International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-2, December 2019

Effective Intrusion Detection System by using LOS Classifier

N. Raghavendra Sai, M.Jogendra Kumar, Pathan Hussain Basha , G.Sai Chaitanya Kumar

Abstract: ***With winning advances like catch of Things, Cloud Computing and Social Networking, mammoth proportions of framework traffic associated information area unit made Intrusion Detection System for sort out security suggests the strategy to look at partner unapproved access on framework traffic. For Intrusion Detection System we are going to call attention to with respect to Machine Learning Approaches. it's accomplice rising field of enrolling which can explicitly act with a decent arrangement of less human affiliation. System gains from the data intentionally affirmation and makes perfect objectives. all through this paper we keep an eye on zone unit going to separated styles of Machine Learning pulls in near and had done relative examination in it. inside the last we keep an eye on territory unit going to foreseen the idea of hybrid development, that might be a blend of host principally and framework based for the most part Intrusion Detection System.***

Keywords: Intrusion Detection, Classification, Machine Learning, User Behavior

I. INTRODUCTION

A couple of interference see molecule system zone unit decide based for the most part as a rule that no doubt won't recognize novel ambushes. Additionally, rule based for the most part for the chief half technique is time genuine inclination to the encoded guideline physically and it amazingly place trust in the past information of the outstanding attacks. during this technique, we tend to tend to foreseen framework based for the most part for the preeminent half interference distinguishing proof structure (NIDS) abuse AI methodology. NIDS should be AN instrument or a structure application that screen a framework traffic and event occurring during a} very estimation system. In compose security interference see molecule system accept a genuine activity to separate horribly stunning assortments of attacks.

Revised Manuscript Received on December 30, 2019.
***** Correspondence Author
Dr.N. Raghavendra Sai*, Assoc. Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India. Email: nallagatlaraghavendra@kluniversity.in.
M.Jogendra Kumar, Asst.Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India.Email: jogendra@kluniversity.in.
Dr.Pathan Hussain Basha, Assoc.Professor, Department OF CSE ML Engineering College Email: phussain786@gmail.com
G.Sai Chaitanya Kumar, Assoc.Professor, Department OF CSE Sri Vani Group Of Institutions Email: sai.dmca@gmail.com

The AI technique are adjusted addition the ambush disclosure execution. all through this paper Network interference acknowledgment structure is foreseen with the system of LOS assessment and Support vector machine (SVM). This foreseen technique was taken a stab at KDDCup dataset and ambush revelation exactitude is diverged from choice tree and straightforward mathematician counts. huge improvement of web use in nearness raise the need with respect to the best way to deal with confirm get ready from completely unexpected types of achievable attacks. this answer is deficient to absolutely cautious framework and web application against the perils and vulnerabilities from advanced ambushes. that the framework security could be a huge amount of critical than the standard framework security. There region unit differed security progressions area unit out there to take a gander at the framework mainly based generally structure at any rate there district unit still some on the Q.T. threats. Interference revelation structure is unimaginably coordinated to shape security of electronic framework. regularly interference disclosure structure requesting into two sorts one is Host fundamentally Intrusion Detection System (HIDS) and various one is Network basically based Intrusion Detection System (NIDS).

II. LITERATURE SURVEY

1. Lucas P (Lucas P, 2004) has analyzed the present place of employment of information mining and Bayesian systems in biomedicine and human administrations. Bayesian frameworks and distinctive probabilistic graphical models have started to create as strategies for finding styles in medicinal strength information and in addition as a purpose behind the depiction of the vulnerabilities shrouded clinical fundamental initiative. At the undefined time, frameworks from AI square measure being used to unravel helpful specialty and human administrations issues.

2. Abdelghani Bellaachia and Erhan Guven (AbdelghaniBellaachia&et.al., 2005) present AN examination of the desire for the survivability pace of chest disorder patients using information mining methodology. they need analyzed the exactitude of 3 information mining techniques: Naïve Thomas Bayes, the back-caused neural framework, and the C4.5 elective tree estimations. The delayed consequence of the tests coordinated exhibits that the C4.5 elective tree count is a great deal of genuine than the others.

3. SabbaghA&Darlu P (Sabbagh AN, et.al, 2006) have incontestible that information mining methods, for instance, diverse spatial property decline and neural frameworks, appear as encouraging gadgets to strengthen the office of genotyping tests in hereditary science with an authoritative goal of pre-screening patients for singular treatment assurance with least genotyping work.

4. Mohammed Ambusaidi and Priyadarshi Nanda. (2016) [1] spoke to flexible common information highlight decision procedure that is useful to choose ideal alternatives inside the technique for arrangement. this strategy cut back the system cost and increment the precision.

5. Jiong Zhang et al. (2008) [2] irregular woods recipe is apply in 3 methods like oddity strategy, abuse system, and half and half discovery procedure. during this work the half and half identification system that will build the superior since it consolidating the focal points or qualities of every method peculiarity discovery procedure and abuse location strategy. the most bit of leeway of this work is it construct designs precisely by maintaining a strategic distance from the standard physically.

6. Tooth Yie Leu et al. (2017) [3] this strategy is utilized to discover corporate official vulnerabilities, assaults, dangers at director call guidance viewpoint, during this work expository and information handling systems square measure utilized. the most bit of leeway of this work is prevent the framework from the corporate official dangers, vindictive exercises and assaults adequately.

7. Kriangkrai Limthong ANd Thidarat Tawsook (2012) [4] led a test to check and investigation of connection between interim based alternatives of system traffic by exploitation 2 methodologies of AI: k-closest neighbor and credulous Thomas Bayes. during this work they battle to help arrange executive and analysts to select right equation of AI for their work.

8. S. Yeldi, S. Gupta, R Ingle et al. (2009) [5] portrays the interruption discovery framework by exploitation the ruler protea, this technique fundamentally focus on the honeypots possibilities in a scholastic setting moreover focus on beneficial setting. The ruler protea is staggeringly useful to pull in and redirect to assaulter from their objective or objectives by imitating genuine assistance.

III. PROPOSED METHODOLOGY AND DISCUSSION

The proposed system take input in the form of labeled and unlabeled data containing the threats and vulnerabilities. data per-processing is preprocessed over the training and test data and important features can classify or distinguish into class. Model for classification is trained using SVM classifier. LOS analysis is used to attack recognition, where the intrusion detected on the test data and decision making is performed to take decision for intrusion detection or normal flow of data.

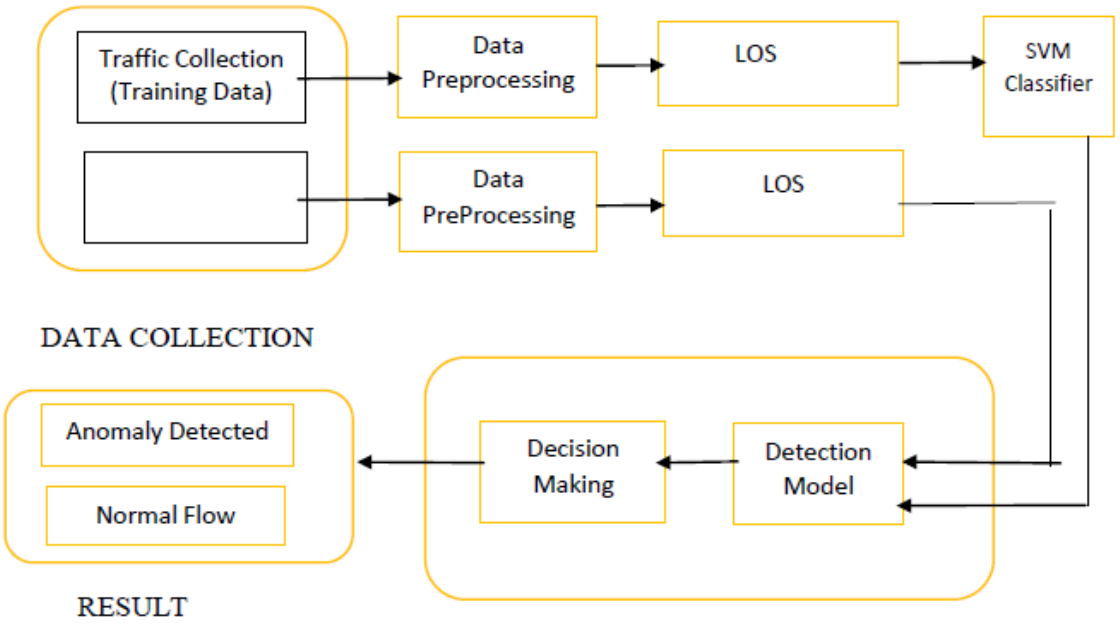


Fig 1: Proposed Methodology

LOS Classifier:-

The proposed calculation of LOS has been executed alongside SVM for contrasting the two. For evaluation of assaults over the system our LOS calculations have been actualized in R-programming and have been tried on KDD CUP 99 dataset.

1) Initialize LR pruned Data as info.
2) Calculate separate capacity in light of preparing estimations of assault for every individual preprocessed information.
3) Selection of individual case.
4) Perform coordinating of a couple of people.
5) Perform remove task.
6) Calculate the target work for the recently made dataset.
7) If it is fulfilled, stop the task.
8) Otherwise, rehash stage 3.
9) Return from KDDCUP 99 dataset that mirrors the properties of assault.

Data Set Used

The KDD cup'99 dataset used in this experiment, dataset 10 percent KDD Cup'99 is used as training dataset and KDD Cup'99 is used as testing data set to detect attack occurred in the network. The whole data set consist of four types of attacks.

Denial of Service(DOS): This is type of cyber attack in which attacker attack on network and try to making network resource inaccessible to its intended users.

Probing: Probing is the attack in which attacker scan networking device or machine to find out the machine's vulnerabilities or weakness.

Root to Local (R2L): Attacker trying to gain access of user authentication of users machine, as attacker does not have any information or account on user machine.

User to Root (U2R): This attack aim to acquire access permission from normal user to gain super user benefits.

Performance Evaluation and Analysis:-

Following the measure have been used for measuring the performances of Proposed IDS

Metrics	Predicted	Predicted
	Normal	Attack
Normal	True negative (TN)	False Positive (FP)
Attack	False negatives (FN)	True positive (TP)

Table 1. Parameters For Performance Estimation

Detection rate (True Positive Rate, TPR): The likelihood of an alarm being raised when malicious activity is observed.

$$TPR=TP/ TP+FN$$

Where, where (TP + FN) is the total number of intrusions

False alarm rate (False Positive Rate, FPR): The probability of an alert being raised when normal activity is observed.

$$TPR=TP/TP+FN$$

Where, (FP + TN) is the total number of normal activities

Sensitivity (also called Recall or True Positive Rate): Sensitivity is that the proportion of actual positives that are properly known as positives by the classifier.

Sensitivity = TP / (TP +FN)

Specificity (also called True Negative Rate): Specificity relates to the classifier's ability to spot negative results. Contemplate the instance of a medical checkwont to determine a definite illness. The specificity of the check is that the proportion of patients that don't to own the illness and can with success check negative for it. In different words:

Specificity: TN / (TN+FP)

Precision: This is a measure for retrieved instances that are relevant. In other words:

Accuracy =Precision: TP/(TP+FP)

F- Measure: - A measure that consolidates precision and recall is the symphonic mean of precision and recall, the traditional F-measure or balanced F-score. F- Measure that mixes precision and recall is the symphonic mean of precision and recall is known as F-measure. This is also prominent as the F1 measure, as a result of recall and precision are evenly weighted. [65]

F-measure = 2*Precision*Recall /Precision+Recall

IV. EXPERIMENTAL RESULTS

In this experiment the accuracy of attack detection is increased by using the LOS and the SVM classifier. Table 2 shows the result table of 10 percent KDD Cup'99 dataset which is used as training dataset and KDD Cup'99 is used as testing data set. result of two data set is mentioned in the form of accuracy, precision, recall and f1-score. The attack detection accuracy of SVM is 91% which is listed in following table 3.

Approaches	Accuracy	Precision	Recall	F1-Score
SVM	0.91	0.88	0.91	0.89
Decision tree	0.89	0.86	0.89	0.87
Naive Bayes	0.82	0.81	0.82	0.81

Table 2. Result Table

The figure 5 shows the performance analysis of NIDS where as comparison of three classification methods such as SVM, decision tree and naive bayes is shown with respect to precision, recall and f1-score. The accuracy of the SVM classifier is 91% which is more than the decision tree and naive bayes classifiers. The comparison of the SVM, decision tree and naive bayes is shown in following graph.

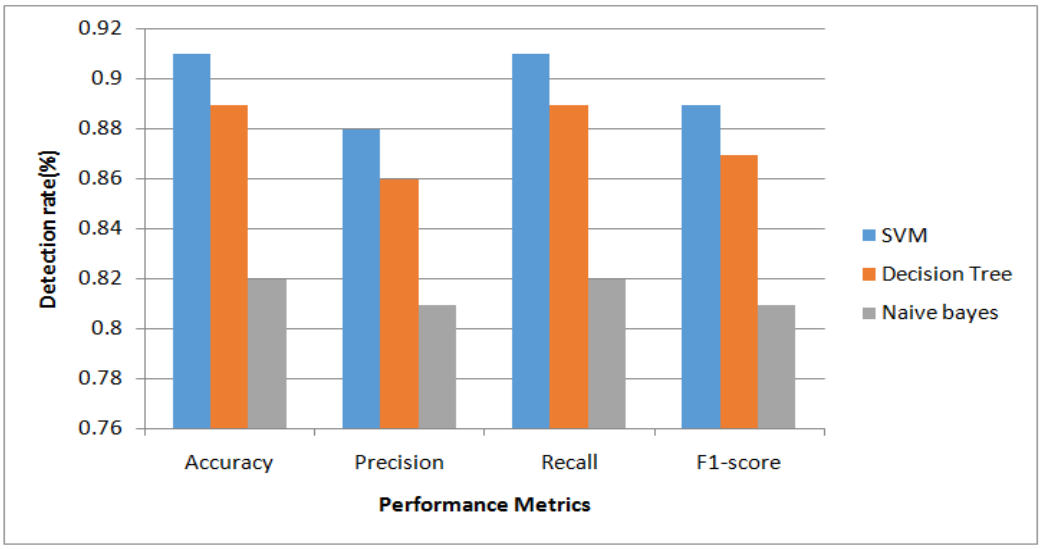


Fig 2: Performance Analysis

The outcome in the wake of playing out a test on the KDD dataset by both the calculation is appeared in the figures beneath. Figure 8.13 demonstrates the Accuracy diagram of the proposed calculation of LOS. It tends to be inferred that the general exactness of the calculation is over 95 percent with the exception of the U2R which relies upon the accessible preparing set from KDD dataset. The calculation is been tried on the 40 percent preparing set of genuine KDD dataset. The motivation to take 40% of the preparation set is the execution time. On the off chance that the proposed calculation will be tried taking all dataset as preparing set outcome will be greatly improved.

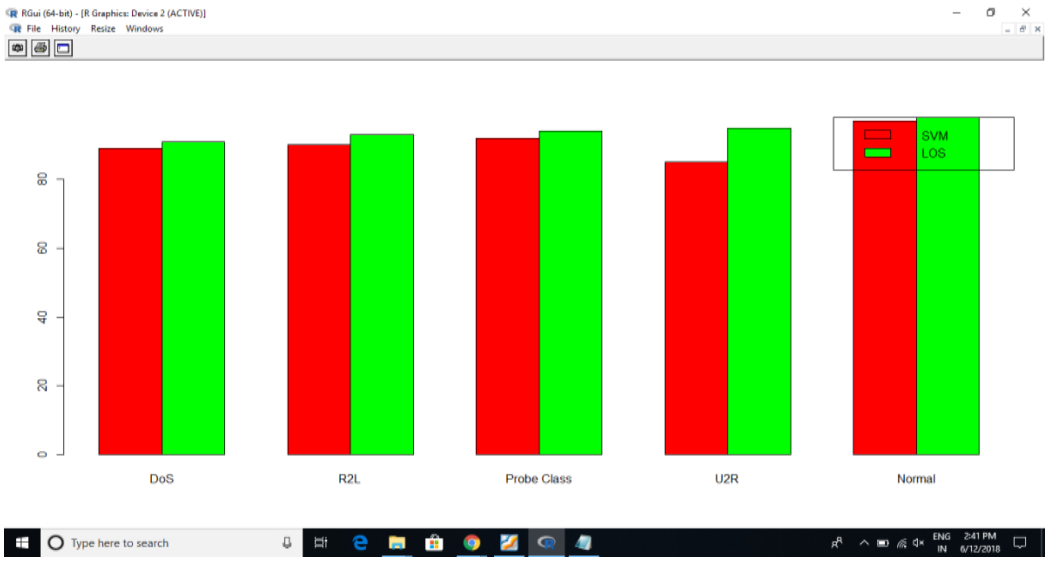
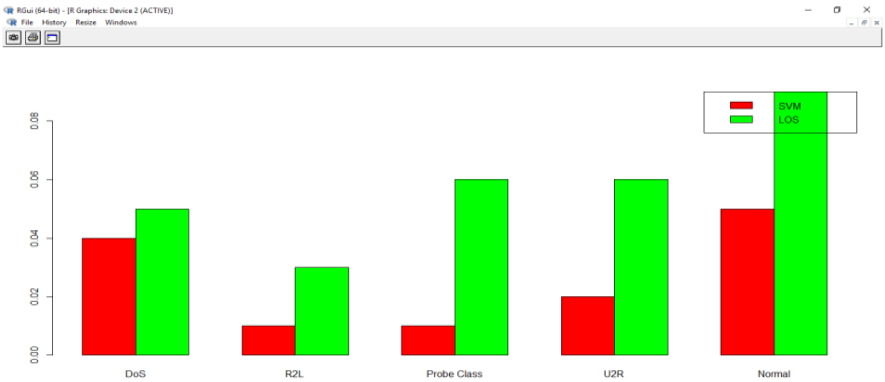


Fig 3: Accuracy Using A) Svm And B) Los

Figure 6 (a) represents the graph of the accuracy of matched labels of classified classes. Figure 6 (b) shows the accuracy of using LOS. Here X-axis represents a different class of attacks and the Y-axis represents the percentage. On X-axis Normal class, Denial of Services (DoS), User-to-Remote (U2R), User to root (R2L) and Probing class has been defined. From above graph, it has been clear that the LOS algorithm is quite capable of clustering the points as accurate as possible as and better than SVM as we can see it is able to classify DOS more accurately. As we can see that most of the classes matched are more than 95% accurate.

Class of Attack	SVM %	LOS %
DOS	89	91
R2L	90	93
Probe Class	92	92
U2R	85	90
Normal	96	97

Accuracy: Table 3. shows the values of Accuracy for SVM- and LOS.



DetectionRate:-

Fig 4 : Detection Rate With A) Svm And B) Los

Figure 7 (a) shows the detection rate of SVM and Figure 7 (b) shows the detection rate of LOS. Table 5 shows the detection rate values of SVM and LOS. The over-all Detection Rate of the proposed algorithm for the given dataset is coming out to be more than 0.95.It shows the detection rate of individual classes of attacks with the proposed algorithm for the given dataset.

In this research the accuracy of attack detection is increased by using the LOS and the SVM classifier. Table 2 shows the result table of 10 percent KDD Cup'99 dataset which is used as training dataset and KDD Cup'99 is used as testing data set. result of two data set is mentioned in the form of accuracy, precision, recall and f1-score. The attack detection accuracy of SVM is 91% which is listed in following table 2.

Detection:

Class of Attack	SVM %	LOS %
DOS	0.04	0.05
R2L	0.02	0.04
Probe Class	0.02	0.06
U2R	0.04	0.08
Normal	0.05	0.09

Table 4 Shows The Values Of Detection For Svm- And Los

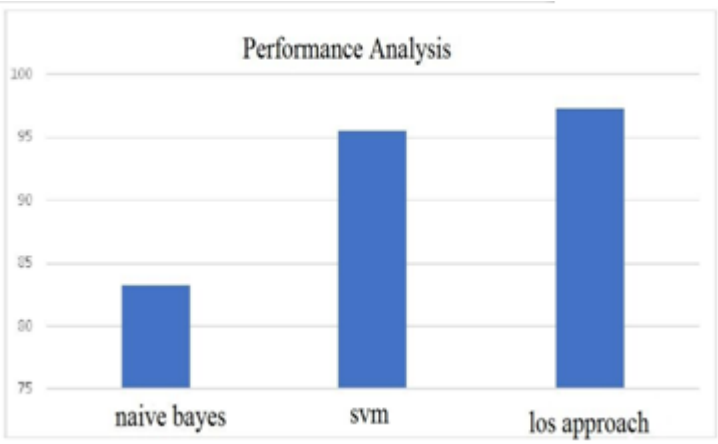


Fig 5: Performance Analysis Of A) Naïve Bayes B) Svm And C) Los

Figure 8 (a) shows the performance analysis of naïve Bayes and Figure 8 (b) shows the performance analysis of SVM Figure 8.(c) shows the performance analysis of LOS. Table 8.15 shows the detection rate values of NaiveBayes, SVM,and LOS. The over-all Performance analysis of the proposed algorithm for the given dataset is coming out to be more than 98%.

Classification Methods	Training Dataset Accuracy (%)	Tested Dataset Accuracy (%)
Navibayes	81.35	83.56
SVM	91.36	95.45
LOS Classifier	98.5	98.63

Table 5 Comparison Of Proposed Los Classifier With Svm And Naïve Bayes

V. CONCLUSION

In this exploration, we propose a cream gathering methodology by consolidating the "close by SVM classifiers" with determined backslide strategies; for instance by a parcel and-vanquish strategy. The computational container neck of the creamer procedure proposed here is still in the SVM. The principal goals of the assessment is growing the distinguishing proof exactness while avoiding the fake positive alert.

From the result it is seen that our proposed philosophy LOS gives more attack disclosure precision when appeared differently in relation to SVM ,Decision tree and guiltless bayes.

The proposed arrangement of LOS (LOGISTIC REGRESSION ONE CLASS SVM) has been applied to the plan of KDDCUP 99 dataset nearby the SVM for the relationship reason. The present assessment indicated raising the display of Network Intrusion structure by the execution of LOS Classifier and in addition to check the ampleness of these classifier appeared differently in relation to the present classifiers and with discard the disfigurements by misuse our new foreseen classifiers. The LOS computation achieved the most vital request precision appeared differently in relation to other chase methodology examined in this work, while the SVM achieved the least false positive rate.

VI. FUTURE STUDY

As a future work, additional analysis to be done on the foremost expert methodology to actualize the planned models during a real system condition. different info mining ways can be investigated, and break away at gathering new info that would be more and more useful assault categories. henceforward Future Study is haunted by utilizing the mixture approach or responsive methodology within the higher than zones to boost the preciseness, discovery within the classifiers during a superior method. It will likewise be planned to feature another model or to boost this model to manage some outstanding condition. Next,future works could examine for any Classifier models to accomplish considerably additional preciseness in discovering the assaults within the system.

REFERENCES

1. Mohammed a. Ambusaidi, Member, Priyadarsi Nanda and Zhiyun Tan, “Building an Intrusion Detection System Using a Filter-BasedFeature Selection Algorithm", IEEE transactions on computers, vol. 65,no. 10,october 2016.
2. Jiong Zhang, Mohammad Zulkemine, and Anwar Haque, “Random-Forest-Based Network Intrusion Detection System", IEEE Transactions In Systems, Man, and Cybernetics, Part C (Applications and Reviews) (Volume: 38, Issue: 5, Sept. 2008).
3. Fang-Yie Leu, Kun-Lin Tsai, Member, IEEE, Yi-Ting Hsiao, and Chao-Tung Yang, “An Internal Intrusion Detection and Protection System by Using Data Mining and Forensic Techniques" , IEEE systems journal, vol. 11, no. 2, june 2017.
4. Kriangkrai Limthong and Thidarat Tawsook, “Network Traffic Anomaly Dtetection using Machine Learning Approach" , IEEE conference nov 2012.
5. Sujata Yeldi, Sweta Gupta, Tanmay Ganacharya, Shirish Doshi, Dhanashree Bahirat, Prof, Rajesh Ingle, Anandamoy Roychowdhary, PICT, “Enhancing Network Intrusion Detection System with Honeypot", TENCON 2003. Conference on Convergent Technologies for the Asia-Pacific Region.
6. Juliette Dromard, Gilles Roudière, and Philippe Owezarski, “Online and Scalable Unsupervised Network Anomaly Detection Method", IEEE transactions on net work and service management, vol. 14, no. 1, march 2017.
7. Shuai Zhao, Mayanka Chandrashekhar, Yugyung Lee, Deep Medhi, "Real-Time Network Anomaly detection System using Machine Learning", IEEE International Conference june 2015.
8. Yogita B. Bhavsar, KalyaniC.Waghmare,” Intrusion Detection System Using Data Mining Technique: Support Vector Machine”, International Journal of Emerging Technology and Advanced Engineering IJETAE 2013.
9. Mostaque Md. Morshedur Hassan,” Network Intrusion Detection System Using Genetic Algorithm and Fuzzy Logic”, International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE2013).
10. PrabhdeepKaur, ShevetaVashisht,”Mingle Intrusion Detection System Using Fuzzy Logic”, International Journal of Engineering and Advanced Technology (IJEAT 2013).
11. S. Devaraju and S. Ramakrishnan,” Performance Comparison for Intrusion Detection System Using Neural Network With KDD Dataset”, ICTACT JOURNAL ON SOFT COMPUTING 2014.

AUTHORS PROFILE:

Dr.N.Raghavendra Sai, has received his Masters Degree from Acharya Nagarjuna University and Ph.D received from Bharathiar University Coimbatore in the Year 2019 . He is dedicated to Teaching field from the past 12 years .He has guided 36 P.G Students and 50 U.G Students .His research area is Data Mining network Security and included Artificial Intelligence .At Present He is working as Assoc.Professor in the Department of Computer Science and Engineering KL University Vaddeswaram . He is highly Passionate and Enthusiastic about his teaching.

Mr. M Jogendra Kumar, He received Masters from Acharya Nagarjuna University. Presently he is working as Assistant Professor in the Department of CSE, KL University, Vijayawada, Andhra Pradesh. He published a good number of papers in national and International journals. His current research interests are Image Processing & Network security

Dr.Pathan Hussain Basha, has his BTechUG) Degree received from Andhra University,M.Tech(P.G) and Ph.D received from Rayalseema University Kurnool in the Year of 2018 .He is dedicated to Teaching field from the past 12 years.He has guided 26 P.G Students and 40 U.G Students .His research area is communication networks and included Artifical Intelligence .At Present He is working as Assoc.Professor and HOD in Malineni Lakshmaiah Engineering College ,Singrayakonda,Prakasam (Dt),AP,India . He is highly Passionate and Enthusiastic about his teaching.

G.Sai Chaitanya Kumar, currently working as Assoc Prof in the Department of CSE in Sri Vani College of Engineering and Technology and also Research scholar in JNTU-Hyderabad. He has 10 years of Teaching Experience. He did his Master Degree M.Tech (CSE) in Acharya Nagarjuna University. He has participated in various International Conferences and workshops held at different places. His area of interest includes ,Image Processing, Data mining ,Information retrieval