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

Fuzzy Controlled Network Intrusion Detection System (FC-NIDS)

Neeraj Kumar, Upendra Kumar

Abstract*: Intrusion Detection System (IDS) is the nearly all imperative constituent of computer network security. IDSs are designed to comprehend intrusion attempts in incoming network traffic shrewdly. It deals with big volume of data containing immaterial and outmoded features, which lead to delay in training as well as testing procedures. Therefore, to minimize the false alarm and computation complexity, the features selection technique for intrusion detection has been implemented. In this paper PCA (Principal Component Analysis) and Fuzzy Inference System (FIS) have been used on kdd99 dataset to develop FC-NIDS model. PCA is used to select the attacked features to minimize the computational work, while FIS is used to develop a fuzzy inference system for accuracy in prophecy using MATLAB. The results of the experiment are tested on UCI data sets as a standard bench-mark. It has been found efficient for true prediction of intrusion as well as to reduce the false alarm rate. The proposed fuzzy logic controller IDS (FC-NIDS), is passable to covenant with signature and anomaly based attacks to get enhanced intrusion detection, decreases false alarm and to optimize complexity.

Keywords: Intrusion, Confusion Matrix, FDR, FLC, IDS, KDD, PCA, Precision, Recall, SVM.

I. INTRODUCTION

Intrusion Detection System (IDS):

Today the era of Information Technology every person is fully rely on exchanging their information using electronic gazette via internet like: business, education, science etc. digitization of word. A system which provides a complete solution in detection of malicious activities occurrence over computer networks as well as early and true intrusion detection. Huge data on Network environment, around the world is filled with uncertainty about hackers and malicious threats. An Intrusion Detection System (IDS) related to rapid detection of malicious network traffic and computer misuse but in fact till we are failed to develop a perfect and honest model. IDS Network is a model of identifying targeted malicious activities in the calculation and giving potential feedback. [1].

- IDSs collect the past experiences from various other sources like systems, computer networks in the form of information, analyze them and IDS get trained to solve further such security problems.
- IDSs serve three vital security functions; monitor, detect and respond to unauthorized activity.
- IDS can also response automatically (in real-time) to a security breach event such as logging off a user, disabling a user account and launching of some scripts.

IDSs have the capabilities to monitor every hardware and software automation in detect the unauthorized access of network resources. IDSs generates alert when it detects security incidents. These alerts prove helpful for network administrator to take appropriate action to prevent the misuse and unauthorized access.

A. Benefits of IDS:

- It provides strong assistance in keeping an eye on activities happening over firewall, routers, servers and various others security issues.
- Network administrator is capable in taking further right action confidently with the help of inputs receives from IDS, so accurate tuning possible security management by providing nice user friendly interface.
- All-inclusive attack come with signature database against which customers can be matched information from the system.
- Generate a details report in the form of data files after identifying about any violation or any kind of changes taken place over network.

An attempt to crash or misuse a system is termed as an ‘intrusion’. An intrusion normally exploits a specific vulnerability, so it has to be detected as early as possible. An intrusion detection system is a mechanism for detecting such intrusions. IDS examine network traffic and Computer systems to detect attacks and abnormal activities then activate security alarm in view of that.

B. Detection Techniques:

Intrusion detection techniques are classified as two major categories: Signature-based Detection, Anomaly-based Detection:

- **Signature-based Detection**:

A signature is a population of pattern or string that matches a known attack or threat. Compare pattern alongside captured events is the process for recognizing potential intrusion. Due to the use of knowledge collected by specific attacks and system susceptibility, this technique is also known as knowledge-based intrusion detection or Misuse Detection [2].

Revised Manuscript Received on December 30, 2019.
* Correspondence Author
Neeraj Kumar*, Research Scholar, Computer Science & Engineering, Birla Institute of Technology, Mesra, India. javaneeraj@gmail.com
Dr. Upendra Kumar, Department of Computer Science & Engineering, Birla Institute of Technology, Mesra, India. upendrakr@bitmesra.ac.in

Advantages:

Misuse detector is very efficient in detecting attacks without prompting false alarms Misuse detectors can be identifying specifically to specially designed intrusion tools and techniques. Misuse detector systems provide an easy-to-use tool for administrators to monitor their systems, even if they're not security specialists, it is the simplest and effective way in detecting known attacks [2, 3].

Disadvantages:

Misuse detectors can be used to identify only that kind of intrusion which is previously known attacks. Due to this reason the system should has to be updated with the occurrence of every newly attack type signatures. Misuse detectors are intended to detect attacks that have signatures introduced to the system only. When a eminent attack turns slightly and receives a version of that attack, then the detector is unable to be recognizable this variant of the same attack and ineffective to detect unknown attacks and variants of known attacks. It is hard to keep signatures up to date. It's time consuming based techniques [2, 3]

- **Anomaly-based Detection:**

Deviation from "normal" or "expected behavior" may be termed as trained data-set: "which are periodically collected from close monitoring the frequent activities, host, network or pipe, are comes under anomaly. Normal or expected behavior profile can be either static or dynamic. Such profiles are for many attributes, like: too many attempt for failed logging, duration of services, utilization of processor, protocol uses etc. Often it is also known as Behavior-based Detection. Example of attack comes under this type are like: attempted break-in, penetration by genuine user, Trojan horse, Denial-of-Service (DOS) etc. [2, 3].

Advantages:

In comparison to signature based, anomaly based found more effective, particularly in detection of such attacks which are not listed into known attack data-set i.e. detection of new attacks. Potency of this technique is not only to predict new type of attack for further assistance for the administrator but also with this method we can also capable to detect signature based. Anomaly based method can be applicable in prediction of signature information used by misuse-based IDS [2, 3].

Disadvantages:

In this technique it is very difficult to fire alarm in time due to changing of all parameters profile frequently. The properties of all concern data get changed at the time of rebuilding of profile. Hence disadvantage of this method is getting many false alarms due to lack of user and network behavior information. So this approach requires a bulky training data sample which consists of system event log in order to build normal behavior profile [2, 3].

Both above detection techniques are complementary to each other, because in both of the methods first they focus to detect for the known type attack then they proceed further for unknown attacks.

B. Related Work:

Various domains in which authors have been proposed to detect intrusion. Most of the authors worked onto the classification, multidimensional reduction of data, data filtering & noise reduction (redundant, duplicate and irrelevant), Detection Time, minimization of false detection alarm.

These are achieved by applying various techniques like: Data Mining: Outlier detection [13-16] for classification, association rules for data filtering [18]. At the same time decision tree [18] for rule generation were extremely successful in discovering known attacks. Data Mining found violation of the privacy of the data, complex computation on data [11, 13]. K-Map is useful for distinction of the data but it is not supported for large volume of data. Support Vector Machine (SVM) [19-23] [25-32]: is beneficial for classification and reformation of data, but at the same time SVM only cannot powerfully in recognizing about new data. Given a set of trained-data-set which require more computational time for big-data [4,12]. In deep learning or Machine Learning are more exhilarating to detect unknown attacks like in neural Network it can achieve through incorporating the Self Organizing Map (SOM) [5-10].

Based on the review conducted, some of the findings are listed below:

- Human Role in the Operation of IDS is still dependent particularly in case of new kind of attack identification. Security expert need to monitor new attack detection techniques and formulate new rules and finally the security analyst after processing huge number of alert manually on regular basis.

  Fuzzy Logic based techniques found best in detection of intrusion for as compared to Data Mining, K-Map [21], MLP [25-30], Random Forest [25-30], SVM [32,33], Neural Network [31-33] and dimensionality reduction [35-37].

C. Dataset as benchmark:

KDD_99 is an intrusion dataset, which is managed through MIT Lincoln Labs. This dataset we had used as benchmark. They had provided KDD into three categories: ‘‘Corrected Dataset’’, ‘‘Full Dataset’’ and ‘‘10 % Dataset”. Every dataset category consists of 41 features (in which 7 are discrete and 34 are continuous) with millions of sample events as records. There are out of 38 types of attacks with 24 in training set and 14 in test set but they are classified into various classes of type: DoS, R2L, Probe, U2R and Normal.

II. PROPOSED FEATURE REDUCTION METHOD

Considered various issues to fix the research gaps are like: Amount of alert generated, Accuracy of detection, Performance Time factor, Future Attacks and Human Intervention.

*“**In creation of an efficient and accurate attack predictive model feature selection plays a vital role for the same**”.*

Feature selection turns into important when the numbers of features are too large. There is no need to use every feature in achieving our desired result. Hence our proposed IDS model chooses only those features that are really important. Subset of feature selection method is chosen in our proposed work and it applied into proposed model. As a result, it found better results than complete set of feature samples.

Before modeling NIDS analysis done on dataset and following observations found in the form of benefits of feature selection:

After analysis of kdd99 data set, observed that into 10% _dataset and full_dataset out of eighteen attacks, records related to eleven types of attacks. So decide to do uses corrected kdd 99 dataset for training and rule generation, where as other two kdd99 datasets 10% dataset and full_dataset usage for testing purpose.

- **Benefits of feature reduction:**

Data Reduction:

By selecting only useful features, we can avoid meaningless calculations on the useless features.

Reduces Over Fitting:

It trims down the over fitting situation. Avoid redundant data means there will no situation less opportunity to make decisions based on noise.

Complexity:

It reduces the classification as well as prediction time.

Improves Accuracy:

It improves the accuracy in detection of intrusion because it has less non relevant data.

Reduces Training Time:

When there is a non-redundant and non-noisy dataset then obviously training as well as testing time will reduce. Resultant reduction in getting false alarm and less computational work need to do.

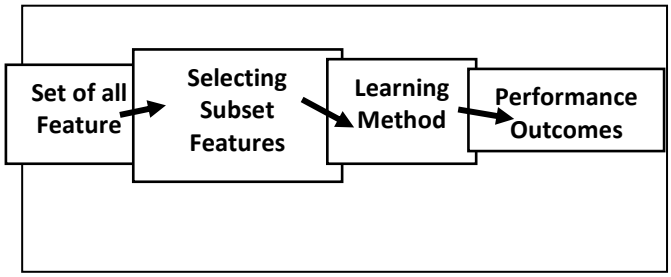

Fig. 1. Filter Method for feature selection model

- **For feature extraction find useless features by following feature engineering**

Step1: Data Cleaning:

Avoid duplication and redundancy from dataset.

Step2: Scaling of Dataset:

We do the normalization for every feature of dataset so that every feature in the dataset has similar difference in the results. Normalization is taking place using Max_Min formula.

$$Value_n = \frac{\text{value} - \text{Min (value)}}{\text{Max(value)} - \text{Min (value)}} \quad (1)$$

Step3: PCA:

Orthogonal dimensionality reduction is taking place in case of PCA. Scaling and transformation need to be done to switch a group of dataset under annotations of possibly interrelated variables into a group of dataset with their values of linearly not correlated variables. Each principal component, need to inspect the magnitude as well as direction of coefficients of the original data variables. Obtaining the larger absolute coefficient value means it is an important feature and it will have impact on result. With the help of such analysis method we will able to get the linear grouping of a set of variables which have highest variance and remove its effect. So repeat this successively.

The traditional approach to PCA is to perform the Eigen decomposition on the covariance matrix Σ. That is matrix of d x d . Covariance between two features of the given dataset important and it is calculating for every elements of covariance matrix.

Calculation of Covariance between two features:

$$\sigma_{jk} = \frac{i}{n-1}(x_{ij} - \acute{x}_j)(x_{ik} - \acute{x}_k) \quad (2)$$

Through the following matrix equation, we can summarize:

Covariance Matrix:

$$\sum \frac{1}{n-1}\left((X - \acute{x})^T (X - \acute{x})\right) \quad (3)$$

Where $\acute{x}$ the mean vector is d-dimensional vector where each value in this vector represents the sample mean of a feature column

$$\acute{x} = \sum_{k=1}^{n} x_i \quad (4)$$

By applying PCA, found best result on selected 30 features of corrected_kdd _dataset including 03 discrete feature viz. protocol_type, service and flag.

III. PROPOSED FUZZY CONTROLLED NETWORK INTRUSION DETECTION SYSTEM (FC-NIDS)

A model proposed which is shown in Fig. 2 below:

- **Multi Class-Classification Model (FC-NIDS):**

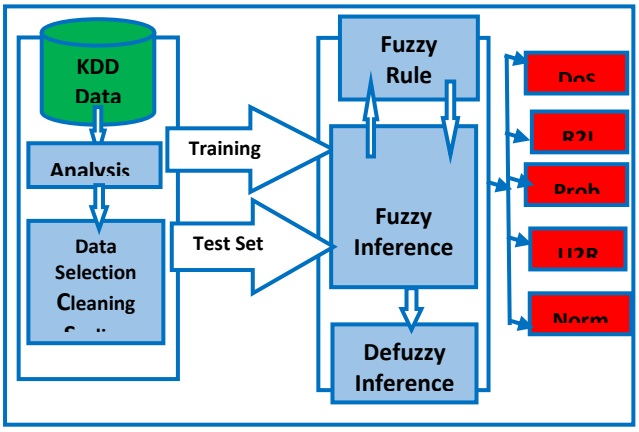

Fig. 2. Multi Class-Classification Model (FC-NIDS)

In the proposed model normalization of the dataset using Min_Max formula & equation (1). Then selection of feature done through PCA (equation (2) and (3) etc.). Train the FIS as per supervised dataset available for true and accurate prediction. Further testing can be done by using FLC (Fuzzy Logic Controller) by applying Fuzzy Rule as per the Membership for each feature.

First we did binary classification by normal and anomaly. This is further classified into various into categories like DoS, R2L, Probe, U2R and Normal. After that it will further classified into sub categories like:

Attacks fall into four main categories:

- DOS: denial-of-service, e.g. sync flood etc.
- R2L: unauthorized access from a remote machine, e.g. guessing password etc.
- U2R: illegal access to local super user i.e. root and its privileges for example buffer_overflow, loadmodule, perl, rootkit etc.
- Probe: inspection and other probing surveillance, e.g., ipsweep, nmap, portsweep, satan, mscan, saint, portscanning etc.

It is important to note that the test data is not from the same probability distribution as the training data, and it includes specific attack types not in the training data. This makes the task more practical. Some intrusion experts consider that nearly all original attacks are variants of known attacks and the "signature" of identified attacks can be sufficient to catch original variants. The datasets contain a total of 24 training attack types, with an additional 14 types in the test data only [19].

Table I: Type of Classes and their corresponding Attacks

Sr. No.	Attack Class (Categories)	Attack Type (Sub Categories)
1	**DoS**	back, land, neptune, pod,smurf, teardrop
2	**R2L**	ftp_write, guess_passwd, imap, multihop, phf, spy, warezclient, warezmaster
3	**Probe**	ipsweep, nmap, portsweep,satan, mscan, saint
4	**U2R**	buffer_overflow, loadmodule, perl, rootkit
5	**Normal**	

A. Fuzzy Logic Controller:

Controller designed to solve the especially under unclear situation. On the basis of past data from data warehouse provide 80% uses for provide the training and rest have used for testing.

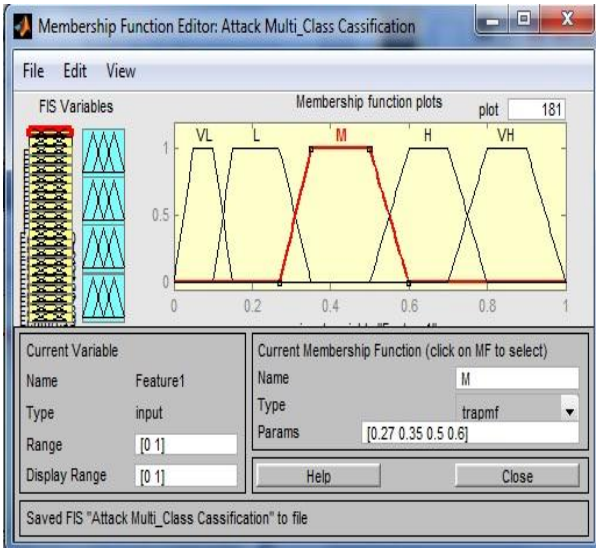

Fig. 3. Membership Function of each Input Features:

In this propose work multi-dimensional feature reduction done through PCA, then split the corrected_kdd_dataset into two halves. First half is used to train the dataset and second half is use for testing. Initially there were will be 311029 records, when we performed PCA, records reduce to 72291 after scaling. 35% of Scaled_corrected_kdd_dataset of i.e 25507 records are used for testing.

For each selected feature as input variables considering VL, L, M, H and VH as linguistic term membership function is there and for every input Fuzzification is taken place to find out corresponding degree of truthiness using Trapezoidal Method. Total 30 features selected included three discrete features are used for classification.

$$\boldsymbol{Trapezoidalrule(x:m,n,p,q)} = \boldsymbol{max}\left(\boldsymbol{min}\left(\frac{x-m}{n-m}, 1, \frac{q-x}{q-p}\right), 0\right)$$

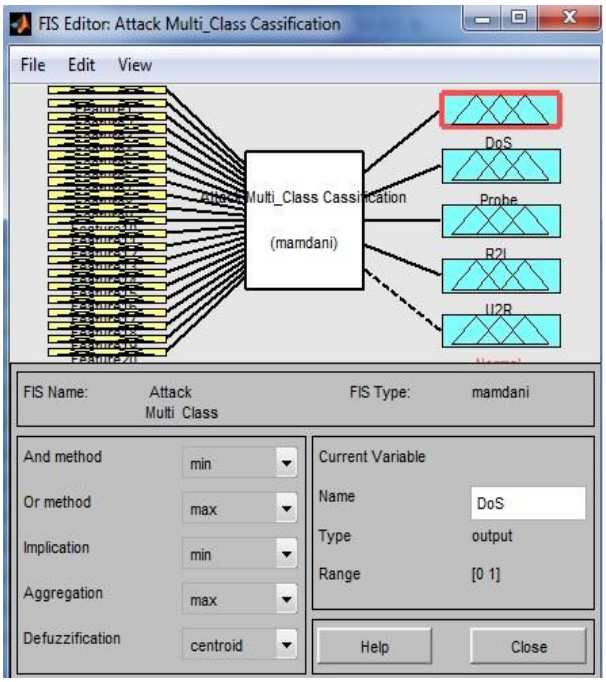

Fig. 4. Fuzzy Inference System

Using Fuzzy Rule and FRB (Fuzzy rule base) dataset are first classified into Normal and Attack. If attack then what category types, further sub-categories

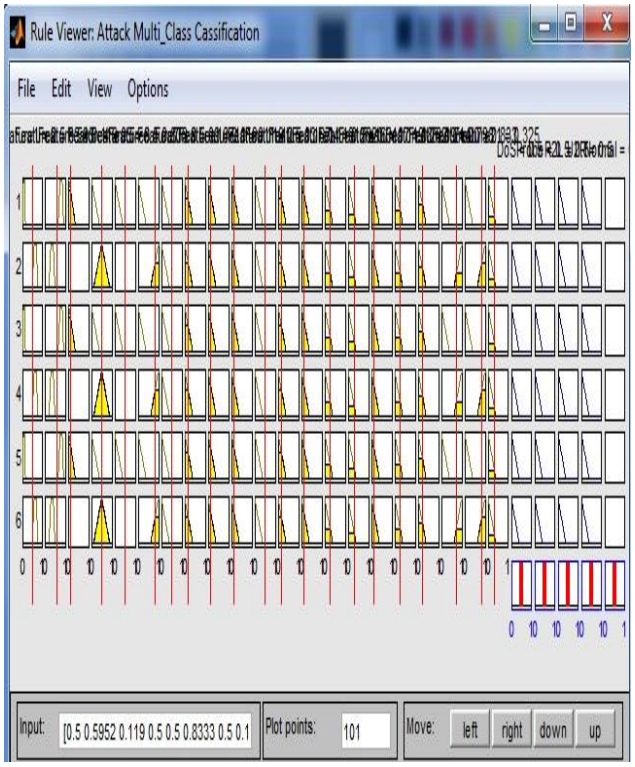

Fig. 5. Fuzzy Rule Base (FRB)

IV. EXPERIMENTAL DETAIL AND RESULT ANALYSIS

Experiment done on Intel Core i3 Processor 4 GB DDR4 RAM, Windows operating System.

A. Binary Classification:

Tab1e II: Classification & assignment time and accuracy in Attack Detection

	10 percent dataset scaled	Full dataset scaled	Corrected dataset
Classified (seconds)	0.656	0.604	0.807
Assigned Labels (Seconds)	0.182	1.457	0.12
Accuracy (Percent)	92.366	75.488	99.896

Experiment result shown into Table II accuracy rate i.e. DR (detection rate) to detect normal and attack is up to **99.896%** in case of Binary classification.

Confusion matrix:

	Class 1 Normal	Class 2 Attack
Class 1 Normal	T_P	F_N
Class 2 Attack	F_P	T_N

Class1: -Positive
Class2: -Negative

Table III: Confusion Matrix

Type\ Dataset	10percent dataset		Full Dataset (Scaled)		Corrected Dataset	
	Normal	Attack	Normal	Attack	Normal	Attack
Normal	55992	1762	258820	3358	29331	47
Attack	9352	78480	260139	552675	33	47880

Definition of the Terms:

- Positive (P) : Observation is positive
- Negative (N): Observation is not positive
- True Positive (T_P): Observation is positive, and is predicted to be positive.
- False Negative (F_N): Observation is positive, but is predicted negative.
- True Negative (T_N): Observation is negative, and is predicted to be negative.
- False Positive (F_P): Observation is negative, but is predicted positive.

Precision:

To get the value of precision we divide the total number of correctly classified positive examples by the total number of predicted positive examples. High Precision indicates an example labeled as positive is indeed positive (small number of F_P).

Precision is given by the relation:

$$\boldsymbol{Precision} = \frac{T_P}{T_P+F_P} \qquad (5)$$

The precision is intuitively the ability of the classifier not to label as positive a sample that is negative

Recall:

Recall can be defined as the ratio of the total number of correctly classified positive examples divide to the total number of positive examples. High Recall indicates the class is correctly recognized (small number of F_N).

$$\boldsymbol{Recall} = \frac{T_P}{T_P+F_N} \qquad (6)$$

The recall is intuitively the ability of the classifier to find all the positive samples.

Accuracy Ratio:

Ratio between total numbers (true positive and true negative) of instances detected by the system to the total number of instances present in the dataset

$$\boldsymbol{AR} = \frac{T_P+T_N}{T_P+T_N+F_P+F_N} \times 100 \qquad (7)$$

Table IV: Testing 10 percent, full and corrected kdd dataset

		Precision	Recall	F1-Score	Support	FDR
10 percent Dataset (Scaled)	Attack	0.86	0.97	0.91	57754	0.11
	Normal	0.98	0.89	0.93	87832	0.03
	Avg/ Total	0.93	0.92	0.92	145586	0.07
Full Dataset (Scaled)	Attack	0.5	0.99	0.66	262178	0.32
	Normal	0.99	0.68	0.81	812814	0.01
	Avg/ Total	0.87	0.75	0.77	1074992	0.16
Corrected Dataset	Attack	1	1	1	29738	0
	Normal	1	1	1	47913	0
	Avg/ Total	1	1	1	77291	0

Here result into the Table IV shows that less in FDR getting very less means getting less false alarm. Precision and Recall getting high means maximum attack truly detected. Means normal_data as normal & attack_data as attack detected.

B. Multi Class Classifications:

Table V: Test Result on Kdd_Corrected_test_data_set

Attack Sub class	Attack Present	Detected	DR %	Precision	Recall
apache2.	273	271	99	0.99	1
back.	133	133	100	1	0.99
buffer_overflow.	7	4	50	0.5	0.57
ftp_write.	0	0	0	0	0
guess_passwd.	450	441	98	0.98	1
httptunnel.	47	47	100	1	0.85
imap.	0	0	0	0	0
ipsweep.	48	46	94	0.94	0.96
land.	1	1	100	1	1
loadmodule.	1	0	0	0	0
mailbomb.	106	106	100	1	1
mscan.	357	347	97	0.97	0.97
multihop.	5	0	0	0	0
named.	8	6	67	0.67	0.5
neptune.	6759	6759	100	1	1
nmap.	26	26	100	1	1
normal.	15702	15702	100	1	1
perl.	1	0	0	0	0
phf.	15	14	93	0.93	0.87
pod.	53	51	96	0.96	0.92
portsweep.	250	250	100	1	1
processtable.	5	2	40	0.4	0.4
ps.	6	0	0	0	0
rootkit.	122	104	85	0.85	0.89
saint.	287	273	95	0.95	0.96
satan.	6	4	62	0.62	0.83
sendmail.	320	320	100	1	0.99
smurf.	55	25	45	0.45	0.42
snmpgetattack.	117	117	100	1	0.99
snmpguess.	3	0	0	0	0
sqlattack.	0	0	0	0	0
teardrop.	334	321	96	0.96	0.98
udpstorm.	1	1	100	1	1
warezmaster.	4	0	0	0	0
worm.	0	0	0	0	0
xlock.	5	5	100	1	0.4
Total/Average	25507	25371	99.467		

False positive rate:

Ratio between total numbers of normal instances classified as attack and the total number of normal instances and misclassified as attack instances.

$$FPR = \frac{F_P}{F_P + T_N}$$

F1-Score:

Since we have two measures (Precision and Recall) it helps to have a measurement that represents both of them. We calculate an F-score which uses Harmonic Mean in place of Arithmetic Mean as it punishes the extreme values more.

The F-score will always be nearer to the smaller value of Precision or Recall.

$$F1 - Score = 2 \times (precision \times recall)/(precision + recall)$$

Experiment result as per our proposed work on to the corrected_data_set_scaled as shown into the Table 5 above. Above Table shows out of 25507 (including 38 types of attacks and normal) records detected 25371 correctly.

DR **(Detection Ration)** = **99.467%**

Hence the resultant found an enhanced detection ration 99.467% improved the accuracy of prediction about the attack.

Table VI: Comparison with Binary and Proposed Multi Classifications

Classification Type	*Binary Classification*			*Multi Class Classifications*
Evaluation Parameter\ Kdd dataset	10 percent dataset scaled	full dataset scaled	Corrected dataset	Proposed FC-IDS (corrected data set)
Classified (in sec)	0.656	0.604	0.807	0.516
Assigned Labels (in sec)	0.182	1.457	0.12	0.031
Accuracy (in percent)	92.37	75.49	99.89	99.467

A. Comparison with other Multi-classifiers:

Table VII: Precision (in %) on various classifiers:

Category/ Classifier	*FC-NIDS*	*MLP*	*Random Forest*	*SVM*
Normal	99.50	98.00	97.50	83.00
Dos	99.00	97.00	98.00	98.50
Probe	39.00	30.00	22.00	28.00
R2L	7.00	4.00	19.00	10.00
U2R	4.00	8.00	0.00	8.00

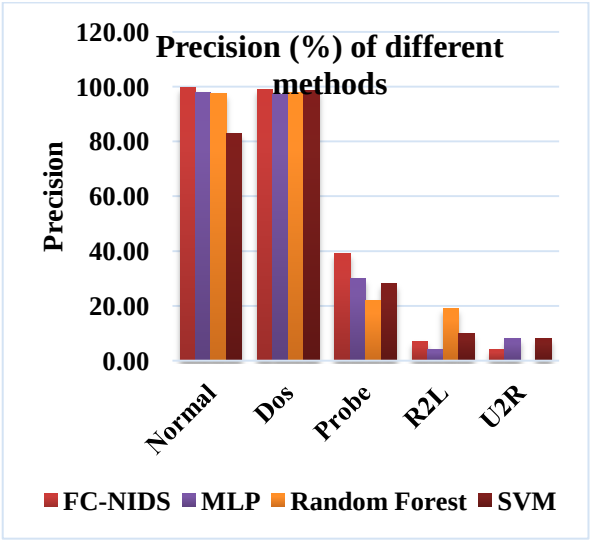

Fig. 6. Precision (%) of different classification methods

Table VIII: Recall of various classifiers for classes

Category/ Classifier	*FC-NIDS*	*MLP*	*Random Forest*	*SVM*
Normal	81.00	95.00	94.00	83.00
Dos	98.00	99.00	99.00	97.00
Probe	75.00	73.00	80.00	83.00
R2L	80.00	23.00	76.00	77.00
U2R	6.00	35.00	0.00	0.00

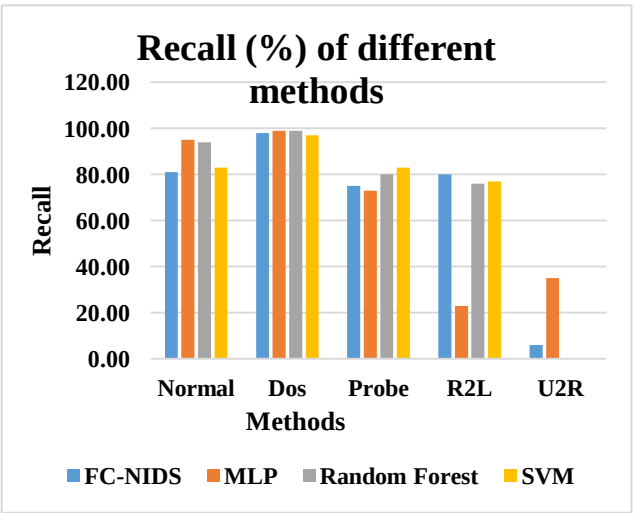

Fig. 7. Recall (%) of different classification methods

Table VII and Table VIII and by their corresponding graphical representation in Fig. 6 and Fig. 7 shows comparison among various multi-classifies like Multi-Layer Perception (MPL), RF (Random Forest) and SVM (Support Vector Machine) with our proposed work FC-NIDS method as per working model Figure 2. To detect all 38 types of attacks, experiment done of corrected test dataset and found that improved in particularly Normal, DoS and Probe in respect to precision and found good in DoS Probe, R2L and U2R in respect to recall.

V. CONCLUSION AND FUTURE WORKS

In this paper, an evolutionary feature selection using fuzzy inference approach for intrusion detection have been discussed. In respect to minimize colossal computational work along with time complexity, PCA had been used for prediction of IDS. A fuzzy decision-making engine has been developed to make the system more authoritative for the detection of attacks, using the fuzzy inference approach. High recall has directed classes were correctly recognized with minimum number of F_N, high precision indicates labeled as positive with minimum number of F_P, and FDR is less in prediction. Hence, through experiment result found that the proposed model is effective towards various intrusion detection in computer networks with enhanced in accuracy, decreased false alarm and optimized time complexity for classification. Further research makes an effort in dimensionality reduction in which feature selection required to be done in such a manner that for every types. Moreover, the effort will be continued to study about fitness of every features to optimized attack feature selection using genetic algorithm to influence more parameters of the fuzzy inference module.

REFERENCES

1. Aydın, M. A., Zaim, A. H., & Ceylan, K. G. (2009). A hybrid intrusion detection system design for computer network security. Computers & Electrical Engineering, 35(3), 517-526, Elsevier.
2. Modi, C., Patel, D., Borisaniya, B., Patel, H., Patel, A., & Rajarajan, M. (2013). A survey of intrusion detection techniques in cloud. Journal of Network and Computer Applications, 36(1), 42-57, Elsevier.
3. Butun, I., Morgera, S. D., & Sankar, R. (2014). A survey of intrusion detection systems in wireless sensor networks. IEEE communications surveys & tutorials, 16(1), 266-282.
4. Dastanpour, A., & Mahmood, R. A. R. (2013). Feature selection based on genetic algorithm and SupportVector Mmachine for intrusion detection system. In The Second International Conference on Informatics Engineering & Information Science (ICIEIS2013) (pp. 169-181). The Society of Digital Information and Wireless Communication.
5. Pachghare, V. K., Kulkarni, P., & Nikam, D. M. (2009, July). Intrusion detection system using self-organizing maps. In Intelligent Agent & Multi-Agent Systems, 2009. IAMA 2009. International Conference on (pp. 1-5). IEEE.
6. Shah, B., & Trivedi, B. H. (2012). Artificial neural network based intrusion detection system: A survey. International Journal of Computer Applications, 39(6), 13-18.
7. Ahrabi, A. A. A., Navin, A. H., Bahrbegi, H., Mirnia, M. K., Bahrbegi, M., Safarzadeh, E., &Ebrahimi, A. (2010). A New System for Clustering and Classification of Intrusion Detection System Alerts Using Self-Organizing Maps. International Journal of Computer Science and Security (IJCSS), 4(6), 589-597.
8. Landress, A. D. (2016, March). A hybrid approach to reducing the false positive rate in unsupervised machine learning intrusion detection. In Southeast Con, 2016 (pp. 1-6). IEEE.
9. Sommer, R., & Paxson, V. (2010, May). Outside the closed world: On using machine learning for network intrusion detection. In Security and Privacy (SP), 2010 IEEE Symposium on (pp. 305-316). IEEE.
10. Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials, 18(2), 1153-1176.
11. Hodge, V., & Austin, J. (2004). A survey of outlier detection methodologies. Artificial intelligence review, 22(2), 85-126.
12. Lazarevic, A., Kumar, V., & Srivastava, J. (2005). Intrusion detection: A survey. In Managing Cyber Threats (pp. 19-78). Springer, Boston, MA.
13. Branch, J. W., Giannella, C., Szymanski, B., Wolff, R., & Kargupta, H. (2013). In-network outlier detection in wireless sensor networks. Knowledge and information systems, 34(1), 23-54.

14. Liu, B., Xiao, Y., Cao, L., Hao, Z., & Deng, F. (2013). SVDD-based outlier detection on uncertain data. Knowledge and information systems, 34(3), 597-618.
15. Aggarwal, C. C. (2015). Outlier analysis. In Data mining (pp. 237-263). Springer, Cham.
16. Kim, G., Lee, S., & Kim, S. (2014). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Systems with Applications, 41(4), 1690-1700 Elsevier.
17. Nahar, J., Imam, T., Tickle, K. S., & Chen, Y. P. P. (2013). Association rule mining to detect factors which contribute to heart disease in males and females. Expert Systems with Applications, 40(4), 1086-1093.
18. Lin, S. W., Ying, K. C., Lee, C. Y., & Lee, Z. J. (2012). An intelligent algorithm with feature selection and decision rules applied to anomaly intrusion detection. Applied Soft Computing, 12(10), 3285-3290.
19. Stolfo, J., Fan, W., Lee, W., Prodromidis, A., & Chan, P. K. (2000). Cost-based modeling and evaluation for data mining with application to fraud and intrusion detection. Results from the JAM Project by Salvatore, 1-15.
20. Ashfaq, R. A. R., Wang, X. Z., Huang, J. Z., Abbas, H., & He, Y. L. (2017). Fuzziness based semi-supervised learning approach for intrusion detection system. Information Sciences, 378, 484-497.Elsevier.
21. Ambusaidi, M. A., He, X., Nanda, P., & Tan, Z. (2016). Building an intrusion detection system using a filter-based feature selection algorithm. IEEE transactions on computers, 65(10), 2986-2998.
22. Kim, G., Lee, S., & Kim, S. (2014). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Systems with Applications, 41(4), 1690-1700.Elsevier.
23. Nayak, J., Naik, B., & Behera, H. S. (2015). Fuzzy C-means (FCM) clustering algorithm: a decade review from 2000 to 2014. In Computational intelligence in data mining-volume 2 (pp. 133-149). Springer, New Delhi.
24. Varma, P. R. K., Kumari, V. V., & Kumar, S. S. (2018). A Survey of Feature Selection Techniques in Intrusion Detection System: A Soft Computing Perspective. In Progress in Computing, Analytics and Networking (pp. 785-793). Springer, Singapore.
25. Elhag, S., Fernández, A., Altalhi, A., Alshomrani, S., & Herrera, F. (2017). A multi-objective evolutionary fuzzy system to obtain a broad and accurate set of solutions in intrusion detection systems. Soft Computing, 1-16.
26. Amato, F., Cozzolino, G., Mazzeo, A., & Vivenzio, E. (2017, June). Using Multilayer Perceptron in Computer Security to Improve Intrusion Detection. In International Conference on Intelligent Interactive Multimedia Systems and Services (pp. 210-219). Springer, Cham.
27. Chandrashekhar, A. M., & Raghuveer, K. (2014). Amalgamation of K-means clustering algorithm with standard MLP and SVM based neural networks to implement network intrusion detection system. In Advanced Computing, Networking and Informatics-Volume 2 (pp. 273-283). Springer, Cham.
28. Sheikhan, M., & Jadidi, Z. (2014). Flow-based anomaly detection in high-speed links using modified GSA-optimized neural network. Neural Computing and Applications, 24(3-4), 599-611.
29. Creech, G., & Hu, J. (2014). A semantic approach to host-based intrusion detection systems using contiguous and discontiguous system call patterns. IEEE Transactions on Computers, 63(4), 807-819.
30. Esmaily, J., Moradinezhad, R., & Ghasemi, J. (2015, May). Intrusion detection system based on Multi-Layer Perceptron Neural Networks and Decision Tree. In Information and Knowledge Technology (IKT), 2015 7th Conference on (pp. 1-5). IEEE.
31. He, J., &Zheng, S. H. (2014). Intrusion detection model with twin support vector machines. Journal of Shanghai Jiaotong University (Science), 19(4), 448-454.
32. Manandhar, P., &Aung, Z. (2014, September). Towards practical anomaly-based intrusion detection by outlier mining on TCP packets. In International Conference on Database and Expert Systems Applications (pp. 164-173). Springer, Cham.
33. Tama, B. A., & Rhee, K. H. (2015, August). Performance analysis of multiple classifier system in DoS attack detection. In International Workshop on Information Security Applications (pp. 339-347). Springer, Cham.
34. Ramakrishnan, S., & Devaraju, S. (2017). Attack's feature selection-based network intrusion detection system using fuzzy control language. International Journal of Fuzzy Systems, 19(2), 316-328.
35. Rodda, Sireesha. "Network Intrusion Detection Systems Using Neural Networks." Information Systems Design and Intelligent Applications. Springer, Singapore, (2018): 903-908.
36. Aljawarneh, Shadi, Monther Aldwairi, and Muneer Bani Yassein. "Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model." Journal of Computational Science 25 (2018): 152-160
37. Zhang, Y., Li, Y., Zhang, T., Gadosey, P. K., & Liu, Z. (2018). Feature clustering dimensionality reduction based on affinity propagation. Intelligent Data Analysis, 22(2), 309-323.

AUTHORS PROFILE

Neeraj Kumar:
Research Scholar at Birla Institute of Technology mesra, Extention Center Patna & presently working as a Principal Incharge, Purnea College of Engineering, Purnea, Govt. of Bihar, under Department of Scienece & Technology, Patna. Life Time Member of CSI. His research work involves using Data Mining, Network Security, Machine Learning techniques improve to the performance of any algorithm in Computer Vision. Contributed papers in the field of Data Mining, Security, Machine Learning, Distrbuted System, MANET, Soft Computing. He had nineteen years of experience in teaching and research.

Dr. Upendra Kumar:
Received his Master in Technology and Doctor in Philosophy from Birla Institute of Technology,Mesra Ranchi in the Department of Computer Science and Engineering respectively. He is currently working as an Assistant Professor in the same Department from last 6 years.His research field is related with Computer Network Security.He has published more than 25 research papers in Journal of repute and Conferences and apart of this he is also reviewer of several Journals of Computer Science including Journal of Computational Science, Elsevier and members of International Conference of ACM, IEEE,Springer.