International Journal of Electrical and Computer Engineering (IJECE)
Vol. 10, No. 2, April 2020, pp. 1963~1971
ISSN: 2088-8708, DOI: 10.11591/ijece.v10i2.pp1963-1971

A novel ensemble modeling for intrusion detection system

Pullagura Indira Priyadarsini[1], G. Anuradha[2]
[1]Department of Information Technology, Vardhaman college of Engineering, India
[2]Department of Computer Science & Engineering, V. R. Siddhartha Engineering College, India

Article Info

Article history:

Received Apr 26, 2019
Revised Oct 28, 2019
Accepted Nov 7, 2019

Keywords:

Artificial neural network (ANN)
Intrusion detection system
K-nearest neighbor
Modeling
Support vector machine

ABSTRACT

Vast increase in data through internet services has made computer systems more vulnerable and difficult to protect from malicious attacks. Intrusion detection systems (IDSs) must be more potent in monitoring intrusions. Therefore an effectual Intrusion Detection system architecture is built which employs a facile classification model and generates low false alarm rates and high accuracy. Noticeably, IDS endure enormous amounts of data traffic that contain redundant and irrelevant features, which affect the performance of the IDS negatively. Despite good feature selection approaches leads to a reduction of unrelated and redundant features and attain better classification accuracy in IDS. This paper proposes a novel ensemble model for IDS based on two algorithms Fuzzy Ensemble Feature selection (FEFS) and Fusion of Multiple Classifier (FMC). FEFS is a unification of five feature scores. These scores are obtained by using feature-class distance functions. Aggregation is done using *fuzzy union* operation. On the other hand, the FMC is the fusion of three classifiers. It works based on Ensemble decisive function. Experiments were made on KDD cup 99 data set have shown that our proposed system works superior to well-known methods such as Support Vector Machines (SVMs), K-Nearest Neighbor (KNN) and Artificial Neural Networks (ANNs). Our examinations ensured clearly the prominence of using ensemble methodology for modeling IDSs, and hence our system is robust and efficient.

Corresponding Author:

Pullagura Indira Priyadarsini,
Department of Information Technology,
Vardhaman College of Engineering,
Nagarguda Shamshabad Road, Kacharam, Hyderabad, Telangana 501218, India.
Email: indupullagura@gmail.com

1. INTRODUCTION

In this digital age, maintaining information through online businesses and social networks remain insecure. Numerous intruders both human and robot, are gaining unauthorized access to information. Also their illusive nature in the internet has increased complexity in detecting intrusions. Mostly prevailing Intrusion Detection Systems (IDSs) have shown chaotic performance in identifying different attacks [1]. It is certainly possible to get a stable and accurate decision for all the attacks by unifying the decisions of multiple classifiers [2, 3]. Therefore merging multiple IDSs is not a great concern, in terms of computation and best solutions can be achieved. With better analysis of data using ensemble learning, all the attacks can be identified. This integration most probably improves predictive accuracy. An Ensemble of classifiers has arisen as a feasible solution to the class imbalance problem [4].

Feature selection is of utmost significance for any learning algorithm which when poorly done (i.e., a poor set of features is selected) may lead to problems associated with incomplete information, noisy or irrelevant features. The learning algorithm applied is slackened gratuitously due to higher dimensions of the feature space, and also undergoing lower prediction accuracies by learning irrelevant information. Constructive feature selection methods generate better classification accuracies [5, 6]. The crucial objective

of feature selection is to attain a feature space with (1) low dimensionality, (2) retention of sufficient information [7]. On operating, applicable feature selection methods produce simplified models which are easy to interpret and reduce training time and also augment the generalization ability.

In the former works, machine learning methods employed a single learning model. Still, it has been witnessed that multiple prediction models can be utilized for solving the same problem. Therefore an approach, known as ensemble learning, was built on the statement that combining the output of multiple experts is better than using the output of any single expert [8]. Ensemble learning has been efficaciously realistic to classification problems and is also a mechanism for boosting other machine learning functions such as feature selection. In feature selection terminology, the individual selectors in an ensemble are called as base selectors. If the base selectors are all of the same kind, the ensemble is termed as homogeneous.

In this paper, we have built a novel ensemble model for Intrusion Detection System using Fuzzy Ensemble Feature Selection (FEFS) algorithm and Fusion of Multiple Classifier (FMC) algorithm. FEFS is done as; examining the prevalence of different feature selection methods, the unification of five methods is done to obtain a strong feature set which is indeed beneficial for better classification. The technique accustomed to joining the outputs is based on fuzzy logic. Its main perspective is to select the most optimistic features in KDD cup 99 dataset. KDD Cup 99 [9] is an eminent intrusion evaluation dataset and is a classic example of large-scale datasets. A Fusion of Multiple Classifier (FMC) is for the process of classifying attack and normal data, through the unification of Support Vector Machine (SVM), K nearest neighbor classifier (KNN) and Artificial Neural Network (ANN). Then by this ensemble classification method, we have achieved better accuracy and lower False Alarm Rate (FAR). This paper is being prepared in a subsequent way. In section 2, related works were described. Methodology for construction of Ensemble modeling is discussed in a detailed manner in section 3. Then in section 4, total experiments made and results attained were discussed specifically. The Last section specifies the conclusions and discussions.

2. RELATED WORK

Ensemble feature selection procedures utilize an idea analogous to ensemble learning for classification [10]. There are several works done, constructing ensemble feature selection techniques, for the selection of the optimal feature set [11]. Olsson et.al have specified ensemble of multiple feature ranking methods that combine three generally used filter based feature ranking techniques like information gain, document frequency thresholding, and the chi-square method mainly for text classification problems. In recent works, Wang et.al has integrated ensemble of six filter based rankers and accomplished notable results [12]. Basant Subba et.al has applied two statistical methods namely Linear Discriminant Analysis (LDA) and Logistic Regression (LR) which were useful successfully to develop new intrusion detection models [13]. In [14] Afef Ben Brahim et.al has developed a robust feature aggregation technique for combining the results of three diverse filter methods. This aggregation technique is relied on determining feature algorithms confidence and conflict with the other ones in order to assign a reliability factor controlling the final feature selection.

Because of the imbalanced distribution of classes in the KDD cup 99 dataset, the results cannot be precise. Recent studies have shown a solution which is to incorporate Ensemble learning. The major challenges and opportunities with the imbalanced data set were clearly given in [15]. Ensemble learning is effectively implemented on classification problems [16, 17]. Bukhtoyarov et al. [18] have developed ensemble based on Genetic Programming known as (GPEN) to categorize the input intrusions as Probe or non-Probe attacks, with nine of the 41 features. Borji [19] has given an ensemble methodology using four base classifiers viz. SVM, k-NN, ANN, and decision trees. In the works done in [20], a new ensemble approach is proposed for effective intrusion detection. This ensemble approach is the grouping of attribute selection, multiclass SVM and k-NN classifier. Besides, an Incremental Particle Swarm Optimization is also embedded as an ensemble classifier for boosting the classification accuracy in their works. In this Perspective, ensemble learning and various fusion methods [21, 22] are considered to have a potential development in classifier's performance we have made the proposed investigations.

3. PROPOSED WORK

Figure 1 describes the proposed ensemble modeling architecture of Intrusion Detection System. It is incorporated with two different phases. First one, which performs Feature selection named, Fuzzy Ensemble Feature Selection (FEFS). Next is, classification phase named, Fusion of Multiple classifiers (FMC) which is employed for classifying the data as attack and normal.

Figure 1. Proposed novel ensemble IDS model

3.1. Fuzzy ensemble feature selection (FEFS)

Merging feature selection methods were executed to achieve stable and robust outputs. An Ensemble can be made by usage of the aggregation operations. This is achieved by considering the advantages of five filtering methods such as Canberra distance, City block distance, Euclidean distance, Cebyshev distance, and Minkowski distance. Fuzzy logic is applied for aggregating the five filters. The main thought behind employing fuzzy logic is backtracking. In contrast, some of the features may be left in the traditional methods where certain threshold is exploited. Hence weights are allocated to all values. Aggregation of all the filters is done by making use of *fuzzy union* operation of the fuzzy sets. On the data set, Euclidean distance is computed for all the features. Now for the same data set cebyshev distance, Canberra distance, City block distance, and Minkowski distance are calculated for all the features. All these values are fuzzified. Then Aggregator is applied. It is shown by the FEFS structure in Figure 2.

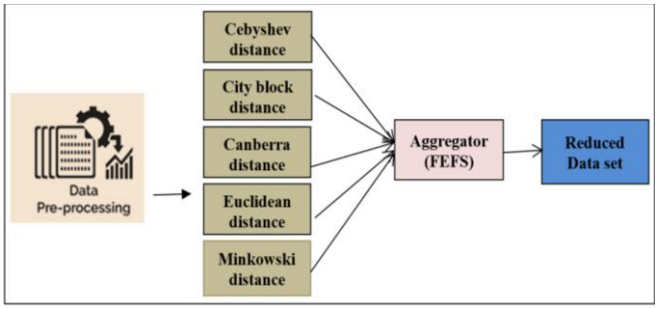

Figure 2. Proposed FEFS construction

For an input pattern {I, J} where I= number of instances and J=number of features i.e., {F_a, F_b,F_c…F_{an}, F_{ao}}. For any particular feature $F_i \in J$, then the Euclidean distance is computed as

$$\text{De} = \sqrt{\sum_{i=1}^{n} (\text{xi} - \text{ci})^2} \quad (1)$$

For any particular feature $F_i \in J$, then the cebyshev distance is computed as

$$\text{Dce} = \max mod(xi - ci) \quad (2)$$

For any particular feature $F_i \in J$, the Canberra distance is computed as

$$\text{Dca} = \sum_{i=1}^{I} \frac{mod\ (xi-ci)}{xi+ci} \quad (3)$$

Likewise, city block distance is also computed for the same data set. It is given as

$$\text{Dcb} = \sum_{i=1}^{I} mod(xi - ci) \quad (4)$$

For any particular feature $F_i \in J$, then the minkowski distance is computed as

$$D_m = \left(\sum_{i=1}^{n} \left(x_i - c_i\right)^{\lambda}\right)^{1/\lambda} \quad (5)$$

Where x_i is an individual feature in J and c_i is the class label. For all the features from {F_a, F_b, Fc,…,F_{an}, F_{ao}}, and in minkowski distance, we have taken λ is equal to 3. Euclidean distance is calculated. Similarly, cebyshev distance, Canberra distance, city block distance and minkowski distance is computed for all the features.

Hence from the above (1), (2), (3), (4) and (5) we get five sets of values. Then the conversion of these values into fuzzy is made. This is known as fuzzification. They are termed as fuzzy sets namely f_ca_1, f_ci_1 f_eu_1 f_mi_1, f_ce1. They are said to be feature scores. The procedure of transformation is done using trapezoidal membership function. A special case of trapezoidal is L-Function. Presume *y* is the element to be transformed then f_y will be (i.e. fuzzy conversion for y) $\frac{y-a}{b-a}$.Here '*a'* and '*b'* is minimum and maximum values in the whole set. Transformation is done after applying all the filters on all the features. Feature score calculation is shown in line 9 to line 13 in the algorithm given in Figure 3.

Fuzzy Ensemble Feature Selection (FEFS) Algorithm:

Input: *KDD data set, i, classlab, can_dist, cit_dist,euc_dist,mink_dist,ceby_dist, , f_ca_1, f_ci_1, f_eu_1, f_mi_1, f_ce_1, N=41, F_i, FS*
Output: *FS*
Start:
1. Take KDD dataset.
2. Apply pre-processing techniques like normalization, converting symbolic attributes to numeric;
//Feature selection//
3. For i=1 to N do
4. ceby_dist = calculate (feature, classlab);
5. cit_dist = calculate (feature, classlab);
6. eucl_dist = calculate (feature, classlab);
7. can_dist = calculate (feature, classlab);
8. mink_dist = calculate (feature, classlab);
End For
9. f_ca_1=Convert can_dist to f_can_dist;
10. f_ci_1=Convert cit_dist to f_cit_dist;
11. f_eu_1= Convert eucl_dist to f_cit_dist;
12. f_mi_1= Convert mink_dist to f_cit_dist;
13. f_ce_1= Convert ceby_dist to f_cit_dist;
14. for i= 1 to N do
15. F_i = {f_ca_i U f_ci_i U f_eu_i U f_mi_i U f_ce_i}
End for
16. FS=0;
17. If (F_i == 1);
18. Add them to FS;
Return resultant FS;
// End feature selection//
End

Figure 3. FEFS algorithm

Then for each feature their feature scores {f_ca_1, f_ci_1 f_eu_1 f_mi_1, f_ce_1} are combined using Aggregator. Here *fuzzy Union* operation is utilized for combining them. The *fuzzy union* operation will return the maximum of all the membership values obtained from all five feature scores [23]. It is shown in line 15 of the Figure 3 Find those features whose F_i =1. For instance, consider a feature F_j .To this feature, five filters is applied. The fuzzy logic is applied to each of the filters. Then, they are transformed to fuzzy values. Then after getting five feature scores for the feature F_i, they were aggregated by operating *fuzzy union* on them. Then F_i will be a single value. The whole process is done for all the remaining features. Finally, all the features whose membership value is equal to 1 are selected as the best feature set. It is shown in line 17 of Figure 3.

3.2. Fusion of multiple classifiers (FMC)

The merging of multiple classifiers can be firm and predict better than single classifiers [24]. The proposed FMC is based on majority voting method over individual base classifier which improvises detection of attacks. An FMC algorithm is developed based on three individual classifiers. They are 1, K-Nearest Neighbor (kNN) classifier, 2.Support Vector Machine (SVM) and 3.Artificial Neural Network (ANN). All the three base classifiers is an expert in a specific region of the predictor space because they treat the attribute space under different theoretical basis [25]. The three classifiers could be joined in such a manner in order to yield an ensemble majority voting classifier that is superior to any of the individual rules.

At this level, the result of FEFS is taken and provided to the FMC algorithm. The structure of the proposed FMC is depicted below in Figure 4. KDD dataset is a dataset with *n* no. of tuples and *α* no. of features. The class label is termed as *classlab*.It can either be "0" or "1". The whole process is summarized in

the FMC algorithm depicted in Figure 5. Feeding the preprocessed data to K-NN classifier is done. It is given in step2. Again it is fed to the SVM and ANN respectively. Therefore three local decisions Y1, Y2, Y_3 are produced. Then the consequences of three base classifiers are fused. Each local decision Y_i will be either "0" or "1". Here "0" means attack and "1" means non-attack. Then the fusion of local decisions from three base classifiers can be obtained by using the ensemble decisive function i.e. Majority voting method. Suppose the final decision from the ensemble classifier *Y* is defined as

$$\sum_{t=1}^{T} d_{t,J} = max_{j=1}^{C} \sum_{t=1}^{T} d_{t,j}$$

Where $d_{t,j} \in \{0,1\}$,t=1,2,…T and j = 1,2..C.Where T is the number of classifiers and C is the number of classes. Here we have considered two classes and three classifiers. Then Y chooses the class that receives the highest number of votes.

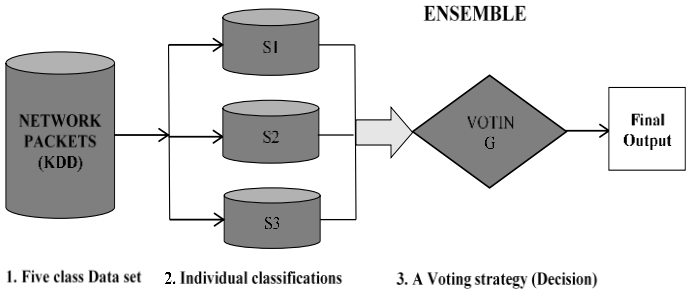

Figure 4. Proposed FMC structure

Algorithm for Fusion of Multiple classifiers
Input: *KDD data set, classlab= {y1, y2}, y_1=0, y_2= 1, Y_1, Y_2, Y_3, Y*
Output: *Evaluation Measures: accuracy, FAR, DR, specificity, precision*
Step 1: *Pre-process the data set by obtaining packet parameters.*
Step 2: *Feed the pre-processed data set to the KNN classifier.*
Step 3: *Repeat step 2 for SVM, ANN* ***//build three models//***
Step 4: *For every input data instance,*
there are three local decisions Y_1, Y2, Y3 from three classifiers.
Step 5: *The local decision y_i is labeled as y1 or y2*
Step 6: *If (y_i ="DOS"‖"PROBE"‖"R2L"‖"U2R")*
Step 7: *Y_i= y1;*
Step 8: *Else*
Step 9: *Y_i =y2;*
Step 10: ***Fusion***: $Y = max_{j=1}^{3} \sum_{t=1}^{T} Yi$
Step 11: *End*

Figure 5. FMC algorithm

4. EXPERIMENTS CONDUCTED & RESULTS OBTAINED

Experiments were made on the KDD cup 99 dataset. The researchers in their works have used the portion of the dataset from the KDD cup 99 data set for building IDSs not including the complete train or test dataset [26]. So, we have taken a subset of KDD cup 99 containing 14207 records and call as "KDD dataset". The size of the dataset is taken in proportion to the relative size of the KDD cup 99 dataset and R2L,U2R records are taken as usual from the original data set.

4.1. Data preprocessing stage

The KDD cup 99 data set which is a raw data set is taken for conducting investigations on the proposed approach. Appropriate preprocessing techniques were applied. The data in the above-mentioned dataset are converted to numeric. Discretization of continuous variables is made to the data set. Symbolic values of three features have been given numeric values ranging from 1 to N. Interquartile range (IQR) also been applied to remove noise and outliers in the data set. A subset of KDD cup 99 data set is taken for experimentations. It has classes with same proportions as in KDD cup 99. Therfore it is named as KDD dataset. It has 14207 instances with 3000 Normal instances, 10000 DoS instances, 574 probe instances,

401 R2L, and 52 U2R instances. All the five classes in the KDD data set are assigned numeric values. They will be assigned as "0" for U2R, R2L, Probe, DoS and "1" for Normal. The 41 consecutive Features are labeled as F_a, F_b, F_c, $F_{d,}$…F_{ao} respectively.

4.2. Applying FEFS algorithm to the dataset

Then Proposed FEFS algorithm is applied to the KDD data set which uses fuzzy logic strategy to get the best feature subset. For the KDD dataset, FEFS algorithm is applied (as described in the earlier section). The fuzzy union of all the obtained scores is done for each feature. The feature score of 41 features after applying aggregator is {**1**, 0.4, **1**, 0.9, **1**, **1**, 0.8, 0.9, 0.9, **1**, **1**, 0.9, 0.8, 0.8, 0.8, 0.9, **1**, 0.8, 0.9, **1**, **1**, 0.9, 0.9, **1**, 0.8, 0.8, 0.8, 0.9, 0.6, 0.9, 0.8, 0.7, **1**, 0.5, 0.9, 0.9, 0.9, 0.8, 0.8, 0.9, **1**}. Accordingly, we have chosen totally 12 features. The corresponding features are F_a, $F_{c,}$ F_e, F_f, F_j, F_k, F_q, F_t, F_u, F_x, F_{ag} and F_{ao}. These are the features selected as a result of FEFS algorithm. Now the reduced data set is fed to the Proposed FMC algorithm.

4.3. Applying FMC

In the complete experiments conducted we have used 10 fold cross validation for analyzing the proposed Novel Ensemble model. The 10 fold cross validation is also referred to as rotation estimation. It is a recommended method over the holdout method and leave-one-out methods for estimating a classifier. The dataset has been split at random into ten parts of the equivalent size. Every part is kept out in turn and the training is conducted on the remaining nine parts, then the testing is made on holdout set. The training is made totally 10 times on different training sets and lastly, the average of ten error rates is considered for attaining complete error estimate. Four different experiments were made to indicate the results. 1. With FEFS outputs given to SVM, 2.With FEFS outputs given to ANN, 3.With FEFS outputs given to K-NN and 4. The Proposed Novel Ensemble Model (FEFS+FMC).At the testing part, instances of the KDD data set are fed to the suggested FMC process by leaving their class-label to which they exist. This ensemble classifier gives the network traffic data either as normal or as an attack. We performed our experiments using Java 1.8 and R data mining software tool. Finally, the results are visualized and recorded. To determine the statistical significance of our results, we compare our proposed method with features selected with individual classifiers.

In the ideal situation, some parameters like accuracy, the true positive rate should have maximum values while others like the number of features, error, should have the least amount. However in exceptional circumstances, some parameters may have more effect than the others, so weight has to conform accordingly. The target metrics for classification are listed below in Table 1.

Table 1. Evaluation metrics for classification

SNo.	Parameter Name	Targeted Values
1	Accuracy	Maximum
2	Number of Features	Minimum
3	True Positive Rate	Maximum
4	False Positive Rate	Minimum
5	Precision	Maximum
6	Recall	Maximum
7	F-Measure	Maximum
8	Receiver Operating Characteristic	Maximum

Comparison of performance of all the four experiments on the KDD dataset using the Accuracy rate, Detection Rate (DR), FAR, Precision of the proposed novel ensemble model is illustrated in Figures 6-9 respectively. The proposed ensemble approach implements significantly better than well-known individual methods such as Support Vector Machines (SVMs), K-Nearest Neighbor (KNN) and ANN. The overall relative improvement of accuracy, the Detection Rate for the proposed approach is high, and also the False Alarm Rate has been decreased.

The classification models are evaluated using the area under the ROC curve (AUC) performance metric. AUC is widely used, providing a general idea of the predictive potential of the classifier. A higher AUC is better, as it indicates that the classifier, across the entire possible range of decision threshold, has a higher true positive rate. From certain studies, it is proved that AUC has lesser variance and is more consistent than remaining performance metrics (such as precision, recall, F measure) [27]. The ROC obtained for the proposed model is shown below in Figure 10. The AUC is 0.9 which is pretty good. The results summarized for the KDD data set is interpreted in the Table 2. The proposed model has achieved 0.9, 0.95, 0.96 and 0.9 of precision, recall, F-measure and ROC area respectively.

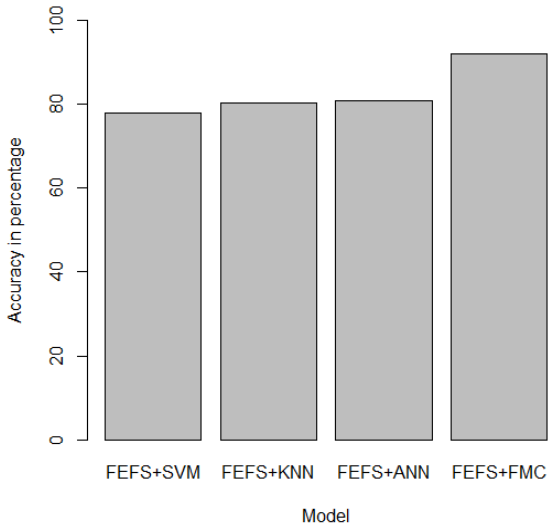

Figure 6. The performance comparison of accuracy rate obtained with FEFS+SVM, FEFS+KNN, FEFS+ANN, FEFS+FMC (proposed model)

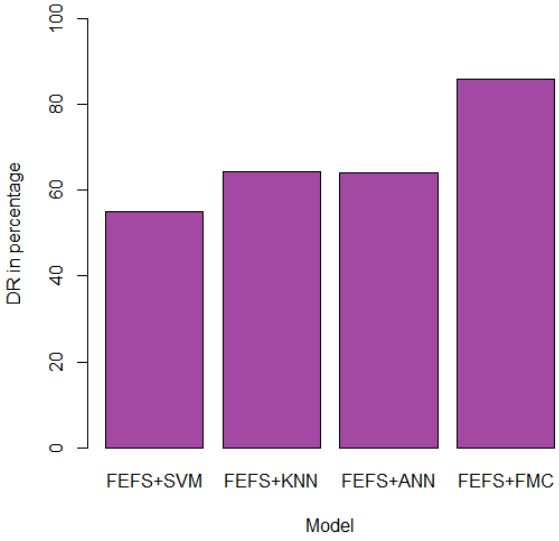

Figure 7. The performance comparison of detection rate obtained using FEFS+SVM, FEFS+KNN, FEFS+ANN, FEFS+FMC (proposed model)

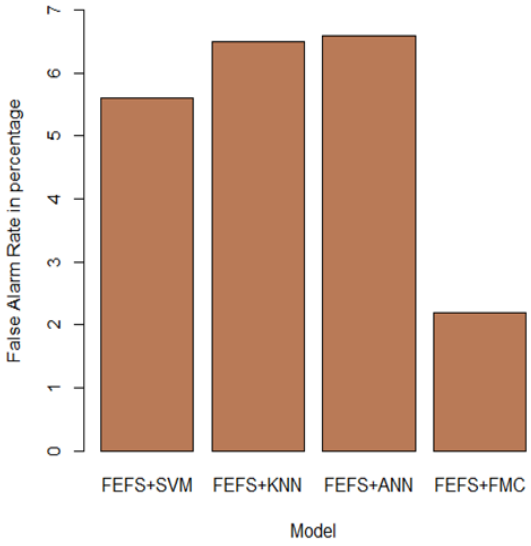

Figure 8. The performance comparison of FAR obtained using FEFS+SVM, FEFS+KNN, FEFS+ANN, FEFS+FMC (proposed model)

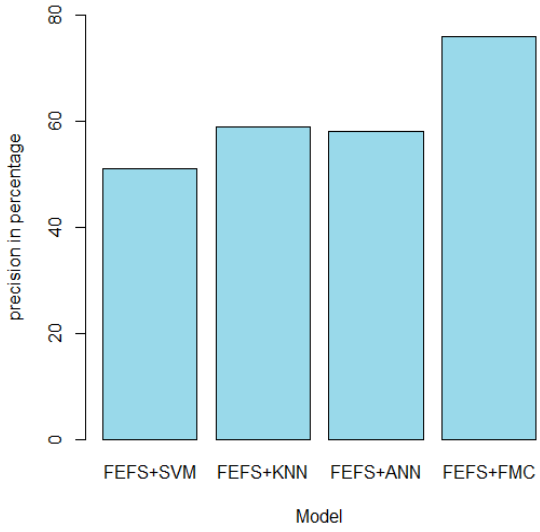

Figure 9. The performance comparison of precision obtained using FEFS+SVM, FEFS+KNN, FEFS+ANN, FEFS+FMC (proposed model)

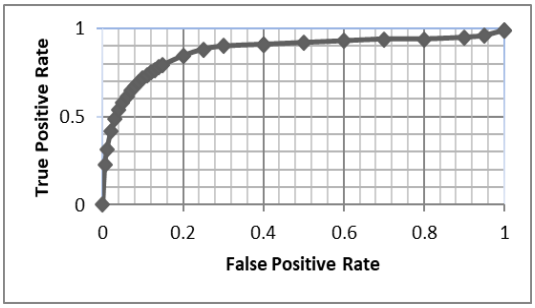

Figure 10. The ROC for the proposed model (FEFS+FMC)

Table 2. The performance of four experiments in terms of precision, recall, F-measure, and ROC area

S.No.	Model	Precision	Recall	F-measure	ROC Area
1.	FEFS+KNN	0.77	0.76	0.7	0.75
2.	FEFS+SVM	0.8	0.8	0.79	0.77
3.	FEFS+ANN	0.74	0.7	0.7	0.6
4.	FEFS+FMC	0.9	0.95	0.96	0.9

5. DISCUSSIONS AND CONCLUSIONS

This research introduced a novel ensemble architecture designed for IDS. It is based on two algorithms Fuzzy Ensemble Feature selection (FEFS) and Fusion of Multiple Classifier (FMC). FEFS is an ensemble of five scores. These scores are obtained by using feature-class distance functions. Aggregation is done using *fuzzy union* operation. An FMC is the fusion of three classifiers. It works based on Ensemble decisive function. Experiments were made on KDD cup 99 data set have shown that our proposed system works superior to well-known methods such as Support Vector Machines (SVMs), K-Nearest Neighbor (KNN) and Artificial Neural Network (ANN). Our examinations ensured noticeably the prominence of using ensemble methodology for modeling IDSs. And consequently, our system is robust and proficient. Since all the reflected performance measures could be improved, such systems will be beneficial in numerous real-world applications. Our experiential results are indicating that ensemble learning is effective in enhancing attack detection rate and lessening the FAR. Performance comparisons were made on the proposed framework versus other base classifier methods with the reduced feature set. The AUC is 0.9 which is pretty good. The Proposed model has achieved 0.9, 0.95, 0.96 and 0.9 of precision, recall, F-measure and ROC area respectively. Since current IDSs are unable to detect all kinds of new attacks because they are designed to restricted applications on the limited environment. Thus, indeed there is a necessity of safeguarding the networks from known attacks and parallel should take measures to identify new and unseen, but possible system abuses, by emerging novel reliable and efficient IDSs. The area of future research includes improvements for machine learning methods to detect novel/unseen attacks.

REFERENCES

[1] R. Sommer, V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," in *IEEE Symposium on Security and Privacy*, pp. 305-316, 2010. DOI: 10.1109/SP.2010.25

[2] M.Govindarajan, "Evaluation of Ensemble Classifiers for Intrusion Detection," *World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering*, vol. 10(6), 2016. doi.org/10.5281/zenodo.1124579

[3] A. A. Aburomman, M.B. Ibne Reaz, "A novel SVM-kNN-PSO ensemble method for intrusion detection system," *Appl. Soft Comput. J*., 2015. http://dx.doi.org/10.1016/j.asoc.2015.10.011.

[4] P. Kang, and S. Cho, "EUS SVMs: Ensemble of under sampled SVMs for data imbalance problems," *ICONIP* 2006, *Neural Information Processing*, pp 837-846, 2016. https://doi.org/10.1007/11893028_93.

[5] Shamsul Huda, John Yearwood, Herbert F. Jelinek, Mohammad Mehedi Hassan, Giancarlo Fortino, Michael Buckland, "A hybrid featureselection with ensemble classification for imbalanced healthcare data: A case study for brain tumor diagnosis," *IEEE*, 2016, DOI 10.1109/ACCESS.2016.2647238, IEEE Access.

[6] B. Seijo-Pardo, I. Porto-Diaz, V. Bolon-Canedo, A. Alonso-Betanzos, "Ensemble feature selection: Homogeneous and heterogeneous approaches," *Knowledge-Based Systems*, 2016. doi: 10.1016/j.knosys.2016.11.017.

[7] Meisel W. S., "Computer-oriented approaches to pattern recognition," Academic Press, New York, 1972.

[8] L. I. Kuncheva, C. J. Whitaker, "Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy," *Machine learning*, vol. 51(2), pp.181-207, 2003. https://doi.org/10.1023/A:1022859003006

[9] Mahbod Tavallaee, Ebrahim Bagheri, Wei Lu, and Ali A. Ghorbani, "*A detailed analysis of the KDD CUP 99 data set,"* Proceedings of the IEEE Symposium on Computational Intelligence in Security and Defense Applications, 2009. DOI: 10.1109/CISDA.2009.5356528.

[10] T. G. Dietterich, "*Ensemble methods in machine learning*," in Proceedings of the First International Workshop on Multiple Classifier Systems, UK: Springer-Verlag, pp. 1-15, 2000. https://doi.org/10.1007/3-540-45014-9_1.

[11] J. O. S. Olsson and D.W. Oard, "Combining feature selectors for text classification," *CIKM* '06: Proceedings of the *15thACM* international conference on Information and knowledge management, pp. 798-799, 2006. DOI: 10.1145/1183614.1183736.

[12] H. Wang, T. M. Khosh Goftaar, and K. Gao, "*Ensemble feature selection technique for software quality classification,*" In Proceedings of the 22nd International Conference on Software Engineering and Knowledge Engineering, pp. 215-220, 2010. https://doi.org/10.1007/s40747-017-0060-x.

[13] Basant Subba, Sushanta Karmakar, "Intrusion detection systems using linear discriminant analysis and logistic regression," in *INDICON*, IEEE, 2015. DOI: 10.1109/INDICON.2015.7443533.

[14] Afef Ben Brahim, Mohamed Limam, "Robust ensemble feature selection for high dimensional data sets," *IEEE*, 2013. DOI: 10.1109/HPCSim.2013.6641406

[15] Haibo He, Edwardo A. Garcia, "Learning from imbalanced data," *IEEE Transactions on knowledge and data engineering*, vol. 21(9), 2009. DOI: 10.1109/TKDE.2008.239.
[16] Meghna Ghosh, Prabu P, "Empirical analysis of ensemble methods for the classification of robocalls in telecommunications," *IJECE*, vol. 9(4), 2019. doi:http://doi.org/10.11591/ijece.v9i4.pp%25p.
[17] Komal Kumar N, R. Lakshmi Tulasi, Vigneswari D, "An ensemble multi-model technique for predicting chronic kidney disease," vol. 9(2), 2019. DOI: 10.11591/ijece.v9i2.
[18] V. Bukhtoyarov, V. Zhukov, "Ensemble-distributed approach in classification problem solution for intrusion detection systems," *Intelligent Data Engineering and Automated Learning-IDEAL*, Springer, pp. 255-265, 2014. DOI:https://doi.org/10.1007/978-3-319-10840-7_32.
[19] A. Borji, "*Combining heterogeneous classifiers for network intrusion detection*," in Proceedings of the Annual Asian Computing Science Conference, Springer, pp. 254-260, 2007. DOI:https://doi.org/10.1007/978-3-540-76929-3_24.
[20] M. Rajasekaran and A. Ayyasamy, "*A novel ensemble approach for effective intrusion detection system,"* Second International Conference on Recent Trends and Challenges in Computational Models, 2017. DOI: 10.1109/ICRTCCM.2017.27.
[21] Dymitr Ruta and Bogdan Gabry, "An overview of classifier fusion methods," *Computing and Information Systems*, pp. 1-10, 2000.
[22] Aizhong Mi, Lei Wang, and Junyan Qi**,** "A multiple classifier fusion algorithm using weighted decision templates**,"** *School of Computer Science and Technology*, Henan Polytechnic University, 2016. doi: http://dx.doi.org/10.1155/2016/3943859.
[23] T. J. Ross, "Fuzzy logic with engineering applications (3rd edition)," *John Wiley & Sons*, New Jersey, USA.
[24] Hastie T., Tibshirani R., & Friedman J., "The Elements of Statistical Learning," New York, Springer-Verlag, 2009.
[25] Gonen M., "Alpaydin, E Multiple kernel learning algorithms," *Journal of Machine Learning Research*, vol. 12, pp. 2211-2268, 2011.
[26] Chebrolu S., Abraham A., and Thomas J.P, "(2004) Hybrid feature selection for modeling intrusion detection systems," *Neural Information Processing*, pp. 1020-1025, 2010. http://dx.doi.org/10.1007/978-3-540-30499-9_158
[27] G. Kumar and K. Kumar, "The use of artificial-intelligence-based ensembles for intrusion detection: a review," *Applied Computational Intelligence and Soft Computing*, p. 21, 2012. Doi:10.1155/2012/850160

BIOGRAPHIES OF AUTHORS

Dr. Pullagura Indira Priyadarsini is an Associate Professor in Dept. of Information Technology, Vardhaman College of Engineering (Autonomous), Hyderabad, Telangana State, India. Earlier she worked as Head of the Department of Computer Science & Engineering in Chalapathi Institute of Technology, Mothadaka, AP, India. She has conducted various conferences and workshops and acted as a convener. She has guided several projects at M.Tech level. She has attended several Faculty Development Programmes at NITs and IITs in India. She published and presented articles in several reputed journals and conferences. She completed her Doctor of Philosophy from Acharya Nagarjuna University, Guntur in the area of machine learning.Her research interests are machine learning, data mining and network security. She completed her M.Tech (CSE) from Koneru Lakshmaiah college of Engineering, Guntur. She stood first in the Post Graduate level. She completed her B.Tech from RVR & JC College of Engineering, Guntur.

Dr. G. Anuradha received her PhD from Andhra University, Visakhapatnam, Andhra Pradesh. She is working as an Associate Professor at the Department of CSE in VRSEC, Vijayawada, Krishna District. She published 17 papers in national and international journals and more than ten papers published in national and international conferences on data mining, web mining and machine learning.