Artificial Immune Network Clustering approach for Anomaly Intrusion Detection

Murad Abdo Rassam
Universiti Teknologi Malaysia,
Faculty of Computer Science and Information Systems,
81310, Skudai, Johor, Malaysia.
E-mail :*eng.murad2009@gmail.com*

Mohd. Aizaini Maarof
Universiti Teknologi Malaysia,
Faculty of Computer Science and Information Systems,
81310, Skudai, Johor, Malaysia.
E-mail :*aizaini@utm.my*

***Abstract*—Many Intrusion Detection approaches (IDS) have been developed in the literature. Signature based approaches for anomaly detection need to be updated with the latest signatures of unknown attacks and hence being impractical. Anomaly based approaches on the other hand, suffer from high false alarms as well as low detection rates and need labeled dataset to construct the detection profile. In fact this kind of labeled dataset cannot be obtained easily. In this paper, we investigate the application of bio-inspired clustering approach, named Artificial Immune Network, for clustering attacks for intrusion detection systems. To reduce the dimension of the DARPA KDD Cup 1999 dataset, Rough Set method was applied to get the most significant features of the dataset. Then the Artificial Immune Network clustering algorithm, aiNet, has been applied on the reduced dataset. The results show that detection rate was enhanced when most significant features were used instead of the whole features. In addition, it shows that, Artificial Immune Network is robust in detecting novel attacks.**

***Index Terms*—IDS, Feature Reduction, Artificial Immune Network, Clustering.**

I. Introduction

Intrusion Detection Systems (IDS's) as defined in [1], 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. The main function of such tools is to differentiate between normal activities of the system and any other deviations that could be intrusive.

Two main intrusion detection systems approaches have been classified in the literature: anomaly intrusion detection system and misuse intrusion detection system. The former focuses on the unusual activities of patterns and uses the normal behavior patterns to identify any deviation of that behavior. However, the later can recognize only the known attack patterns and uses predefined signatures of attacks.

Many studies in the literature like in [2,16,18] have tackled the IDS problem as a pattern recognition problem or rather classified as learning system. There is a need for removing redundant and irrelative features for learning systems in order to reduce the complexity and increase the accuracy of classification systems [2, 17]. To this end, we need to reduce the representation space of such features in order to cope with the requirements of accurate and inexpensive IDSs.

Bello *et al.* in [3] suggested that feature reduction was necessary to reduce the dimensionality of training dataset. Furthermore, they claimed that feature reduction helps to enhance the speed of data manipulation and improves the classification rate by reducing the influence of noise.

Many anomaly detection systems have been proposed in the literature based on different soft computing and machine learning techniques. 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.

Furthermore, some IDSs mentioned in [9, 24, 25] are based on ensemble or a combination of different learning techniques. All of these techniques in particular have been developed to classify or recognize whether the incoming network access is normal or an attack.

Inspired by biology, computing models can be designed to make the 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. Recently, these bio-inspired systems are getting more attention because of their ability of to adapt naturally with the environment in which they applied. One of these systems is the human immune system which provides the inspiration for solving a wide range of innovative problems [23].

The aim of this paper is to address the impact of the feature reduction in designing the anomaly detection system. In addition, it introduced the use of the bio-inspired artificial immune network algorithm (aiNet) to detect the novel attacks that have not been seen in the training patterns.

doi:10.4304/jait.3.3.147-154

The rest of the paper is organized as follows. Section 2 gives a view on the methods used in this study which are rough set and artificial immune network. Section 3 describes some related works in both areas namely, feature reduction and unsupervised immune network for clustering. In section 4, the experiments using KDD CUP 99 dataset are shown. It also includes an analysis of the results and performance comparison against k-Means method. We conclude the paper in Section 5.

II. Background

A. *Rough Set Theory*

Rough Set can be defined as a mathematical tool for approximate reasoning for decision support and is particularly well suited for classification of objects [4]. It has been stated that, this tool can also be used for feature reduction and feature extraction. The most attractive characteristics of rough set 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. More theoritical definitions about rough eet can be found in [5].

B. *Artificial Immune Network*

Immune network theory has been proposed first by Jerne in [6] and 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.

Network activation and network suppression are two important characteristics of 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. The 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 cells [8]. Fig.1 below gives a view about the immune network.

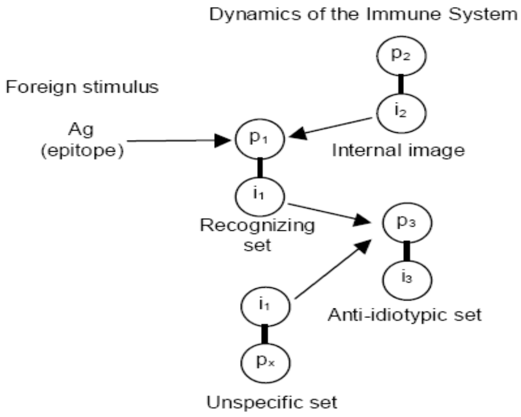

Figure 1. A view on idiotypic Immune Network [8]

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 in 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:
                1. If its affinity is greater than the affinity threshold d σ
                Then.
                        a. Reproduce the antibody proportionally to its affinity
                        b. Each clone undergoes a mutation inversely proportional to its affinity.
                        c. 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 f σ
4. End while
5. Output each cluster in the network
6. Output each cluster in the network
```

III. Related Work

A. Feature Reduction in Intrusion Detection Systems

In the literature, most of approaches for IDSs examine all features of dataset to detect intrusions. In fact, 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 the important input features in the IDS dataset that contribute to the detection process and hence improve the efficiency and the effectiveness of our proposed model.

In [9] Chebrolu *et al.*, have investigated the performance of two feature reduction algorithms involving Bayesian networks (BN) and Classification and Regression Trees (CART). In addition, they also investigated ensemble of BN and CART. Their results indicated that feature reduction is important to design any IDS that is efficient and effective for real world detection systems. The use of Rough Set for feature reduction has been studied by Zhang *et al.* in [4]. Their study showed that this tool is capable of getting the classification rules to determine the category of attack in IDS. In fact, they did not show the features that were implemented in the classification process.

According to [9], data reduction can be achieved by different ways like filtering, data clustering and feature selection. Generally, authors in [10] stated that, the capability of anomaly intrusion detection is often hindered by the inability to accurately classify a variation of normal behavior as an intrusion. In addition, this study also stated that network traffic data is huge, and it causes a prohibitively high overhead and often becomes a major problem in IDS. They also demonstrated that, the elimination of these unimportant and irrelevant features did not significantly lowering the performance of IDS.

Chakraborty in [11] argued that, the existence of these irrelevant and redundant features does generally affects the performance of machine learning or pattern classification algorithms in detecting attacks. Hassan, *et al*., in [12] proved that proper selection of feature set has resulted in better classification performance.

B. Immune Network Clustering

The importance of IDSs is not in its ability to tackle the huge number of vulnerabilities that have been identified in advance but in their ability of detecting unknown number of unexposed vulnerabilities that may not be immediately available to the experts for analysis and inclusion in the knowledge base [13]. In order to cope with this need, authors in [13] introduced an unsupervised anomaly detection based on clustering. They argued that, their approach increase the detection rate of different kinds of unknown attacks.

Generally, 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 [14]. To this end, in order to address these problems an unsupervised anomaly detection approach using artificial immune network is proposed due to the ability of this bio-inspired algorithm to adapt and cluster normal and attacks data without any prior knowledge.

IV. Experiments and Results

In this section we starts by giving some information about the dataset been used to validate our approach. After that, we show the experimental procedure used to implement our approach. The experiments have been done in two phases, first phase; we apply the rough set tool to reduce the features of the dataset. Then feature subset obtained from the first phase is used as input to the second phase, immune network clustering to cluster normal data from attacks.

A. Dataset

KDD Cup 1999 is the dataset that is used to validate the proposed approach. It 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.

The KDD CUP dataset can be classified into four main categories of attacks. A brief description of each class in the subsequent sections.

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

B. Feature Reduction using ROSETTA

For validating our proposed approaches, three different samples of the dataset are 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 is applied on these data samples. Rough Set is implemented using ROSETTA (Rough SET Toolkit for data Analysis) system developed by Ohrn [15].

The experimental procedure can be outlined as follows: First, the raw data samples are transformed into Tables recognized by ROSETTA. After the preprocessing of data samples, each data sample is 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 is used to reduce the data sample features in this study. We are interested in GA, because according to Ohrn [15]; 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 is used to generate the rules using the GA built in algorithm in ROSETTA tool. These rules will be 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 are obtained and shown in the following table.

TABLE 2
THE MOST 8 SIGNIFICANT FEATURES OBTAINED BY ROUGH SET IN THREE DIFFERENT SAMPLES OF DATA.

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 OBTAINED BY ROUGH SET

C	E	F	W	Y	AF	AG	AI

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

TABLE 4
THE CORRESPONDING NETWORK FEATURES AND THE DESCRIPTION OF THE FEATURES FOR THE OBTAINED FEATURES

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	Nnumber 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 also was applied to classify the data in order to evaluate the performance of the classification by rough set classifier before and after feature reduction. The results are shown in the following table.

TABLE 5
THE CLASSIFICATION ACCURACY OBTAINED BY ROUGH SET ON THREE DIFFERENT SAMPLES USING ALL 41 FEATURES.

Type		Sample 1	Sample 2	Sample 3	Mean	StDv
Normal		92.8%	95.2%	79%	92%	0.04
Attack	Prob	94.3%	100%	99.3%	97.9%	0.03
	DoS	99.9%	99.9%	100%	99.9%	0.00
	U2R	46.7%	66.7%	26.7%	46.7	0.20
	R2L	92.5%	84.3%	94%	90.2	0.05

Table 5 shows the result of classifying the data samples using the whole dataset features which are 41 features. From the 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 so the data used by this study was grouped into samples to maintain the original distribution as in the main dataset.

After applying the rough set classifier on the dataset with all 41 features, we also applied it on the dataset with the new reduced feature subset for the same data samples to see the effect of feature reduction.

TABLE 6
THE CLASSIFICATION ACCURACY OBTAINED BY ROUGH SET ON THREE DIFFERENT SAMPLES USING ONLY THE 8 MOST SIGNIFICANT FEATURES.

Type		Sample 1	Sample 2	Sample 3	Mean	StDv
Normal		93.2%	97.5%	88.8%	93.2%	0.643
Attack	Prob	95.5%	94.7%	96.4%	95.5%	0.008
	DoS	99.4%	99.7%	99.3%	99.4%	0.002
	U2R	34.3%	66.7%	80%	60.3%	0.235
	R2L	85%	84.9%	99.3%	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 are less correlated. As a result of that, 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 are compared with the features selected by Chebrolu *et al.* in [9] using Bayesian Networks approach (BN) as shown in Fig. 2 below. 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.

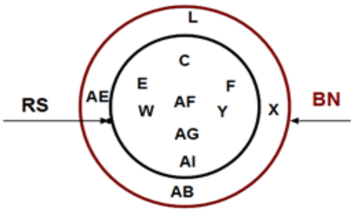

Figure 2. A comparison with BN approach in [9].

C. Immune Network Clustering using aiNet algorithm

In this phase, the same data samples that have been used for feature reduction using rough set are 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 ATTAK 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 is 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.

After setting up the parameters of the aiNet algorithm such that (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), and applying it on the data samples, the results are shown in Table 7.

TABLE 8
CLUSTERING RESULTS OBTAINED BY AINET CLUSTERING

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

Table 8 shows the result of clustering data samples into 5 clusters using aiNet, each class in the data sample is represented by one separate cluster. 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 and 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 clearly.

The results shown in Table 8 can be presented in binary classification format, segregating between normal and anomalies (attacks) as shown in Table 9.

TABLE 9
THE RESULT OF CLUSTERING DATA SAMPLES IN TWO CATEGORIES (NORMAL AND ANOMALIES)

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

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 three sample sets.

TABLE 10
DETECTION RATE AND FALSE POSITIVE RATE FOR THE CLUSTERING PROCESS DONE BY AINET ALGORITHM

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 of the dataset.

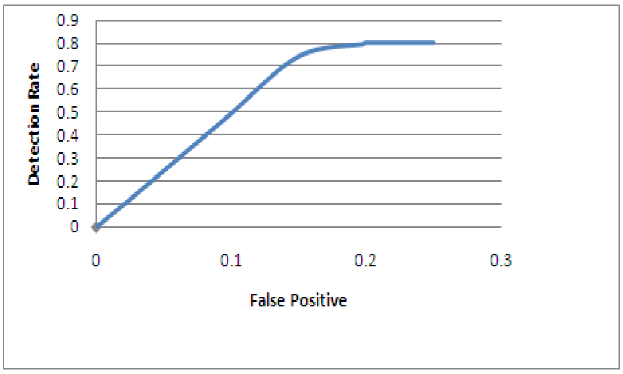

Figure 3. ROC curve for sample1

From Figure 3, we found that 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 figure shows the ROC curve for the performance of K-Means method. It shows the relation between the DR and the FPR.

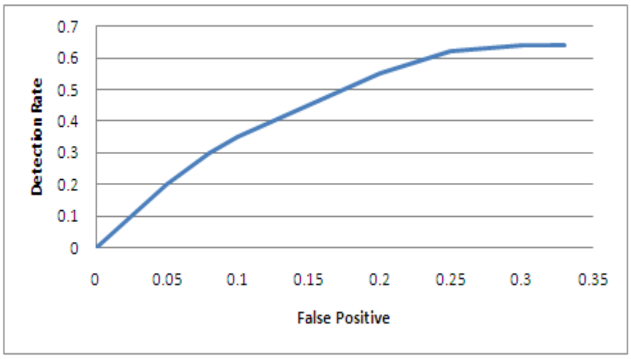

Figure 4. ROC curve for sample1 of group 2 using K-Means.

Fig.4 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.

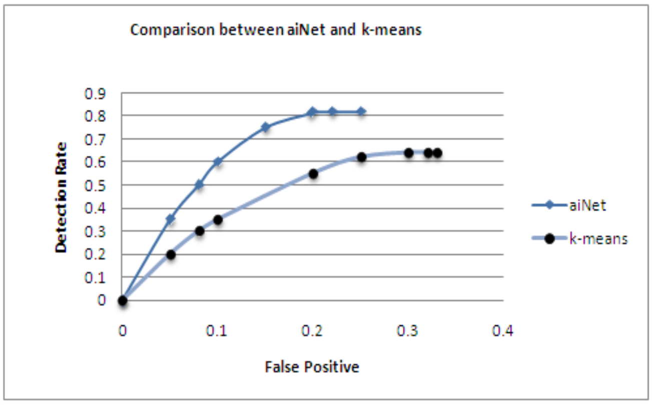

Figure 5. The comparison between the ROC curves of both aiNet and K-Means methods for the same data sample.

In Fig. 5, 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.

Further investigation has been done on the aiNet characteristics show that aiNet is efficient enough in compressing the datasets. Figure 6 below shows the tradeoff between the suppressing threshold σ_S and the output cells that aiNet produces at the end of clustering process.

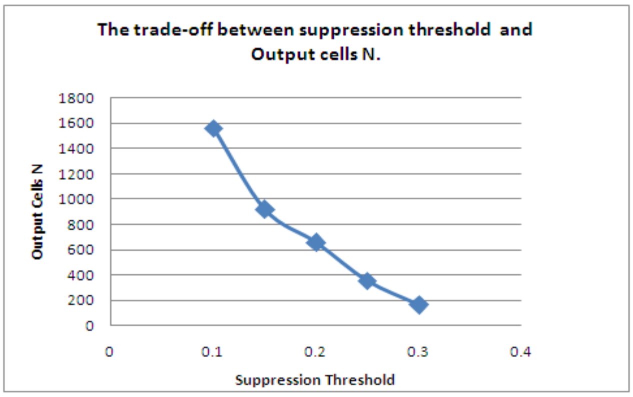

Figure 6. The tradeoff between the suppression threshold and the number of output cells produced by aiNet.

Fig. 6 shows that, beside the ability of aiNet in clustering attacks, it has also the ability to compresses the dataset which make it more suitable for large scale datasets.

V. Conclusion and Future Work

In this paper, the impact of using a proper feature reduction tool has been found to enhance the detection accuracy and reduce the false alarms in the IDSs. In addition, the novel unknown attacks that have not been seen during the training can be detected using a bio-inspired immune network clustering approach. The experimental results show that the accuracy of detection has been improved using rough set for feature reduction. Meanwhile, the problem of detecting novel attacks has been addressed by artificial immune network clustering algorithm (aiNet). To prove the viability of our approach, a comparison with k-Means clustering approach has been done and showed that our approach gives better results in terms of detection accuracy of novel attacks. The findings also show that Immune Network clustering approach is robust in detecting novel attacks in the absence of labels.

To make the usage of aiNet easier, our future work will include the automatic setting of its parameters. Another point to be focused on for future studies is to study the visibility of using a semi supervised approach instead of unsupervised to enhance the accuracy of detection by introduce some labels to the clustering approach.

References

[1] G. Teodoro, J. Dı´az Verdejo, G. Macia´-Ferna´ndez, and E. Va´zquez. “Anomaly-based network intrusion detection: Techniques, systems, and challenges” *Computers and security*, Vol 28, issue 1-2, pp. 18-2,March 2009.

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

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

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

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

[6] N.K. Jerne. “Network theory of the immune system. 1974. *Ann. Immunol”,Paris*, 1974.

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

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

[9] S. Chebrolu, A. Abraham, and J.P. Thomas. “Feature Deduction and Ensemble Design of Intrusion Detection Systems”. *International Journal of Computers and Security* .Vol 24, Issue 2, pp. 295-307, 2004.

[10] A.H. Sung, and S. Mukkamala. “The Feature Selection and Intrusion Detection Problems”. *LNCS, vol. 3321, Springer Hieldelberg,* pp. 468-48,2004.

[11] B. Chakraborty. “Feature Subset Selection by Neuro-rough Hybridization.”. *LNCS, Springer Hiedelberg,* pp. 519-526, 2005.

[12] A. Hassan, M.S. Nabi Baksh, A.M. Shaharoun, and H. Jamaluddin. “Improved SPC Chart Pattern Recognition Using Statistical Feature”. *International Journal of Production Research,* Vol 41 Issue 7, , pp. 1587-1603, 2003.

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

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

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

[16] G. Liu, Z. Yi. and S. Yang.” A Hierarchical Intrusion Detection Model based on the PCA Neural Networks”. *International Journal of Neurocomputing,* Vol 70, pp.1561-1568.2007.

[17] D.M. Farid, N. Harbi, and M.Z. Rahman. “Combining Naïve Bayes and Detection Tree for Adaptive Intrusion Detection”. *Internation Jornal of Network Security & Its Applications (IJNSA),* Vol2-2, pp. 12-25. 2010.

[18] L. Deng, and D.Y. Gao. “Research on Immune based Adaptive Intrusion Detection System Model”. *In Proceedings of IEEE International Conference on Networks Security, Wireless Comunications and Trusted Computing,* pp. 488-491.2009.

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

[20] T. Shon, Y. Kim, C. Lee, J. Moon.” Machine learning framework for network anomaly detection using SVM. *Information Assurance Workshop, IAW'05.*2005.

[21] D. Kim, and J. Park. “Network-Based Intrusion Detection with Support Vector Machines”. *Lecture notes in computer science*, pp.747-756, Springer.2003.

[22] D. Dasgupta.“Immunity-Based Intrusion Detection System: A General Framework”. *In Proc. of the 22nd NISSC.*1999b.

[23] X.Hang and H.Dai. “An Immune Network Approach for Web Document Clustering”. *In Proc of WI*, pp. 278-284. 2004.

[24] Y. Yasami, S. Khorsandi, SP. Mozaffari, and A. Jalalian .”An unsupervised network anomaly detection approach by k-Means” . *In IEEE Symposium on Computers*. 2008.

[25] X. Cheng, Y.P. Chin and S.M. Lim . “Design of multiplelevel hybrid classifier for intrusion detection system using Bayesian clustering and decision trees”. *Elsevie Pattern Recognition Letters,*Vol 29, Issue 7, , Pages 918-924. May 2008

Murad A. Rassam is currently a Ph.D. student in the Information Assurance & Security Research Group (IASRG) at Universiti Teknologi Malaysia in Skudai, Johor, Malaysia. He received his B.Sc. in Information Technology Engineering from Tishreen University, Lattakia, Syria, in 2005. He received his M.Sc. degrees in Computer Science from the Universiti Teknologi Malaysia, Skudai, Johor, Malaysia in 2010. His research interests include wireless sensor network intrusion detection, network intrusion detection, and application of soft computing techniques and machine learning to computer and network security. He is involved as a reviewer for some international journals.

Mohd A. Maarof, Ph.D. Is a Professor at Faculty of Computer Science and Information System, Universiti Teknologi Malaysia (UTM). He obtained his B.Sc. (Computer Science) and M.Sc. (Computer Science) from U.S.A and his PhD from Aston University, Birmingham, United Kingdom in the area of Information Technology (IT) Security. He is currently leading the Information Assurance & Security Research Group (IASRG) at UTM. Currently his research involve in the areas of Intrusion Detection System, Malware Detection, Web Content Filtering, and Cryptography.