

Survey on Fuzzy Based Extreme Learning Machine for Intrusion Detection

Thamizharasi.E¹, P.Salini²

¹(Department of CSE, Pondicherry Engineering College, Puducherry, India)

²(Department of CSE, Pondicherry Engineering College, Puducherry, India)

Abstract: Nowadays, the network based applications are growing rapidly, and therefore the quality of utilizing web contains few dangers of network attacks. Thus, network security must be necessary to produce secure data because of increase in potential network attacks. Intrusion detection is among the most necessary analyses problems in network security. Intrusion Detection System (IDS) is a fundamental tool to monitors and analyze network activities for detecting intrusions and security attacks. But it faces large number of problems to perform efficient intrusion detection. One among the vital problem is, it examines all features within the dataset to find intrusion. Many redundant and irrelevant features occurs, which can decrease the intrusion detection efficiency and also it may take more computational time for the effective response in real time network traffic. In this regards, machine learning techniques are applied to IDS that helps to detect intrusion. This works used to make survey of a different machine learning techniques for classification and also presents a feature selection method to find the well-known or unknown attacks in a network. This in turn will help to boost accuracy and decrease false alarm rate to detect intrusion in the intrusion detection system.

Keywords: Intrusion Detection System (IDS), Machine learning, Accuracy, False Alarm Rate

I. Introduction

The present examination proposes an overview of various Machine Learning algorithms for Intrusion Detection System. There are several applications are available for network security. Network security is the arrangement of technologies explicitly intended to ensure system, data, network from unauthorized access, modifying data, crashing the system, etc. An Intrusion Detection System (IDS) has become a serious research theme as a fundamental apparatus for system network security. As indicated by the network security, intrusion detection could be founded on two general gatherings, network-based intrusion detection and host-based intrusion detection. In network based intrusion detection, the network analyzes approaching system dangers to neighbourhood Local Area Network (LAN), while the host based intrusion detection discover LAN threats originating from the host in a network [1]. The approach of intrusion detection can be arranged into two classifications, namely signature based detection and anomaly based detection. Signature based techniques are intended for recognize recently known attacks is to utilize signature for attacks.

The framework stores best-known attacks signature and trying to find such (marks) signatures in network traffic, if any matches happen, it is known as a signature. It will recognize attacks with less warning (false alarm) rate, yet cannot identify some new sort of attacks that do not have any characterised signatures [2]. Anomaly based detection used to detect the unknown attacks that breach from normal attacks. They are significant because they are able to identify the normal attacks. The primary disadvantage of the anomaly based detection is highest false caution (alarm) rate.

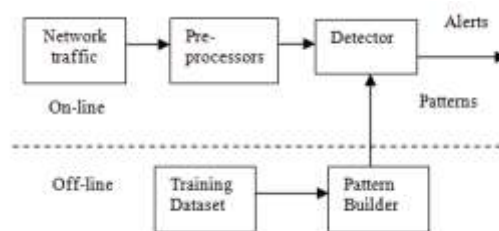


Figure 1: Working Environment of Intrusion Detection

In anomaly detection, assembling a model is typically established by recording ordinary activities traffic within the LAN, when a framework monitors any break from that model, it will be thought-about as an anomaly or an attack. The fuzzy logic framework applies machine learning procedures to assemble designs for

intrusion detection. The working environment of the Intrusion Detection system (IDS) has been showed in Figure 1.

There are two phases in the network framework: off-line and on-line [3]. The system fabricates pattern of interruptions within the off-line section and identify interruptions in the on-line section. In the off-line section, to feed training dataset into the pattern developer module which may assemble the designs of intrusions. The module utilizes the feature selection algorithmic rule, handling unequal intrusions, and assembles the designs (patterns) by random forests within the optimum parameters. Once mining the designs for intrusions, the module yields the patterns because of the input of the Detector module. In the on-line section, the framework catches the packets from network/system traffic. The features for every affiliation were created by the pre-processors from the caught network traffic. At that point, in the detector module, the connections are named diverse intrusions or normal traffic exploitation the patterns in-constructed the off-line section. At long last, the network raises an alarm once it recognizes any intrusion.

The remainder of the paper is organized as detailed below. The Classification of Anomaly Intrusion Detection is discussed in Section II. Some of the classification algorithms of Machine Learning classifiers are explained in Section III. The related work is presented in Section IV. KDD'99 Dataset Description is described in Section V. Summarization of the Classification results of the different Machine Learning classifiers is discussed in Section VI. The paper is concluded in Section VII, which provides a summary and directions for future work.

II. Classification Of Anomaly Detection

Anomaly detection is the general class of intrusion detection which works by distinguishing activities which may change from set up examples for users, or groups of users. Figure 2 shows the Classifiers of Anomaly Intrusion Detection System. According to the kind of handling process related to the “normal behavioural” model of the target framework (system), anomaly detection techniques can be organized into three standards of classifications, for example, statistical-based, knowledge-based and machine learning based [4].

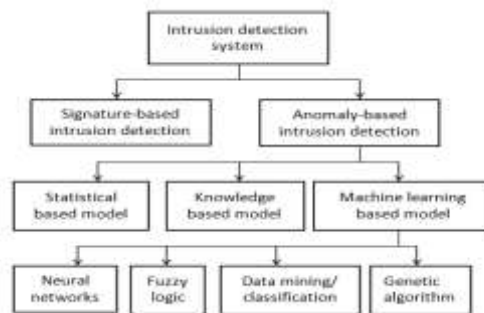


Figure 2: Classification of Anomaly Intrusion Detection System

2.1 Statistical-based IDS

Statistical-based IDS decides an ancient network activity like what to measure reasonably information (data) utilized for the most parts of protocols utilized, ports and gadget typically associate with each other and mindful the administrator or client once traffic is distinguished that is unpredictable.

It's again categorized into single-variable, multi-variable and statistic model. Single-variable model parameters are exhibited as autonomous Gaussian irregular factors so forms (processes) an adequate vary of characteristics for each factor (variable).

The multi-variable model thinks about the connection between two numbers of variables. The statistical model uses an interval timer, combined with an incident counter or resource measure and considers the demand and rest section time of recognitions and their characteristics that are marked as inconsistency (Anomaly) if its likelihood of frequency is essentially excessively low at a given time.

2.2 Knowledge-based IDS

Knowledge based stores information (data) with respect to subject area. Information in knowledge based contains symbolic/numeric representations of master guidelines of judgment in an arrangement that empower the effective engine to perform conclusion upon it. The expert system approach is one among the principal wide utilized knowledge-based IDS plot. Knowledge-based techniques are classified into frame-based model, rule-based model and expert system.

Frame-based model confines a total assemblage of expected information and activities into one structure. Rule based model is changed assortment of the language based generation rules. Expert systems

should arrange the audit information predictable with a gathering of rules (principles). Expert systems can be included into three stages. To begin with, different features and classes are known from the training data. Second, a gathering of classification rules, techniques or parameters are reasoned. Third, the audit information which arranged as required.

2.3 Machine Learning-based IDS

Machine learning techniques are developing an explicit or implicit model. A single characteristic of plans is that the prerequisite for labelled knowledge to prepare the behavioural mode, a strategy that places extreme demands on assets. In a couple cases, the significance of machine learning standards agrees with that for the statistical techniques, notwithstanding the fact that the past is centred overbuilding a model that upgrades its execution on the likelihood of past outcomes. Henceforth, machine learning for Intrusion Detection System has the versatility to change its execution procedures since it acquires new information (data). This component could create it attractive to utilize such plans for all circumstances.

III. Classification Of Machine Learning Classifiers

A few Machine Learning-based models (techniques) have been associated with IDS. Some of the fundamental strategies are explained in following below.

3.1 Neural Networks

Neural Networks was generally used to refer a framework or organic (biological) neurons. In Intrusion Detection System (IDS), neural network has been utilized for every signature (abuse) and anomaly based intrusion detection. In signature based intrusion detection the neural network would gather knowledge from the system/network stream and break down the information for cases of abuse.

In anomaly based intrusion detection the neural networks were demonstrated to recognize measurably noteworthy varieties from the client's perceived conduct likewise develops the ordinary attributes of framework clients.

In neural network the misuse intrusion detection will be authorized in 2 manners by which the chief methodology fuses the neural network part into partner existing framework or knowledgeable framework. This strategy utilizes the neural network to the approaching (data) information for suspicious occasions and moves them to the prevailing and knowledgeable system.

This enhances the capability of the detection system. The second stage procedure utilizes the independent abuse (misuse) detection system. This procedure gets information from the system stream and examines it for abuse intrusion. It has the adaptability to discover the qualities of misuse attacks and build up occurrences that dislike any that are resolved before by the network. It is high level of precision to recognize familiar suspicious occasions. It was utilized to learn confused non-linear input data and output data [4].

3.2 Support vector machine (SVM)

Support vector machine (SVM) is projected by the author vavnik [5], where SVM beginning maps the input vector into a superior dimensional feature area and after that acquire the best splitting hyper-plane within the higher dimensional feature space. Besides, a decision boundary, i.e. the isolating hyper-plane, is set by support vectors rather than the whole training tests as is exceptionally very strong to outliers [5]. Figure 3 shows the SVM model of intrusion detection system (IDS).

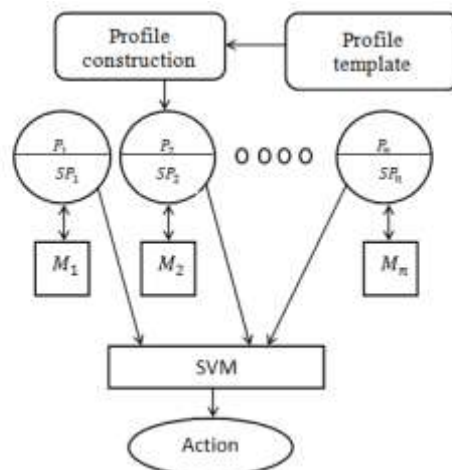


Figure 3: A SVM techniques for intrusion detection system [6].

The support vector machine approach changes data (information) into a feature area F that ordinarily includes an extensive dimensional. It asentrancing to takehints of that SVM general techniquesreliesupon the geometryqualities of the training data, not on the size of the information zone, training SVM results in a quadratic optimisation drawback with certain requirements and one linear equality constraints. Author Vapnik, [6] indicates how preparing a SVM for the pattern recognition drawback end up in the following quadratic optimisation problem.

A client provides a grouping S if S is contained inside the client arrangement for their user. The meaning of this help is given ontheir grounds that are part of the whole range that exists inside the grouping. In the successive example profiling, client day by day exercises was taken as a sequence and a database for every client involves client’s day by day consecutive example was made.

3.3 Fuzzy logic (FL)

Fuzzy systems have incontestable their capacity to determine diverse kinds of issues in various application spaces. Fuzzy systems supported fuzzy if-rules are with progress utilized in a few applications areas [7]. Fuzzy if-then rules aregenerallypicked up from naturalspecialists. As of late, different techniques are immediate for naturally creating and changing fuzzy if-then rules while not exploitation thatassistance of human experts. Genetic algorithms are utilized as the rule generation and their enhancement apparatus inside the style of fuzzy rule-based frameworks [8].

Fuzzy rules have the form:

IF condition THEN subsequent [weight]

Where,

1. The condition could be a complicated fuzzy expression, i.e., a logic expression that utilizesformal fuzzy logic operators and fuzzy expression.
2. The subsequent is an fuzzy expression and after that
3. The load (weight)could be a complex quantity that characterizes the certainty of standard (rule).

3.4 Extreme Learning Machine (ELM)

The extreme learning machine (ELM) aims at avoiding time-costing iterative training process and improving the generalization performance [9].The initial structure of ELM is expressed in Figure 4. As a single hidden layer feed-forward neural networks (SLFNs), the ELM structure incorporates input layer, hidden layer, and output layer [10].

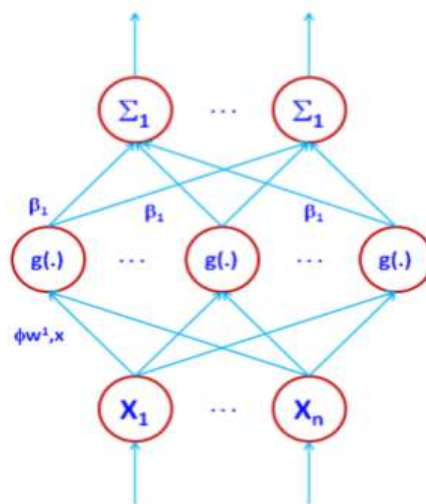


Figure 4: Basic structure of ELM

Different from the standard neural network learning algorithms (such as BP algorithm)randomly setting all the network training parameters and simply generating local optimumresolution, the ELM just sets the no of hidden neurons of the network, randomizes the loads (weights)within the input layer and in this mannerhidden layer correspondingly on grounds that bias of the hidden neurons within the rule of execution process, calculates hidden layer output matrix, and lastly gets weight in-between hidden layer and output layer by utilizing Moore-Penrose pseudo inverse[11] under the criterion of least-squares method.

Because the ELM has the easy network structure and also the concise parameters computation processes, therefore the ELM has the benefits of quick learning speed.

3.5 Genetic Algorithm (GA)

Genetic Algorithm (GA) utilizes the system to actualize the survival and evolution. This thought originates from the “adaptive survival in common creatures” [12]. The algorithmic guideline begins by all over creating an outsized populace of each program. Some kind of wellness live to pass judgment on the execution of every individual populace is utilized. Anlarge kind of iterations are then performed that less activity programs are exchanged by genetic recombination of better activity programs. (i.e.), a program with an occasional fitness live is deleted and doesn’t survive for consecutive laptop iteration.

As of late, researchers have attempted to coordinate these algorithms with IDSs. The imperial system could be a conception learning system supported a distributed genetic algorithmic rule.

3.6 Random Forest (RF)

One of the premier standard ways or frameworks utilized by scientists within exploration of information is random forest. Random forest algorithmic standard is one in all the most straight forward among the characterization formula will classify huge measure of information with precision. Random forest might be a smart apparatus to frame expectations once may consider that they are not asper the law of huge numbers. The presentation of the algorithm may be legitimate to frame their random classifiers satisfactory and reasonable. Random forests are easy to be told and use for talented. Random forest is most appropriate for examining advanced knowledge structures that are inserted in a very exceptionally straight forward record contains under a great many lines, yet anyway perhaps a largenumbers of columns [13].

IV. Related Work

Iftikhar Ahmad et al [10], proposed a machine learning strategies for breaking downlarge data for intrusion detection of network and data frameworks. By utilizing these procedures, specifically, Support Vector Machine (SVM), Random Forest (RF), and Extreme Learning Machine (ELM) are assessed orlooked at. In this ELM well-performs different methodologies in the metrics of accuracy, recall, and precision of full data tests when contrasted with other than SVM and RF. Besides, ELM would be investigatedtowards to assess it as execution in feature selection.

Chi Cheng et al [14], with expanding availabilitywithin networks, the danger of information network systems to external attacks or intrusions hadexpanded immensely. These strategies commonly experience the ill effects of long training times, require parameter standardisation, or don’t perform well in multi-class classification. So they are proposed the utilization of essential and kernel based extreme learning machine for intrusion detection in a network system. It decreases the ideal opportunity for training and gives great adaptability. The basic ELM has marginally lower accuracy when contrastedwith SVM techniques. To expand accuracy and to distinguish this type of attacks, and kernel-based ELM canbe executed.

Nabila Farnaaz et al [15], proposed a Random Forest (RF) classifier model for intrusion detection. RF is a classifier and it performs well when contrasted with various classifiers for effective classification of attacks. Thesetechniquesdeliver low exactness in identifying intrusion and it has low classification error. The test results should be possible by utilizing accuracy, detection rate, false alarm rate, and Mathew’s correlation co-efficient.

Sylvia Lilly Jebarani et al [16], proposed a technique, use extreme learning machine for classification. It is trained by using training dataset and classifies new types of connection records. The features can be classified by three techniques which can be decrease space and time complexity and furthermore reduces the accuracy of the classifier.They proposed another approach, Online Sequence-Extreme Learning Machine (OS-ELM) classifier for intrusion detection that overcomes slow learning limitation of other classifiers and is also capable of solving several classification issues and process massive dataset during a very less time.

M. A. M. Hasan et al. [17], Feature selection will be treated as a pre-preparing step to reinforce the overall framework execution altogether while mining on huge datasets. So that, they centre around two step approach of feature selection supported random forest. The experimental outcomes demonstrate that the random forest primarily based methodology will select most vital and relevant options (features)useful for classification, that it diminishes does not simply the amount of input features and time anyway furthermore willbuilds the classification accuracy.

Masarat et al. [18] exhibited a unique multistep structurepassionate about machine learning techniques to make a capable classifier. In beginning advance, the feature selection strategycanexecuteaddicted to gain proportion of highlights (features) to the creators. Their procedure willupgrade the execution of classifiers that are made dependent on these features. In classifiers hybrid step, canshow an exceptional fuzzy ensemble technique, on these ways that classifiers with part ofexecution and lowworth have extra outcome to form the classifier.

According to the analysis of Huang et al. [15] in ELM the word “Extreme” is used to show the moving from general artificial intelligence learning techniques towards brain like learning which is self-adaptive and very faster. ELM has been applied to severe real world applications and has been shown to get sensible generalization performance at very high learning speed. It is the key strength of considerably low procedure time.

V. Dataset Description

In 1998, Defence Advanced Research Projects Agency (DARPA) together with Lincoln Laboratory at MIT moved the DARPA 1998 dataset for surveying IDS. The DARPA 1998 dataset contains seven weeks of training and moreover day and age of testing information (data).

Altogether, there are thirty eight attacks in preparing information (data) in like manners in testing information (data). The refined adjustment of DARPA dataset that contains only system information itself (i.e. TCP dump data) is named as KDD dataset [20].

The third all inclusive Knowledge Discovery and Data mining (KDD) tools competition were control in colligation with KDD-99, the fifth global meeting accumulated on KDD. It could be a dataset utilized for this third universal knowledge Discovery and data processing tools competition. KDD training dataset comprises of comparatively millions of single association vectors wherever single affiliation vectors comprises of 41 features includes and is set apart as either traditional or an attack, with correctly one explicit attack type.

In KDD’99 dataset, every model represents attribute values of a category within the network information (data) stream, and every category (class) are tagged either ordinary (normal) or attack. The categories in KDD99 dataset are going to be classified into 5 primary classification one normal (traditional) class and 4 fundamental intrusion classes [21].

Table I outlines various attacks falling into four major categories such as Denial of Services (DOS), User to Root (U2R), Remote to Local (R2L) and probes are examined detail in underneath.

TABLE I: VARIOUS CLASSES OF ATTACKS

Classes of Attacks	Attack Name
Denial of Service (DOS) Attacks	Back, Land, Neptune, Pod, Smurf, Teardrop
User to Root (U2R) Attacks	Buffer-overflow, Load-module, Perl, Rootkit,
Remote to Local (R2L) Attacks	Guess-password, Ftp-write, IMAP, PHF, Multi-Hop, Spy, WarezClient, WarezMaster
Probes	Satan, IP-sweep, N-map, Port-sweep

- **Denial of service (DOS) Attacks:** A Denial of service attack is related to attack wherever the attacker constructs some enrolling or memory asset totally possessed or inaccessible to reach to oversee real needs, or reject genuine client ideal to utilize a system.
- **User to Root (U2R) Attacks:** User to Root abuses are a class of endeavours where the attacker start by getting to an ordinary client account on the framework (system) perhaps achieved by finding the passwords, a word reference attack, or social designing) and exploit some powerlessness to achieve root access to the framework.
- **Remote to Local (R2L) Attacks:** A Remote to User attack happens once an aggressor (attacker) who can possibly send packets to a conventional over a network however doesn’t have an record in themachine, makes utilization of some powerlessness to accomplish nearby access as a customer of that machine.
- **Probes:** Probing might be a class of assaults (attacks) wherever an offender looks at a system to accumulate data or findsurely understood vulnerabilities.

VI. Experimental Results

In this area, we summarize the exploratory outcomes to assemble designs for intrusion detection over the KDD’99 datasets. Table II summarizes the classification results of the different machine learning classifiers in regard to accuracy, precision and recall. The datasets are classified into normal and four attack types. The KDD’99 dataset was used for the experiment.

TABLE II: RESULTS OF DETECTION OF ATTACKS.

S.no	Algorithms	Accuracy%	Precision %	Recall %
1	SVM	82.37	91.01	79.18
2	GA	92.64	92.19	81.47
3	NN	81.29	84.09	78.95
4	RF	93.18	97.4	71.42
5	FUZZY	93.24	89.32	83.94

6	ELM	96	98.56	88.27
---	-----	----	-------	-------

The performance of the machine learning classifiers are compared based upon accuracy, precision, and recall was calculated by the measure of True Positive (TP), False Positive (FP), and True Negative (TN) and False Negative (FN) metrics.

Accuracy is that the most important basic live (measure) of the performance of a learning technique. It offers the chance that the algorithms will properly predict positive and negative instances and is computed as:

$$\text{Accuracy} = \frac{TP+TN}{P+N}$$

Precision is outlined because the proportion of positive predictions that created by the classifier that are true. The precision rate straight forwardly influences the performance of the system. The precision is determined by

$$\text{Precision} = \frac{TP}{TP+FP}$$

The Recall rate are additionally a vital value for estimating the execution of the recognition (detection) system and to demonstrate the extent of occurrences having a place with the positive rate that are effectively anticipated as positive. This is otherwise called as sensitivity or true positive rate.

$$\text{Recall} = \frac{TP}{P}$$

The outcomes obviously exhibit that the classification performance of IDS is upgraded by machine learning classifiers dependent on KDD'99 datasets are appeared table II and figure 4.

Figure 5 represents the performance metrics for different classifiers and demonstrates that Extreme Learning Machine have outperformed in characterizing the intrusions with 96% of accuracy, 98.56% of precision and has got 88.27% of recall followed by SVM, GA, NN, RF and FUZZY. The neural network (NN) has get the lowest rate of accuracy (81.29%) and precision (84.09%) and the genetic algorithm (GA) has also got the lowest rate of recall (71.24%) were compared to other classifier algorithms.

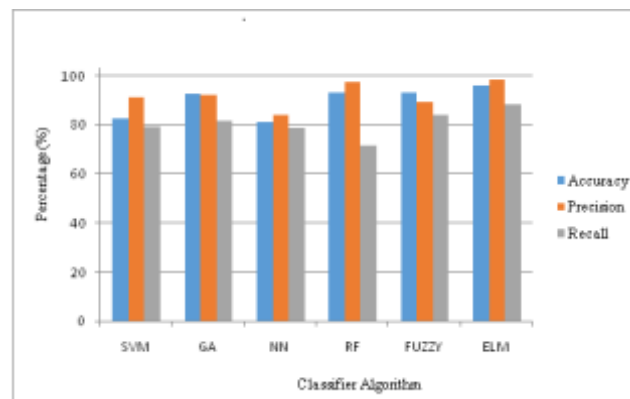


Figure 5: Performance comparison of Machine Learning Classifiers for Intrusion Detection.

VII. Conclusion

In this paper, different machine learning methods are used to discuss for intrusion detection, however there is no methods are concluded as most effective, with help of these methods are growing research are for intrusion detection. In the literature survey many studies used KDD'99 datasets and used different machine learning algorithms, and analysed the performance of different algorithms for anomaly detection.

If the machine learning methods compared based on accuracy, these methods should be trained and tested on same accurate data. So these results based on accuracy are not comparable. In the proposed work, have to combine the hybrid methods of fuzzy and extreme learning machine for Intrusion detection will make it even more time consuming and to reduce false alarm.

References

- [1]. Aburomman, A. A., & Reaz, M. B. I. (2016, November). Survey of learning methods in intrusion detectionsystems. In *Advances in Electrical, Electronic and Systems Engineering (ICAEES), International Conference on* (pp. 362-365). IEEE.
- [2]. Kala, T. Sree, and A. Christy. "A Survey and Analysis of Machine Learning Algorithms for Intrusion Detection System." (2017): 40-46.

- [3]. Zhang, J., Zulkernine, M., &Haque, A. (2008). Random-forests-based network intrusion detection systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 38(5), 649-659.
- [4]. Subaira, A. S., &Anitha, P. (2013). A survey: network intrusion detection system based on data mining techniques. *Int. J. Comput. Sci.Mob.Comput*, 2(10), 145-153.
- [5]. Tsai, C. F., Hsu, Y. F., Lin, C. Y., & Lin, W. Y. (2009). Intrusion detection by machine learning: A review. *Expert Systems with Applications*, 36(10), 11994-12000.
- [6]. Mohammed, M. N., &Sulaiman, N. (2012). Intrusion detection system based on SVM for WLAN. *Procedia Technology*, 1, 313-317.
- [7]. Thakare, S. P., & Ali, M. S. (2012). Introducing Fuzzy Logic in Network Intrusion Detection System. *International Journal of Advanced Research in Computer Science*, 3(3).
- [8]. Cerli, A. A., &Ramamoorthy, S. (2015). Intrusion Detection System by Combining Fuzzy Logic with Genetic Algorithm. *Global Journal of Pure and Applied Mathematics (GJPAM)*, 11(1), 2015.
- [9]. Zhai, M. Y., Yu, R. H., Zhang, S. F., &Zhai, J. H. (2012, July). Feature selection based on extreme learning machine. In *Machine Learning and Cybernetics (ICMLC), 2012 International Conference on (Vol. 1, pp. 157-162).IEEE*.
- [10]. Ahmad, I., Basher, M., Iqbal, M. J., &Raheem, A. Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection. *IEEE Access*, vol.6, (2018).
- [11]. Gao, Y., & Chen, T. (2017). An Efficient Extreme Learning Machine Based on Fuzzy Information Granulation. *International Journal of Emerging Technologies in Learning (iJET)*, 12(06), 161-170.
- [12]. Rafsanjani, M. K., &Varzaneha, Z. A. (2013). Intrusion Detection By Data Mining Algorithms: A Review. *Journal of New Results in Science*, 2(2).
- [13]. Aung, Y. Y., & Min, M. M. (2017, June). An analysis of random forest algorithm based network intrusion detection system. In *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), 2017 18th IEEE/ACIS International Conference on (pp. 127-132). IEEE*.
- [14]. Cheng, C., Tay, W. P., & Huang, G. B. (2012, June). Extreme learning machines for intrusion detection. In *Neural networks (IJCNN), the 2012 international joint conference on (pp. 1-8).IEEE*.
- [15]. Farnaaz, N., &Jabbar, M. A. (2016). Random forest modeling for network intrusion detection system. *Procedia Computer Science*, 89, 213-217.
- [16]. Gao, Y., & Chen, T. (2017). An Efficient Extreme Learning Machine Based on Fuzzy Information Granulation. *International Journal of Emerging Technologies in Learning (iJET)*, 12(06), 161-170.
- [17]. Hasan, M. A. M., Nasser, M., Ahmad, S., &Molla, K. I. (2016). Feature selection for intrusion detection using random forest. *Journal of information security*, 7(03), 129.
- [18]. Sharma, N., & Gaur, B. (2016). An approach for efficient intrusion detection for KDD dataset: a survey. *International Journal of Advanced Technology and Engineering Exploration*, 3(18), 72.
- [19]. Levonevskiy, D. K., Fatkueva, R. R., &Ryzhkov, S. R. (2015, May). Network attacks detection using fuzzy logic. In *Soft Computing andMeasurements (SCM), 2015 XVIII International Conference on (pp. 243-244). IEEE*.
- [20]. Shanmugavadivu, R., &Nagarajan, N. (2011). Network intrusion detection system using fuzzy logic. *Indian Journal of Computer Science and Engineering (IJCSE)*, 2(1), 101-111.
- [21]. Aggarwal, P., & Sharma, S. K. (2015). Analysis of KDD dataset attributes-class wise for intrusion detection. *Procedia Computer Science*, 57, 842-851.