ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

Multilevel and Multi-class Support Vector Machine based on Affinity Propagation Clustering for Intrusion Detection

Shinee Singh[1], S. K. Shrivastava[2]

M.Tech Scholar, Department of Computer Science & Engineering, SATI, Vidisha, India[1]

Professor, Department of Computer Science & Engineering, SATI, Vidisha, India[2]

ABSTRACT: Intrusion detection system is used as a security tool in network and host computer. Intrusion detection system is a software which analyzes the network to detect intrusion and monitor the event occurs in the network.In the existing model"Multilevel hybrid support vector machine and extreme learning machine based on modified k-means for intrusion detection system",the dataset is reduced using modified k-mean clustering. The drawback of modified k-mean is that the number of cluster have to be passed at initial stage and the clustering result also changed if the initial centroid is changed.Our research work eliminates this drawback by using affinity clustering. In this research work the proposed model for intrusion detection is based onaffinity clustering and multilevel SVMclassifier. The proposed work is performed in four stages. In the first stagefeature selection is performed using Pearson, Kendall and Spearman co-relation algorithm.In second stage data is normalized using statistical normalization. In third stage affinity clustering is performed to cluster dataset into different data instances. In last stage multilevel SVMclassifier, Multi-class SVM classifier and multilevel SVM-ELM classifier is used.For evaluating performance,the KDD dataset and NSL dataset have been used. Comparing all three models using both datasets, the proposed multilevel SVM classifier gives highest accuracy and lowest false alarm rate.

KEYWORDS: Intrusion Detection, Machine Learning, statistical normalization affinity propagation, Classification, Accuracy, Detection Rate,Precision,F1-Score, FAR.

I. INTRODUCTION

Intrusion is a set of action that attacks the integrity,confidentiality,availability and authenticity of the network. IDS is a software which analyzes the network to detect intrusion and monitor the event occurs in the network.In the existing model the dataset is reduced using modified k-mean clustering and then classified data using Multilevel hybrid SVM and ELM.The accuracy and detection rate of existing model is 95.75% and 95.17% respectively with FAR of 1.87%.The drawback of modified k-mean is that the number of clusters has to be passed at initial stage only and the clustering result also changed when initial centroid is changed [1]. Our research work eliminates this drawback by using affinity propagation clustering.This research work is performed in major four stages.In the first stage feature selection is performed using Pearson,Spearman and Kendall co-relation algorithm. There are some feature of the dataset which are not relevant for classification.These irrelevant features can be removed without having loss of useful data.Since the feature selection is used to remove the irrelevant feature and select the relevant feature [2].For selecting relevant feature, feature selection uses the co-relation coefficient namely Pearson,Spearman and Kendall.The result of Pearson,Spearman and Kendall can be positive, negative and zero.Select only those features which have a positive result for all the three co-relation coefficients.It also increases the processing of the proposed model [3].

In the second stage data are normalized using statistical normalization. There are some redundant data and any related data which are not in order. To eliminate this problem, the data should standardized.This research work standardizes dataset

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

using statistical normalization.Statistical normalization convert or transform the data derivable from any normal distribution into standard normal distribution with mean zero and unit variance [4].
In the third stage affinity propagation clustering is performed to group the dataset into clusters.AP clustering is used to reduce dataset so as to acquire increased accuracy and detection rate.AP clustering operates on three matrices which are similarity matrices,responsibility matrices and availability matrices.The similarity matrices are the matrices which contain feature relation between the two data points.In our research work the similarity matrices is calculated by square Euclidean distance.The responsibility matrices are calculated by responsibility message which contain evidence should be suited to serve as an exemplar.The availability matrices are calculated by the availability message which gives evidence to be appropriate to choose a data point as an exemplar.The clusters formed as a result of AP clustering is then divided into training data and testing data.these data are then sent into the classification processes [5][6].
In the last stage Multilevel svm classifier,Multi-class svm classifier and Multilevel svm-elm classifier is used for classification. In Multilevel svm model, the binary svm is applied at each level and single attack is detected at each level, this model completed its classification within 4 levels.Support vector machine is a supervised algorithm which means the training dataset is required to train the classifier.The advantage of using SVM is that it gives efficient result and uses the memory efficiently [7].In Multi-class SVM, data is classified at single level only.Firstly the one-class svm classifies normal data and attack data, then Multi-class svm classifies the attack data into one of dos, u2r, r2l and probe [8].In multilevel svm-elm classifier model, the data are classified using svm and elm at alternate levels.The first and third level classification is done by svm classifier and second and fourth level classification is done by elm classifier.In multilevel classification, the dos attack is detected at first level, probe at second level, u2r at third level and r2l at the fourth level.All these stages are implemented in MATLAB and performance is done using the KDD dataset and NSL dataset.
Step by step implementation of this work is described in proposed Algorithm section and all performance results using several performance measurements and comparison result are discussed in experimental result analysis section.In recent years Several works have been done on intrusion detection.All these related work is discussed in the next section.

II. RELATED WORK

WathiqLaftah Al-Yaseen et al. [1] proposes a multilevel hybrid intrusion detection model that uses support vector machine and extreme learning machine to improve the efficiency of detecting known and unknown attacks. A modified K-means algorithm is also proposed to build a high-quality training dataset that contributes significantly to improving the performance of classifiers. The modified K-means is used to build new small training data-sets representing the entire original training dataset, significantly reduce the training time of classifiers, and improve the performance of intrusion detection system. The popular KDD Cup 1999 dataset is used to evaluate the proposed model. Compared with other methods based on the same dataset, the proposed model shows high efficiency in attack detection, and its accuracy (95.75%) is the best performance thus far.
YadigarImamverdiyev [9] discussed that intrusion detection systems are one of the most relevant security features against network attacks. Machine learning methods are used to analyze network traffic parameters in the presence of attack signs. This article discusses the extreme machine learning method for detecting intrusions in network traffic. The experimental results lead to the conclusion of the practical significance of the proposed approach to detect attacks in network traffic.
Bhanu Vrat et al. [10] discussed that detection of anomalies is important requirement to protect a network against the strikers. Detects attacks on a network the analysis of the behavioral model was a important field of study for many researchers application systems in IPv4 and IPv6 networks. For accurate detection of anomalies, it is essential implement and use effective data mining methodology such as machine learning. In this article we considered a model of anomaly detection that uses machine learning algorithms for data mining in a network to detect anomalies present at any time. This the proposed model is evaluated against denial of service Attacks (DOS) in IPv4 and IPv6 networks selecting the most common and obvious features of IPv6 and IPv4 networks to optimize detection. The results also show that the proposed system can detects most IPv4 and IPv6 attacks effectively way.
Khadija Hanifi et al. [11] discussed that network attacks are exceptional cases they are not observed in the normal behavior of the traffic. In this work, to detect network attacks, using the k-means algorithm a new semi-supervised anomaly detection system was designed and implemented. During the training phase, normal samples were split into clusters by applying the k-means algorithm. So in To be able to distinguish between normal and abnormal samples, based on their distance from cluster centers and using a validation data set, a threshold value has been calculated. New samples that are far from cluster centers more than the threshold value is detected as anomalies. We used NSL-KDD- a data set labeled network connection traces - to test ours the effectiveness of the method. The experiments result in NSL-KDD dataset, shows that we have reached an accuracy of 80.119%.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

III. PROPOSED ALGORITHM

The following steps will be used to build the proposed model for intrusion detection:
Step 1: Select the dataset.
Step 2: Convert the symbolic attributes protocol, service, and flag to numerical.
Step 3: Feature selection and Extraction.
Step 4: Separate the instances of dataset into two categories: Normal, DOS, R2L, U2R and Probe.
Step 5: Normalize data to [0,1].
Step 6: Data Clustering
Step 7: The data set is divided as training data and testing data.
Step 8: Train classifier with these new training dataset.
Step 9: Test model with dataset.
Step10: Finally computing and comparing Accuracy and FAR for different classifiers.

Proposed Methodology:
The proposed framework consists of three phases, i.e. Preprocessing, Post Processing Phase and Intrusion Detection Phase. Below each stage is described individually in details.

Pre-processing phase:
This phase purpose is to pre-process the database file.This phase consists two main steps which are as follows: Firstly, the conversion of the symbolic attributes protocol, service, and flag in numerical is done.The symbolic attributes of the dataset are seen in the column 2,3,4 of the dataset respectively.Secondly, the feature selection process is done.The aim of Feature selection is to further select only those features from the database which are relevant for proper classification of the dataset and consequently reduces the feature space dimension so as to reduce complexity by removing irrelevant data. In this research work, for feature selection Correlation Analysis is performed using Pearson, Spearman and Kendall coefficients.

Post-processing phase:
This phase consists of two steps which is statistical normalization and Data clustering.The statistical normalization is the process of converting data derivable from any Normal distribution into standard Normal distribution with mean zero and unit variance.
The statistical normalization is defined as:

$$\text{Xi} = \frac{\upsilon i - \mu}{\sigma}$$

where μ is mean of n values for a given attribute: $\mu = \frac{1}{n}\sum_{i=1}^{n} \upsilon i$

Rho(σ)is its standard deviation.

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\upsilon i - \mu)}$$

However, using statistical normalization, the data set should follow a Normal distribution, that is, the number of sample '*n*' should be large according to the central limit theorem.The statistical normalization does not scale the value of the attribute into [0,1]. It instead ranges 99.9% samples of the attribute into [-3, 3].Statistical normalization not only considers the mean scale of attribute values, but also takes into account their statistical distribution and this may help a lot for the Detection.
In this research work affinity propagation clustering is used to cluster a dataset into a set of clusters. Affinity Propagation is a clustering algorithm that generates a set of 'exemplars' from the dataset. The input of Affinity Propagation is the pairwise similarities between each pair of data points, s[i, j](i, j = 1, 2,...,N).
Affinity Propagation finds the exemplars that maximize the net similarity, i.e. the overall sum of similarities between all exemplars and their member data points. The process of Affinity Propagation can be viewed as a message passing process with two kinds of messages exchanged between data points: responsibility and availability. Responsibility, r[i, j], is a message from data point i to j that reflects the accumulated evidence for how well-suited data point j is to serve as the exemplar for data point i. Availability, a[i, j], is a message from data point j to i that reflects the accumulated evidence for how appropriate it would be for data point i to choose data point j as its exemplar. Initially All responsibilities and availabilities are set to 0. Their values are iteratively updated as follows to compute values:

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

$$r[i, j] = (1-\lambda)\rho[i, j] + \lambda r[i, j]$$
$$a[i, j] = (1-\lambda)\alpha[i, j] + \lambda a[i, j]$$

where λ is a damping factor,ρ[i, j] is propagating responsibility and α[i, j] is propagating availability.

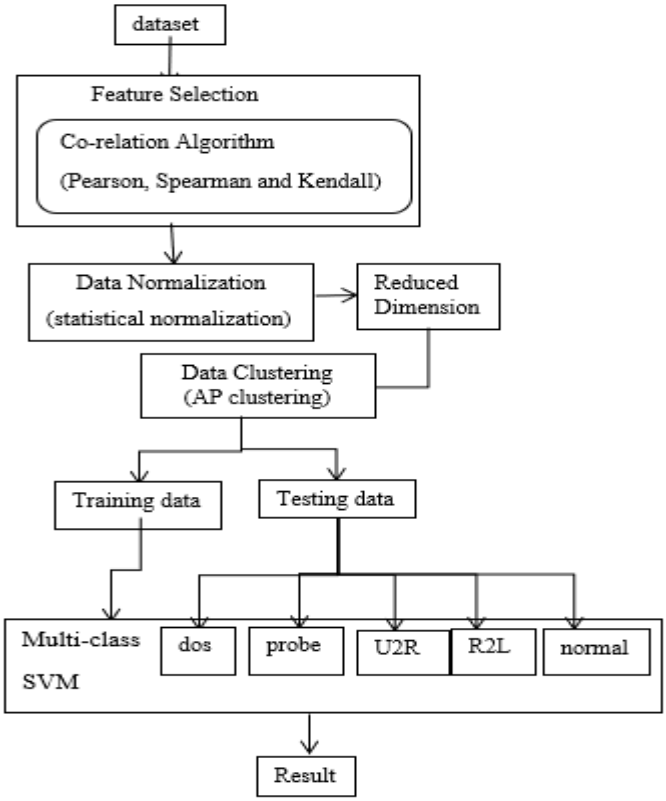

Fig:1 flow diagram of proposed Multi-class SVM classification.

Intrusion Detection phase:

In this phase, the proposed method is designed by combining statistical normalization and AP clustering with only SVM and with SVM-ELM both as a Multilevel model and with Multi-class SVM as a Multi-class model for intrusion detection.Statistical normalization and AP clustering is used to reduce dataset size and form new reduced dataset.Several classifier has been used to design Multilevel and Multi-class IDS.Classifier used in this phase is binary SVM, ELM and Multi-class SVM.

This research work described three models to analyze intrusion detection these models are Multi-class svm, Multilevel svm and Multilevel svm-elm.The aim of this research work is to find out best intrusion detection model with high accuracy, detection rate and low false alarm rate.

In Multilevel svm model, svm is applied at each level of testing data. One attack is detected at each level.In Multilevel svm-elm model, svm is applied at the first and third level and elm is applied at second and fourth level to detect attacks.In Multi-class svm model,there is only single level and Multi-class svm is applied at this level.All attacks are detected at single level itself.The flow diagram of all three described model i.e. Multi-class svm classification model is shown in Fig:1, the Multilevel svm classification model is shown in Fig:2 and Multilevel svm-elm classification model is shown in Fig:3.

Training data and Testing data:

Training dataset is a set of examples used for learning so that it fits into the parameter of the classifier.Testing dataset is set of examples used only to assess the performance of a fully specified classifier.In this research work reduced data formed after clustering is then

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

divided into training and testing dataset for the classifier.The new reduced dataset divides data randomly into training and testing dataset in the ratio 7:3. The 70% data serve as a training data and 30% data serves as a testing data.

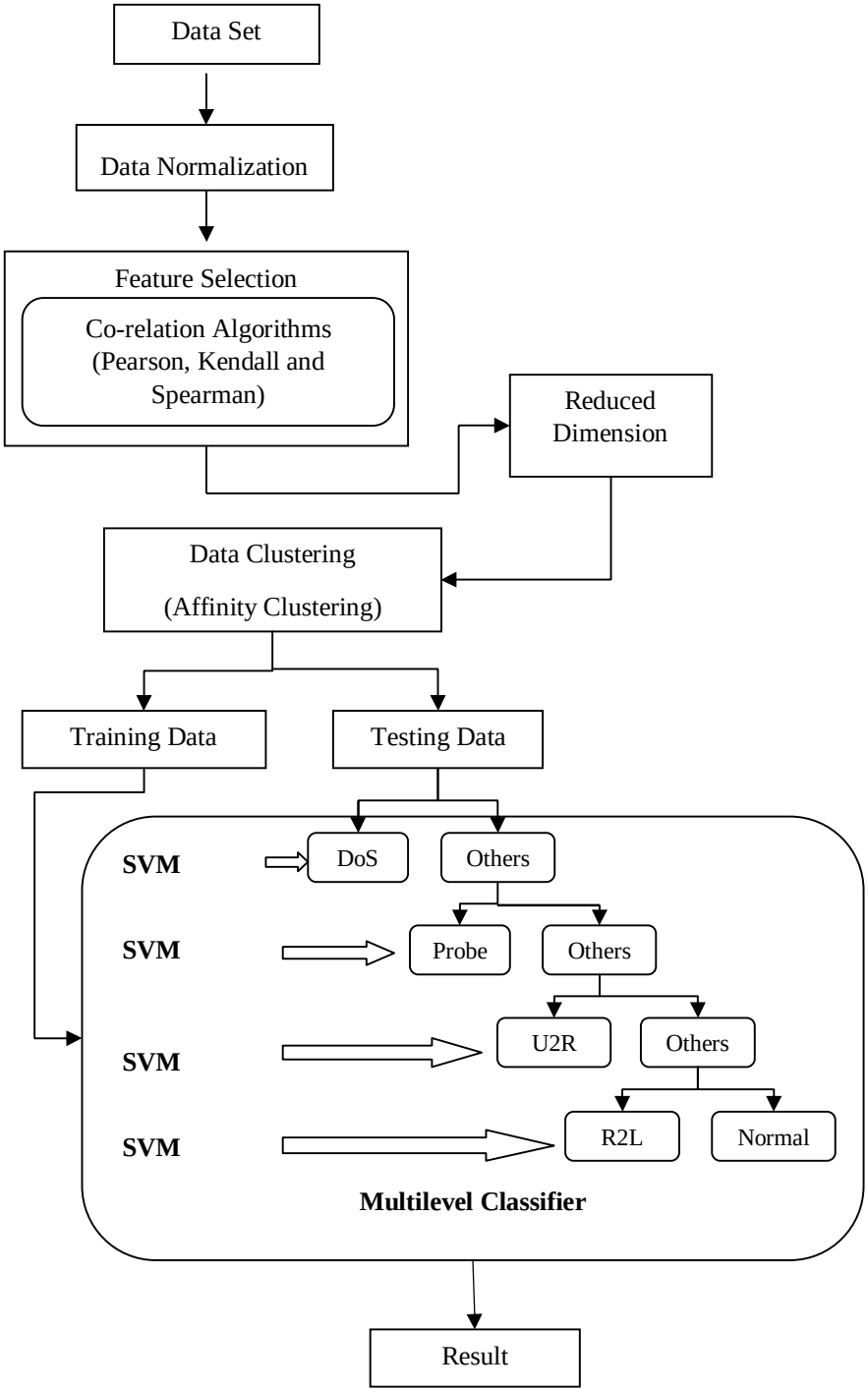

Fig:2 flow diagram of proposed Multilevel SVM classification.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

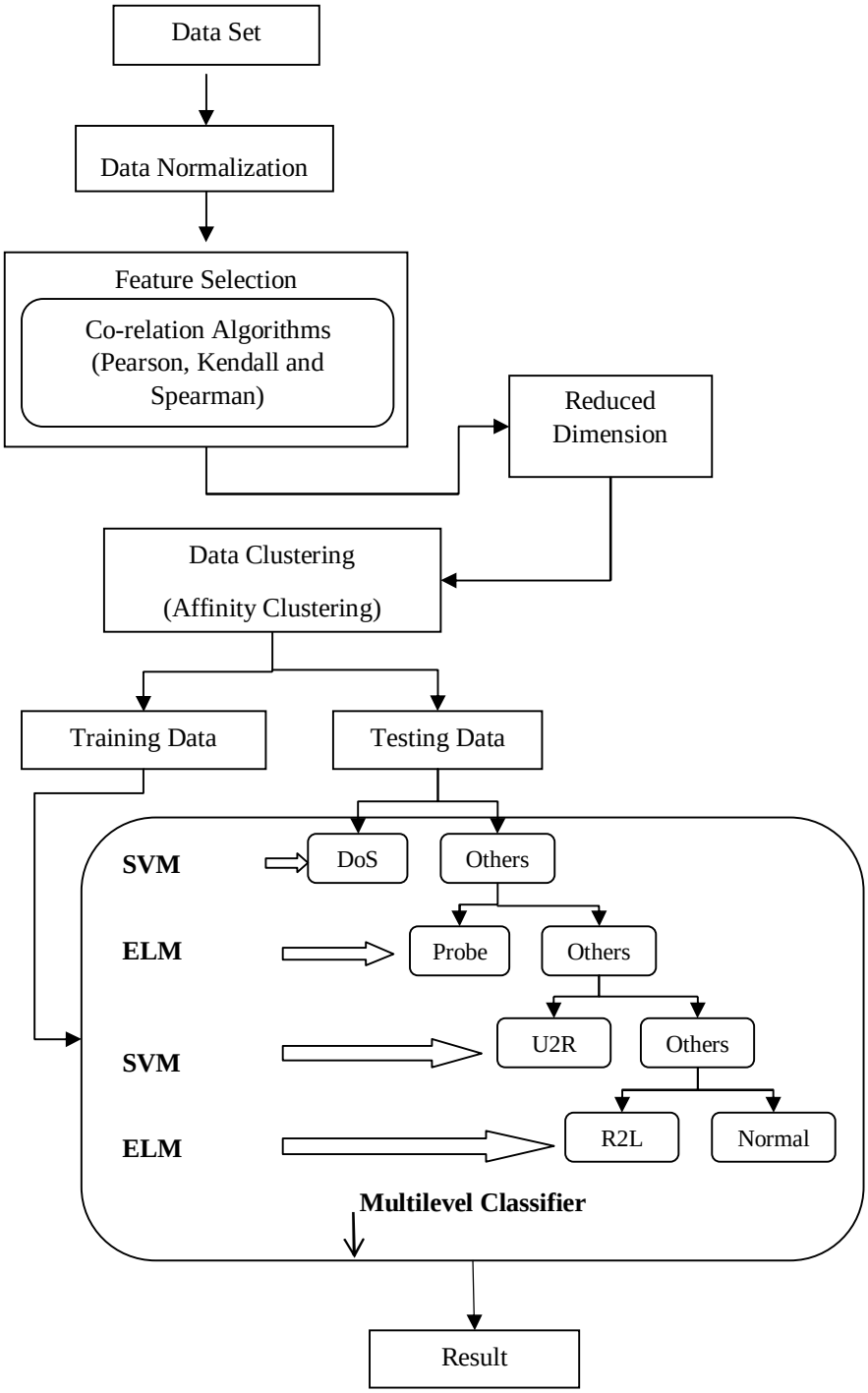

Fig:3 flow diagram of proposed Multilevel SVM-ELM classification.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

IV.EXPERIMENTAL RESULT ANALYSIS

This section evaluates the performance of the proposed models and compares all 3 models.The number of instances between frequency normalization and statistical normalization is also compared.For evaluating the performance of proposed models kdd dataset and NSL dataset has been used.Number of instances between modified k-mean clustering and Affinity propagation clustering are also compared using kdd dataset.All the evaluation is done on WINDOWS 7, MATLAB, running on an Intel core™i3 processor with 8GB RAM. Performance measurement used are accuracy, DR, FAR, recall, precision, F1-score.

Performance: using KDD dataset

Performance Comparison of proposed models, i.e. Multi-class SVM,multilevel SVM and multilevel SVM-ELM using kdd dataset are shown in TABLE:1.

performance	Proposed Multi-class svm	Proposed Multilevel svm	Proposed Multilevel svm -elm	Yaseen Et al
accuracy	85.38	99.06	96.01	95.75
Detection rate	62.99	100	98.07	95.17
False alarm rate	23.0136	1.32	3.86	1.87
recall	62.99	100	98.07	
Precision	63.57	98.57	91.24	
F1-Score	63.28	99.28	94.13	

TABLE:1 performance comparison of models

A key observation from TABLE:1 are as follows:

1. Proposed Multilevel SVM shows higher accuracy, detection rate, recall, precision and F1-Score.
2. False alarm rate (FAR) is lowest for proposed Multilevel SVM.

Bar graph representation of the above table is shown in figure 1(a) and figure 1(b).Figure 1(a) represents the performance comparison between proposed models and existing model. Figure 1(b) represents the performance comparison between all 3 proposed models.

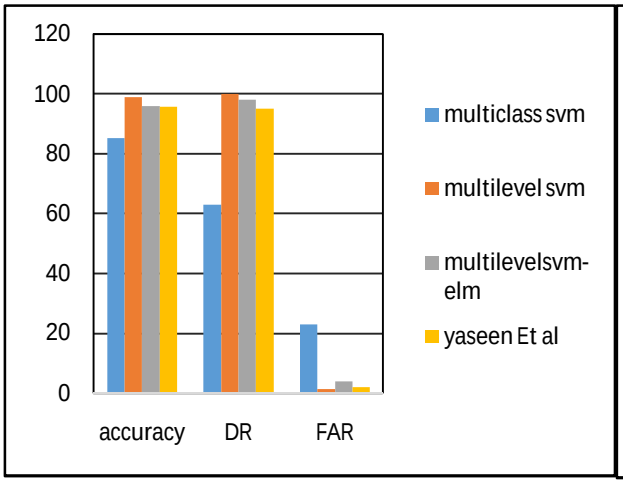

Figure 1(a)

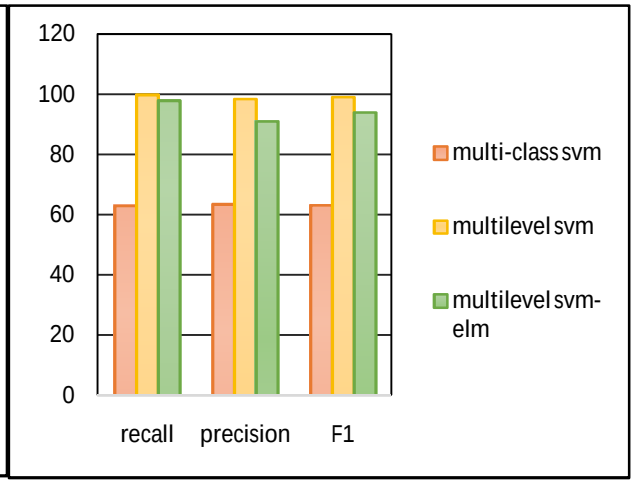

Figure 1(b)

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

TABLE:2 below illustrates the normalization of kdd data-set instances using frequency and statistical normalization.

category	Kdd dataset	Frequency normalization	Statistical normalization
normal	97278	87832	25360
dos	391458	54572	3784
probe	4107	2130	720
U2r	52	52	46
R2l	1126	999	422
total	494021	145585	30332

TABLE:2 number of instances comparison between frequencyand statistical normalization using kdd dataset instances.

A key observation from TABLE:2 are as follows:

- Statistical normalization gives a less number of instances in all categories namely Normal, dos, probe, u2r and r2l.
- Overall instances in statistical normalization are less than the frequency normalization.

The bar graph representation of frequency normalization and statistical normalization instances are shown in figure 2(a). The ROC curve of figure2(b) represents the instances difference between statistical normalization and frequency normalization.

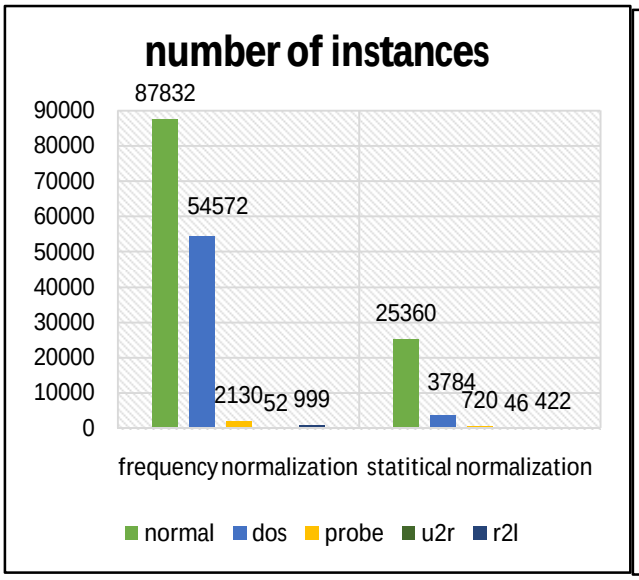

Figure 2(a)

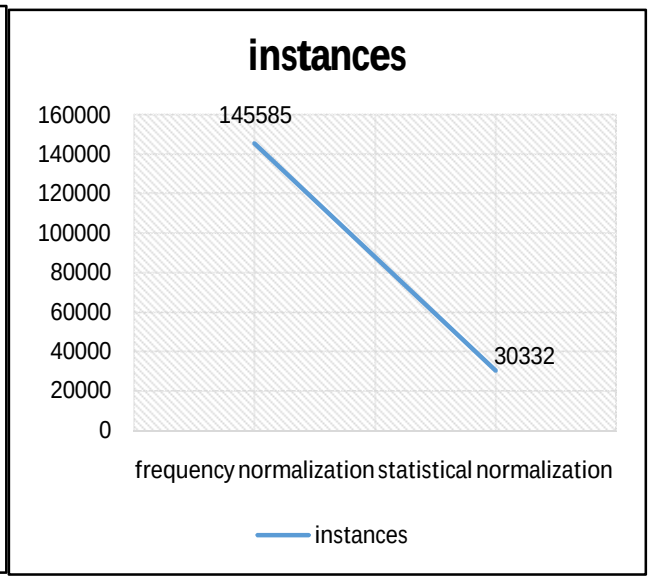

Figure 2(b)

Comparison of number of instances between clustering of existing model (i.e. Modified k-mean) and clustering of proposed model (i.e. Affinity propagation) using kdd dataset is shown in TABLE:3.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

category	10% kdd no. Of instances	Modified k-mean no. Of instances	AP clustering no. Of instances
normal	97278	639	38
dos	391458	140	580
probe	4107	134	53
R2l	1126	51	22
U2r	52	25	46
total	494021	989	739

TABLE:3 Reduced number of instances in modified k-mean,AP clustering.

A key observation from TABLE:3 are as follows:

- Instances of normal, probe, r2l is less in AP clustering.
- Instances of dos and u2r is less in modified k-mean.
- Overall number of instances in AP clustering are less as compared to modified

K-mean clustering.

Bar graph representation of TABLE:3 is shown in figure 3(a) and figure 3(b).Figure 3(a) represents the number of instances between modified k-mean clustering, AP-clustering and kdddataset.The ROC curve for figure 3(a) is shown in figure 3(c), Figure 3(b) represents the number of instances comparison between modified k-mean and AP clustering.The ROC curve for figure 3(b) is shown in figure 3(d).

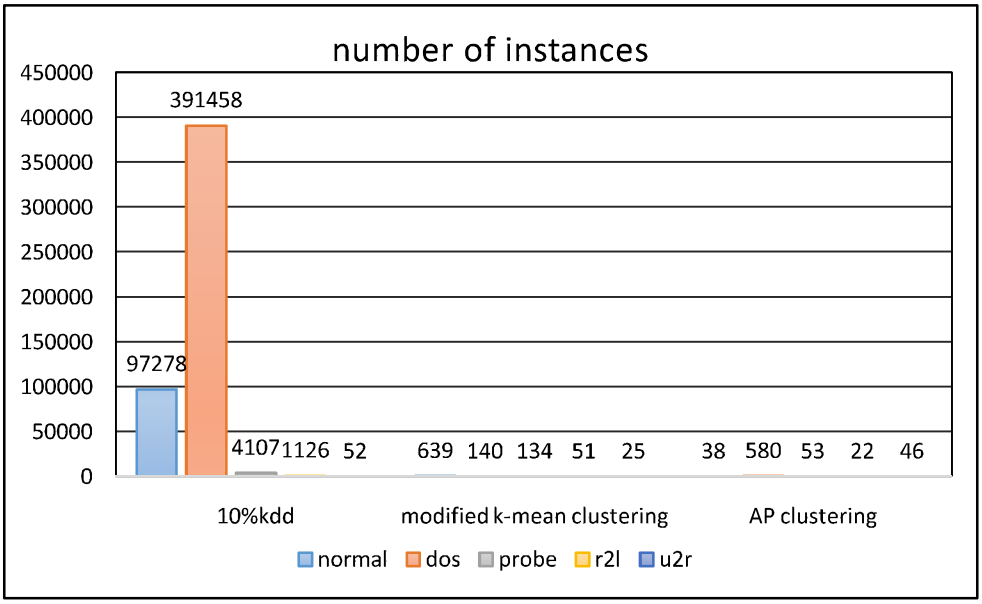

Figure 3(a): number of instances in modified k-mean, AP clustering and the kdd dataset.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

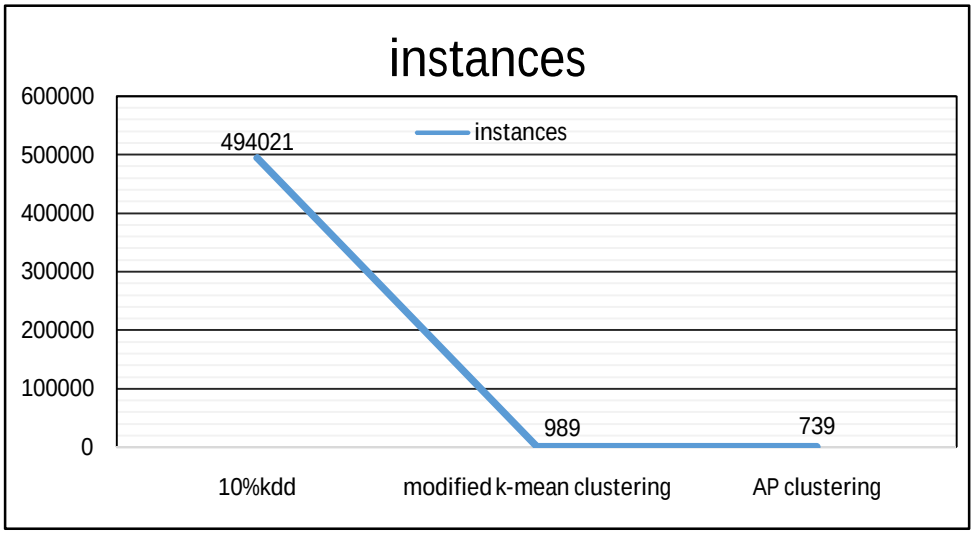

Figure 3(c): ROC curve for instances comparison between kdddataset,modified k-meanand AP clustering.

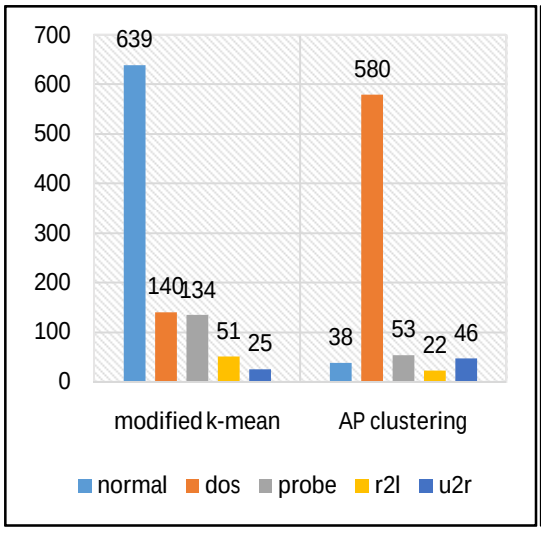

Figure 3(b)

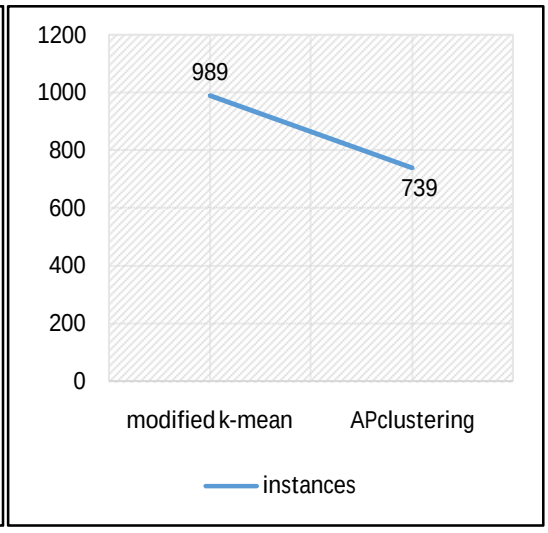

Figure 3(d)

Performance: using NSL dataset

Performance evaluation of the classification algorithm over NSL dataset is shown in TABLE:4.

performance	Multi-class svm	Multilevel svm	Multilevel svm-elm
accuracy	94.42	98.76	94.59
DR	85.94	99.67	98.62
FAR	8.75	3.41	13.69
recall	85.94	99.67	98.62
Precision	86.14	98.88	95.37
F1-score	86.04	99.27	96.82

TABLE:4 performance comparison of classification.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

A key observation from TABLE:4 are as follows:

- Accuracy, detection rate, recall, precision, F1-Score of multilevel SVM classification Achieved best result.
- False alarm rate (FAR) of multilevel classification is less.

Graphical representation of performance of classification is shown in figure 4.

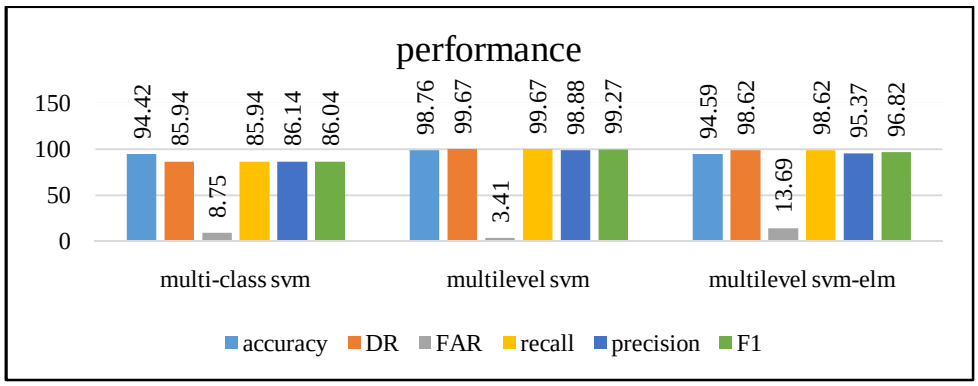

Figure 4: bar graph for performance of classification

Comparison of no. Of instances between frequency normalization and statistical normalization using NSL dataset is shown in TABLE:5.

Category	Nsl dataset	Frequency normalization	Statistical normalization
normal	13449	13449	11913
dos	9235	9233	3988
probe	2289	2288	1267
R2l	208	208	112
U2r	11	11	11
total	25192	25189	17291

TABLE:5 number of instances comparison in frequency and statistical normalization using NSL dataset instances.

A key observation from TABLE 4 are as follows:

- statistical normalization illustrates a less number of instances in each category.
- Total instances in statistical normalization are less.

Graphical representation of normalization instances is shown in figure 5(a) and ROC for representing the normalization instances difference is shown in figure 5(b).

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

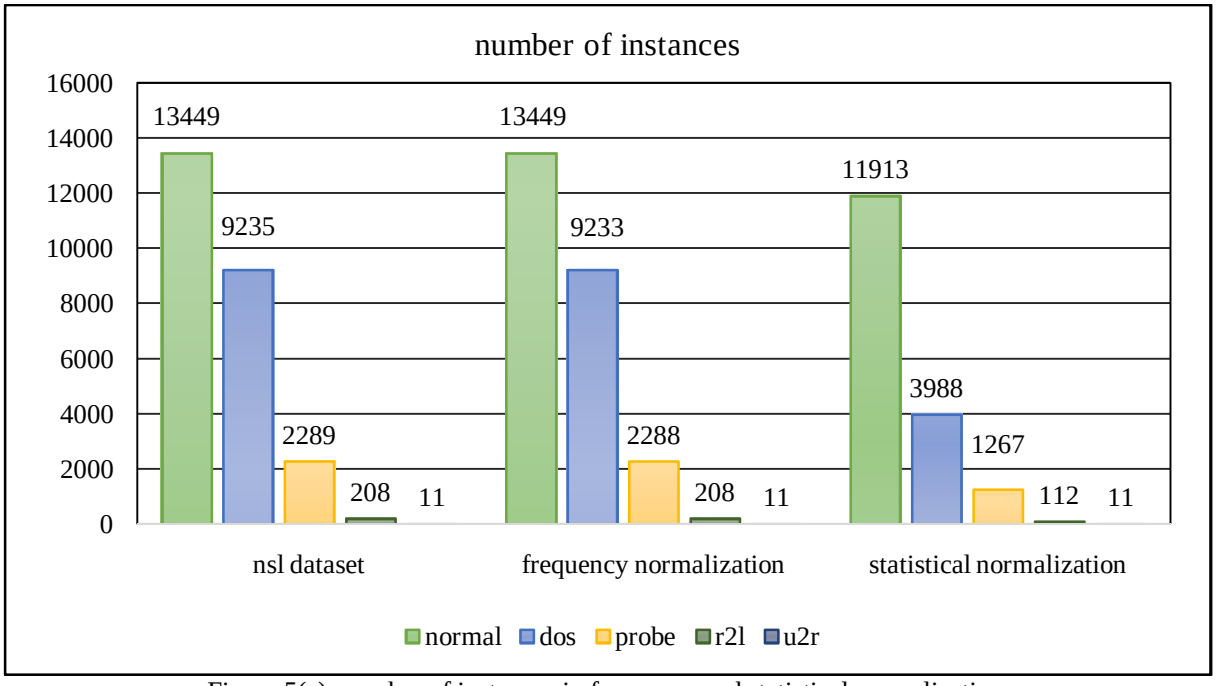

Figure 5(a): number of instances in frequency and statistical normalization.

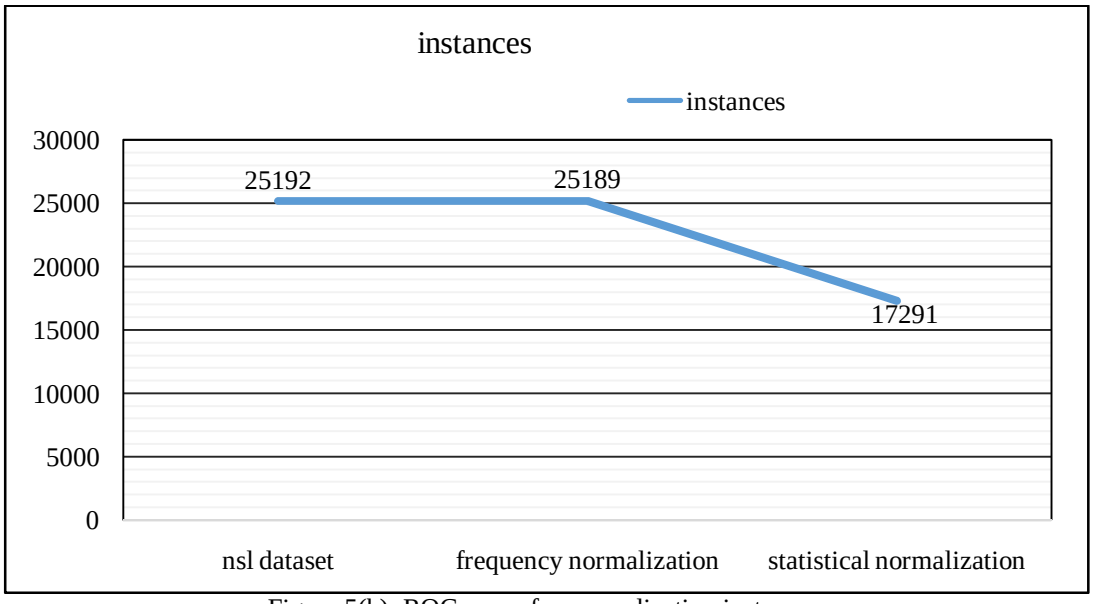

Figure 5(b): ROC curve for normalization instances.

V. CONCLUSION

This research work proposes a multilevel SVM classification intrusion detection system. The proposed model illustrates better performance than Multi-class SVM and Multilevel SVM-ELM models in both datasets. The affinity clustering technique is used to pre-process training dataset and provides high accuracy and detection rate as compared to existing

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

work.The feature selection technique(co-relation technique) used in a proposed model gives the best result. The statistical normalization used in the proposed model to normalize the data gives the best result and improves the detection performance.In general, for the detection with distance based methods such as SVM, statistical normalization is the best choice. According to simulation on KDD-99 dataset, the proposed algorithm achieved approx.. 99.06% accuracy, approx. 100% detection rate, approx. 98.57% precision and approx. 99.28% F1-score.The proposed algorithm achieves approx. 1.32% false alarm rate, which is the lowest among all other models.According to simulation on NSL dataset, the proposed algorithm achieved approx. 98.76% accuracy, approx. 99.67% detection rate, approx. 98.88% precision and approx. 99.27% F1-Score. The lowest false alarm rate of approx. 3.41% is achieved. The proposed system is implemented with the entire training and testing dataset.

In future work the system will be design for the classification of new attacks with enhanced performance with respect to the accuracy, detection rate, the precision, F1-score and false alarm rate.

REFERENCES

1. WathiqLaftah Al-Yaseen, Zulaiha Ali Othman, MohdZakree Ahmad Nazri, "Multilevel Hybrid Support Vector Machine and Extreme Learning Machine Based on Modified K-means for Intrusion Detection System", International Journal in Expert Systems With Applications, Elsevier, 2017.
2. Velmurugan T, DeepalakshmiSuriyamurthi,"Empirical Study Of Feature Selection Methods For High Dimensional Data",Indian Journal of Science and Technology,2016.
3. E.chandraBlessie, E.kosthikeyan,"A Feature Selection Algorithm Using Co-relation Based Method",Journal Of Algorithm And Computational Technology,2012
4. Wu wang, Xiangliangzhang, Sylvain gombault, SveinJ.knapskag,"Attribute Normalization In Network Intrusion Detection",10th international symposium on pervasive system and Network,2009.
5. Preeti Kashyap, Shailendra Kumar Shrivastava, BabitaUjjainiya,"A weighted Seeds Affinity Propagation Clustering For Efficient Document mining",IEEE,2013.
6. Walid Atwa, Kan Li,"Clustering Evolving Data Stream With Affinity Propagation Algorithm",sringer internal publishing switzerland,2014.
7. SumaiyaThaseen Ikram, Ashwani Kumar Cherukuri,"Improving Accuracy Of Intrusion Detection Using PCA And Optimized SVM",Journal Of Computing And Information Technology,2016
8. Hansung Lee, Jiyoung Song, DaiheePark,"Intrusion Detection System Based On Multi-class SVM",Sringer,2005.
9. YadigarImamverdiyev "Anomaly detection in network traffic using extreme learning machine", IEEE, 2016.
10. Athanasios Tsiligkaridis "Anomaly Detection In Transportation Networks Using Machine Learning Techniques", IEEE, 2017.
11. Bhanu Vrat et al "Anomaly Detection in IPv4 and IPv6 Networks Using Machine Learning", IEEE, 2015.
12. Shuai Zhao et al "Real-Time Network Anomaly Detection System Using Machine Learning", IEEE, 2015.
13. Khadija Hanifive Hasan Bank "Network Intrusion Detection Using Machine Learning Anomaly Detection Algorithms" , IEEE, 2016.
14. Prasanta Gogoi, D.K. Bhattacharyya, B. Borah1 and Juga, K. Kalita, "MLH-IDS: A Multi-Level Hybrid Intrusion Detection Method", The Computer Journal, Vol. 57 issue 4, pp. 602-623, 2014.
15. Nutan Farah Haq, MusharratRafni, Abdur Rahman Onik, Faisal Muhammad Shah, Md. Avishek Khan Hridoy and Dewan Md. Farid, " Application of machine Learning Approaches in Intrusion Detection System : A Survey", (IJARAI) International Journal of Advanced Research in Artificial Intelligence, Vol. 4, No. 3, 2015.
16. Shi-JinnHorng, Ming-Yang Su, Yuan-Hsin Chen, Tzong-Wann Kao, Rong-Jian Chen, Jui-Lin Lai, Citra Dwi Perkasa, "A novel intrusion detection system based on hierarchical clustering and support vector machines" Expert Systems with Applications, Elsevier, vol. 38, pp. 306–313, 2011.
17. Garcia-Teodoro, P., "Anomaly-based network intrusion detection: techniques", systems and challenges. Comput. Security vol. 28. issue, pp. 18–28, 2009.
18. Sufyan T Faraj Al-Janabi, Hadeel Amjed Saeed, "A neural network-based anomaly intrusion detection system", IEEE, 2011.
19. J. Ryan, M. Lin, and R. Miikkulainen, "Intrusion Detection with Neural Networks," Conference in Neural Information Processing Systems, 943–949.
20. A. K. Ghosh and A. Schwartzbard, "A Study in Using Neural Networks for Anomaly and Misuse Detection," Conference on USENIX Security Symposium, Volume 8, pp. 12–12, 1999.
21. P. L. Nur, A. N. Zincir-heywood, and M. I. Heywood, "Host-Based Intrusion Detection Using Self-Organizing Maps," in Proceedings of the IEEE International Joint Conference on Neural Networks, pp. 1714–1719, 2002.
22. K. Labib and R. Vemuri, "NSOM: A Real-Time Network-Based Intrusion Detection System Using Self-Organizing Maps," 2000.
23. Sharma, R.K., Kalita, H.K., Issac, B., "Different firewall techniques: a survey", International Conference on Computing, Communication and Networking Technologies (ICCCNT), IEEE, 2014.
24. Meng, Y.-X., "The practice on using machine learning for network anomaly intrusion detection", International Conference on Machine Learning and Cybernetics (ICMLC), vol. 2, IEEE, 2011.
25. SumaiyaThaseen Ikram, Aswani Kumar Cherukuri, "Intrusion detection model using fusion of chi-square feature selection and multi class SVM", Journal of King Saud University –Computer and Information Sciences, 2016.
26. Manjula C. Belavagi and BalachandraMuniyal, "Performance Evaluation of Supervised Machine Learning Algorithms for Intrusion Detection, Procedia Computer Science", Elsevier, 2016.