International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 5 Issue: 4 263 – 267

Multiple Classifier Fusion With Cuttlefish Algorithm Based Feature Selection

K.Jayakumar
Assistant professor
Department of Communication and Networking
Kamaraj College of Engineering and Technology
Virudhunagar, India.
k_jeyakumar1979@yahoo.co.in

S.Karpagam
Assistant professor
Department of Computer Science and Engineering,
Kamaraj College of Engineering and Technology
Virudhunagar, India.
karpagamcse@kamarajengg.edu.in

R.Ashok
Assistant professor
Department of Electronics and Communication Engineering,
Kamaraj College of Engineering and Technology
Virudhunagar, India.
ashokr_online@yahoo.com

Abstract-An intrusion detection system monitors whether the network event is malicious or normal for that network. Intrusion Detection Systems deal with a large amount of data, one of the crucial tasks of IDSs is to keep the best quality of features that represent the whole data and remove the redundant and irrelevant features.. Reducing the redundant information on a network packet shall improve the performance of the IDS A Wrapper based feature selection approach has been designed. The proposed model uses the cuttlefish algorithm (CFA) as a search strategy to ascertain the optimal subset of features and 3 different classifiers are used as a judgement on the selected features that are produced by the CFA. The NSL-KDD Cup 99 dataset is used to evaluate the proposed model. The results show that the feature subset obtained by using CFA gives a higher detection rate with a lower false alarm rate.

Keywords-*Intrusion detection, Cuttlefish algorithm (CFA), multiple-classifier, feature selection.*

*****

I. INTRODUCTION

A. Intrusion Detection

Over the past decades the development in computer networks is exponential. CERT statistics report concludes that the amount of intrusions is also growing every year. To reduce the amount of intruders activities on a network,security policies can be implemented on a network. Devices like firewall only prevent the unauthorized users from accessing the service. Attacker's uses different methods to find vulnerability in the security schemes implemented in a network. Intrusion detection is the process used to detect malicious activity on a network. Monitoring the network manually is a very tedious job Intrusion Detection System was developed to do the job automatically.

Intrusion detection system is classified as Network intrusion detection system (NIDS) and Host intrusion detection system (HIDS) [1].The NIDS collect and examine data captured directly from the network. When packets pass through the network, they are captured and then the types and contents of packets are examined. Relying on the location where the data are collected, a NIDS monitor and analyze the traffic within a network or between a local network (LAN) and the Internet. A HIDS is usually a software running on the confined host, hence the coverage of the HIDS is restricted to only one machine. In order to protect an enterprise network, a HIDS must be installed on each individual system within the internal network. The process of examining the events in a computer system or network, and analyzing them for signs of intrusions is known as Intrusion Detection

The primary classes of intrusion detection methodologies are Misuse Detection and Anomaly Detection. Misuse detection approach aims to encode the knowledge about patterns in the dataflow that are known to corresponding intrusive procedure in the form of specific signatures. Misuse detection is also called as Signature based detection. Misuse detection is fully effective in uncovering a known attack. Anomaly detection refers to the detection of pattern data that do not conform to the expected behavior. It is based on the assumption that all intrusive activities are necessarily anomalous. The overall Intrusion detection functions are

- Monitoring user and system activities
- Identification of attack pattern
- Analyzing system weakness

B. Feature Selection

To trace user policy violations to monitor the network events for intrusion, IDS have to process large amount of data in real time. Dimensionality reduction technique is used to reduce the number of the features. Feature selection is used to find the optimal feature set that represents the whole data without redundant information [2]. The simplest method is to test every possible combination of

features and select the one which gives the best result. For a high dimensional space the exhaustive search needs a high computation timer. Feature selection algorithm is used to find the optimal feature set efficiently with reduced search time.

The feature selection process consists of three steps

1.Subset generation
2.Subset evaluation
3.Stopping criteria.

The feature subset may be generated randomly or heuristic based. Then the generated feature subset is given to a classifier for evaluation. This process is done until the stopping criteria are met. A feature selection algorithm can be seen as the combination of a search technique for proposing new feature subsets, along with an evaluation measure which scores the different feature subsets. The simplest algorithm is to test each possible subset of features and finding the one which minimizes the error rate. This is an exhaustive search of the space, and is computationally intractable for all but the smallest of feature sets. The choice of evaluation metric heavily influences the algorithm, and it is these evaluation metrics which distinguish between the three main categories of feature selection algorithms: wrappers, filters and embedded methods.

A feature selection based on the multiple classifiers can further use the above experimental observation that attack feature can be collected separately in different feature subset. First each feature subset is used independently to perform attack detection. Then the feature is combined to produce the final decision. The effectiveness of multiple classifiers depends on the decision fusion function. Decision fusion function combine the multiple classifier result and give the better classification accuracy. In the proposed system, same feature selection algorithm is used in more than one IDS [3]. But each Intrusion detection system uses different classifier. Decision fusion is used to obtain the best result.

II. RELATED WORK

Adhi Tama and Kyung Hyune Rhee [5], evolutionary algorithm is used for feature selection. They developed particle swarm optimization (PSO) for feature selection and the ensembles of tree-based classifiers (C4.5, Random Forest and CART) perform the classification task. Mohammed A. Ambusaidi and Zhiyuan Tan [6], filter based feature selection could handle linearly and nonlinearly dependant data features. Classification is done by SVM classifier. Though ANN used to detect attacks in IDS but provide the less accuracy due to its design to solve this Anqing Wu and Li Feng [7], The ICA was used to fuse the complex intrusion input and hence attain renowned characteristics (that is, self-determining components, ICs) about the original data. By the use of ICs, the intricate of the ANN structure design could be condensed. Then, the PSO was employed to optimize the structural parameters of the ANN. Adel Sabry Eesa, Zeynep Orman, Adnan Mohsin Abdulazeez Brifcani [8], uses the cuttlefish algorithm (CFA) as a search tactic to determine the best subset of features and the decision tree (DT) classifier as a judgment on the selected features that are produced by the CFA.

III. PROPOSED METHOD

A. Feature selection

Feature selection is a technique for selecting the best subset of features that produce a better characterization of patterns belonging to different classes. Feature selection method are classified into

- Wrapper approach
- Filter approach

Instead of processing data with the whole features to the learning algorithm directly, feature selection for classification will first perform feature selection to select a subset of features and then process the data with the selected features to the learning algorithm. The feature selection phase might be ndependent of the learning algorithm, like filter models, or it may iteratively utilize the performance of the learning algorithms to evaluate the quality of the selected features, like wrapper models [10]. With the finally selected features, a classifier is induced for the prediction phase. Usually feature selection for classification attempts to select the minimally sized subset of features.In this paper, cuttlefish algorithm wrapper based method is used for selecting most relevant feature.In this paper, cuttlefish algorithm a wrapper based method is used for selecting most relevant feature

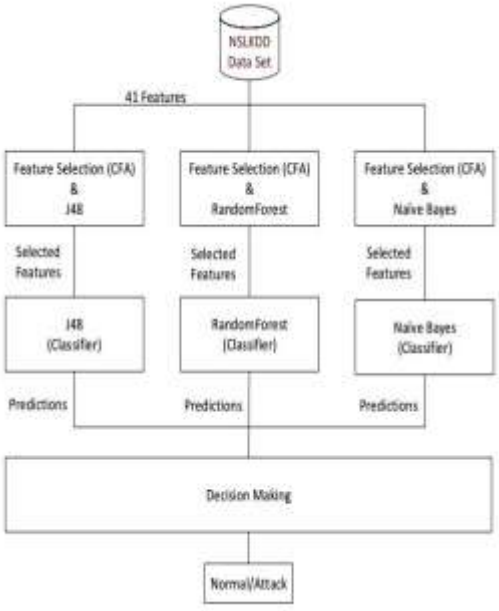

Figure 1. Proposed Architecture

B. Cuttlefish Algorithm

The algorithm was developed based on the techniques used by the cuttlefish to change its skin colour for camouflage. The two main process used in CFA are Reflection and Visibility to generate a new solution. The reflection process mimics or simulates how the cells of cuttlefish reflect the light . Visibility process simulates the visibility of matching patterns.

Initialization

the algorithm starts with a population P of N initial solutions generated randomly, $P[N] = cells[N] = \{p_1, p_2, p_3, \dots, p_N\}$. Each p_i is associated with two subsets: *selectedFeatures* and *unselectedFeatures*. The DT classifier will evaluate each *selectedFeatures* in each p_i, the best solution will be kept in both $AV_{Bestsubset}$ and *bestSubset*. As we mentioned before, the size of *bestSubset* is less than $AV_{Bestsubset}$ by one element, so we have to remove one element from *bestSubset* randomly.

Cases 1 and 2

In these cases, the best k of population P will be used to generate new subsets. This can be done by sorting the P in descending order based on the *fitness* values and choosing the first k subsets from P, where k is a random number between (1, N/2), N is the size of P. After that, the new subsets will be generated from each subset in k subsets using *Reflection* set and *Visibility* set. Where $Reflection_i$ is a set with R features selected randomly from *selectedFeatures* that is associated with p_i, $i = \{1, 2, \dots, k\}$, and $Visibility_i$ is a set with V features selected randomly from *unselectedFeatures* which is also associated with p_i. The combination (union) of $Reflection_i$ and $Visibility_i$ will produce a new subset. The DT will evaluate the new produced subset. If there is any new generated subset better than $AV_{Bestsubset}$, $AV_{Bestsubset}$ will be replaced with it.

Cases 3 and 4

In these two cases, a single feature-exchanging operator is used to produce new solutions from *bestSubset*. A random feature is selected from *selectedFeatures* to be exchanged with another random feature selected from *unselectedFeatures*, where *selectedFeatures* and *unselectedFeatures* are the two subsets associated with *bestSubset*. This operator is repeated t times, where t is an integer constant value; its value is specified by the user. If any new produced solution is better than *bestSubset*, replace *bestSubset* with it.

Case 5

In this case, the $AV_{Bestsubset}$ will be used to produce R subsets of features by removing one element from $AV_{Bestsubset}$ each time. The number of generations (R) is equal to the size of $AV_{Bestsubset}$. Each of the generated group will be evaluated by the DT, if any new subset is better than *bestSubset*, the *bestSubset* will be replaced with that new subset.

Case 6

After best k solutions are used in Cases 1 and 2 to generate new solutions. In this case the remaining solutions of P will be generated randomly. This operator is the same as in the initialization step. The $AV_{Bestsubset}$ will be replaced with the best new generated solution if the new solution is better than it, after DT is used to evaluate the new generation.

The work of these cases is repeated until the stop criterion, such as number of iterations, is satisfied.

C. Multiple-classifier approach

The combination of multiple classifiers has been viewed as a new direction for the development of highly reliable intrusion detection system [4]. The three classifiers used in this system are J48,Random Forest and Naïve Bayesian. The most relevant features are selected by the CFA using all the three used classifiers. These selected features are classified by the different classifier. Classification output given by the classifiers J48, Random Forest and Naive Bayes are used in taking the final decision. Decision fusion is done with the help of majority voting algorithm [9]. Multiple classifiers are used to achieve better accuracy.

IV. RESULT AND DISCUSSION

A. System Setup

The proposed system model is simulated in NetBeans IDE using JAVA on a laptop powered by Dual core AMD A4 3330x APU @ 2.30 GHz and 2 GB DDR3 RAM @ 1333MHz. In the proposed model the NSL-KDD dataset is used to benchmark the system. The system was trained using a training data set with 17342 instances and tested using a test dataset with 25192 instances.

B. Dataset

NSL-KDD data set is used to train and test the proposed model. It has the lower complexity level. This is the improved version of KDD CUP99 dataset. It doesn't contain any redundant records and help the classifier to produce the best result. In NSL-KDD, each record has the 41 attributes and it is classified as either normal or attack.

TABLE 1. NSL-KDD CUP99 DATASET INFORMATION

Class	Train Dataset	Test Dataset
Normal	8317	13449
DoS	6536	9234
R2L	209	209
Probe	2269	2289
U2R	11	11
Total Instances	17342	25192

C. Performance evaluation

In order to estimate the efficiency of the proposed model, we use performance measures: Detection rate, false alarm rate and accuracy. Evaluation parameters are:

- True Positive (TP): detect the attack packet as it is.
- False Positive (FP): detect the normal packet as attack.
- True Negative (TN): no detection of attack packet.
- False Negative (FN): detect the attack packet as normal

i)Attack Detection Rate (ADR) – measure the correctly detected attack.

$$\text{Detection Rate} = \frac{TP}{TP+FN} * 100 \quad (1)$$

ii)False Alarm Rate (FAR) – measures the number of false alarm generated by the proposed model.

$$\text{False Alarm Rate} = \frac{FP}{TN+FP} * 100 \quad (2)$$

iii)Accuracy – shows the proposed model has the ability to correctly detect the attacks.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} * 100 \quad (3)$$

D. Experimental Result

Single Classifier Performance

All experiments were performed using an Intel Core™ i5 processor with 4 GB of RAM under windows7. The evaluation of classifiers such as J48, Naïve Bayes and Random forest with all 41 features (without Feature selection) is done and the result are shown in Table 2 and Fig. 2. From the Table 2 it's found that the accuracy of all the classifiers is too low.

TABLE 2. CLASSIFIERS PERFORMANCE FOR 41 FEATURES

Classifier	DR	FPR	Accuracy
J48	75.21	23	63.97
Naive Bayes	77.9	22.09	55.77
Random Forest	74.77	23	62.98

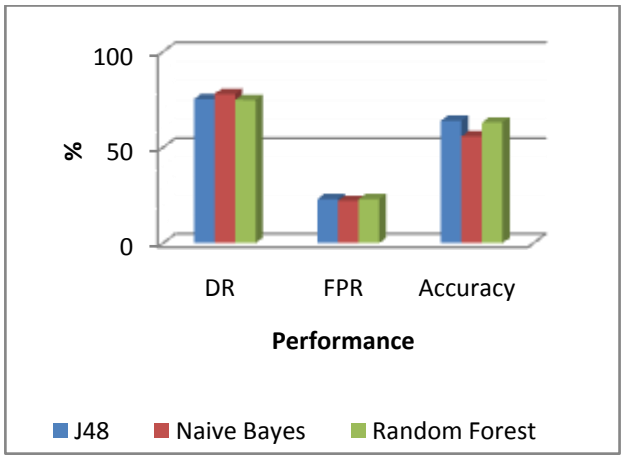

Figure 2. Single Classifier Performance without feature selection

Then same process is done with the selected features using CFA. The results are shown in Table 3and Fig. 3. Using CFA, minimum number of features is selected. For these features above 3 classifier give the better accuracy. For selected features, Naïve Bayes give the better accuracy rate 86.55%. It also gives the better attack detection rate. Using CFA, Naïve Bayes give the better result.

TABLE 3. CLASSIFIERS PERFORMANCE WITH FEATURE SELECTION

Classifier	No of Features	DR	FPR	Accuracy
J48	20	80.32	19	80.85
Naive Bayes	5	86.22	13	86.55
RandomForest	7	82.99	17.1	79.71

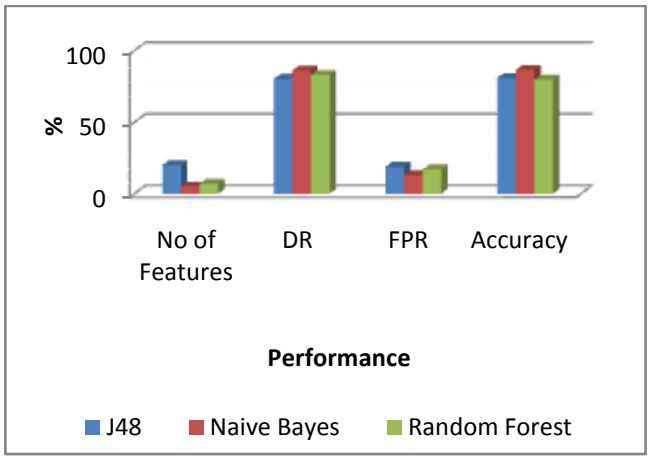

Figure 3. Single Classifier Performance with feature selection

E. Multi classifier Performance

In our proposed model, combining all these classifiers with the majority voting fusion rule it gives the

overall better accuracy rate using minimum number of selected features is shown in Table 4.

TABLE 4. MULTI CLASSIFIER PERFORMANCE

Classifier	No. of Features	DR	FPR	Accuracy
Multi Classifier	20	85.25	14.05	91.23

Table 4.3: Performance of multi classifier using CFA.

CONCLUSION

In this paper, multi classifier fusion with CFA based feature selection is proposed. The cuttlefish algorithm is used to remove the redundant information from the data and to select the best feature set for individual classifiers. The feature set derived from the CFA is used to train the different IDS. The evaluation of the IDS is done using NSL-KDD test dataset. Three experiments had been carried out. In experiment I the classification performance of three classifiers using all features is found out. In experiment II the classification performance of three classifiers using selected features found using CFA is found out. In experiment III hybrid intrusion detection system was created with J48, Random forest and Naïve Bayesian Classifier and decision fusion is done by majority voting rule. The performance of the Experiment III gives very good results.

REFERENCES

[1] Monowar H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita," Network Anomaly Detection: Methods, Systems and Tools, "IEEE Communications Surveys & Tutorials, Vol. 16, No. 1, First Quarter 2014.

[2] F.Amiri, M.Rezaei yousefi, C. lucas, A.Shakery, N. Yazdani, "Information based feature selection for intrusion detection systems", Journals of Network and Computer Application, Vol. 34, Issue 4, 2011.

[3] Meesala Shobha Rani, S. Basil Xavier, "A Hybrid Intrusion Detection System Based on C 5.0 Decision Tree Algorithm and One-Class SVM with CFA", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 3, Issue 6, June 2015

[4] Giorgio Giacinto, Fabio Roli, Luca Didaci, " Fusion of multiple classifiers for intrusion detection in computer networks." Pattern Recognition Letters Vol. 24, pp.1795–1803, 2003.

[5] Bayu Adhi Tama and Kyung Hyune Rhee, "A Combination of PSO-Based Feature Selection and Tree-Based Classifiers Ensemble for Intrusion Detection Systems,"Springer Science Business Media Singapore, 2015.

[6] Mohammed A. Ambusaidi, Zhiyuan Tan, " Building an intrusion detection system using a filter-based feature selection algorithm," IEEE Transactions on Computers, 2014.

[7] Anqing Wu and Li Feng, "Information Fusion and Intelligent Pattern Recognition for Network Intrusion in Industrial Network Systems Based on ICA and PSO–ANN," Green Communications and Networks, Lecture Notes in Electrical Engineering, Springer Science Business Media B.V. 2012.

[8] Adel Sabry Eesa, Zeynep Orman, Adnan Mohsin Abdulazeez Brifcani, "A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems", Expert Systems with Applications, pp. 2670–2679, 2015.

[9] Jayakumar Kaliappan, Revathi Thiagarajan, and Karpagam Sundararajan, "Fusion of Heterogeneous Intrusion Detection Systems for Network Attack Detection,"Hindawi Publishing Corporation, Scientific World Journal, 2015, Article ID 314601.

[10] Jayakumar Kaliappan, Revathi Thiagarajan, and Karpagam Sundararajan, "Intrusion Detection using Artificial Neural Networks (ANN) with best set of features," The International Arab Journal of Information Technology, Vol. 2015,