International Journal For Technological Research In Engineering
Volume 4, Issue 8, April-2017 *ISSN (Online): 2347 - 4718*

INTRUSION DETECTION RATE USING FEATURE SELECTION AND MODIFIED DENDRITIC CELL ALGORITHM

Neeraj Yadav[1], Abhinav Jain[2]
S.R Group Of Institutions (College Of Science & Engineering) Jhansi.

ABSTRACT: In this paper, we proposed a feature selection and feature reduction method based on a modified DCA algorithm. The proposed algorithm select multiple feature for reduction and the reduce feature set participant in the process of detection. The reduce feature of network file is classified by DCA classification algorithm. In DCA algorithm if the size of data is increasing the selection of attribute process raises problem related to feature selection. For solving this problem Dumpster belief function is used to increase the biased value of feature and feature subset selection [1,2].In this paper, we proposed a very simple and fast feature selection method to eliminate features with no helpful information which result faster learning in process of redundant feature omission. We compared our proposed method with three most successful feature selection algorithm, including Correlation Coefficient, Least Square Regression Error and Maximal Information Compression Index [3].For the validation and performance evaluation of proposed algorithm MATLAB software and KDDCUP99 dataset 10% is used. This dataset contains approx. 5 lacks number of instances. The process of result shows better classification and reduce time instead of another feature reduction.

Keywords: - IDS, NB, SVM, Belief Function, DCA.

I. INTRODUCTION

The Internet has become a major surrounding for disseminating malicious codes, in particular, through a web application. Internet Worms spread through computer networks by probing, attacking and infecting remote computers automatically. Computer security is defined as the protection of computing systems against threats to confidentiality, integrity, and availability. Confidentiality means that information is disclosed only according to policy, integrity means that information is not destroyed or corrupted, and that the system performs correctly, availability means that system services are available when they are needed. Security threats come from different sources such as natural forces, accidents, failure of services and people known as intruders. There are two types of intruders: the external intruders who are unauthorized users of the machines they attack, and internal intruders, who have permission to access the system with some restrictions. Intrusions, in particular, web based ones, have become increasing threats for important information Due to the rapidly increasing unauthorized activities on the network, Intrusion Detection System (IDS) as a component of defines-in- depth is very necessary because traditional firewall techniques cannot provide complete protection against intrusion. Presently the main threats against network and information security are the attack on the network infrastructure. Intrusion Detection (ID) is an active and important explore area of network security. There are several methods used to implement intrusion detection such as statistical analysis, expert systems, and state transition approaches, etc., and these several approaches are based on the immune system were proposed in recent years .The goal of Intrusion Detection is to detect unauthorized access and abuse of computer systems by both system insiders and external intruders, and secure the system integrity, confidentiality, usability and availability.

II. LITERATURE WORK

A rich literature of intrusion detection focuses on feature reduction using soft computing and neural network. Many of these techniques are based on principal of component into a set of class classification problems. Despite the success of these techniques reported in different domains for various types of applications, such as text document classification, and speech recognition, most of these techniques are mainly proposed for learning from relatively balanced training data. However, in much application, the training data can be often intrusion, where some classes of data have a small number of samples compared to the other classes, and in which it is important to accurately classify the minority cases.

B.Senthilnayaki ,Dr.K. Venkatalakshmi [1]explain, the rapid advancement of computer networks has led to many security problems by malicious users to the modern computer systems. Hence, it is necessary to detect illegitimate users by monitoring the unusual user activities in the network. In this paper, we propose an Intrusion Detection System (IDS) which uses a genetic algorithm based feature selection approach and a Support vector machine based classification algorithm.

Chung-Ming Ou, Yao-Tien [2] use an agent-based artificial immune system (ABAIS) to apply intrusion detection systems (IDS). A multi agent-based IDS (ABIDS) inspired by the danger theory of human immune. ABAIS is an intelligent system with learning and memory capabilities.

Matzinger. P, explains the feature reduction method with using classifier and the details are Synthetic Minority Oversampling Technique (SMOTE) is applied to the training dataset [5]. A feature selection method based on Information Gain is presented and used to construct a reduced feature subset of NSL-KDD dataset. Random Forests are used as a classifier for the proposed intrusion detection framework. Empirical results show that Random Forests classifier with SMOTE and information gain based feature selection gives better performance in designing IDS that is efficient and effective for network intrusion detection.

III. PROPOSED METHODLOGY

In this paper, we proposed a feature selection and reduction based intrusion detection system. The process of feature reduction and selection improves the detection and classification ratio of intrusion detection system. The feature selection process used for finding common feature for attacker participant and feature reduction processes used for unwanted feature for those who are not involved in the attack and normal communication. Dendritic cell algorithm (DCA) is used for the reduction of feature. The DCA function work on common feature correlation and generates similar and dissimilar pattern with the help of modified algorithm. The reduction process reduces the large number of attribute and improves the detection of intrusion detection system.

Modified DCA:
Input: S = number of input string
Output: D = set of string as normal and abnormal
DBF= Dempester belief function

Begin
Create an initial population of belief function (BFs), D
Create a set to contain migrated BFs, M
For all data items in S do
Create a set of BFs randomly selected from D,P
For all BFs in P do
Add data item to BFs collected list
Update abnormal, check entropy condition
Update concentrations of output string
Migrate the BF from D to M and create a new BFinD if concentration of string as matched
End
End
If entropy (normal && abnormal)
Abnormal=high
Normal=low
Pass DBF (abnormal)
For all BFs in DBF do
Set BF to be semi-mature if output concentration of semi-mature unmatched
Otherwise set as matched

Denial of Service Attacks	Back, land, neptune, pod, smurf, teardrop
User to Root Attacks	Buffer_overflow, loadmodule, perl, rootkit,
Remote to Local Attacks	Ftp_write, guess_passwd, imap, multihop, phf, spy, warezclient, warezmaster
Probes	Satan, ipsweep, nmap, portsweep

End
For all data items in S do
Calculate number of times data item is presented by a mature BF and unmatched
Add data item to labelled set M
End
End
METHDOLOGY STEP

In this section we discuss the steps of methodology for improved intrusion detection using DCA function. In this proposed model we used some steps they are following:-

Step1: With the help of Hybrid algorithm we estimate the nature of attack.

Setp2: Demister-Belief Theory is used to compute the probability of evidences that indicate support,which showsstrings are normal and abnormal.
Step3: After the detectioncalculate the entropy of the string, if its entropy is high, treat as abnormal string. On the basis of calculated entropy we find the intruder. Higher entropy, is regarded as the "intruder", and alarm is raised.

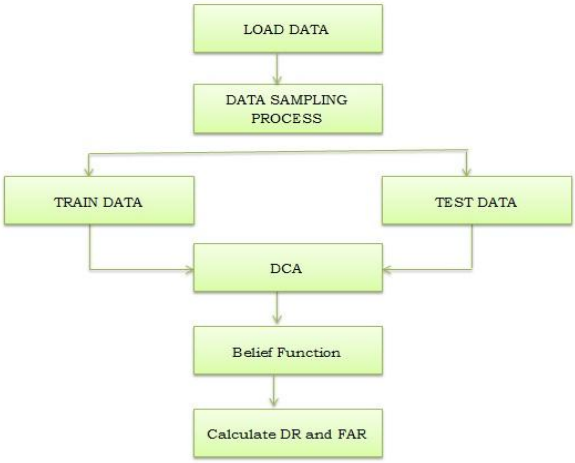

Figure 1: Proposed Model For Feature Based Intrusion Detection.

IV. EXPERIMENTAL RESULT ANALYSIS

In this paper, we perform the experimental process of the proposed classification algorithm for intrusion detection. The proposed method is implemented in Matlab 8.1.0 and tested with very reputed data set from UCI machine learning research center. In the research work, we have measured detection accuracy, true positive rate, false positive rate, true negative rate and finally false negative rate error of classification method. To evaluate these performance parameters I have used KDDCUP99 datasets from UCI machine learning repository namely intrusion detection dataset. Out of these datasets, we create five data set in total number of instant is 7000 and create five different model set. These are number of attacks falling into following categoriesas shown in table 5.1.
We have used parameters i.e. - Accuracy, Precision, Recall for data sets. So we can calculate the false positive and false negative rate of IDS, which are performance indicators of IDS. Precision measures the proportion of predicted

positives/negatives which are actually positive/negative. Recall is the proportion of actual positives/negatives which are predicted positive/negative. Accuracy is the proportion of the total number of prediction that were correct or it is the percentage of correctly classified instances. Below we are showing how to calculate these parameters by the suitable formulas. And also, below we are showing the graph for that particular data set.

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

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

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

FPR= $\frac{FP}{FP+TN}$, FNR= $\frac{FN}{FN+TP}$

For evaluation of performance, we used a different number of ratios of dataset for classification of intrusion data.

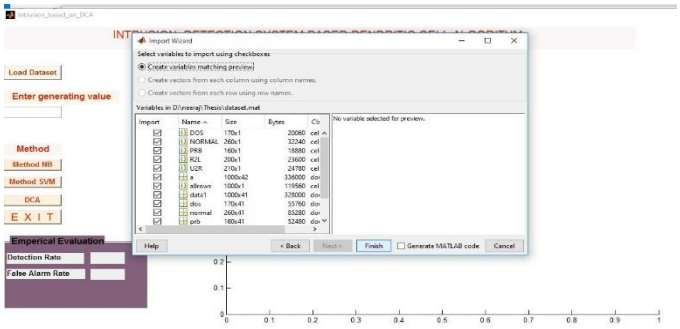

Figure 2: Loading Dataset

Figure shows that data selection windows of all type the data type and initially load the data set for intrusion detection classification

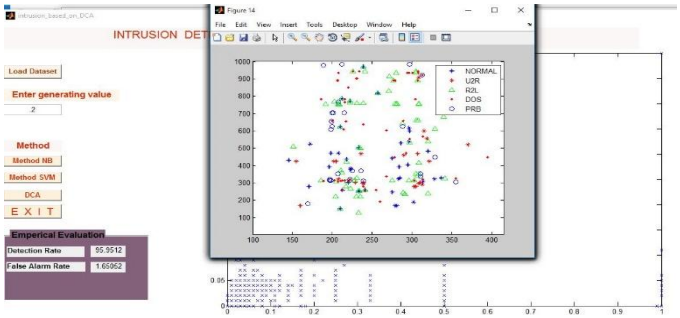

Figure 3: Data Uploading Process Of The Method M DCA For Generating Result Value 0.2.

Figure shows that data uploading process for intrusion data classification of the method M DCA for generating result value 0.2.

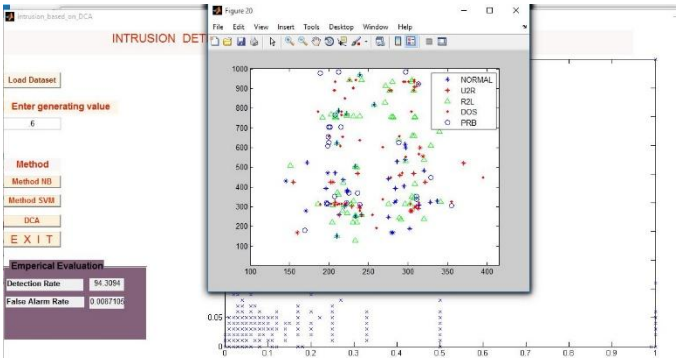

Figure 4:Data Classifying In Attack Categories M DCA For Generating Result Value 0.6.

Figure shows data classifying in attack categories such as Normal, U2R, R2L, DOS and Probe by M DCA method for generating result value 0.6.

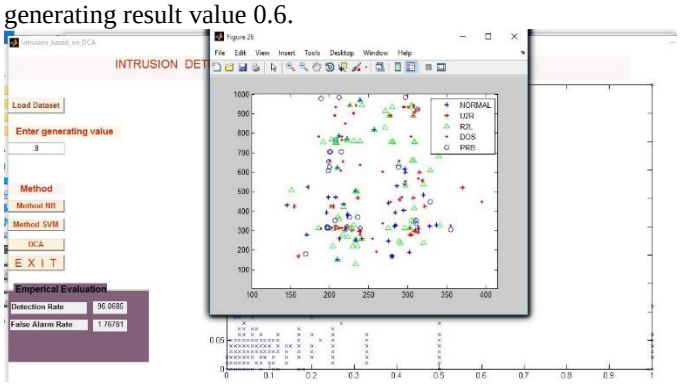

Figure 5: Data Uploading Process Of The Method M DCA For Generating Result Value 0.8.

Figure shows that data uploading process for intrusion data classification of the method M DCA for generating result values 0.8.

Table 1: Shows the performance evaluation of classification.

Metric		DR	FAR
0.2	Method NB	90.9597	4.3543
	Method SVM	91.4417	5.9149
	Method M DCA	95.9512	1.6505
0.6	Method NB	89.2706	2.6652
	Method SVM	89.7999	4.2731
	Method M DCA	94.3094	0.0087
0.8	Method NB	91.3341	4.7186
	Method SVM	91.5590	6.0321
	Method M DCA	96.0685	1.7678

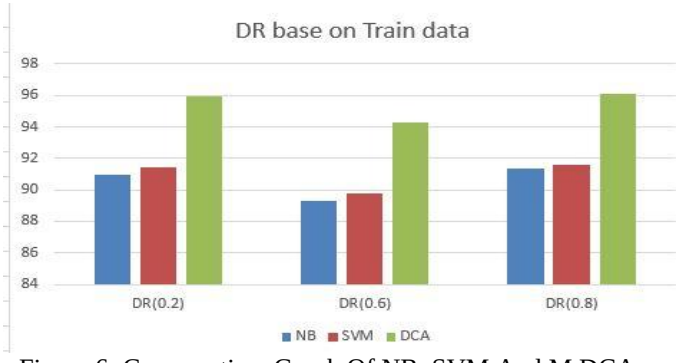

Figure 6: Comparative Graph Of NB, SVM And M DCA For The DR.

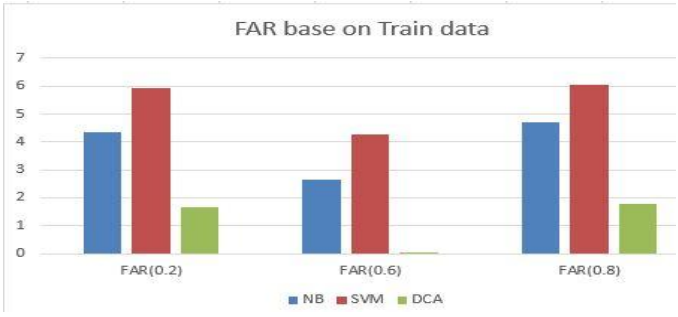

Figure 7:Comparative Graph Of NB, SVM And M DCA For The FAR.

Figure 6,7 shows that the comparative result graph for the intrusion detection classification on the basis of NB, SVM and M DCA for the generating value 0.2, 0.6, 0.8, and also shows that our proposed method DCA gives the better classification detection rate and low false alarm rate.

V. CONCLUSION AND FUTURE WORK

In this we proposed a feature based intrusion data classification technique. The reduction process of feature attribute is performed by BF function along with feature correlation factor. The proposed method work as feature reducers and classification technique, because of this reduction of feature attribute, the execution time of classification also decreases. This decrease time increases the performance of intrusion detection system. Our experimental process gets some standard attribute set of intrusion file such as pot_type, service, sa_srv_rate, dst_host_count, dst_host_sa_srv_rate. These feature attribute are most important attribute in domain of network traffic area. The classification rate achieved in these attribute is 98 %.

In this paper reduction computational time of feature selection process is main objective With proposed technique, consumed time of each algorithm with different reject threshold measured is increased. As evaluation result shows, although FFR (Fast Feature Reduction) cannot defeat other methodologies in accuracy of classification and accuracy didn't changed very much, but in speed FFR outperformed all other feature selection method with great differences. We used DCA classifier with BF for developing efficient and effective IDS

REFERENCES

[1] B.Senthilnayaki , Dr.K. Venkatalakshmi , Dr. A. Kannan "Intrusion Detection Using Optimal Genetic Feature Selection and SVM based Classifier" 3rd International Conference on Signal Processing, Communication and Networking (ICSCN), IEEE 2015

[2] Chung-Ming Ou, Yao-Tien Wang "Intrusion detection systems adapted from agent-based artificial immune systems" Fuzzy Systems (FUZZ), 2011 IEEE International Conference,

[3] Chung-Ming Ou , "Host-based intrusion detection systems adapted from agent-based artificial immune systems" Expert Systems with ApplicationsVolume 39, Issue 1, January 2012, Pages 129–141

[4] Debar H, Wespi A (2007), "Intrusion Detection Based On Immune Dynamical Matching Algorithm", LNCS 2212, pp 85-103.

[5] Matzinger. P, "Intrusion Detection Using Random Forests Classifier With Smote And Feature Reduction", Annual Review in Immunology, vol.12, 2004, pp. 991-1045.

[6] Junmin Zhang, Junmin Zhang "A Novel Intrusion Detection Model Based On Danger Theory" Computational Intelligence and Industrial Application, PACIIA . Pacific-Asia Workshop, IEEE 2008

[7] P. Matzinger, "Research on Immune Based Adaptive Intrusion Detection System Model" Annual Review in Immunology, vol. 12, pp. 991–1045, 1994.

[8] AbdelkaderAlem, YoucefDahmani, AllelHadjali, "On the use of Belief Functions to improve High Performance Intrusion Detection System" 12th International Conference on Signal-Image Technology & Internet-Based Systems, IEEE 2016

[9] Haraszti, Z Townsend, "Decision Tree Classifier For Network Intrusion Detection With GA-Based Feature Selection", International Conference on 1998, pp: 1443 - 1450 vol.3.

[10] MahbodTavallaee, EbrahimBagheri, Wei Lu, and Ali A. Ghorbani "A Detailed Analysis of the KDD CUP 99 Data Set" Proceedings of the 2009 IEEE Symposium on Computational Intelligence Đn Security and Defense Applications (CISDA 2009)

[11] AdetunmbiA.Olusola., AdeolaS.Oladele. and DaramolaO.Abosede "Analysis of KDD '99 Intrusion Detection Dataset for Selection of Relevance Features", Proceedings of the World Congress on Engineering and Computer Science 2010 Vol I WCECS 2010, October 20-22, 2010, San Francisco, USA

[12] John Zhong Lei and Ali Ghorbani "Network Intrusion Detection Using an Improved Competitive Learning Neural Network" in Proceedings of the Second Annual Conference on Communication Networks and Services Research IEEE.

[13] N Wing, Rocky Chang and Daniel Yeung "Dimensionality Reduction For Denial Of Service Detection Problems Using Rbfnn Output Sensitivity" in Proceedings of the Second International Conference on Machine Learning and Cybernetics, Wan, 2-5 November 2003.

[14] Deepak Rathoreand Anurag Jain "a novel method for intrusion detection based on ecc and radial bias feed forword network" in Int. J. of Engg. Sci. & Mgmt. (IJESM), Vol. 2, Issue 3: July-Sep.: 2012.

[15] AnshulChaturvedi and Prof. VineetRichharia "A Novel Method for Intrusion Detection Based on SARSA and Radial Bias Feed Forward Network (RBFFN)" in international journal of computers & technology vol 7, no 3.

[16] Mohammad Behdad, Luigi Barone, Mohammed Bennamounand Tim French "Nature-Inspired Techniques in the Context of Fraud Detection" in ieee transactions on systems, man, and cybernetics part c: applications and reviews, vol. 42, no. 6, november 2012.

[17] Alberto Fernandez, Maria Jose del Jesus and Francisco Herrera "On the influence of an adaptive inference system in fuzzy rule based classification system for imbalanced data-sets" in Elsevier Ltd. All rights reserved 2009.

[18] P. Garcia-Teodoro, J. Diaz-Verdejo, G. Macia-Fernandez and E.Vazquez "Anomaly-based network intrusion detection: Techniques, Systems and challenges" in Elsevier Ltd. All rights reserved

2008.

[19] Terrence P. Fries "A Fuzzy-Genetic Approach to Network Intrusion Detection" in GECCO 08, July12–16, 2008, Atlanta, Georgia, USA.

[20] ZoranaBankovic, DusanStepanovic,SlobodanBojanic and Octavio Nieto-Taladriz "Improving network security using genetic algorithm approach" in Published by Elsevier Ltd 2007.

[21] Mrutyunjaya Panda and ManasRanjanPatra "network intrusion detection using naive bayes" in IJCSNS International Journal of Computer Science and Network Security, VOL.7 No.12, December 2007.

[22] AnimeshPatcha and Jung-Min Park "An Overview of Anomaly Detection Techniques: Existing Solutions and Latest Technological Trends" in Computer networks 2007.

[23] RenHui Gong, Mohammad Zulkernine and PurangAbolmaesumi "A Software Implementation of a Genetic Algorithm Based Approach to Network Intrusion Detection" in IEEE 2005.

[24] Jonatan Gomez and DipankarDasgupta "Evolving Fuzzy Classifiers for Intrusion Detection" in IEEE 2002.

[25] Francisco Herrera "Genetic fuzzy systems: taxonomy, current research trends and prospects" in Springer-Verlag 2008.

[26] Adel NadjaranToosi and Mohsen Kahani "A new approach to intrusion detection based on an evolutionary soft computing model using neuro-fuzzy classifiers" in Elsevier B.V. All rights reserved 2007.

[27] G. Shafer, A Mathematical Theory of Evidence, Princeton, University Press, Princeton, NJ, 1976. E. Blanzieri and A. Bryl "Asurvey of learning-based techniques of email spam filtering" Artif. Intell. Rev., vol. 29, no. 1, pp. 63–92, 2008.

[28] WunHwa Chen, ShengHsun Hsu, and HwangPinShen "Application of SVM and ANN for intrusion detection" Computers & Operations Research, Vol.32, 2005, pp. 2617–2634.

[29] M. Wu, R. C. Miller, and S. L. Garfinkel "Do security toolbars actually prevent phishing attacks" in Proc. SIGCHI Conf. Human Factors Comput. Syst., 2006, pp. 601–610.

[30] S. X. Wu and W. Banzhaf "The use of computational intelligence in intrusion detection systems: A review" Appl. Soft Comput., vol. 10, no. 1, pp. 1–35, 2010.

[31] Engen, J. Vincent, and K. Phalp "Enhancing network based intrusion detection for imbalanced data" Int. J. Knowl.-Based Intell. Eng. Syst., vol. 12, no. 5–6, pp. 357–367, 2008.

[32] K. Shafi, T. Kovacs, H. A. Abbass, andW. Zhu "Intrusion detection with evolutionary learning classifier systems" Nat. Comput., vol. 8, no. 1, pp. 3–27, 2009.

[33] Y.Yang and S. A. Elfayoumy "Anti-spam filtering using neural networks and Baysian classifiers" in Proc. IEEE Int. Symp. Comput. Intell. Robot. Autom., 2007, pp. 272–278.

[34] J. Zhang, G. Yang, L. Lu, M. Huang, and M. Che "A novel visualization method for detecting DDoS network attacks" Vis. Inf. Commun., pp. 185–194, 2009.

[35] P. Espejo, S. Ventura, and F. Herrera "A survey on the application of genetic programming to classification" IEEE Trans. Syst.,Man, Cybern. C, Appl. Rev., vol. 40, no. 2, pp. 121–144, Mar. 2010.

