IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 5, May 2022

DOI: 10.17148/IJARCCE.2022.11529

Enhance Network Aggression Classification Using Neural Network and SVM

Vijay Kumar Uikey, Sushma Kushwaha

Department of Computer Science & Engineering, Swami Vivekanand college of Science and Technology, Bhopal (M.P)

Assistant Professor, Department of Computer Science & Engineering, Swami Vivekanand college of Science and Technology

Abstract: The objective of this dissertation is to propose an improved ensemble classifier method based on Neural Network and Gaussian Support Vector Machines, for cyber-attack classification problem. Previous work done [1] using hybrid techniques for Cyber attack classification was suffering while working with less amount of data also the structure of hybrid technique is very complex. The improved ensemble classifier is built using two different types of classification techniques. In the proposed method base classifier is the Support Vector Machine and another one is Neural Network classifier. The process of ensemble is done by bagging process, which uses multiple kernel function. The multiple kernels are Gaussian in nature.

Keywords: Cyber,Neural network, SVM, Optimize, GSVM.

I. INTRODUCTION

Computer security is protection of computer systems against different types of threats like privacy, integrity, and accessibility. Confidentiality is nothing but disclosing information according to rule. Integrity means that information is not damaged and that the system performs correctly, availability means that system services are available when needed. Threats come from different sources such as natural forces, accidents, failure of services and people known as intruders. Intruders are of two types, the external intruders who are unlawful users of the machines and internal intruders are those who have permission to access the system with some restrictions. Traditional prevention techniques like user authentication, data encryption, avoiding programming errors and firewalls are first way of defense for computer security. Intrusion detection is therefore required as an additional requirement for protecting systems. Intrusion detection is useful not only in detecting intrusions correctly, but also provides information about timely actions to be taken. With the incredible growth of network-based services and sensitive information on networks, network security is becoming more and more importance than before. Intrusion detection techniques are the last way of defense against computer attacks succeeding to secure network architecture design, firewalls, and private screening. Despite the overabundance of intrusion prevention techniques available, attacks against computer systems are still successful. Thus, intrusion detection systems (IDSs) play a vital role in network security.[2] Symantec in a fresh report uncovered on the increase in number of fishing attacks under theft of important information, such as credit card numbers, passwords, and other financial information are on a hike, going from 9 million attacks in June2013 to over 33 millions in approximately one year. Answer to this is the use of network intrusion detection systems (NIDS) that detect attacks by observing various network activities. It is therefore important that such systems are correct in identifying attacks, easy to train and generate as few false positives. Internet is one of the main ways of communication now days. Different types of internet application are available whose use is in increase. Rapid Increase in usages of network applications also increases security risks. The main security goals which need to be considered while working in a network are confidentiality, Availability, Authentication, Integrity and Non-repudiation.

II. BACKGROUND

Ensemble classification technique plays a vital role in data mining for classification of data. The performance of individual classifier is not so good. So it is not a nice way to utilize only one method or algorithm to solve a particular problem because every algorithm has strength with different limitations. So the best idea is use strengths of one method over the limitations of another algorithm. So techniques of applying algorithms in such way are called ensemble of classifiers. COB (core, outlier, and boundary) method quantitatively measures the accuracies of best part voting ensembles for binary classification.[3]. Good ensemble methods are that in which each individual classifiers are accurate and different. But ensemble methods are mixture of predictions made by a set of individual classifiers. For

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 5, May 2022

DOI: 10.17148/IJARCCE.2022.11529

experimental purpose of COB three different ensemble methods bagging, random forests, and a randomized ensemble, two different individual classifiers and three different machine learning algorithms decision trees, k-nearest neighbors, and support vector machines are used.

ENSEMBLE CLASSIFIER WORKING IS DESCRIBED BELOW

An ensemble classifier constructs a set of 'base classifiers' from the training data
Methods for constructing an Ensemble Classifier
1. Manipulating training set
2. Manipulating input features
3. Manipulating class labels
4. Manipulating learning algorithms

1. Manipulating training set
Resampling using sampling distribution multiple training sets are created, these sampling distribution it determine what is the possibility of an example being selected for training, it varies from one trial to another .By using a particular learning algorithm classifiers are built from training set
Examples: Bagging & Boosting

2. Manipulating input features
For forming training set subset of input feature are taken subset of input features chosen to form each training set, this subset can be chosen randomly or by input given by domain experts, generally for unnecessary features. Random Forest is an example which uses DT as its base classifiers

3. Manipulating class labels
If the numbers of classes are large then the data is changed into binary class problem by randomly partitioning it into two different classes like A1 and A2. Re-labeled examples are used to train a base classifier. By repeating the classification method and model building steps a number of times ensemble of base classifiers is obtained. Example – error correcting output coding
4. Manipulating learning algorithm
If Learning algorithm is applied number of times on same training set ,it results in different models and is automatically manipulated .Example – ANN can produce different models by changing network topology or the initial weights of links between neurons
Example –by introducing randomness into the tree growing procedure ensemble of DTs can be constructed – instead of selecting the best divide attribute at each node, we at random choose one of the top k attributes
First 3 approaches are generic – can be applied to any classifier, fourth approach depends on the type of classifier used, Base classifiers can be generated sequentially or in parallel

SAMPLING TECHNIQUES

Sampling methods consider the class twist and property of the dataset as a whole. However, machine learning and data mining often face nontrivial datasets, which often show characteristics and properties at a local, rather than global level. It is noted that a classifier enhanced through global sampling levels may be insensible to the peculiarities of different components in the data, resulting in a suboptimal performance. Many highly skewed data sets are huge and the size of the training set must be reduced in order for learning to be reasonable hence sampling is used. Under sampling and Over sampling both strategy can be applied in any learning system, since they act as a pre-processing phase, allowing the learning system to receive the training instances as if they belonged to a well-balanced data set. Hulse suggest that the utility of the resampling methods depends on a number of factor, including the ratio between positive and negative examples, additional characteristics of data, and the nature of the classifier. Under-sampling may throw out useful data, while over-sampling artificially increases the size of the data set and as a result, worsens the computational burden of the learning algorithm.

IIIDATASET KDD CUP 99 DESCRIPTION

KDD'99 [4] has been the most widly used data set for the evaluation of anomaly detection methods. This data set is prepared by Stolfo et al. [5] and is built based on the data captured in DARPA'98 IDS evaluation program [6]. DARPA'98 is about 4 gigabytes of compressed raw (binary) tcpdump data of 7 weeks of network traffic, which can be processed into about 5 million connection records, each with about 100 bytes. The two weeks of test data have around 2

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 5, May 2022

DOI: 10.17148/IJARCCE.2022.11529

million connection records. KDD training dataset consists of just about 4,900,000 single connection vectors each of which contains 41 features and is labeled as either normal or an attack, with exactly one specific attack type.[7]
KDD CUP 1999 data was the data set used for the Third International Knowledge Discovery and Data Mining Tools Competition. In our experiments, we trial the data only from the training data set and use in both the training and testing stages. A connection is a sequence of TCP packets containing values of 41 features and labeled as either normal or an attack, with exactly one specific attack type. A complete listing of Features and details are in KDD CUP 1999 data. There are 22 attack types in the training dat. the attacks in the training data are grouped into broad classes; each neuron is then labeled as on behalf of one of these classes. Specifically, the four wide classes of attack type defined by MIT Lincoln Labs are used, as stated below:
1) Denial-of-Service (DoS): These are attacks designed to make some service accessible through the network unavailable to legitimate users.
2) Probe: A Probe is a reconnaissance attack designed to uncover information about the network, which can be exploited by another attack.
3) Remote-to-Local (R2L): This is where an attacker with no privileges to access a private network attempts to gain access to that network from outside, e.g. over the internet.
4) User-to-Root (U2R): The attacker has a legitimate user account on the target network. However, the attack is designed to escalate his privileges so that he can perform unauthorized actions on the network.

Kdd cup 99 features can be categorized in to three categories
1. **Basic features** :It leads to implicit delay in connection and it encapsulates all the attribute that can be collected from the TCP/IP connection
2. **Traffic features:** Includes feature that are collected from window interval and are divided in to following two categories
2.1 Same "host" feature:examine only theconnections in the past 2 seconds that have thesame service as the current connection. And calculate statistical related protocol behavior service.
2.2 Same "Service" feature:examine only theconnections in the past 2 seconds that have thesame service as the current connection.
3. **Content feature:** unlike most of the DoS and Probing attacks, the R2L and U2R attacks don't have anyintrusion frequent sequential patterns. This is becausethe DoS and Probing attacks involve many connectionsto some host(s) in a very short period of time; howeverthe R2L and U2R attacks are embedded in the dataportions of the packets, and involve only a single connection to detect such type of attack we need some feature to look for suspicious behavior These are called as content feature.

IV RELATED WORK

With increase in the use of network based services among people, security of data has become a big problem. Data is continuously flowing through in the network system; this flowing data also contains sensitive data, which needs to be secured. The data which is flowing through, is in huge amount .It is not an easy task to make such large amount of data secure. Since it contains number of features which needs to be extracted, also time required to process on such large amount of data, is large .The problem is to find an answer to these question: "What features need to be taken into consider when calculating or examining whether the activity is malicious or not?" .Based on previous research on IDS, it is clear that any one of the techniques alone cannot detect everything but the grouping of the both is the most capable approach. Even though IDS have been researched for more than 20 years, we do not have an answer to the question of what features should be monitored. So far different kinds of methods and algorithms have been developed for anomaly detection but the focus has been on making them more efficient. The problems which motivated to work on this dissertation are,
1. Processing large amount of data.
2. Pre-processing of data takes large amount of time
3. Rate of false alarm generation is high how to reduce it.
4. Selection of as base classifier, since some data mining classifier are ambiguous situation for selection of base classifier
6. Detection of dynamic feature evaluation.

V PROPOSED ALGORITHM

The proposed method works in three tier architecture. The first module is the preprocessing module in which the KDDcup 99 dataset is applied as an input . The second tier is the feature extractor module where feature are extracted by applying a feature extraction method like transformation, reduction etc it extracts the feature in the form of feature

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 5, May 2022

DOI: 10.17148/IJARCCE.2022.11529

space which are then finally transferred to the third module that is the classification module in in this ensemble classifier are present, which extracts feature in the form of feature space these are given to Neural network where feature allocation is done ,finally the feature allocated are transferred to the SVM ,each vector in the SV's neighborhood has its weights adjusted to become more like the SV. Vector closest to the SV are altered more than the vector furthest away in the neighborhood. Above step are repeated for enough iteration for convergence.Calculating the SV is done according to the Euclidean distance among the node's weights (W_1, W_2, … , W_n) and the input vector's values (V_1, V_2, … , V_n).The new weight for an attribute is the old weight, plus a fraction (L) of the difference between the old weight and the input vector… adjusted (theta) based on distance from the SV. Finally the attack is classified.
Proposed application, initially data which has to be classified is entered in the load dataset entry. In the second stage generating value is provided, which is less than 1.After entering this ,next step is to select the method by which the classification is to be done. Three Classification are available ensemble classifier, hybrid classifier, and the proposed method that is improved ensemble technique you can select any one at a time, result shows that the proposed method gives the best result in terms of false positive, true positive, true negative, false negative, precision rate detection rate and recall rate.

The proposed method works in the following step.
Step1: Initially input Cyber-attack data passes through preprocessing function and extract feature part of Cyber-attack data in form of traffic type.

Step2: the extracted traffic feature data converted into feature vector.

Step 3: In phase of feature mapping in feature space of NN create a fixed class according to the group of data.

Step 4: steps of processing of NN.
1. Initialize Gaussian hyper plane margin.
2. Choose a random vector from training data and present it to the NN.
3. The weight of the plane support vector is estimated. The size of the vector decreases with each iteration.
4. Each vector in the SV's neighborhood has its weights adjusted to become more like the SV. Vector closest to the SV are altered more than the vector furthest away in the neighborhood.
5. Repeat from step 2 for enough iteration for convergence.
6. Calculating the SV is done according to the Euclidean distance among the node's weights (W_1, W_2, … ,W_n) and the input vector's values (V_1, V_2, … , V_n).
7. The new weight for an attribute is the old weight, plus a fraction (L) of the difference between the old weight and the input vector… adjusted (theta) based on distance from the SV.

Step 5: After processing of support vector finally cyber-attack data are classified

VI RESULT PERFORMANCE EVALUATION AND COMPARISON

PERFORMANCE PARAMETERS

Earlier application of isolated feature reduction on dataset has much greater Accuracy, than later by integrating both feature reduction and Improved ID3 Methods. Also there is a considerable enhancement in the true positive and true negative detection ratio and minimizes in false positive and false negative ratio .Thus this gives the direct improvised accuracy in the result. Basis the result of confusion matrix (true positive, true negative, false positive, false negative).We are showing the consequence for the following parameters i.e. -Accuracy, Precision, Recall for data sets.
Precision- Precision measures the proportion of predicted positives/negatives which are actually positive/negative.
Recall -It is the proportion of actual positives/negatives which are predicted positive/negative.
Accuracy-It 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.

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

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

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

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

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 5, May 2022

DOI: 10.17148/IJARCCE.2022.11529

3.10.2 PERFORMANCE EVALUATION

In this section show the selection of variable no. of attribute for the process of the classification algorithm and improved ensemble method. The variable no. of attribute differs the classification rate and classification time. The evaluation parameter corresponding to attribute shown in given below table.

Method Name	Value	TYPES OF ATTACK	TPR	TNR	FPR	FNR	DETECTION RATE	PRECISION RATE	RECALL RATE
IMPROVED ENSEMBLE	0.1	NORMAL	3.654	0.843	1.748	0.751	95.80	85.02	83.97
		DOS	3.513	1.843	1.738	0.741	93.83	81.97	80.97
		PROBE	2.313	1.853	0.738	1.131	94.83	84.97	81.97
		U2R	3.543	0.854	0.853	1.851	95.67	85.97	84.97
		R2L	3.093	0.698	0.408	1.846	92.83	86.94	82.94

Table 1.1: Shows that the performance evaluation of TPR, TNR, FPR, FNR, Detection rate, Precision rate and Recall rate for Improved Ensemble method, and the input value is 0.1.

In this section discuss the valuation of result in terms of TPR, TNR, FPR,FNR, Detection rate, Precision rate and Recall rate for the method Ensemble, Hybrid ensemble method and Improved ensemble method with using input value is 0.1, and 0.5.

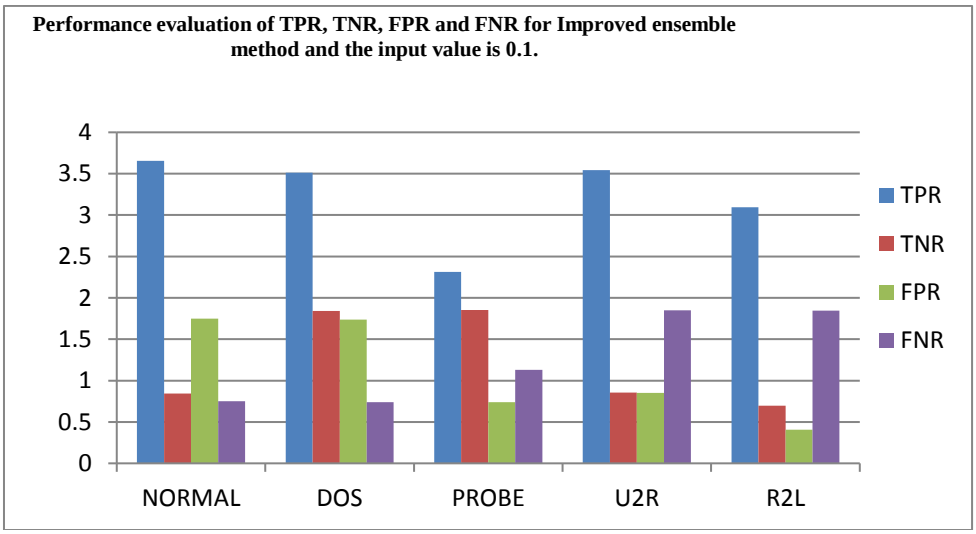

Figure 1.1 : Shows that the performance evaluation of TPR, TNR, FPR and FNR for the Improved ensemble method and the input value is 0.1.

VII CONCLUSION AND FUTURE WORK

In this paper, we have proposed a improved ensemble method, based on NN and Gaussian Support Vector Machines, for cyber-attack classification. Experiments with the KDD Cup 1999 Data show that SVM-NN can providebetter generalization ability and effectively classifies cyber-attack data. Moreover, the modified algorithms proposed in this desecration outperform simple classifier and hybrid ensemble classifier in terms of precision and recall. Specifically, accuracy of the modified algorithms can be increased due to future allocation of NN, and reduces feature subset increases the accuracy of classification. From our experiments, the NN-SVM can detect known attack types with high accuracy and low false positive rate which is less than 1%. The proposed method classified attack and normal data with very high accuracy. The classification ratio shows better result in compression of both classifier that is simple ensemble classifier and hybrid ensemble classifier

A hybrid technique of cyber attack classification suffered from problem of low range data. Theproposed model overcomes this limitation. The proposed classification has some limitations, discussed here.

1. classifier suffred from the data imbalancing problem.
2. Due to the enormous amount input data, the data processing and classification is slow.
3. For the improvement of margin error rate, optimization technique can be applied for balancing of data.

DOI: 10.17148/IJARCCE.2022.11529

REFERENCES

[1] Shailendra Singh, Sanjay Silakari "An Ensemble Approach for Cyber Attack Detection System: A Generic Framework" 14th ACIS, IEEE 2013. Pp 79-85.
[2] Mohammad A. Faysel , and Syed S. Haque "Towards Cyber Defense: Research in Intrusion Detection and Intrusion Prevention Systems" IJCSNS, vol-7 2010. Pp 316-325.
[3] Bimal Kumar Mishra,Hemraj Saini "Cyber Attack Classification using Game Theoretic Weighted Metrics Approach" World Applied Sciences Journal 7, 2009. Pp 206-215.
[4] X. Li et al., "Smart Community: An Internet of Things Application," IEEE Commun. Mag., vol. 49, no. 11, 2011, pp. 68–75.
[5] HoaDinhNguyen , Qi Cheng "An Efficient Feature Selection Method For Distributed Cyber Attack Detection and Classification" IEEE 2013. pp 1-6.
[6]"A two-stage technique to improve intrusion detection systems based on data mining algorithms" FatmaH.Mohammed L,IEEE Modeling, Simulation and Applied Optimization (ICMSAO), 2013 5th International Conference
[7] "Evaluation of Different Data Mining Algorithms with KDD CUP 99 Data Set "Safaa O. Al-mamory Firas S. Jassim Journal of Babylon University/Pure and Applied Sciences/ No.(8)/ Vol.(21): 2013 2663
[8] "Fast Feature Reduction in Intrusion Detection Datasets "ShafighParsazad*, Ehsan Saboori**, Amin Allahyar* MIPRO 2012, May 21-25,2012, Opatija, Croatia
[9] V. Bapuji, R. Naveen Kumar2,Dr. A. Govardhan, S.S.V.N. Sarma "Soft Computing and Artificial Intelligence Techniques for Intrusion Detection System" Vol 2, No.4, 2012, pp 24-33.
[10]. "Intrusion detection using random forests classifier with smote and feature reduction" by AbebeTeshfun,D.lalithaBhaskari By 2013 international conference on Cloud &Ubiquitous computing and engineering technologies
[11]"Adaptive network intrusion detection learning: attribute selection and classification" Dewan Farid, J_er^ome Darmont, Nouria Harbi, Huu Hoa Nguyen, Mohammad Zahidur Rahman, HAL Id: hal-00503951 https://hal.archives-ouvertes.fr/hal-00503951 Submitted on 19 Jul 2010.
[12] "Decision tree classifier for network intrusion detection with ga-based feature selection" byAnnie S Wu,Gary Stein,Being Chang,kieu A hua,proceding of 43rd conference of south regional conferenceISSN 1-5959B-059
[13] "To reduce the false alarm in intrusion detection system using self organizing map Ritu Ranjani Singh, Neetesh Gupta, Shiv Kumar, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-1, Issue-2, May 2011
[14] Xu Li, Inria Lille, Xiaohui Liang, Xiaodong Lin, Haojin Zhu "Securing Smart Grid: Cyber Attacks,Countermeasures, and Challenges" IEEE Communications Magazine IEEE 2012. Pp 38-46.
[15] "Network intrusion detection using an improved competitive learning neural network " Fredericton, NB, Canada; DOI: 10.1109/DNSR.2004.1344728 Conference: Communication Networks and Services Research, 2004. Proceedings. Second Annual Conference on Source: IEEE Xplore
[16] "Dimensionality Reduction for Denial of Service Detection Problems Using Rbfnn" Output Sensitivity Ng, W.W.Y. Chang, R.K.C. Yeung, D.S. 2003
[17] Abhishek Jain And Ashwani Kumar Singh "Distributed Denial Of Service (Ddos) Attacks - Classification And Implications"journal of Information and Operations Management vol-3 2012. Pp 136– 140.
[18] "A Novel Method for Intrusion Detection Based on SARSA and Radial Bias Feed Forward Network (RBFFN)"Anshul Chaturvedi, Prof. Vineet Richharia
[19**]** "Adaptive network intrusion detection learning: attribute selection and classification" DewanFarid, J_er^omeDarmont, NouriaHarbi, HuuHoa Nguyen, Mohammad Zahidur Rahman, HAL Id: hal-00503951 https://hal.archives-ouvertes.fr/hal-00503951Submitted on 19 Jul 2010.
[20] Shailendra Singh, Sanjay Agrawal, Murtaza,A. Rizvi and Ramjeevan Singh Thakur " Improved Support Vector Machine for Cyber Attack Detection" WCECS IEEE 2011. Pp 1-6.

