ISSN 2277-3061

A Novel Method for Intrusion Detection Based on SARSA and Radial Bias Feed Forward Network (RBFFN)

Anshul Chaturvedi, Prof. Vineet Richharia
Reseach Scholar, Department of Computer Science & Engg., LNCT, Bhopal
anshulchaturvedi03@gmail.com
Head of the Department of Computer Science & Engg., LNCT, Bhopal
vineet_rich@yahoo.com

ABSTRACT

The Internet, computer networks and information are vital resources of current information trend and their protection has increased importance in current existence. Any attempt, successful or unsuccessful to finding the middle ground the discretion, truthfulness and accessibility of any information resource or the information itself is measured a security attack or an intrusion. Intrusion compromised a loose of information credential and trust of security concern. The mechanism of intrusion detection faced a problem of new generated schema and pattern of attack data. Various authors and researchers proposed a method for intrusion detection based on machine learning approach and neural network approach all these compromised with new pattern and schema. Now in this paper a new model of intrusion detection based on SARAS reinforced learning scheme and RBF neural network has proposed. SARAS method imposed a state of attack behaviour and RBF neural network process for training pattern for new schema. Our empirical result shows that the proposed model is better in compression of SARSA and other machine learning technique.

Indexing terms/Keywords

Intrusion detection, SARSA, SARSA-RBF, KDD CUP 99.

Council for Innovative Research
Peer Review Research Publishing System
Journal: INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY
Vol 7, No 3
editor@cirworld.com
www.cirworld.com, member.cirworld.com

ISSN 2277-3061

INTRODUCTION

Reinforced learning and neural network play an important role in intrusion detection. The property of reinforced learning is multi-stage and multi-agent uses differentiate different attack and anomaly category in intrusion [1]. Intrusion detection system algorithms can be categorized into three types: supervised learning, unsupervised learning and reinforced learning. A supervised learning is a technique that builds detection rule/model by learning pattern from provided information. The supervised learning normally has high detection rate and low false alarm rate. On the other hand, this technique can detect only known pattern. For that reason, it is not secure enough because in actuality there are many new and mysterious attacks in the network [2]. The second type of algorithms is an unsupervised learning technique. It is able to learn new/ unknown attacks without training information. However, it often has relatively lower detection rate and having high false alarm rate. A reinforced learning implies as multi-state modeling technique for intrusion detection. The learning rate of reinforced mechanism implies the training pattern of unknown attack of intrusion in RBF neural network. The current scenario of intrusion detection system suffered from detection rate and false alarm generation. The problem of detection and false alarm generation arise due to large features attribute of intruder file[3]. No any process of algorithm find during survey of intrusion detection system based reinforced learning work directly on dynamic feature reduction of intruder file. The feature reduction is important issues in improving of detection rate of intrusion detection system. All feature attribute of anomaly file are not involved in reaction of action, so we reduce those attribute and improve the efficiency of method of intrusion detection system. In this paper we discuss a hybrid method for feature reduction using reinforced learning with SARAS learning factor and radial bias function network (RBF)[4].

A great advantage of hybrid method is without learning of parameter work a complete system and reduces feature of anomaly file. Radial bias function network (RBF) to identify important input features for intrusion detection. Through identifying the important inputs and redundant inputs, a classifier can achieve the reduced problem size, faster training and more accurate results. Feature extraction is an important issue in intrusion detection. Of the large number of features that can be monitored for intrusion detection purpose, which are truly useful, which are less significant, or which may be useless? The question is relevant because the elimination of useless features (the so-called audit trail reduction) enhances the accuracy of detection while speeding up the computation, thus improving the overall performance of IDS. In this paper, we focus on network intrusion detection for unknown attack types meaning that the approach is able to detect new or unknown type of attacks in the network. In particular, the network intrusion detection system should be able to identify normal network activity and classify attack types. We are interested in designing an IDS technique using SARSA and RBF neural network. The SARSA is reinforced learning technique and RBF neural network able to learn new attacks by itself. Moreover, this technique has high detection rate and robust. Therefore, we apply the Q- learning factor approach for SARSA intrusion detection system i.e. the data is detected right after it arrived to the detection system[5]. We evaluate our IDS in terms of, detection rate and alarm generation rate. The rest of paper is organized as follows. In related work of IDS then discuss methodology, some experimental result and finally followed conclusion and future scope.

RELATED WORK

In survey, numbers of anomaly detection systems are study based on many different machine learning techniques. Some studies apply single agent learning technique, such as neural networks, genetic algorithms, support vector machines, etc. On the other hand, some systems are based on combining different learning techniques, such as hybrid or ensemble techniques. In particular, these techniques are developed as classifiers, which are used to classify or recognize whether the incoming Internet access is the normal access or an attack.

In 2011, Z. Muda et al. [6] proposed network detection solution by combining supervised learning technique and unsupervised learning technique. They used K-Means algorithm for unsupervised learning and Naive Bayes algorithm for supervised learning. The first step of algorithm is using K-Means algorithm to group data to normal or attack. Then, use Naïve Bayes algorithm to classify the obtained result into attack type. The KDD99 dataset was used to evaluate the performance of this algorithm. The detection rate was improved to 99.6 percent. However, this solution is not practical for real network because K-Means algorithm requires more time to process huge data in real networks which could lead to bottleneck problem and system clash.

In 2009, T. Komviriyavut et al. [1] proposed a real-time detection approach. They used packet sniffer to sniff network packets in every 2 seconds and pre-processed it into 12 features and used decision tree algorithm to classify the network data. The output can be categorized into 3 types which are DoS, Probe and normal. The result shows that this algorithm has 97.5 percent of detection rate. This technique is fast and able to use in real network. However, it was not designed to detect unknown attacks.

N. Ngamwitthayanon and N. Wattanapongsakorn [2] proposed Fuzzy-Adaptive Resonance Theory (ART) in network anomaly detection with feature-reduction dataset. The Adaptive Resonance is one type of neural network algorithm. The main algorithm is ART algorithm while Fuzzy is used to simplified network structure of ART. Moreover, they applied feature reduction method to KDD99 dataset [7]. This approach can offer 98.07 percent detection rate and use only 14 features of KDD99's 41 features. A Dependable Network Intrusion Detection System (DNIDS) based on the Combined Strangeness and Isolation measure K-Nearest Neighbor (CSIKNN) algorithm. The intrusion detection algorithm analyses different characteristics of network data by employing two measures: strangeness and isolation. But in general the K-NN still needs intensive computations. The Unsupervised Anomaly Detection Using an Optimized K-Nearest Neighbors Algorithm can work without the need for massive sets of pre-labeled training data. A k-nearest neighbors algorithm to detect anomalies in network connections, as well as the optimization necessary to make the algorithm feasible for a real-world system [8]. The development of anomaly based intrusion detection systems during the recent years. As several supervised and unsupervised clustering techniques were optimized resulting in more elegant techniques that provided

more detection accuracy and lower false alarm rate. Moreover, the newly proposed techniques tend to avoid the creation of unnecessary neurons in the training process to faithfully represent data inputs as applied in hierarchical clustering. Furthermore, this restriction in creating neurons significantly contributes in reducing the complexity of the training process and producing more accurate topologies. Since, our main concern in our research is to increase the quality of clustering and attacks classification for larger scope of attacks. Additionally, increasing the identification rate of novel patterns in the training process as well. Intrusion Detection System (IDS) is an important detection that is used as a countermeasure to preserve data integrity and system availability from attacks. The work is implemented in two phases; in first phase clustering by K-means is done and in next step of classification is done with k-nearest neighbors and decision trees. The objects are clustered or grouped based on the principle of maximizing the intra-class similarity and minimizing the interclass similarity. This paper proposes an approach which makes the clusters of similar attacks and in next step of classification with K nearest neighbors it detect the attack types. This method is advantageous over single classifier as it detect better class than single classifier system [3].

PROPOSED METHODOLOGY

The proposed methodology of intrusion detection based on reinforced learning and RBF neural network for classification of attack in off line intrusion data. The proposed method work in dual mode first SARSA make a policy for detection and different the category of attack and finally RBF neural network classified all these state in separate group of data [5]. The process of classification improves the detection rate of intrusion.

SARSA-RBF

The algorithm will randomly find learning rate of SARSA and pick the initial sate of cluster dataset. Then, we use pattern learning concept from RBF neural network algorithm to improve the policy in training phase. Then, we will use the policy to classify dataset in testing phase. The pseudo code of the SARSA-RBF algorithm can be given below.

```
Initial policy ();
while {
for each datasetset {
for each policy{
for each attribute{
event = SARSA();
totalevent = totalevent + event;
}
If (total event > Q factor)
class is attack;
else
class is normal;
}
Compare the label class with test class data
}
Calculate optimal state for next process of classification
Stored_Pattern()
Selection-state ()
Voting-process ()
}
```

SARSA Algorithm

1) In our algorithm, we use five state policy for classification of attack in intrusion detection process to form a detection policy. The five state includes five parameters which are a, b, c ,d and e. The algorithm calculates the event status of being attacked from the parameters as shown below.

```
If (a=matched pattern) && (b=unmatched pattern){
Event = dataset – a/(b-a)
}
Else if ( b= matched pattern and  c= unmatched pattern){
Event = dataset – b/(c-b)}
Else if (c=matched pattern and d=unmatched pattern){
Event = d – dataset/(d-c)
Else if (d=matched pattern and e=unmatched pattern){
Event = e– dataset/(e-d)
Else if (e=matched pattern and a=unmatched pattern){
Event = a– dataset/(a-e)
}
Else
{
Event =0
}
```

2) After finding a real state of data we apply RBF neural network for clustering and classification process of categories the data set into normal and attack categories.

Input: number of state X== {a; b,……..,e}
Output: centers of clusters

ISSN 2277-3061

```
Variable
C: number of clusters
cj : center of the j-th cluster
nj : number of patterns in the j-th cluster
di j : distance between xi and the j-th cluster
begin
C =1; c1  x1;n1 :=1;
for i :=2 to P do /* for each pattern */
for j :=1 to C do /* for each cluster */
compute di j;
if di j _R0 then
/* include xi into the j-th cluster */
cj  (cjnj +xi)=(ni+1);
ni :=ni+1;
exit from the loop;
end if
end for
if xi is not included in any clusters then
/* create a new cluster */
C :=C+1;
cC  xi;
nC :=1;
end if
end for
end
```

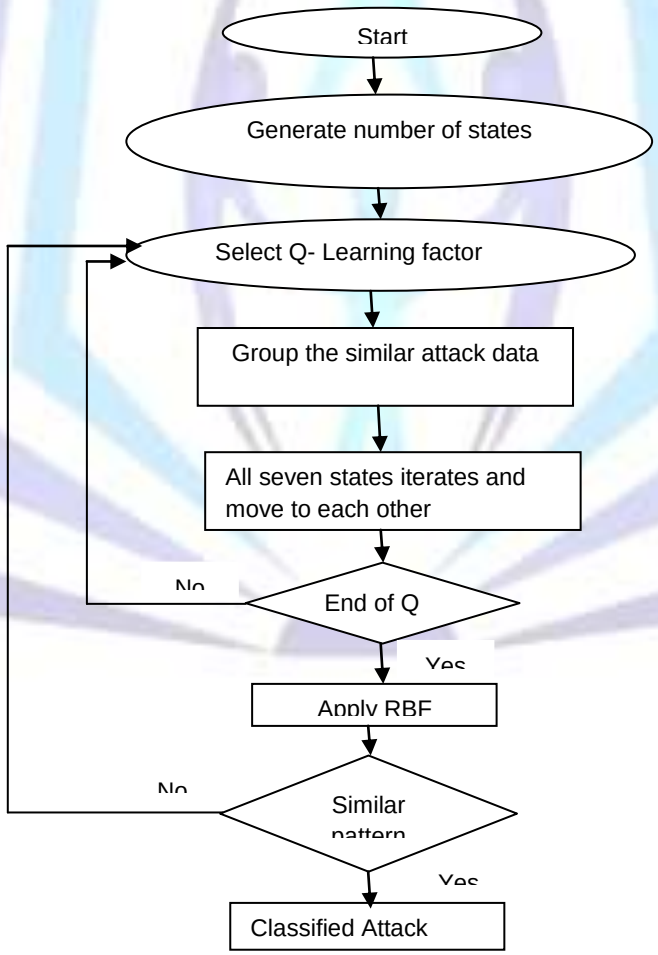


Fig 1: Proposed model for intrusion detection based on SARSA-RBF

EXPERIMENTAL ANALYSIS

We implement our intrusion detection system with MATLAB 7.8.0 and perform experiments in our personal computer with 2.67 GHz Intel core i5 CPU 750 and 4 GB RAM. We use 4000 records of normal data and 3000 records of attack data. The record of attack contains 1000 of DoS and 500 of Probe attack types. We use 1,000 records for each type of DoS attack which are Smurf, UDP-flood, HTTP-flood and Jping. They were generated from closed LAN network with attack generator namely Smurf.c, NetTool5 and Jping.c . We use 500 records for each type of Probe attack. Port scan and Host scan were generated by NetTool5, Connect attack is generated by Host Scan 1.6 . Other 10 types of attack were generated using NMap Win 1.3.1 which are SYN Stealt, FIN Stealt, UDP Scan, Null Scan, IP Scan, Window Scan, RCP Scan, Advanced Port Scan, Xmas Tree and ACK Scan.

Work evaluation on the basis of following parameters

Precision- Precision measures the proportion of predicted positives/negatives which are actually positive/negative.

$$\text{Precision} = \frac{TP}{TP+FP}$$

Recall -It is the proportion of actual positives/negatives which are predicted positive/negative.

$$\text{Recall} = \frac{TP}{TP+FN}$$

Accuracy- It is the proportion of the total number of predictions that were correct or it is the percentage of correctly classified instances.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FN+FP}$$

False-negative rate (FNR) - It is the percentage that attacks are misclassified from total number of attack records.

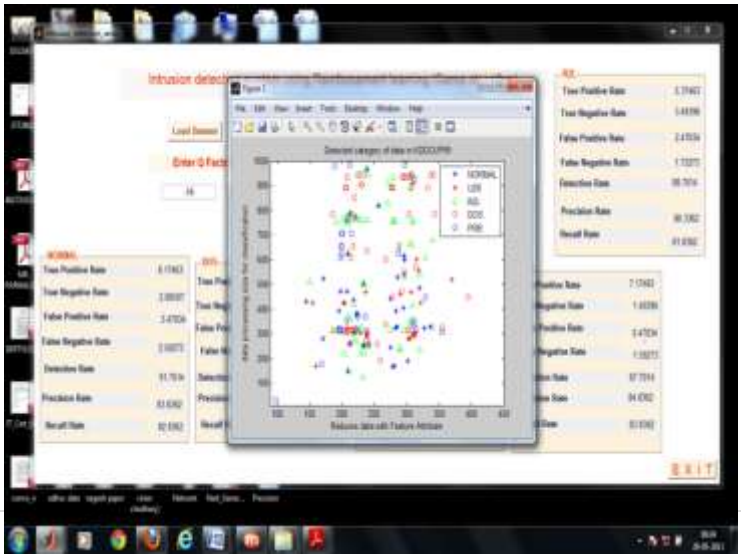
$$FNR = \frac{FN}{FN+TP}$$

False-positive (FPR)- It is the percentage that normal data records are classified as attacks from total number of normal data records.

$$FPR = \frac{FP}{FP+TN}$$

Table 1. It gives the information about result analysis of different combination of dataset with both algorithms SARSA and SARSA-RBF

Metric		Detection rate (%)	Precision (%)	Recall (%)
Data-Set 1	SARSA	93.14	88.56	85. 34
	SARSA-RBF	98.14	97.32	95.21
Data-Set 2	SARSA	89.90	84.32	83.23
	SARSA-RBF	95.23	92.14	91.21
Data-Set 3	SARSA	91.34	86.14	85.11
	SARSA-RBF	95.12	93.21	91.13
Data-Set 4	SARSA	92.22	88.21	87.66
	SARSA-RBF	97.13	94.52	93.67



ISSN 2277-3061

Fig 2: Shows that classification of attack in different categories such as normal ,dos,prob,u2r and r2l data according to predefined policy of state according to SARSA technique.

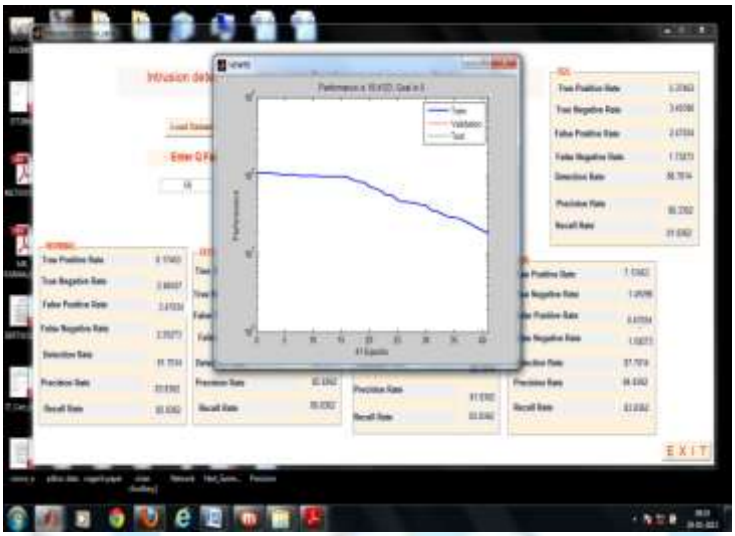


Fig 3: Shows that training pattern of RBF neural network for classification of data according to policy make by SARA technique. Here total number of neurons value is 400 and time of iteration is 1000 second

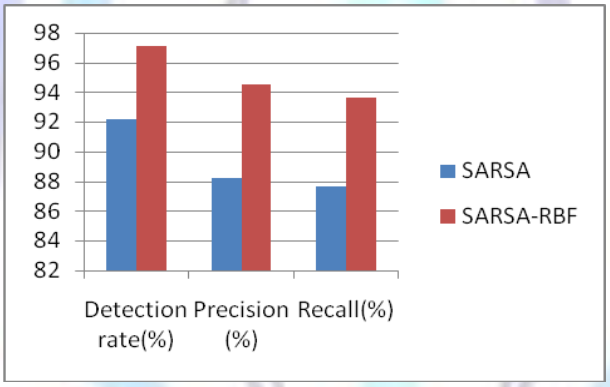


Fig 4: Gives the information of result analysis of first combination of KDDCUP99 data set the total instant value is 7000.

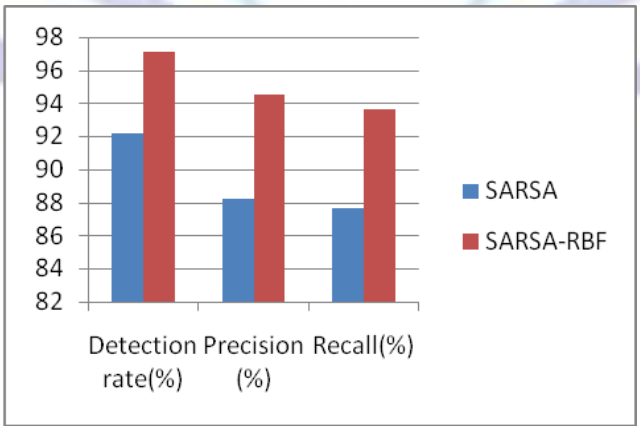


Fig 5: Gives the information of result analysis of second combination of KDDCUP99 data set the total instant value is 4000 attack data and 3000 normal data.

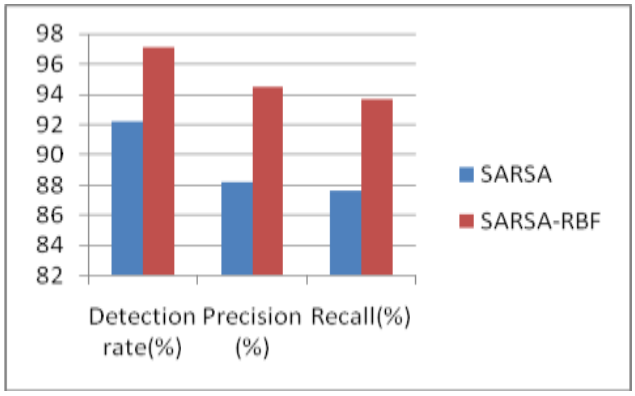


Fig 6: Gives the information of result analysis of second combination of KDDCUP99 data set the total instant value is 5000 attack data and 2000 normal data.

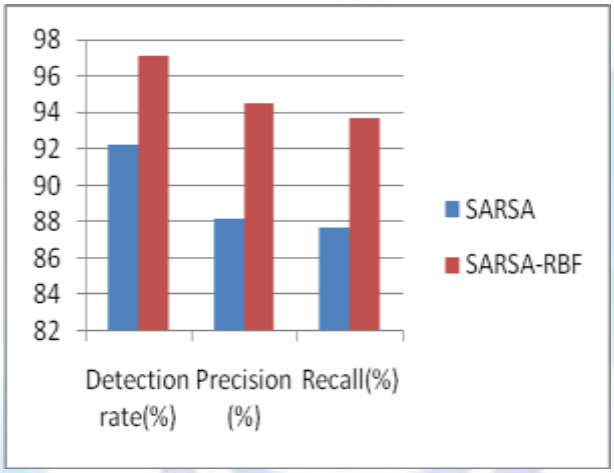


Fig 7: Gives the information of result analysis of second combination of KDDCUP99 data set the total instant value is 2000 attack data and 5000 normal data.

CONCLUSION AND FUTURE WORK

In this paper proposed a method for intrusion detection based on SARSA and RBF neural network. The proposed method classified attack and normal data of KDDCUP99 is very accurately. The proposed method work in process of making policy of SARSA learning par diagram. The learning process of Q factor and RBF training process makes very efficient classification rate of intrusion data. Our empirical result shows better performance in compression of SARSA and another machine learning approach technique for intrusion detection process. In future we will reduce the iteration process of RBF neural network for speed classification and detection of intrusion.

REFERENCES

[1] T. Komviriyavut, P. Sangkatsanee, N. Wattanapongsakorn, and C. Charnsripinyo, “Network intrusion detection and classification with decision tree and rule based approaches”, 9th International Symposium on Communications and Information Technology (ISCIT), 2009, pp. 1046-1050.

[2] N. Ngamwitthayanon and N. Wattanapongsakorn, “Fuzzy-ART in network anomaly detection with feature-reduction dataset” Proceedings - 7th International Conference on Networked Computing, INC2011 , art. no. 6058956 , pp. 116-121, 2011.

[3] Lei Li, De-Zhang Yang, Fang-Cheng Shen,“A Novel rule based Intrusion Detection System”, 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT), China, Vol. 6, pp. 169-172, 9-11 July.

[4] Liwei (Vivian) Kuang,“DNIDS: A Dependable Network Intrusion Detection System”, ACMSIGKDD, 1(2), pp. 67-75.

[5] Hitoshi Iima and Yasuaki kuroe “ Swarm Reniforced leaning Algorithm based on SARSA Method” in SICE Annual Confrence 2008. pp. 2045 – 2049.

[6] Z. Muda, W. Yassin, M.N. Sulaiman, and N.I. Udzir, “Intrusion detection based on K-Means clustering and Naïve Bayes classification”, 7th International Conference on Information Technology in Asia: Emerging Convergences and Singularity of Forms (CITA), 2011.

[7] Jiankun Hu, Xinghuo D. Qiu, Hsiao-Hwa Chen,“A Simple and Efficient Hidden markov model Scheme for Host-based anomaly Intrusion Detection”, Network IEEE, Vol. 23, Issue 1, pp. 42-47, 0890-8044, Jan-Feb 2009.

[8] Marimuthu, A., Dr. A.Shanmugam,"Intelligent Progression for anomaly Intrusion detection", 6th International Symposium on Applied Machine Intelligence and informatics, Coimbatore, pp. 261 - 265, 21-22 Jan 2008.

[9] Ms. P J. Pathak, S S. Dongre,“Attack Detection by Clustering and Classification Approach”, International Journal of Advanced Research in Computer Science and Electronics Engineering, Vol. 1, No 2, pp. 115-118, April 2012.
[10] S.Devaraju, Dr. S.Ramakrishnan,“Performance analysis of intrusion Detection system using various neural networks clasifiers”, International Conference on Recent Trends in Information Technology (ICRTIT), India, pp. 1033-1038, IEEE 3-5 June 2011.
[11] Anshul Chaturvedi and Vineet Richhariya “A Review of Intrusion Detection Based on Instant Based Learning Technique” in International Journal of Computer Science and Technology, Jan - March 2013.
[12] John Zhong Lei, Ali Ghorbani,“Network Intrusion Detection Using an Improved Competitive Learning Neural Network”, Proceeding. Second Annual Conference on Communication Networks and Services Research (CNSR), Canada, pp. 190-197, 2004 IEEE.
[13] Hai-hua gao, Hui-Hua Yang, Xing-Yu Wang,"Ant colony optimization based network intrusion feature selection and detection”, International Conference on Machine Learning and Cybernetics, China, Vol. 6, pp. 3871-3875, 2005 IEEE.
[14] Y I Shakhatreh, K A Bakar,“A Review of Clustering Techniques Based on Machine learning Approach in Intrusion Detection Systems”, International Journal of Computer Science Issues (IJCSI), Vol. 8, Issue 5, No 3, pp. 373-381, Sep 2011. Available: http://ijcsi.org/papers/IJCSI-8-5-3-373- 381.pdf
[15] Wing W. Y. Ng, Rocky K. C. Chang, Daniel S. Yeung, "Dimensionality reduction for denial of service detection problems using rbfnn output sensitivity”, Proceedings of the Second International Conference on Machine Learning and Cybernetics, Wan, 2003 IEEE.
[16] P. Natesan, P. Balasubramanie, G. Gowrison,"Improving Attack Detection Rate in Network Intrusion Detection Using Adaboost Algorithm with Multiple Weak Classifiers”, Journal of Information & Computational Science, Vol. 9 (8), pp. 2239–2251, August 2012. [Online] Available: http://www. joics.com
[17] V.Venkatachalam S.Selvan,"Intrusion Detection using an Improved Competitive Learning Lamstar Neural Network”, International Journal of Computer Science and Network Security, Vol. 7 No. 2, pp. 255-263, February 2007 Available:http://www.paper.ijcsns.org/07_book/200702/200702B11. pdf
[18] Chung-Ming Ou, Yao-Tien Wang, C.R. Ou,"Intrusion Detection Systems Adapted from Agent-based Artificial Immune Systems”, IEEE International Conference on Fuzzy Systems, Taiwan, pp. 115-122 , 27-30 June 2011.
[19] Li Rui, Luo Wanbo,“Intrusion Response Model based on AIS”, International Forum on Information Technology and Applications (IFITA), China, Vol. 1, pp. 86-90, 18-16 July 2010 IEEE.
[20] YUAN Hui, LIU Jian-yong,“Intrusion Detection Based on Immune Dynamical Matching Algorithm”, International Conference on E-Business and E-Government (ICEE), China, pp. 1342-1345, 7-9 May 2010.
[21] Lei Deng De-yuan Gao,“Research on Immune based Adaptive Intrusion Detection System Model”, International Conference on Networks Security, Wireless Communications and Trusted Computing, Vol. 2, pp. 488-491, 25-26 April 2009 IEEE.
[22] Junmin Zhang, Yiwen Liang,“A Novel Intrusion Detection Model Based on Danger Theory”, Pacific-Asia Workshop on Computational Intelligence and Industrial Application, Vol. 2, pp. 67-871, 19-20 Dec 2008 IEEE.
[23] Haidong Fu, Xiuo Yuan, Liping Hu,“Design of a Fourlayer Model Based on Danger Theory and AIS for IDS", International Conference on Wireless Communications, Networking and Mobile Computing, Wuhan, pp. 6337-6340, 21-25 Sepetmber 2007 IEEE.
[24] Baoyi WANG Shaomin ZHANG,“A New Intrusion Detection Method Based on Artificial Immune System”, IFIP International Conference on Network and Parallel Computing Workshops, Baoding, pp. 91-98, 18-21 September 2007 IEEE.