measures, DR and FPR.

5 Conclusion

In this paper the investigation on capability of using only significant features together with the feasibility of using bio-inspired algorithm to detect unknown attacks were studied. The empirical results suggest that the problem of low detection rate can be addressed by using Rough Set feature reduction. Meanwhile, the problem of detecting novel attacks can be addressed by artificial Immune Network clustering method (aiNet). A comparison with k-Means clustering method was done to evaluate the capability of aiNet relative to the common existing clustering method to detect novel attacks. The results revealed that detection rate was improved by using significant features. Furthermore, the finding also shows that Immune Network clustering method is robust in detecting novel attacks in the absence of labels.

6 Future Work

To make the usage of aiNet easier, our future work will focus on the automatic setting of its parameters. In addition, this work mainly focused on the unsupervised clustering of attacks using aiNet. We plan to develop a new semi-supervised approach in which some labels of the data features will be used as guidance for the clustering process.

References

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

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

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

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

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

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

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

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

[9] Siraj, M., Maarof, M.A., Hashim, S.Z., "Intelligent Alert Clustering Model for Ntwork Intrusion, Analysis", *International Journal of Advances in Soft Computing and its Applications*., Vol. 1, No.1,July 2009.

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

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

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

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

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

[15] Leung, K. and Leckie, C. "Unsupervised Anomaly Detection in Network. Intrusion Detection Using Clusters". *Appeared at the 28th Australasian Computer Science Conference,* The University of Newcastle, Australia, (2005)..

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

[17] Liu, G., Yi, Z. and Yang, S. "A Hierarchical Intrusion Detection Model based on the PCA Neural Networks", *International Journal of Neurocomputing,* (2007), *Vol(70), pp.1561-1568.*

[18] Farid, D.M., Harbi, N. and Rahman, M. Z. "Combining Naïve Bayes and Detection Tree for Adaptive Intrusion Detection". *International Jornal of Network Security & Its Applications (IJNSA),* (2010), *Vol(2-2), pp. 12-25.*

[19] Deng, L. and Gao, D. Y. "Research on Immune based Adaptive Intrusion Detection System Model". *In Proceedings of IEEE International Conference on Networks Security, Wireless Comunications and Trusted Computing,* (2009), *pp. 488-491*.

[20] A.K. Ghosh, J. Wanken, and F. Charron. "Detecting anomalous and unknown intrusions against programs", Proceedings of the 1998 Annual Computer Security Applications Conference (ACSAC'98), December 1998.