Enhanced Intrusion Network System using Fuzzy –K-Mediod Clustering Method

Jaskirat Singh, Sanjay Singla

Abstract**: *Intrusion detection scheme (IDS) is software applications that are used for monitoring of the network to recognize the malicious activity in the system. With the advent in technology and internet, the incident of the intruder activity on the system has been increased. Hence, the protection and security of the system become an essential approach because attackers utilize different kinds of attack methods to hack the useful information. However, various intrusion detection methods and algorithm have been developed to detect various types of the attacks. Some issues in existing research are excesses of information with different volumetric data. It became difficult to detect intrusion in large amount of data in a computer network.. Hence, the proposed research on different machine learning approach is used for network intrusion detection. In addition, clustering method implemented to segment the features using k mean clustering and k-medoids clustering algorithm. Moreover, implement an enhanced Fuzzy k medoids clustering approach for recognition of intrusion and faults on the network. Fuzzy k-mediod clustering helps in the evaluation of the maximum degree matrix. Experimental analysis is done by evaluating and comparing the parameters using Precision, Recall, Accuracy ,FAR and FRR.***

Keywords* : *Clustering process, Malicious attack, Machine learning approach, Fuzzy k-mediod clustering.

I. INTRODUCTION

In the world of the computer era, there has been wide advancement of the digital technology that gives rise to threats and attacks [1]. Moreover, due to the growing demand for the online services, cyber crime has also been increased. In the present world, a malicious activity has become common among people. So, novel controlled and detection techniques are required to be developed [2] [3]. With the expansion in the amount of the distribution and interconnection of the data, security of the network has become an essential concern. Commonly, the kind of the security attacks on the network is hacker threat, viruses, and denial of services (DOS) attacks [4]. Hence, some security detection methods are required to be implemented to improve the performance of network [5].

Some of the data security methods are intrusion detection method, firewall method, communication security and prevention against anti-virus and hacker method [6]. The intrusion detection method is generally the method of searching the illegal access (intrusion) by examining the

in bound traffic [7]. The study and surveillance of vulnerabilities and the configuration of the uncommon actions in network is intrusion detection method. Generally, software security tools have been used for monitoring of the system traffic and detection of the distrustful action used in the network intrusion system [8]. Intrusion detection system (IDS) is the method that focused on detection of the attacks, monitoring information about intrusion and removes them and also creates security techniques in an actual timed environment is known as security recognition or intrusion detection method. Some of the challenges faced in the network intrusion detection system or inappropriate data; multiple arrangements of data that affect the system, and compressed information due to bad infrastructure [9] [10]. Though there are some issues, some of the applications are identical of external and internal attacks, monitoring the system and broadcasting and reply to threats of system, enhance the security system.

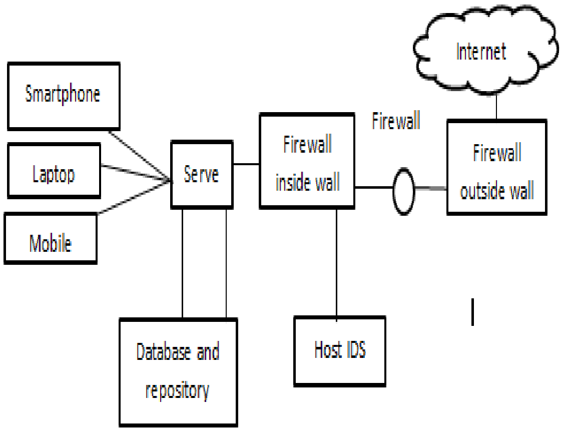

Fig.1 Architecture of intrusion detection system [11]

Intrusion detection system is of various types which are described as host based, network based and hybrid based [12].

a) Host-based IDS: In this method, single host can be protected from threats and required to be fixed on preserving host. Host based-IDS is utilized for the observation of the data and information element of data objects in standard functioning method.

b) Network-based IDS: This method is helpful in determining different threats that take place during illegal access and failure during the login process.

Revised Manuscript Received on October 30, 2019.
***** Correspondence Author

Jaskirat Singh*, Department of Computer Science and Engineering at GGS college of Modern technology(IKGPTU), kharar(Mohali), Punjab, India. Email: jaskiratsaini94@gmail.com.

Dr Sanjay Singla is a Professor and Dean Academics, at GGS college of Modern technology(IKGPTU), kharar(Mohali), Punjab, India. Email: da.ggscmt@gmail.com.

c) Anomaly -based IDS: This method is utilized at the time of the network failure. The failure may occur due to the alterations in the power elements. Wireless system problems of the hardware and software may cause a modification in hop anomalies. Unexpected distinctions in information signal and announcement issue is used for intrusion detection in the network. Information anomaly may occur due to an interruption in the system network [13].

d) Hybrid based IDS: This technique is the grouping of the hybrid and anomaly detection technique. It is used for the addition of the neural network that depends on pattern matching method. Moreover, this method is utilized in the identification of the predictable and unpredictable attacks. Though, both the technique presents filtration and arranges the safety alerts to decrease the number of alarm messages to the system manager [14] [15].

In existing research, classification methods were used for detection of anomaly based intrusion utilizing machine learning technique. They applied various machine learning methods along with data entropy computation using database Kyoto 2006 and performance was analyzed using these techniques. This research described the specific database through the machine a learning method that presents more than 90% accuracy, recall and precision. Conversely, ROC(receiver operating curvature) parameters were used to search the radiated based function whose performance was more appropriate as compared to other algorithms. In the proposed research, the detection and classification of intrusion detection using a cluster based method with dataset KDD 1999 and KYOTO 2006+. During pre-processing phase undesired information was extracted from the database. After that, separation of the related data into clusters. In next procedure, segmentation of the knowledge database is divided into groups through k-mediod mechanism. This model computes the average data value of database of objects in data records. Moreover, the fuzzy k mediod method is utilized for the execution of the implementation and construction of the position matrix. Experimental analysis is done using the parameters FAR, FRR, accuracy, Precision and Recall.

II. LITERATURE SURVEY

Tiwari, S., Roy, S. S., Charaborty, S and Kumar, A. et al.,2013 [16] proposed research on the detection of the hybrid method for the network intrusion detection. The research described the resistance against the malicious threats on the system. Detection observes the security of the sensor network because of the restraints on energy, storage and the capacity of the sensor and many other influences. In this research, described an intrusion detection system by utilising the neural network, uneven dataset and firefly approach. ***Asif, M. K., Khan, T. A., Taj, T. A., Naeem, U and Yakoob, S et al ., 2013 [17]*** overviewed the technological factor of the intrusion detection and also features by calculation and investigation of the intrusion system , its issues and future challenges. Presently, intrusion detection determines the applications and domains of the system. The benefits of the intrusion detection were used in army services, commercial from the last few decades. Intrusion detection methods were developed to improve the network system. ***Moon, S. Y., Kim, J. W and Cho, T. H. (2014 [18]*** presented a research on an efficient routing technology for detection and prevention of network intrusion method in wireless sensor network. Moreover, prediction of the threats in wireless sensor network was done through some prevention techniques. They developed a reliable route protocol for the atmosphere in which there was prevention and detection methods of the network intrusion. Experimental analysis was done to reduce the data transmission overhead and energy consumption over the network and comparison analysis was done with the existing methods. ***Al-Jarrah, O and Arafat, A. et al ., 2014 [19]*** utilised the intelligent scheme to detect the network threats using TDDN neural network. The system catches the data packets in actual environment provides data packets pre-processing using two pipes. The planned scheme contains data packet catch engine, pre-processor, pattern detection, classifier and investigating. The extraction of the features in pre-processing stage for the port scanning and host curve attacks and placed the features of TDDN and presents outcome recognise probable attack conduct in determined amount of the threats. They had testing features in the real time environment where maximum ability in detection of the threats. Moreover, scheme was verified through DARPA method database with detection value up to 100%. ***Shone, N., Ngoc, T. N., Phai, V. D and Shi, Q. et al., 2018[20]*** presented research on new deep learning method for the detection of intrusion. They proposed research on un-supervised feature learning method. Moreover, they determined the classifier method of deep learning that was developed using NDAE. Along with that, they projected the classifier which applied in graphical processing unit (GPU) and compute the method using standard KDD and NSL-KDD database. An experimental result was observed the method for improving NID method and compared that with existing approach. ***Yu, D et al., 2018[21]*** analysed the principles of the different detection techniques and implement the data alphanumerical analysis for the intrusion detection. However, it builds traffic numerical analysis method of network irregular intrusion and launches network intrusion (NI) digital signal method though combination of the digital signal technique. The related matched filtration method was utilised to filter the network intrusion (NI) digital signal for improvement of the outcome of signal to noise ratio. Hence, time frequency analysis method was used for extraction of the amount of features of network irregular intrusion, and adapted correlated spectra analysis technique was utilised to determine the intrusion detection. Experimental outcomes were demonstrated the technique with maximum accuracy and robust anti-distortion capability that assured the system security.

III. RESEARCH METHODOLOGY

In this section, elaborate the proposed work with methods. Study and Analysis the various machine learning techniques in Network Intrusion detection. Develop clustering approach to divide the properties with the help of K-means clustering algorithm and K_mediod Clustering. Implement improved clustering algorithm (Fuzzy-KMediod algorithm) to detection the intrusion or network faults.

Evaluate and Compare the performance parameters are Precision, Recall, Accuracy and Error Rate.

1.Collect dataset from the online source (UCI Machine Learning Repository) Site.
2. Upload the dataset (.xls and .csv).
3. Data pre-processing (to remove the unwanted data in the uploaded dataset file).
4. Clustering Approach Implement
    4.1 Kmeans
    4.2 K-mediod
    4.3 Fuzzy-K-mediod Clsutering
5. To divide the data into two groups or clusters.
6. Analysis the Intrusion Detection
    6.1 Normal Data
    6.2 Intruded Data
7. Performance Metrics
8. Comparison
9. Stop.

Explanation in proposed work

Step 1. Upload the dataset from the given KDD 1999 and KYOTO 2006+ dataset in intrusion Detection Systems.

Step 2. Pre-processing phase to remove the unwanted data in the given datasets.

Step 3. Clustering Process :The clustering process is the main task of separating the given relevant data points into a number of cluster or groups such that all data values in the similar to other data values in the similar cluster than those in other groups.

Step 4. K mediod: -medoids is generally location of data objects of the group of clusters with less mean displacement related to other data objects. The approach based on the segmentation of the database in to groups. Medoids is the group of the definite database with finite set of information that may not be same to decreased data values. The K-medoids Clustering is the process of the partitioning of the clusters through clustering algorithm. The k- medoids algorithm is more prone to noise and data points are required in the k- medoids. The medoids is the object of the cluster where all the objects are different between the mean and the medoids. The medoids algorithm calculates the average value of the data sets of the objects in the data items.

Step 5. Fuzzy K-mediod: The selection of objects is clusters called as medoids and fuzzy k mean clustering is called as fuzzy k medoids [19]. Fuzzy k mediod clustering technique is used for execution and production of the degree matrix. The selection of the maximum degree members is determined in this approach [20].In this clustering mechanism there is partitioning algorithm where the number of clusters is recognised in advance and then partitioned in to the number of the clusters. In fuzzy k-medoids clustering, this is suitable for the large database.

Step 6. Performance metrics:- Evaluate and compare the performance metrics such as FAR, FRR, accuracy, Precision and Recall.

IV. RESULT DISCUSSION

The protection against the intrusion attacks or an unauthorized access by an intruder has been done by the software system. The intrusion detection technique is used for the building of the classifier technique proficient of the modification among the bad connections is called as assaults or interruption and standard connections. The intrusion detection method is DARPA in 1998 that was established and accomplished by MIT Lincoln Laboratory. The main objective was to investigate on intrusion detection method. [44]. The new trained data were about four gigabyte of the compress double TCP dump data through 7week network system traffic. This was controlled in 5 million linking measures. The two weeks of the testing data consists 2 million scheduled at the similar time. The connection is sequence of the TCP information data frames initialised and stops at certain time period where data flows in direction of IP address of source to destination by regular protocol. Every connection is labelled as equal normal and along special attack. Every connection score consist of 100 bytes [45].

The different threats occur are:-

- Denial of Service: Threats that take place during flow of information.
- R2L:An unauthenticated access by sensor tools such as guess password.
- Unauthorised2root: Attacks may occur due to unauthorised access.
- Investigation: Searching and kind of finding such as scanned post.

In this research, the main goal is detection of the attacks using fuzzy k-medoids clustering approach. The interface describes the database, information pre-processing stage and hybridization of clustering to recognise and determine the performance of threat. Subsequently, compute the performance metric and comparison is done with existing techniques.

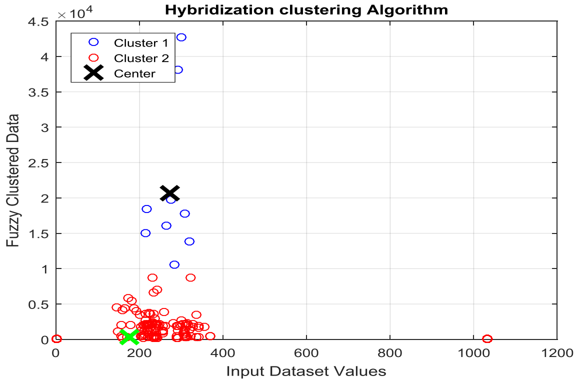

Fig 1. Hybrization (Fuzzy-Kmediod) Clustering Process

Fig 1. It is data mining method which includes conversion of raw database in to information based format. In real world, actual dataset is inadequate, decreasing in specific exceptions. It is technique of solving issues. It contains raw information for future image processing. The figure describes the graphical demonstration that recognise all matrix data and plotting the values in graphical format.

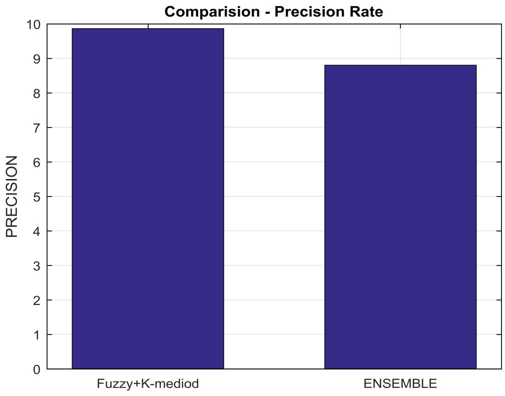

Fig 2.Comparison –Precision

Fig 2. describes the comparison among planned and existing research in precision value. In planned research, an enhanced precision rate using Fuzzy k-mediod and present parameter rate is 9.8%.

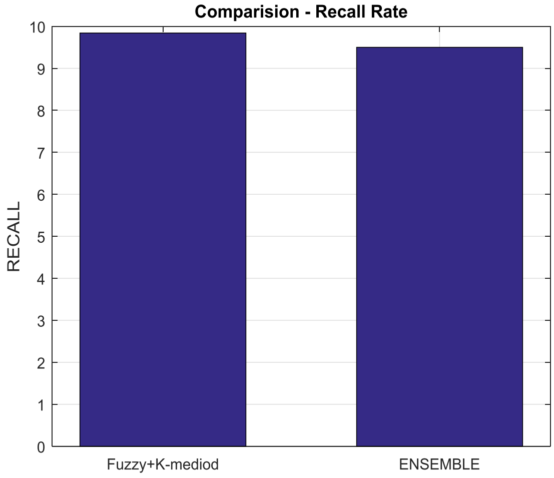

Fig 3. Comparison – Recall Rate

Fig 3. describes the comparison among the proposed and existing research in recall value. In planned research, an improved recall value using Fuzzy k-mediod and present parameter rate is 9.840%.

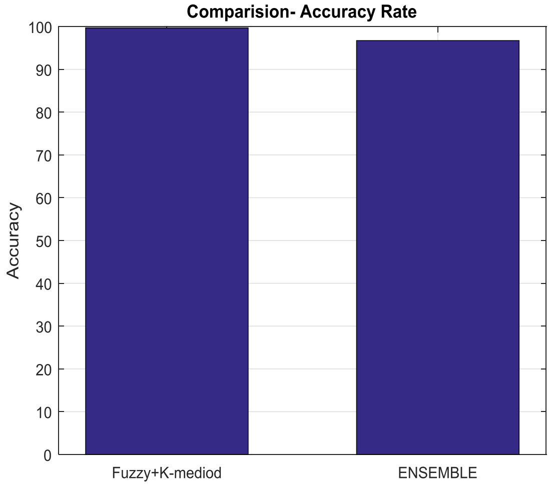

Fig 4. Comparison – Accuracy Rate

The given figure 4.2.4 describes the comparison among the proposed and existing research in recall value. In planned research, an improved recall value using Fuzzy k-mediod and present parameter rate is 99%.

Table 1. Proposed Parameters

Parameters	Values
Far	0.0014
MSE	0.0014
Frr	0.0016
Precision	9.8
Recall	9.840
Accuracy	99
Roc	0.9986

Table 2. Compative Analysis

Parameters	Proposed Work (FuzzyKmediod)Clustering	Existing Work (Ensemble)
Accuracy	99	96.7
Precision	9.8	8.8
Recall	9.840	9.5

The table 1 calculate the parameter metric using accuracy rate, precision, recall, false acceptance rate, false rejection rate and accuracy rate. The table 2 describes the comparison among planned and existing research using accuracy rate, precision, recall value.

IV. CONCLUSION AND FUTURE SCOPE

In conclusion, Intrusion detection scheme may be used for monitoring the malicious activity or network changes. The detection method helps in the detecting the modifications in the network. An intrusion detection system is used to detect the different kinds of the malicious behaviour that may affect the security of the computer system. Hence, various detection methods have been developed in this research to detect the intrusion using machine learning techniques. Firstly, dataset KDD 1999 and KYOTO 2006+ uploaded for intrusion Detection Systems. Then, distorted data are removed from dataset through Pre-processing phase. Clustering is performed for separation of information into number of groups due to similarity of clusters. K-medoids and fuzzy k-mean clustering is used for detection of intrusion system. K-medoids clustering is the method of the segmentation of the clusters by clustering approach. This method is more prone to noise and determine the average value of the data sets of the attributes in the data items. Novel Fuzzy k-medoids method implemented to divide the data into groups of cluster format based on the mean, distance, index and iterations. It used for the selection of the instances for implementation and presenting the degree matrix. The clustering approach leads to segmentation of the clusters in complex structure. But Fuzzy K mediod resolve the clustering and confusion matrix problems. With Novel method to improve the accuracy rate, precision and recall values as compared with the existing algorithms (Ensemble).

Future scope is emphasis on the enhancement of the processing time of detection of intrusion using deep learning approach. In addition, artificial neural network can be applied on dataset for evaluating the performance and compared with the proposed detection approach. Moreover, the performance of the anomaly detection technique can be improved using deep learning approach with neural network.

REFERENCES

1. Liao, H. J., Lin, C. H. R., Lin, Y. C and Tung, K. Y. (2013), Intrusion detection system: A comprehensive review, *Journal of Network and Computer Applications*, vol. *36*(1), pp.16-24.
2. Smaha, S. E. (1988),Haystack: An intrusion detection syste,. In *[Proceedings 1988] Fourth Aerospace Computer Security Applications* ,vol 2(3), 37-44, IEEE.
3. Zhang, B., Liu, Z., Jia, Y., Ren, J. and Zhao, X. (2018), Network Intrusion Detection Method Based on PCA and Bayes Algorithm.,*Security and Communication Networks*, *2018* ,Vol 2 (4).
4. Park, K., Song, Y., & Cheong, Y. G. (2018), Classification of attack types for intrusion detection systems using a machine learning algorithm, In *2018 IEEE Fourth International Conference on Big Data Computing Service and Applications (BigDataService)* .,vol. 2(3).pp, 282-286.,IEEE.
5. Zaman, M. and Lung, C. H. (2011), Evaluation of machine learning techniques for network intrusion detection. In *NOMS 2018-2018 IEEE/IFIP Network Operations and Management Symposium* ,vol. 3(2), pp. 1-5. IEEE.
6. Mukherjee, B., Heberlein, L. T. and Levitt, K. N. (1994), Network intrusion detection. *IEEE network*, ,vol. *8*(3), pp. 26-41.
7. Sommer, R and Paxson, V. (2010), Outside the closed world: On using machine learning for network intrusion detection.,In *2010 IEEE symposium on security and privacy* , vol. *2(1)*,(pp. 305-316, IEEE.
8. Sperotto, A., Schaffrath, G., Sadre, R., Morariu, C., Pras, A. and Stiller, B. (2010), An overview of IP flow-based intrusion detection.,*IEEE communications surveys & tutorials*, vol *12*(3), pp.343-356.
9. Catania, C. A. and Garino, C. G. (2012). Automatic network intrusion detection: Current techniques and open issues. *Computers & Electrical Engineering*,vol . *38*(5), pp.1062-1072.
10. Hu, W., Hu, W and Maybank, S. (2008). Adaboost-based algorithm for network intrusion detection. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. *38*(2), pp.577-583.
11. Anwar, S., Mohamad Zain, J., Zolkipli, M. F., Inayat, Z., Khan, S., Anthony, B. and Chang, V. (2017).,From intrusion detection to an intrusion response system: fundamentals, requirements, and future directions. *Algorithms*, vol *10*(2),pp. 39.
12. Sato, T. and Fukase, M. A. (2003), Reconfigurable hardware implementation of host-based IDS, In *9th Asia-Pacific Conference on Communications (IEEE Cat. No. 03EX732)* ,Vol. 2, pp. 849-853,IEEE.
13. Samrin, R. and Vasumathi, D. (2017), Review on anomaly based network intrusion detection system. In *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)* ,vol. 3(1),pp. 141-147, IEEE.
14. Joo, D., Hong, T and Han, I. (2003), The neural network models for IDS based on the asymmetric costs of false negative errors and false positive errors, *Expert Systems with Applications*, vol. *25*(1), pp.69-75.
15. Pan, S., Morris, T. and Adhikari, U. (2015), Developing a hybrid intrusion detection system using data mining for power systems., *IEEE Transactions on Smart Grid*, vol. *6*(6), pp. 3104-3113.
16. Tiwari, S., Roy, S. S., Charaborty, S. and Kumar, A. (2013), A novel hybrid model for network intrusion detection, In *2013 International Conference on Green Computing, Communication and Conservation of Energy (ICGCE)*, pp. 685-688, IEEE.
17. Asif, M. K., Khan, T. A., Taj, T. A., Naeem, U. and Yakoob, S. (2013), Network intrusion detection and its strategic importance, In *2013 IEEE Business Engineering and Industrial Applications Colloquium (BEIAC)* ,pp. 140-144,IEEE.
18. Moon, S. Y., Kim, J. W and Cho, T. H. (2014), An energy-efficient routing method with intrusion detection and prevention for wireless sensor networks, In *16th International Conference on Advanced Communication Technology* ,pp. 467-470, IEEE.
19. Al-Jarrah, O. and Arafat, A. (2014), Network Intrusion Detection System using attack behavior classification, In *2014 5th International Conference on Information and Communication Systems (ICICS)* , vol2(1), pp. 1-6), IEEE.
20. Shone, N., Ngoc, T. N., Phai, V. D and Shi, Q. (2018), A deep learning approach to network intrusion detection.,*IEEE Transactions on Emerging Topics in Computational Intelligence*, vol.*2*(1), pp.41-50.
21. Yu, D. (2018), Research on Anomaly Intrusion Detection Technology in Wireless Network, In *2018 International Conference on Virtual Reality and Intelligent Systems (ICVRIS)* ,vol. 2(2) ,pp. 540-543, IEEE.

AUTHORS PROFILE

Jaskirat Singh is an M.tech Student in Department of Computer Science and Engineering at GGS college of Modern technology(IKGPTU), kharar(Mohali), Punjab, India.He has received his B.Tech degree in Computer Science and Engineering from Indo Global college of Engineering and Technology(IKGPTU), Abhipur(Mohali), India in the year of 2016. He is doing his research in the field of Intrusion Network System, titled as "Enhanced Intrusion Network System using Fuzzy-Kmediod Clustering Approach". Beside this, his other research interests include artificial intelligence, data mining, machine learning and data analysis.

Dr Sanjay Singla is a Professor and Dean Acadamics at GGS college of modern technology(IKGPTU), Kharar(Mohali), Punjab, India.He is M.tech, P.hd, MISTE(I), MIACSIT(Singapore), MIAENG(Taiwan).