International Journal of Scientific Research & Engineering Trends
Volume 6, Issue 1, Jan-Feb-2020, ISSN (Online): 2395-566X

Robust Intrusion Detection System using SVM and Artificial Neural Network

Rajkumar Pandey, Dr. Shiv Shakti Shrivastava
Department of Computer Science & Engineering,
Rabindranath Tagore University - Bhopal, MP, India.

Abstract - **With large increase of internet user's network security is important issues in today era. As variety of users have different requirement, so proper identification of safe network is required. This work focus on the network intrusion detection using SVM and neural network. Here SVM classify network behavior into two class first is safe and other is unsafe. Once unsafe network is identified then trained neural network identified attack type of the input sessions. So Whole work is divide into two modules, first is separation of safe and unsafe session from the dataset using SVM. Then in second module identification of type of intrusion is done in unsafe network by EBPNN. Experiment is done on real dataset and obtained results are better than previous works on different parameters.**

Keywords- **Anomaly Detection, ANN, Clustering, Genetic Algorithm, Intrusion Detection.**

I. INTRODUCTION

Providing web network protection to many web services on the internet, distinctive network foundations, communications arrange abundant means that has been taken similar to encryption, firewall, and virtual confidential network and so forth systematize Intrusion detection structure is a notable advance among those. Intrusion detection field increases up out of a large amount of current couple of years and built up a immense deal which utilizes the assembled information from a variety of sort of interruption assault, based on those distinguishing business and open source training items emerge to harden your network to improve safety of the miscellaneous correspondence, service providing networks. As the amount of network customers and machine are growing step by step to give different sort of directions and smoothness for the efficiency of the world. Be that as it may, some unapproved customers or exercises from diverse sorts of attackers which may hidden attack or external attack keeping in mind the final aim to hurt the running structure, which are recognized as programmers. The elementary consideration procedure of such sort of programmer and gatecrashers is to slash down cumbersome networks and web directions. Because of growth in eagerness of network safety of diverse forms of attack, frequent scientists has added their keenness for their field and broad collection of gatherings and in accumulation result has been formed by them, with a precise end objective to provide protected directions to the end customers. Along with diverse forms of attack interruptions is a sort of attack that build up a business scheme. Interruption detection structure is accessible for the insurance from disruption attack.

From the above exchange this work can lock the chief spot of the network Intrusion detection framework is to recognize all imaginable intermission which execute malicious movement, computer assault, extend of infections, computer abuse, and so forth so a Internet intermission recognition structure investigations distinguishing data packages as well as monitor that movement over the Internet for such sort of spiteful act. So the smooth operation of universal network distinguishing server needs to resolve largely network which go about as network Intrusion detection framework that screen every one of the parcels developments and distinguish their behavior with the harmful exercises.

II. RELATED WORK

Yogitha et. al. [1] presented interruption discovery framework with Support Vector Machine (SVM). Confirmation is ended by organizing surveys on NSL-KDD Cup'99 information gathering which is reformer type of KDD Cup'99 data index. By using this NSLKDD Cup'99 information gathering they have reduced spacious time essential to shape SVM exemplary by attainment appropriate pre-training on information gathering. In this organization SVM made bunching of information. By compulsion suitable part gathering attack location rate is opened up and false positive rate (FPT) is pointed. In this planned work writer has used Gaussian Circular Basis.

A.R. Jakhale, et. al [2] In this exertion the writer depicts a anamoly discovery framework and its two phases chiefly

are training and testing. The slipping window and gathering is familiar to tending the web network move by pulling out the recurring examples using computations. The estimation is as authentic and used as a division of regular monitoring. The standard multi-design communicable computation has elevated location rate. At long last, boost the recognition rate and compact the fake aware rate.

Research by Jiefei, Lobo and Russo [3] discovers the occasion of Multi-way steered assault where an attack is separated and sent over dissimilar courses to attempt to deception an IDS framework. This is unfair feasible due to multi way TCP (MPTCP) which enables communication to classes finished frequent ways between a foundation and objective.

Barolli et al [4] investigates the consumption of IDS using neural network for giving IDS arrangement in a Tor (The Onion Router) manage. Examinations did use a Tor server and client with back engendering NN to replicate exchanges over the Tor organizes while infectious for assessment. The structure planned is a ready ANN with data taken from Wireshark, at that position the server and client data are examined, and distinctions will identified an interruption or abuse. The conclusion from testing was productive in giving feasible accuracy when charged in the test situation.

ChuanLong [5] In this article, writer examine how to present an interruption recognition framework in light of thoughtful learning, and this exertion offer a thoughtful knowledge approach for interruption recognition using recurrent neural networks (RNN-IDS). In addition, this exertion inspect the execution of the model in balancing categorization and multiclass arrangement, and the amount of neurons and characteristic learning rate impacts on the implementation of the planned show. This effort compare it and those of J48, artificial neural network, arbitrary woodland, bolster vector machine, and further machine knowledge approach planned by history analysts on the standard information directory index.

III. METHODOLOGY

Whole work is dividing into different modules base on the steps of calculation from the user query to final output on the screen. In fig. it is seen that there are two different modules. first is separation of various class data into separate cluster with the help of SVM. Then in second module learning of clustered data was done where error back propogation neural network was used for this module.

1. **Pre-Processing**
   Here information likes type of protocol, socket type, etc. are removed. As presence of these information increases confusion for the clustering process. This can be understand as let raw data session is
   {0,tcp,ftp_data,SF,491,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,2,0.00,0.00,0.00,0.00,1.00,0.00,0.00,150,25,0.17,0.03,0.17,0.00,0.00,0.00,0.05,0.00,normal,20}
   After applying pre-processing session will be:
   {0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,2,0.00,0.00,0.00,0.00,1.00,0.00,0.00,150,25,0.17,0.03,0.17,0.00,0.00,0.00,0.05,0.00,normal}

2. **Training of SVM (Support Vector Machine):**
   As dataset consist of sessions which has normal behavior of network as well as abnormal behavior of network. Separation of session in these two category is done by SVM. Here pre-processing is done for the sessions where unnecessary information is remove and data is arrange for training.

   In this phase both feature vector of safe and unsafe mode is pass in the SVM, here the use of linear separation training of the SVM is done. The goal of linear classification problem is to obtain two parallel hyperplanes as shown in Fig. 1, $wx + b=1$ and $wx' + b=-1$, where w and b are the classification parameters obtained during the training process. Both
   Hyperplanes separate the training data of the two classes such that the distance between those hyperplanes is maximized.

3. **Session Separation:**
   In this step of first module trained SVM is use for the separation of session in two category first is safe mode and other is unsafe mode. So session from the testing dataset is pass one by one in the trained SVM. Output of the SVM is two class + or -.

4. **Error Back Propagation Neural Network**
   In this module cluster dataset of sessions are utilize to train the EBPNN. Then trained dataset is utilize for the testing of unknown attack session.

5. **Input Feature**
   As cluster dataset provide the sessions in group as per the SVM so it has to divide into two group first act as neural network input while other act as desired output. Considering first input feature vector which consist of numeric values where clustered numeric values are arrange in the input matrix. While second desired output vector consist of the class of session which was obtained from the genetic algorithm. This can be understand by below example where

6. **Training of Error Back Propagation Neural Network (EBPNN):** Here feature vector obtained are used as the input in the neural network while desired output make proper weight adjustment in the

network. So with fix number of iteration or epochs work will get trained neural network

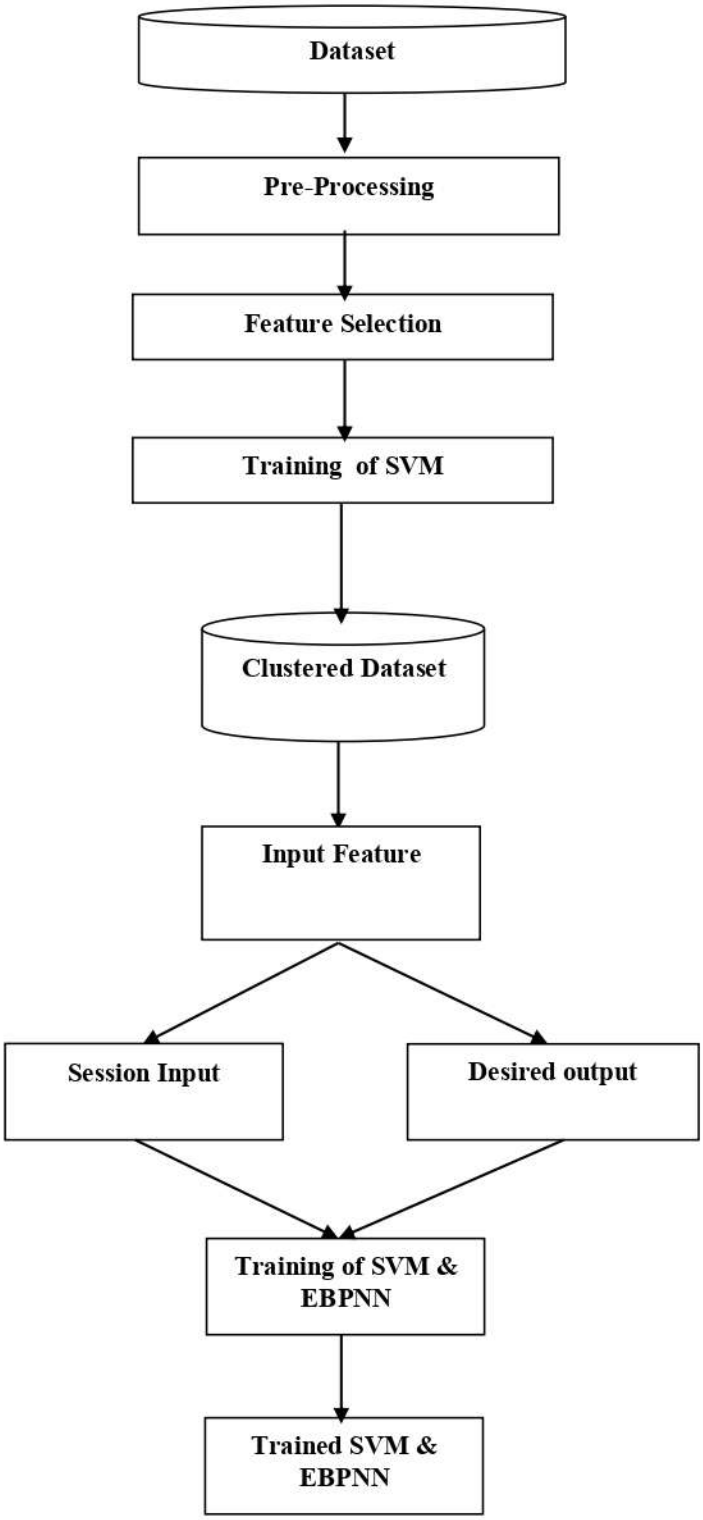

Fig.1. Proposed work first module.

7. Proposed Algorithm: Neural Network Algorithm

Input: D // Dataset
Output: EBPNN // Trained Neural Network

1. PD←Pre_Process(D) // Preprocessed Dataset
2. CD←SVM(PD) // CD: Clustered Dataset
3. Loop 1:n // n : number of session in the dataset
4. S←CD[n]
5. [X D]←Input_Feature(S) // X input session feature and D desired output
6. EndLoop
7. EBPNN←Initialize(In, Hn, On,) // In: neurons in inputs layer, Hn Number of Hidden layer, On: number of output layer
8. Loop 1:n
9. Loop 1: iter // iter: number of iterations
10. EBPNN←Train(EBPNN, X[n], D[n])
11. EndLoop

8. Testing of EBPNN

In this step input session is preprocess as done in the training module, similarly feature vector is create for input in the neural network. Finally feature vector is input in the EBPNN which give output. Now analysis of that output is done that whether specified class is desired one or not.

IV. EXPERIMENT AND RESULTS

Data Set For the evaluation of the whole work the dataset is NSL KDD [12] about which previous chaper has already explained and the collection of the all evaluating vectors look like. Where numeric terms are use for feature learning and at the end of each vector it has the corresponding class. The pre-processing step and its requirement have been already explained.

1. Evaluation Parameter

To test our result this work use following measures the accuracy of the, that is to say Precision, Recall and F-score. These parameters are depend on the TP, TN, FP and FN.

$$\Pr ecision = \frac{True\_Positive}{True\_Positive + False\_Positive}$$

$$\mathrm{Re}call = \frac{True\_Positive}{True\_Positive + False\_Negative}$$

$$F\_Score = \frac{2 * \Pr ecision * \mathrm{Re}call}{\Pr ecision + \mathrm{Re}call}$$

2. Results

Table I: Precision value based comparison of IDS.

Data-Set Size	Intrusion Detection System	
	RNN [11]	SVM&EBPNN
2000	0.879694	0.934478
4000	0.879532	0.974698
8000	0.877113	0.953863

Above table 1 shows that proposed hybrid model SVM and EBPNN has increased the precision value by7.91 %. This enhancement was achieved by use of SVM for initial classification of normal and attack sessions.

Table II: Recall value based comparison of IDS.

Data-Set Size	Intrusion Detection System	
	RNN [11]	SVM&EBPNN
2000	0.978754	0.968997
4000	0.978571	0.986621
8000	0.977572	0.987619

Above table 2 shows that proposed hybrid model SVM and EBPNN has increased the recall value by 2.83%. This enhancement was achieved by use of EBPNN for further identification of intrusion classes. Use of al feature values for training of neural network has improved the parameter values.

Table III: F-measure value based comparison of IDS.

Data-Set Size	Intrusion Detection System	
	RNN [11]	SVM&EBPNN
2000	0.926584	0.951425
4000	0.978571	0.980623
8000	0.924622	0.960692

Above table 3 shows that proposed hybrid model SVM and EBPNN has increased the F-measure value by 2.17%. This enhancement was achieved by use of SVM for initial classification of normal and attack sessions.

Table IV: Accuracy value based comparison of IDS.

Data-Set Size	Intrusion Detection System	
	RNN [11]	SVM&EBPNN
2000	0.927	0.95
4000	0.926	0.9796
8000	0.923857	0.958429

Above table 2 shows that proposed hybrid model SVM and EBPNN has increased the accuracy value by 3.84%. This enhancement was achieved by use of EBPNN for further identification of intrusion classes. Use of al feature values for training of neural network has improved the parameter values.

Table V: Execution time (Seconds) value based comparison of IDS.

Data-Set Size	Intrusion Detection System	
	RNN [11]	SVM&EBPNN
2000	31.4567	29.0813
4000	44.5671	42.4338
8000	63.5533	62.3971

Above table 2 shows that proposed hybrid model SVM and EBPNN has reduced the time value by 4.23%. This enhancement was achieved by use of EBPNN for further identification of intrusion classes. Use of al feature values for training of neural network has improved the parameter values.

V. CONCLUSION

Detection of intrusion in a network is an important issue as number of researchers have proposed different model for its detection. Tis paper has proposed a neural network based model for prediction of session status as normal or intrusion. As two class partitions by SVM model was quit perfect so detection of normal condition from other is identified by this SVM. While to learn about the class of intrusion if session is malicious than EBPNN was used. This two step classification increases the accuracy of the work. Experiment was done on real datast NSL-KDD while comparison was done by existing methods. Results shows that proposed SVM&EBPNN model has increase the precision by 7.91%, while accuracy was enhance by 3.84%. In future researcher can reduce training feature vector.

REFERENCES

[1]. Yogita B. Bhavasar, Kalyani C. Waghmare "Intrusion Detection System Using Data Mining Technique: Support Vector Machine" 2013 International Journal of Emerging Technology and Advance Engineering volume 3, Issue 3, March 2013.

[2]. A.R. Jakhale, G.A. Patil, "Anomaly Detection System by Mining Frequent Pattern using Data

Mining Algorithm from Network Flow", International Journal of Engineering Research and Technology, Vol. 3, No.1, January 2014, ISSN. 2278-0181.

[3]. Aljurayban, N.S.; Emam, A. (21-23 March 2015). Framework for Cloud Intrusion Detection System Service. Web Applications and Networking (WSWAN), 2015 2nd World Symposium on, p1-5

[4]. Barolli, Leonard; Elmazi, Donald; Ishitaki, Oda, Tetsuya; Taro; Yi Liu, Uchida, Kazunori. (24-27 March 2015). Application of Neural Networks for Intrusion Detection in Tor Networks. Advanced Information Networking and Applications Workshops (WAINA), 2015 IEEE 29th International Conference on, p67-72.

[5]. Koushal Kumar, Jaspreet Singh Batth "Network Intrusion Detection with Feature Selection Techniques using Machine-Learning Algorithms" International Journal of Computer Applications (0975 – 8887) Volume 150 – No.12, September 2016

[6]. R. Karthik, Dr.S.Veni, Dr.B.L.Shivakumar "Improved Extreme Learning Machine (IELM) Classifier For Intrusion Detection System" International Journal of Engineering Trends and Technology (IJETT) – Volume-41 Number-2 - November 2016

[7]. Premansu sekhara rath, 2manisha mohanty, 3silva acharya, 4monica aich "optimization of ids algorithms using data mining technique" International Journal of Industrial Electronics and Electrical Engineering, ISSN: 2347-6982 Volume-4, Issue-3, Mar.-2016

[8]. Mohammadreza Ektefa, Sara Memar, Fatimah Sidi, Lilly Suriani Affendey "Intrusion Detection Using Data Mining Techniques", 978-1-4244-5651-2/10/$26.00 ©2010 IEEE

[9]. YU-XIN MENG," The Practice on Using Machine Learning For Network Anomaly Intrusion Detection" Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong, 978-1-4577-0308-9/11/$26.00 ©2011 IEEE

[10]. Liu Hui, CAO Yonghui "Research Intrusion Detection Techniques from the Perspective of

[11]. Machine Learning" 2010 Second International Conference on MultiMedia and Information Technology 978-0-7695-4008-5/10 $26.00 © 2010 IEEE

[12]. Chuanlong Yin , Yuefei Zhu, Jinlong Fei, And Xinzheng He. "A Deep Learning Approach For Intrusion Detection Using Recurrent Neural Networks" current version November 7, 2017. Digital Object Identifier 10.1109/ACCESS.2017.2762418

[13]. https://github.com/defcom17/NSL_KDD/blob/master/Original%20NSL%20KDD%20Zip.zip