Journal of Soft Computing Paradigm (JSCP) (2019)
Vol.01/ No. 02
Pages: 69-79
http://irojournals.com/jscp/
DOI: https://doi.org/10.36548/jscp.2019.2.002

ENHANCED SOFT COMPUTING APPROACHES FOR INTRUSION DETECTION SCHEMES IN SOCIAL MEDIA NETWORKS

Dr. A. Sathesh,

Assistant Professor, Department of Electronics and Communication Engineering,

Eritrea Institute of Technology,

Eritrea.

Email: Sathesh4you@gmail.com

Abstract: The soft computing methods play a vital role in identifying the malicious activities in the social network. The low cost solutions and the robustness provided by the soft computing in the identifying the unwanted activities make it a predominant area of research. The paper combines the soft computing techniques and frames an enhanced soft computing approach to detect the intrusion that cause security issues in the social network. The proffered method of the paper employs the enhanced soft computing technique that combines the fuzzy logic, decision tree, K means -EM and the machine learning in preprocessing, feature reduction, clustering and classification respectively to develop a security approach that is more effective than the traditional computations in identifying the misuse in the social networks. The intrusion detection system developed using the soft computing approach is tested using the KDD-NSL and the DARPA dataset to note down the security percentage, time utilization, cost and compared with the other traditional methods.

Keywords: Intrusion Detection, Soft Computing, Fuzzy Logic, Decision Tree, Evolutionary Algorithm and Machine Learning

1. INTRODUCTION

The intrusion detection plays a vital role in computer and network security and has become a pre dominant area of research in the network and the computer security due to growing number internet users in the everyday life. The failure to detect he intruders would result in the data loss, unauthorized usage or misuse of data’s though various detection approaches for the intrusion is put forward to solve the issues in the security the approaches used does not provide a satisfactory performance that results with an increased detection rate and decreased false alarm rate. The figure.1 below shows the basic step involved in the intrusion detection system to identify the malware in the network.

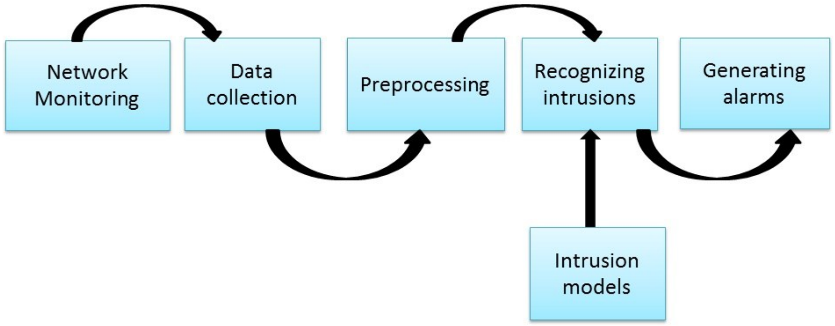


Fig.1 A General Intrusion Detection System

The intrusion occurring in the network is an illegal/ unlicensed activity that misuses the network resources that are provided for the other authorized users and causes security threats to the network. Identifying the intrusion relies on the defenders who possess a clear understanding of the working of the attacks. Some of the popular intrusions that causes vulnerability to the network are as follows

Asymmetric Routing: The network set up that holds the configurations for the asymmetric routing are vulnerable to these types of attacks. The intrusion actually occurs by utilizing the more than one route to cause an overall attack in the important segment of the network.

Butter Flow Attacks: This intrusion replaces the normal content in the set of commands in the memory that will be executed later as a segment of intrusion and cause problem in the network by denying services to the authorized users.

Protocol Specific Attacks: The protocols that does not hold authentication are liable of getting attacked and causing malfunctions in the regular network activities.

Traffic flooding: In this type of attack the intruder frames heavy traffic load towards network, causing inadequacy in screening and congested network environment, enabling the intruder to execute the hidden attack.

There are many other attacks such as CGI, Trojans, and Worms etc. that provides threats to the security of the network by affecting its activities or misusing its data.

So the paper proposes the enhanced soft computing technique that combines the fuzzy logic, decision tree, K means -EM and the machine learning in preprocessing, feature reduction, clustering and classification respectively to develop a security approach that is more effective than the traditional computations in identifying the misuse in the social networks.

The remaining paper is arranged with the section2 providing the detail of soft computing techniques utilized in developing the intrusion detection system. Section 3 presenting the proposed architecture of the intrusion detection, section 4 analyzing and comparing the results and section 5 holding the conclusion followed by the references.

2. RELATED WORKS

The soft computing approaches are most predominant in the intrusion detection nowadays the Langin, et al [1] in his paper provides the involvement of the soft computing techniques in identifying the intruder in the network and Singh, et al [2] provides the “review based on the soft computing in the malware detection”. To further improvise the detection by basically improving the feature extraction and classification Ahmad et al [3] puts for the “soft computing techniques to have optimized intrusion detection”

Jadhav, etal [4] in his paper “presents the approach for detecting of the intrusion using the data mining” Shamshirband et al [5] put forth the "Co-FAIS: cooperative fuzzy artificial immune system for detecting intrusion in wireless sensor networks." Panda et al [6] utilizes the “hybridized algorithm in the detection of the intrusions in the network” by employing dual classifier stage.

The author Dash, et al [7] presents the “study reports on the two new hybrid algorithms GS and PSO and their utilization in the training of ANN to identify the intruders” “an adaptable intrusion detection system integrating the artificial immune system and the computation intensive nature of the soft computing is developed by the author Sanyal, et al [8] to identify the intrusion in the given network”.

The survey of Butun et al [9] holds the importance of the intrusion detection system in the adhoc networks the author initially provides the survey of the IDS for the wireless sensor networks followed by the IDS developed for the mobile adhoc networks.

Sharma et al [10] puts forth an “enhanced approach of fuzzy C-means clustering for anomaly detection.” Kumar, et al [11] elaborates the survey on techniques of soft computing engaged in the intrusion detection. And Sangve, et al [12] put forth the “Anomaly based improved network intrusion detection system using clustering techniques." S. Smys et al has proposed a "Trust-based intrusion detection and clustering approach for wireless body area networks."

3. ENHANCED SOFT COMPUTING APPROACH FOR INTRUSION DETECTION

To attain a satisfied performance in terms of heightened detection rate and reduced false alarm rate the proposed method in the paper utilizes the enhanced soft computing approach that combines the sniffer module [5] for monitoring the data packets that enter into the network this utilizes the fuzzy rule based system as preprocessor analyze the features of the packet, along with the decision tree for feature reduction followed by a clustering based on the Kmeans-EM and finally with the classification based on the SVM (support vector machine) separate the normal from the attacked data. The fig.2 below shows the step in the proposed model.

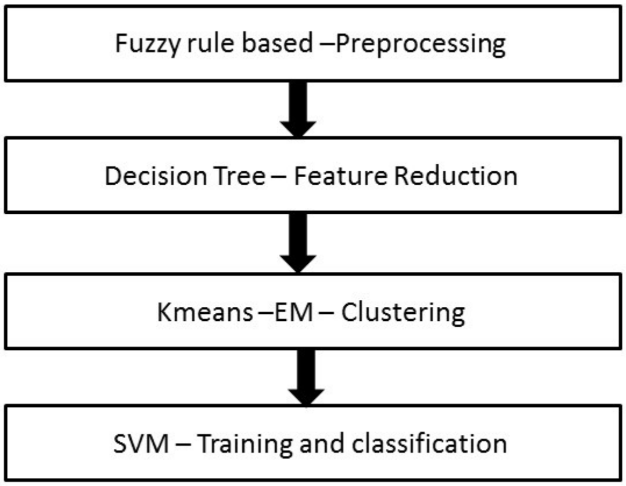
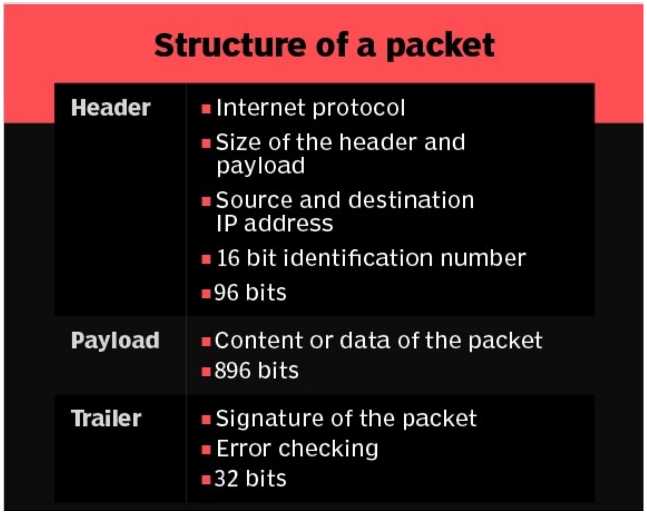

Fuzzy rule based –Preprocessing
Decision Tree – Feature Reduction
Kmeans –EM – Clustering
SVM – Training and classification

Fig.2 steps involved in the proposed process

The sniffer module captures the packets in the network traffic and analyzes each packet components. The data packets are monitored online and conveyed to its next stage. During the heavy traffic the data flow the module saves the outcomes in the log file. The packet features are of the capture data packets are preprocessed employing the fuzzy rule based analyzer [5] that analyses the packet based on the packet features. The fig .3below shows the packets features that are utilized by the fuzzy interference to preprocess the data packets. The preprocessed output are fed into the feature reducer that is based on the decision tree.

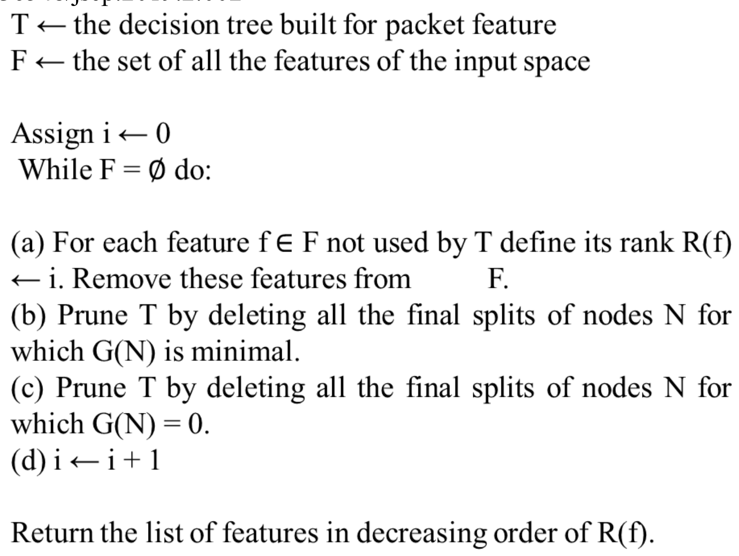


Structure of a packet

Header	▪ Internet protocol ▪ Size of the header and payload ▪ Source and destination IP address ▪ 16 bit identification number ▪ 96 bits
Payload	▪ Content or data of the packet ▪ 896 bits
Trailer	▪ Signature of the packet ▪ Error checking ▪ 32 bits

Fig .3 Packet Features

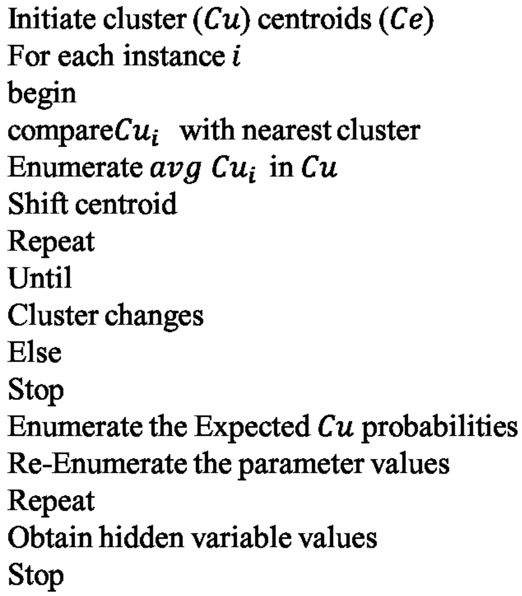
The algorithm shown below in figure.4 provides the essentials steps involved in the decision tree [1] in for reduction of features to the identifying the appropriate intruders who remains as menace to the network activity.



```
T ← the decision tree built for packet feature
F ← the set of all the features of the input space

Assign i ← 0
 While F = ∅ do:

(a) For each feature f ∈ F not used by T define its rank R(f)
← i. Remove these features from       F.
(b) Prune T by deleting all the final splits of nodes N for
which G(N) is minimal.
(c) Prune T by deleting all the final splits of nodes N for
which G(N) = 0.
(d) i ← i + 1

Return the list of features in decreasing order of R(f).
```

Fig .4 Decision Tree algorithm for Feature Reduction

The extracted features of the data set obtained are clustered based on the similarities using the k-means –EM (Expectation Maximization Algorithm) clustering to minimize the number of the data sets used in training in order to bring down the complexities in the computation and processing. The steps below in the fig.5 shows the K-Means - EM in the process of clustering to reduce the number of dataset used in training



```
Initiate cluster ($Cu$) centroids ($Ce$)
For each instance $i$
begin
compare$Cu_i$ with nearest cluster
Enumerate $avg\ Cu_i$ in $Cu$
Shift centroid
Repeat
Until
Cluster changes
Else
Stop
Enumerate the Expected $Cu$ probabilities
Re-Enumerate the parameter values
Repeat
Obtain hidden variable values
Stop
```

Fig.5 K-Means-EM Clustering

The K-Means-EM Clustering [12] procedure scopes to reduce the squared function. Where the squared function is represented as shown in the equation (1)

$$s(v) = \sum_{i=1}^{Cu} \sum_{j=1}^{Cu_i} (\| Ed \|)^2 \qquad (1)$$

Where, $- Ed = x_i - y_j$ is the Euclidean distance of the x_i and y_j , Cu_i is the number of data instances in the cluster i, and Ce the total cluster centers. The clusters framed enables to detect the characteristics that belongs to normal or abnormal categories.

The Expectation Maximization Algorithm that proceeds as two phase initially enumerates the expectations of the cluster probability followed by the re-estimation of the parameter values that were obtained from the results of the first stride. The outputs obtained at the second step is once again fed back as input to the first step. This process continues until the results converge. So the EM helps to assign the values to the hidden variables and process the parameters based on the data.

Finally the data set obtained are applied for training of the SVM classifier to categorize the intrusion the network. The algorithm below in fig. 6 shows the steps involved in SVM classifier to segregate between the normal and the abnormal data set.

[redacted]

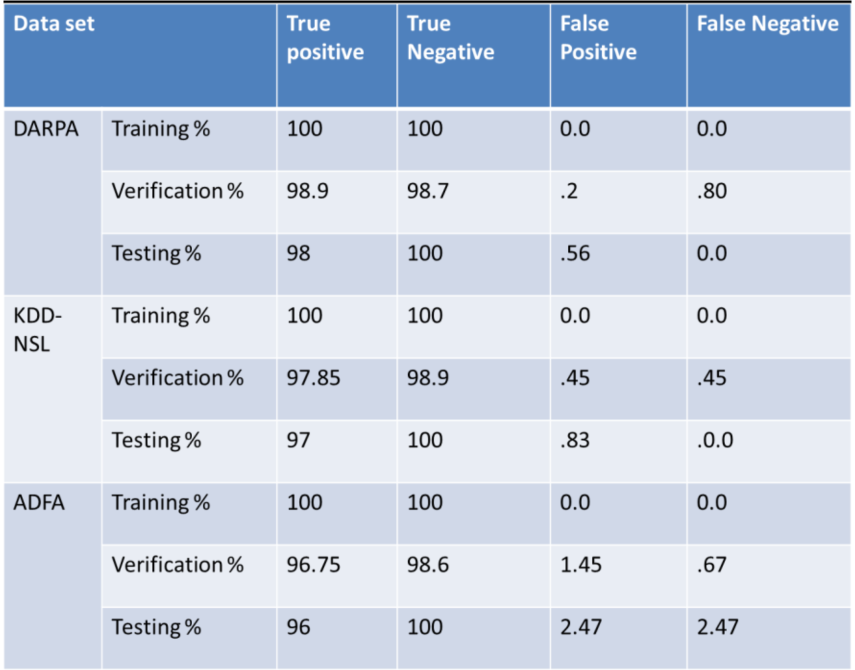
Fig.6 SVM classifier

4. RESULT ANALYSIS

The proposed enhanced soft computing approach is tested using the KDD-NSL and the DARPA data set and was found to have a high accuracy, prediction on training and testing. The training dataset holds both the patterns of the inputs as well as the outputs that are associated with the input scoping to minimize the false alarms and heighten the detection rate. The proffered system is trained using the 80% of the dataset. In the remaining data set 10% are used in cross validation and 10% in testing.

The testing of the system is segregated into two phases as the verification and generalization, the verification process test whether the system is properly trained using the training dataset, the successful training of system would lead to the output that closely resembles the input [3].

The analysis of the proffered system in terms of true positive, false positive, false negative and true negative for the training verification and the testing phase for using different dataset is shown below in the table.1



Data set		True positive	True Negative	False Positive	False Negative
DARPA	Training %	100	100	0.0	0.0
	Verification %	98.9	98.7	.2	.80
	Testing %	98	100	.56	0.0
KDD-NSL	Training %	100	100	0.0	0.0
	Verification %	97.85	98.9	.45	.45
	Testing %	97	100	.83	.0.0
ADFA	Training %	100	100	0.0	0.0
	Verification %	96.75	98.6	1.45	.67
	Testing %	96	100	2.47	2.47

Table.1 Analysis of the Proposed System

The fig.6 below shows the comparison of the detection rate, false alarm rate, time complexity and cost for the proposed system along with the other existing approaches.

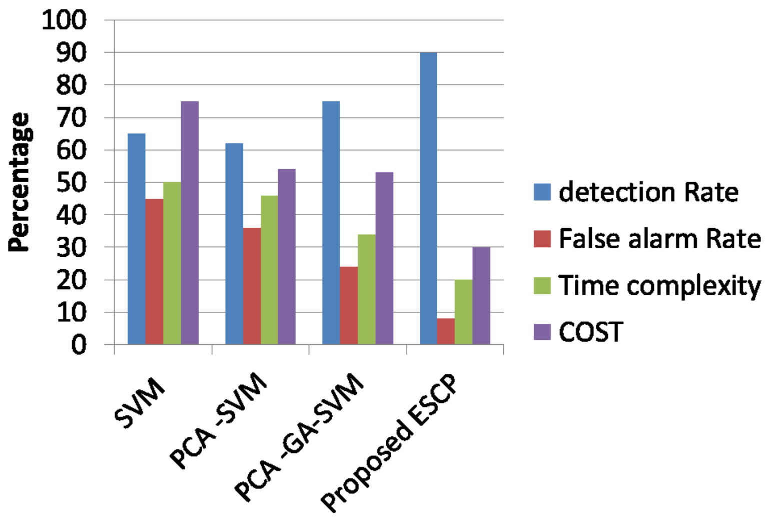


Fig.7 Performance Comparison

The method enables an optimized intrusion detection, which is enriched with the capability to bring down the amount of features and maximize the detection rates. The fig.7 included evinces that the proposed enhanced soft computing approach (ECSP) shows an improved performance compared to the other methods of intrusion detection.

5. CONCLUSION

The paper put forth the enhanced soft computing approach that combines the fuzzy rule based preprocessing , decision tree based feature reduction , K-Means –EM clustering and the support vector machine based classification to identify the intrusions in the network. The clustering and the feature reduction methods enable to minimize the data set used in training and thus reduce the time complexity involved in the training. The testing of the proposed system using the KDD-NSL data set and the DARPA data set show an enhanced accuracy and precision in training and the testing as well as the considerable minimization in the cost and the time complexity.

References

[1] Langin, Chet, and Shahram Rahimi. "Soft computing in intrusion detection: the state of the art." *Journal of Ambient Intelligence and Humanized Computing* 1, no. 2 (2010): 133-145.

[2] Singh, Raman, Harish Kumar, and R. Singla. "Review of soft computing in malware detection." *Special issues on IP Multimedia Communications* 1, no. 1 (2011): 55-60.

[3] Ahmad, Iftikhar, Azween Abdullah, Abdullah Alghamdi, and Muhammad Hussain. "Optimized intrusion detection mechanism using soft computing techniques." *Telecommunication Systems* 52, no. 4 (2013): 2187-2195.

[4] Jadhav, R. J., and U. T. Pawar. "Data mining for intrusion detection." *International Journal of Power Control Signal and Computation* 1, no. 4 (2005): 45-48.

[5] Shamshirband, Shahaboddin, Nor Badrul Anuar, Miss Laiha Mat Kiah, Vala Ali Rohani, Dalibor Petković, Sanjay Misra, and Abdul Nasir Khan. "Co-FAIS: cooperative fuzzy artificial immune system for detecting intrusion in wireless sensor networks." *Journal of Network and Computer Applications* 42 (2014): 102-117.

[6] Panda, Mrutyunjaya, Ajith Abraham, and Manas Ranjan Patra. "A hybrid intelligent approach for network intrusion detection." *Procedia Engineering* 30 (2012): 1-9.

DOI: https://doi.org/10.36548/jscp.2019.2.002

[7] Dash, Tirtharaj. "A study on intrusion detection using neural networks trained with evolutionary algorithms." *Soft Computing* 21, no. 10 (2017): 2687-2700.

[8] Sanyal, Sugata, and Manoj Rameshchandra Thakur. "A Hybrid Approach towards Intrusion Detection Based on Artificial Immune System and Soft Computing." *arXiv preprint arXiv:1205.4457* (2012).

[9] Butun, Ismail, Salvatore D. Morgera, and Ravi Sankar. "A survey of intrusion detection systems in wireless sensor networks." *IEEE communications surveys & tutorials* 16, no. 1 (2013): 266-282.

[10] Sharma, Ruby, and Sandeep Chaurasia. "An enhanced approach to fuzzy C-means clustering for anomaly detection." In *Proceedings of First International Conference on Smart System, Innovations and Computing*, pp. 623-636. Springer, Singapore, 2018.

[11] Kumar, Koushal, and Simranjit Singh. "Intrusion Detection Using Soft Computing Techniques." (2016).

[12] Sangve, Sunil M., and Uday V. Kulkarni. "Anomaly based improved network intrusion detection system using clustering techniques." *International Journal of Advanced Research in Computer Science* 8, no. 7 (2017).

[13] Anguraj, Dinesh Kumar, and S. Smys. "Trust-based intrusion detection and clustering approach for wireless body area networks." *Wireless Personal Communications* 104, no. 1 (2019): 1-20.