A Literature Survey and Comprehensive Study of Intrusion Detection

Sravan Kumar Jonnalagadda
Assistant Professor in Department of CSE
DMSSVH College of Engineering
Machilipatnam

Ravi Prakash Reddy I, Ph.D
Professor & Head in Department of IT
GNITS ShaikPet
Hyderabad

ABSTRACT

With the rapid expansion of computer usage and computer network the security of the computer system has became very important. Every day new kind of attacks are being faced by industries. As the threat becomes a serious matter year by year, intrusion detection technologies are indispensable for network and computer security. A variety of intrusion detection approaches be present to resolve this severe issue but the main problem is performance. It is important to increase the detection rates and reduce false alarm rates in the area of intrusion detection. In order to detect the intrusion, various approaches have been developed and proposed over the last decade. In this paper, a detailed survey of intrusion detection based various techniques has been presented. Here, the techniques are classified as follows: i) papers related to Neural network ii) papers related to Support vector machine iii) papers related to K-means classifier iv) papers related to hybrid technique and v) paper related to other detection techniques. For comprehensive analysis, detection rate, time and false alarm rate from various research papers *have been taken.*

Keywords

Intrusion detection, clustering, classifier, detection rate, false alarm rate

1. INTRODUCTION

By means of the extensive application of computer networks, the quantity of attacks has developed widely, and numerous hacking tools and intrusive methods have been appeared. Inside a network, one way of dealing with suspicious actions is by utilizing an intrusion detection system (IDS). By investigating different data records watched in processes on the network [1] [2] Intrusion detection attempts to detect computer attacks. To work out network security problems it is one of the essential ways. The two key signs to assess intrusion detection systems (IDS) [3] are Detection precision and detection stability. A lot of research has been done in order to improve the detection precision and detection stability.

Commonly, Intrusion explained an act of encroaching or infringing the reliability, confidentiality or avoiding the accessibility of a resource [4]. Through internet, Intrusions Detection Systems discovers illegal or malicious assaults over a computer system which happens mainly. By the safety and hope of a system, these assaults can be compromised. To perceptive files, these harasses can acquire quite a few forms like network attack against vulnerable services, data driven attacks on applications, host based attacks such as privilege escalation, illegal logins and access. IDSs can be categorized as misuse detectors or anomaly detectors by sorting out broadly based on their models of detection. Mishandling detectors rely on understanding the models of known attacks [5, 6], whereas irregularity detection develops user profiles as the basis of detection, and sorts the distinctiveness of the unexpected from the normal ones as incursion [5, 6, 7, and 8].

Conversely, the quantity and kinds of the intrusions increase radically as the speed and difficulty of networks expand quickly, particularly when these networks are unlock to the public Web. Therefore, it is becoming hard for any presented intrusion detection system to suggest a trustworthy repair with the varying technology and the exponential growth of Internet traffic it has been created that a behavioral model exists in the attacks that can be educated from former study. Various machine learning (ML) algorithms, such as Neural Network [9], Support Vector Machine [10], Genetic Algorithm [11], Fuzzy Logic [12, 13], and Data Mining [14], etc have been broadly used to the huge volume of complex and active dataset to detect known and unknown intrusions so as to identify intrusion activities.

2. SURVEY OF INTRUSION DETECTION BASED ON DIFFERENT TECHNIQUES

This section presents an extensive study over the various intrusion detection classifier techniques and other techniques. A number of research papers regarding to intrusion detection are discussed below and are widely classified into i) papers related to Neural network ii) papers related to Support vector machine iii) papers related to K-means classifier iv) papers related to hybrid technique and v) paper related to other detection techniques.

2.1 Neural network based intrusion detection

A brief review of two techniques related with neural network based intrusion detection is discussed in this section. In 2009 a lot of papers have been presented to represent the neural network based intrusion detection. Some of the papers have been discussed below. The following approach was presented in the year 2009. The concept of anomaly detection and use both neural network (NN) and decision tree (DT) for intrusion detection has been improved by Marjan Bahrololum *et al.* [15]. At the same time DTs were extremely victorious in discovering known attacks, NNs were more exciting to detect unknown attacks. They designed the system using together with DT and mixture of unsupervised and supervised NN for Intrusion Detection System (IDS). Known attacks were familiar with a quick implementation time by concerning DT. For collecting attacks into smaller categories, unknown attacks was identified by pertaining the unsupervised neural network based on hybrid of Self Organizing Map (SOM) and supervised NN based on Back propagation for complete grouping.

In the same year 2009 M. Bahrololum *et al.* published a paper to plan the system using a hybrid of misuse and irregularity detection for training of normal and attack packets respectively[16]. The used method for attack training was the mixture of unsupervised and supervised Neural Network (NN) for Intrusion Detection System. Attacks was categorized into smaller categories taking into consideration their similar features by the unsupervised NN based on Self Organizing Map (SOM), and followed by unsupervised NN based on Back propagation was utilized for grouping. Known packets were recognized fast by misuse approach and unknown attacks will be able to spot by this method.

2.2 Support vector machine based intrusion detection

A brief review of support vector machine classifier related is discussed in this section. In the period 2007-2012, a lot of papers have been presented to represent the Support vector machine based intrusion detection. Some of the papers have been discussed below. A revise for improving the training time of SVM has been presented by Latifur Khan *et al.* [17], particularly when contracting with large data sets using hierarchical clustering analysis in 2007. For gathering, they utilized the Dynamically Growing Self-Organizing Tree (DGSOT) algorithm since it had verified to triumph over the disadvantages of traditional hierarchical clustering algorithms (e.g., hierarchical agglomerative clustering). Among two classes, clustering analysis assisted discover the boundary points, which were the most capable data points to coach SVM. Using the clustering arrangement created by the DGSOT algorithm, they offered an approach of amalgamation of SVM and DGSOT, which in progress with a first training set and enlarge it slowly. In terms of precision loss and training time gain by means of a single bench-mark real data set they match up to their approach with the Rocchio Bundling technique and casual choice.

In advanced to the above approach in the year 2011, Iftikhar Ahmad,Azween *et al.* have proposed a paper to surmount presentation issues an optimized interference detection mechanism by means of soft computing techniques [18]. The KDD-cup dataset was applied that was a benchmark for assessing the safety identification mechanisms. To change the key in models into a feature space the Principal Component Analysis (PCA) was applied. Selecting of a suitable quantity of principal components was an important problem. As an alternative of using conventional method, Genetic Algorithm (GA) was applied in the optimum choice of principal components accordingly. The Support Vector Machine (SVM) was employed for categorization reason. In addition, a proportional study was prepared with presented approaches. Therefore, the technique presented optimal interference detection mechanism was proficient to minimize amount of features and maximize the identification rates.

S. Ganapathy *et al.* [19] have proposed an intelligent multi level classification technique for effective intrusion detection in Mobile Ad-hoc Networks in the same year. The algorithm used a combination of a tree classifier which used a labeled training data and an Enhanced Multiclass SVM algorithm. Moreover, an effective preprocessing technique had been proposed and implemented in this work in order to improve the detection accuracy and to reduce the processing time. As a way of dealing with conditions for independent labeling of regular traffic where class distribution does not present the imbalance necessary for SVM algorithms, an approach has been presented by Carlos A. Catania et al. [20] in 2012. In such case, the self-governing labeling process was made by SNORT, a misuse-based intrusion detection system. The use of the planned autonomous labeling approach not only outperforms presented SVM alternatives that was shown by the experiments conducted on the 1998 DARPA dataset however attains developments over SNORT itself also under some attack distributions.

2.3 K-means algorithm based intrusion detection

In this section, we discuss the different papers that utilize k-means algorithm. In 2003-2004 some papers presented to represent the K-means algorithm based intrusion detection. Some of the papers have been discussed below. In the year 2003, a K-means based clustering algorithm, named Y-means, for incursion detection has been offered by Yu Guan *et al.* [21]. Y-means surmounts two failings of K-means: quantity of clusters dependency and degeneracy. The original number of clusters was no longer serious to the collecting results in the Y-means algorithm. A suitable number of clusters were divided by a data set routinely. This was one of the benefits of the Y-means algorithm for intrusion detection. The unprocessed log data of information systems can directly be applied as training data with-out being physically labeled was the another advantage.

In the very next year 2004, using the K-means clustering algorithm a technique have built up by K. M. Faraoun and A. Boukelif [22] to improve the learning capacities and decrease the computation strength of a competitive learning multi-layered neural network. Through a back propagation learning means the recommended model used multi-layered network structural design. To decrease the amount of examples to be offered to the neural network, the K-means algorithm was initially used to the training dataset by automatically choosing a most favorable set of samples. The acquired results showed that the suggested technique executes specially in terms of both precision and computation time when pertained to the KDD99 dataset match up to a normal learning schema that utilized the full dataset.

2.4 Hybrid technique based intrusion detection

In the period 2007-2012, a lot of papers have been presented to represent the hybrid technique based intrusion detection. Some of the papers have been discussed below. For categorizing irregular and normal activities in a computer network, a dynamic electronic circuit, and a motorized mass-beam system, Shekhar R. Gaddam *et al.* [23] have offered a method to flow k-Means grouping and the ID3 decision tree learning methods and this paper has been published in the year 2007. By means of Euclidean distance resemblance the k-Means grouping method first divided the training cases into k clusters. On every cluster, an ID3 decision tree on behalf of a density region of normal or anomaly instances has been constructed. By studying the subgroups inside the cluster the decision tree on every cluster purified the decision boundaries. The conclusions of the k-Means and ID3 methods were united using two rules to get a concluding decision on classification. The two rules are: 1) the Nearest-neighbor rule and 2) the Nearest-consensus rule. Testing were executed by them on three data sets: 1) Network Anomaly Data (NAD), 2) Duffing Equation Data (DED), and 3) Mechanical System Data (MSD), which enclosed measurements from three separate application domains of computer networks, an electronic circuit applying a forced Duffing Equation, and a mechanical system, correspondingly.

In 2008 a paper was presented by S. Anil Kumar and Dr. V. NandaMohan. An approach which was a grouping of three techniques containing two machine-learning paradigms has been built up [24]. In order to organize an effective intrusion detection system, K-Means Clustering, Fuzzy Logics and Neural Network techniques were arranged. The existence of high rate of false alerts makes unnecessary interference of human analyst even out of the numerous problems in the traditional techniques of Intrusion Detection Systems. To differentiate the nature of such alerts and commence sufficient actions, the human analysts in turn carry out a serious analysis continually. In order to remove the unnecessary interference of human analyst in such events, the approach offered exposed the advantage of converging K-Means-Fuzzy-Neuro techniques. The technique was examined using mass of background knowledge sets in DARPA network traffic datasets. The experimental results leave outstanding development in reducing the false alarms as well to increased capacity to confine intrusion packets that were no alike to the ones in the training datasets.

In the year 2010, to work out the problem and aid IDS get higher detection rate, less false positive rate and stronger constancy based on ANN and fuzzy clustering an approach, named FC-ANN has been offered by Gang Wang *et al.* [25]. The common process of FC-ANN was as follows: initially fuzzy clustering system was utilized to produce dissimilar training subsets. Based on different training subsets, dissimilar ANN models were educated to put together different base models consequently. At last, to summative those results a meta-learner, fuzzy aggregation module, were utilized. Investigational results on the KDD CUP 1999 dataset proved that their offered approach, FC-ANN, outperforms BPNN and other famous methods such as decision tree, the naïve Bayes in terms of detection accuracy and finding constancy.

In the very next year, a hybrid plan that united the advantages of deep belief network and support vector machine has been brought in by Mostafa A *et al.* [26]. Hybridization plan has been related and a submission of interference detection imaging had been selected to see their capability and correctness to categorize the interference into two products: normal or attack, and the attacks fall into four classes; R2L, DoS, U2R, and Probing. In order to cut the dimensionality of the feature sets they employed deep belief network at first. This was pursued by a support vector machine to categorize the interference into five product; Normal, R2L, DoS, U2R, and Probing. Tests on NSL-KDD dataset were offered by them to assess the performance of their approach and demonstrated that the on the whole accuracy proposed by the utilized approach was high.

In 2012, An Advancement in the performance of network intrusion detection systems has been investigated to find out the possibility by means of assembling algorithms by Iwan Syarif *et al.* [27]. An ensemble of three different methods such as bagging, boosting and stacking was used with the purpose of improving the accuracy and reducing the false positive rate. A variety of four different data mining algorithms, naïve bayes, J48 (decision tree), JRip (rule induction) and iBK (nearest neighbour) were used as base classifiers for those ensemble methods. Their experiment proved that the prototype which activates four base classifiers and three ensemble algorithms achieves in detecting known intrusions, but it has been unsuccessful in detecting intrusions with the accuracy rates. Considerably improvement in the accuracy is not able to be reached by the usage of bagging, boosting and stacking.

A hybrid clever approach using combination of classifiers so as to formulate the result wisely has been presented in the same year by Mrutyunjaya Panda *et al.* [28]. This method enabled the overall performance of the resultant model. The supervised or un-supervised data filtering with classifier or cluster first on the whole training dataset is followed and then the output was given to another classifier to classify the data. This was the common practice. 2-class classification approach along with 10-fold cross validation method is used to produce the final classification results in conditions of normal or else intrusion. Their proposed method was competent by way of high detection rate and low false alarm rate. This was proved by the results of experiments on NSL-KDD dataset which is an improved version of KDDCup 1999 dataset.

2.5 Other classifier based intrusion detection

Here, we discuss about the different papers of various intrusion detection techniques. In the year 2003-2012 a lot of papers have been presented to represent the classifier based intrusion detection. Some of the papers have been discussed below.

In the year 2003, a model recognition approach to set of connections intrusion detection based on the combination of manifold classifiers has been developed by Giorgio Giacinto *et al.* [29]. Five decision mixture methods were evaluated by experiments and their presentations were compared. The potentialities of classifier fusion for the improvement of competent incursion detection systems were assessed and argued. The statement results explained that the MCS approach offers an improved trade-off among generalization abilities and false alarm generation than that offered by an individual classifier educated on the general feature set.

Suseela T *et al* have presented a multilevel hierarchical Kohonen Net (K-Map) [30] for an intrusion detection system in 2005. Every level of the hierarchical map was modeled as a straightforward winner-take-all K-Map. The computational effectiveness is the major advantage of this multilevel hierarchical K-Map. Statistical anomaly detection methods such as nearest neighbor approach, K-means clustering or probabilistic analysis that employed distance computation in the feature space to identify the outliers. However this multilevel hierarchical K-Map’s approach does not engage expensive point-to-point computation in grouping the data into clusters. The reduced network size is an additional advantage. The categorization potential of the K-Map on chosen dimensions of data was used to set for detecting anomalies. Sub-sets selected aimlessly that enclose both attacks and normal records from the KDD Cup 1999 benchmark data were used to guide the hierarchical net.

In 2007, an evolutionary soft computing approach for intrusion has been commenced by Adel Nadjaran Toosi and Mohsen Kahani [31] and was effectively explained its utility on the training and testing subset of KDD Cup 99 dataset. For interference detection, the ANFIS network was utilized as a neuro-fuzzy classifier. Without the help of human experts ANFIS was talented of generating fuzzy rules. Furthermore, to find out the number of rules and membership functions with their initial locations for enhanced classification subtractive clustering had been used. To create the system more prevailing for attack detection using the fuzzy inference approach, a fuzzy decision-making engine was built up. In the end, they planned a system to utilize genetic algorithms to optimize the fuzzy decision-making engine. The planned system was successful in discovering various intrusions in computer networks were shown by the experimentation results.

Again in the same year 2007, an approach to network interference detection, based merely on a hierarchy of Self-Organizing Feature Maps has been inspected by H. Gunes Kayacik *et al.* [32]. Establishing immediately how far such an approach was full in practice was their principle interest. For doing this, the KDD benchmark dataset from the International Knowledge Discovery and Data Mining Tools Competition was utilized. Extensive analysis was conducted In order to concentrate on the importance of the features used, the division of training data and the complication of the architecture. Using unsubstantiated learning in comparison to results statement formerly, they explained that most excellent presentation was obtained by means of a two-layer SOM hierarchy, based on all 41-features since the KDD dataset.

Yang Li and Li Guo have presented a paper Based on TCM-KNN (Transductive Confidence Machines for K-Nearest Neighbors) machine learning algorithm and dynamic education based training data selection scheme a supervised network intrusion detection method in the year 2007 [33]. It was successfully identified anomalies with elevated detection rate, low false positives under the condition of utilizing much less chosen data as well as selected features for training in association with the traditional managed intrusion detection techniques. The recommended method was more tough and successful than the state-of-the-art intrusion detection methods which were explained by a chain of experimental results on the familiar KDD Cup 1999 data set.

In advance to the above concept in the year 2009, a network based anomaly detection system that makes use of a hierarchy of SOMs has been offered by Saroj Kumar Panigrahy *et al.* [34]. By means of a controllable rate of false alarms the system was set up to identify very soon over 60% of the attacks. Even though the result of this job was construed with warning, it was recommended that the arrangement offered carry out comparably to few of the superior systems that took part in the DARPA Intrusion Detection Evaluation. The system was also not at all trained on the complete training dataset, sense that it could not have had a opening to learn the full sort of usual behavior and it was not tested on the full test dataset, i.e., it may not have come across few of the more not easy attacks.

In development of the above concept in the year 2010 by means of enhanced self adaptive Bayesian algorithm (ISABA), an approach to the vigilant classification to lessen false positives in intrusion detection has been offered by Dewan Md. Farid and Mohammad Zahidur Rahman [35]. The recommended approach used to the security domain of anomaly based network intrusion detection, which properly categorized dissimilar types of attacks of KDD99 benchmark dataset with elevated arrangement rates in small reply time and decrease false positives with restricted computational assets.

In 2010, an intrusion detection system was constructed by Muna M. Taher Jawhar and Monica Mehrotra [36] using hamming and MAXNET Neural Network for recognizing attack class in the network traffic. One more approach based on Multilayer Perceptrons (MLP) network has been derived and the evaluation of the system is done by comparing the results of the two approaches. The results of experiments confirmed that the designed models were capable in terms of accuracy and computational time of real word intrusion detection. Defense Advanced Research Projects Agency (DARPA) intrusion detection evaluation datasets provided the necessary training and testing data. The utilization of Self Organizing Maps for building an Intrusion Detection System is well described by Mr. Vivek A. Patole *et al.* [37] in the same year. The system architecture and the flow diagram for the SOM have been explained. The advantages and disadvantages of the algorithm have also been presented. Their real experiments proved that even a simple map, when trained on usual data, could detect the anomalous features of both buffer overflow intrusions to which they are exposed.

Also, an interference detection system with Bayesian probability has been developed by Hesham Altwaijry and Saeed Algarny [38]. To categorize possible intrusions, the structure developed was an adolescent Bayesian classifier that was utilized. The structure was taught a priori by means of a subset of the KDD dataset. The capability of Bayesian classifier was to distinguish the interference with a better detection rate. With Conditional Random Fields and Layered Approach, Kapil Kumar Gupta et al. [39] had lectured two topics of Accuracy and Efficiency. They revealed high attack detection precision was obtained by means of Conditional Random Fields and high competence by applying the Layered Approach. Their recommended system based on Layered Conditional Random Fields outperforms other familiar means such as the decision trees and the naive Bayes was shown by the experimental results on the benchmark KDD ’99 intrusion data set. Detection accuracy for our method is confirmed by statistical Tests which induced higher confidence level. In conclusion, it has been shown that the system was healthy and is capable of handling noisy data with no compromise on performance.

In development to the above approach in the year 2011, for independent rule creation, a multi-modal genetic algorithm solution has been discovered by Todd Vollmer et al. [40]. Relatively than the development of rules to give a solution for interference detection this algorithm spotlight on the process of making rules once interference had been recognized. The algorithm was explained on irregular ICMP network packets (input) and Snort rules (output of the algorithm). According to a fitness value output rules were arranged and any replacements were detached. These rules were precise to the packets and formed only four false positives from 33,804 test packets that were shown by testing. In advance to the above technique in 2011, different ways such as to categorize normal and intrusive actions or to extract interesting intrusion models, Machine learning techniques are often used to interference detection problems. Besides providing interference classification capacities self learning rule-based systems can ease area specialists from the complicated task of hand crafting signatures.

A genetic-based signature knowledge system that was adaptively and energetically learns signatures of both standard and interfering actions from the network traffic has been developed by Kamran Shafi, Hussein A. Abbass [41] to this end. The assessment of their systems to actual time network traffic which was incarcerated from a university departmental server was completed by them. By combining real background traffic with attacks replicated in a controlled environment an attitude was developed to put together fully labeled interference detection data set. Proper for a managed learning classifier system and other associated machine learning systems apparatus were improved to preprocess the unrefined network data into quality vector format. The signature extraction system was then pertained to this data set and the results were argued. Detecting payload based attacks they illustrated that even easy feature sets were assisted.

In the very next year in 2012, Iwan Syarif et al. [42] have illustrated the compensation of utilizing the variance detection approach over the mishandling detection technique in detecting unknown network intrusions or attacks. When applied to anomaly detection it also examined the presentation of different grouping algorithms. We have five different clustering algorithms: k-Means, improved k-Means, k-Medoids, EM clustering and distance-based outlier detection algorithms were utilized. Their testing showed that mishandling detection techniques, which executed four dissimilar classifiers (naïve Bayes, rule induction, decision tree and nearest neighbor) unsuccessful to detect network traffic, which enclosed a large number of unknown interferences.

In the same year in 2012, an intangible model for identifying and mitigating Distributed Denial-of-Service (DDoS) attacks and its incomplete achievement has been offered by Sajal Bhatia et al. [43]. To identify DDoS attacks, to distinguish them from like looking FEs, and to bring about source IP based mitigation strategies upon attack identification an assembly of network traffic and MIB server load data analysis was used in the mold. The testing and presentation assessment of the suggested model was performed by means of artificial network traffic, intimately on behalf of real-world DDoS attacks and FE traffic, a produced using a software-based traffic generator developed. Prasanta Gogoi et al. [44] have suggested an actual dataset to modernize this critical inadequacy. A test bed has been set up by them to begin network traffic of both attack as well as standard nature by means of attack tools. The network traffic in sachet and flow format was incarcerated by them. To produce a featured dataset the incarcerated traffic was sorted out and preprocessed. For investigate purpose the dataset was made accessible. They have High-level study of the KDD Cup 1999 and NSL-KDD datasets which are offered by them.

3. COMPREHENSIVE ANALYSIS AND DISCUSSIONS

This section presents a comprehensive analysis of various methods such as neural network based, support vector machine based, k-means based, hybrid technique based and other techniques for intrusion detection. As we have discussed earlier, major classification is done based on intrusion detection with respect to the detection rate, time and false alarm rate achieved by the different methods.

3.1 Neural network based intrusion detection

Table: 1 shows the comparison of intrusion detection using neural network technique. From the table, we can find that M. Bahrololum *et al.* [16] achieved detection rate of 93.8%.

Table 1: Comparison of intrusion detection using neural network

S. No	Authors	Method/Algorithms	Detection rate (%)	False alarm rate
1	M. Bahrololum *et al.* [15]	Neural network based on Hybrid Self Organizing Map (SOM)	91.3 %	-
2	M. Bahrololum *et al.* [16]	Neural network based on K Self Organizing Map (K- SOM)	93.8 %	-

3.2 Support vector machine based intrusion detection

Table: 2 shows the comparison of some intrusion detection using support vector machine technique. From the table, it is found that Ahmad *et al.* [18] achieved a better detection rate 99.6% and false alarm 0.4% compared to others. Also, S. Ganapathy *et al.* [19] attained 98.51% detection rate which is slightly down when compared to Ahmad *et al.* [18] approach.

Table 2: Comparison of intrusion detection using support vector machine

S.No	Authors	Method/Algorithms	Detection rate (%)	False alarm rate
1	Carlos A. Catania *et al.* [20]	Support vector machine	87.02 %	-
2	Ahmad *et al.* [18]	Support vector machine	99.6%	0.4%
3	S. Ganapathy *et al.* [19]	Support vector machine	98.51 %	-

3.3 K-means based intrusion detection

Table: 3 shows the comparison of some intrusion detection using support vector machine technique. From the table, it is found that K. M. Faraoun and A. Boukelif [22] achieved a better detection rate 92% compared to Yu Guan *et al.* [21] technique [89.9%]. On the other hand, false alarm rate is found to be 6.21% in K. M. Faraoun and A. Boukelif method when compared to Yu Guan *et al* method.

3.4 Hybrid technique based intrusion detection

Table: 3 shows the comparison of intrusion detection using two different techniques. From the table, it can be found that Iwan Syarif *et al.* [27] achieved a better detection rate values compared to others. In Mrutyunjaya Panda *et al.* [28] the detection rate is found to be 99.5% and the false alarm rate is found to be 0.1 which is lower when compared to other techniques. Technique Shekhar R. Gaddam *et al.* [23] achieved a better false alarm rate values compared to others and the detection rate is found to be 96.24%. In Gang Wang *et al.* [25] technique the detection rate is found to be 96.71% and in Mostafa A *et al.* [26] technique the detection rate is found to be 92.84% which is very low when compared to the technique proposed by Iwan Syarif *et al.*

Table 3: Comparison of intrusion detection using hybrid techniques

S. No	Authors	Method/ Algorithms	Detection rate (%) (Accuracy)	False alarm rate (%)
1	Mrutyunjaya Panda *et al.* [28]	Decision trees, principal component analysis, SPegasos (Stochastic variant of Piramol estimated sub-gradient solver in SVM)	99.5%	0.1%
2	Shekhar R. Gaddam *et al.* [23]	K-Means and Clustering and ID3 Decision Tree Learning Methods	96.24%	0.03%
3	Gang Wang *et al.* [25]	Fuzzy C-means-Artificial Neural Network	96.71%	-
4	Mostafa A *et al.* [26]	Support vector machine and Deep Belief Network	92.84%	-
5	Iwan Syarif *et al.* [27]	Naive Bayes, J48 (decision tree), JRip (rule induction) and iBK (nearest neighbour),	99.80%	-

3.5 Other technique based intrusion detection

Table: 4 shows the comparison of intrusion detection using various techniques. In the period 2003-2012, several intrusion detection based techniques were developed. From the table, it is found that Dewan Md. Farid and Mohammad Zahidur Rahman [35] achieved a better detection rate value (99.82%) compared to others. Adel Nadjaran Toosi and Mohsen Kahani [31] have attained 95.3% detection rate and the false alarm rate is 1.9%.This is lower when compared with previous technique. In Muna M. Taher Jawhar and Monica Mehrotra [36] technique the detection rate is 95% and the false alarm rate is found to be 4.94%%. In 2007, H. Gunes Kayacik *et al.* [32] have achieved the detection rate of 99.8% and the false alarm rate of 10.6%.. In 2010 Hesham Altwaijry and Saeed Algarny [38] have achieved 99.03% detection rate. In next year, Kamran Shafi, Hussein A. Abbass [41] have attained 97.65% detection rate. At the same time, Suseela T *et al.* [30] achieved 99.63% detection rate and the false alarm rate is 0.34%. Iwan Syarif *et al.* [42] has achieved detection rate is 99.56% and the false alarm rate of 0.40% in the year 2012.

Table 4: Comparison of intrusion detection using various algorithm/techniques

S.No	Authors	Method/Algorithms	Detection rate (%)	False alarm rate
1	Adel Nadjaran Toosi and Mohsen Kahani [31]	Neuro-fuzzy classifier	95.3%	1.9%
2	Dewan Md. Farid and Mohammad Zahidur Rahman [35]	Improved Self Adaptive Bayesian Algorithm	99.82 %	-
3	Muna M. Taher Jawhar and Monica Mehrotra [36]	Hamming and MAXNET	95%	4.94%
4	H. Gunes Kayacik *et al.* [32]	Kohonen's Self-Organizing Feature Map (SOM) algorithm	99.8%	10.6%
5	Iwan Syarif *et al.* [42]	Unsupervised clustering	99.56 %	0.40%
6	Hesham Altwaijry and Saeed Algarny [38]	Naive Bayesian classifier	99.03 %	-
7	Kamran Shafi, Hussein A. Abbass [41]	genetic-based learning	97.65 %	-
8	Suseela T *et al.* [30]	Single-Layer Winner-Take-All Kohonen Map	99.63 %	0.34%
9	Kapil Kumar Gupta *et al.* [39]	Conditional Random Fields and Layered Approach	98.62 %	17%
10	Prasanta Gogoi *et al.* [44]	Packet Network Traffic Feature Extraction	99.29 %	0.71%

4. CONCLUSION

Network Intrusion Detection System is a latest kind of defense technology which is one of the vibrant areas in network security. In recent years many techniques are available for intrusion detection. In this paper, a detailed survey of important techniques based on intrusion detection is presented. Also the classification of the techniques based on neural network, k-means, hybrid techniques, support vector machine etc., is provided.

Detection rate and false alarm rate are considered for comprehensive analysis.

5. REFERENCES

[1] Anderson, J. P.” Computer security threat monitoring and surveillance,” Technical Report, Fort Washington, PA, USA, 1980.

[2] Endorf. C, Schultz, E., & Mellander, J.”Intrusion detection and prevention,” California: McGraw-Hill, 2004.

[3] Silva, L. D. S., Santos, A. C., Mancilha, T. D., Silva, J. D., & Montes, A, “Detecting attack signatures in the real network traffic with ANNIDA,”Expert Systems with Applications, vol.34, no.4, pp.2326–2333, 2008.

[4] Heady R., Luger G., Maccabe A., and Servilla M. “The architecture of a Network level intrusion detection system,” Technical Report, CS90-20, Dept. of Computer Science, University of New Mexico, Albuquerque, NM 87131,1990.

[5] Denning D. “An Intrusion-Detection Model,” IEEE Transactions on Software Engineering, Vol. SE-13, No. 2, pp.222-232, 1987.

[6] Kumar S., Spafford E. H. “An Application of Pattern Matching in Intrusion Detection,” Technical Report CSD-TR-94-013. Purdue University, 1994.

[7] Ryan J., Lin M-J., Miikkulainen R. (1998) “Intrusion Detection with Neural Networks,” Advances in Neural Information Processing Systems, Vol. 10, and Cambridge, MA: MIT Press.

[8] Terran lane, Carla E. Brodley, Temporal Sequence Learning and Data Reduction for anomaly Detection, Vol. 2, No. 3, pp. 295- 331,August 1999.

[9] Cannady J., “Artificial neural networks for misuse detection,” Proceedings of the ’98 National Information System Security Conference (NISSC’98), Arlington: Virginia Press, pp. 443-456, 1998.

[10] Shon T, Seo J, and Moon J, “SVM approach with a genetic algorithm for network intrusion detection,” Proceedings of the 20th International Symposium on Computer and Information Sciences (ISCIS 05), Berlin: Springer Verlag, pp. 224-233, 2005.

[11] Yu Y, and Huang Hao, “An ensemble approach to intrusion detection based on improved multi-objective genetic algorithm,” Journal of Software, Vol.18, No.6, June 2007, pp.1369-1378.

[12] J. Luo, and S. M. Bridges, “Mining fuzzy association rules and fuzzy frequency episodes for intrusion detection,” International Journal of Intelligent Systems, pp. 687-703, 2000.

[13] Pavel Kromer, Jan Platos, Vaclav Snasel, Ajith Abraham,” Fuzzy Classification by Evolutionary Algorithms,” IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp.313 - 318, 2011.

[14] W.K. Lee, and S. J. Stolfo, “A data mining framework for building intrusion detection model,” Proceedings of the IEEE Symposium on Security and Privacy, Oakland, CA: IEEE Computer Society Press, pp. 120-132, 1999.

[15] Marjan Bahrololum, Elham Salahi, Mahmoud Khaleghi,"An Improved Intrusion Detection Technique based on two Strategies Using Decision Tree and Neural Network,” Journal of Convergence Information Technology,Vol.4, No.4, December 2009.

[16] M. Bahrololum, E. Salahi and M. Khaleghi, “Anomaly intrusion detection design Using Hybrid of Unsupervised and supervised neural Network,” International Journal of Computer Networks & Communications (IJCNC), Vol.1, No.2, July 2009.

[17] Latifur Khan, Mamoun Awad, Bhavani Thuraisingham,"A new intrusion detection system using support vector machines and hierarchical clustering,” Journal of VLDB Journal, vol.16, pp.507-521, 2007.

[18] Iftikhar Ahmad,Azween Abdullah,Abdullah Alghamdi,Muhammad Hussain,"Optimized intrusion detection mechanism using soft computing techniques,"Telecommun System,2011.

[19] S. Ganapathy, P. Yogesh, and A. Kannan,” An Intelligent Intrusion Detection System for Mobile Ad-Hoc Networks Using Classification Techniques,” Advances in Power Electronics and Instrumentation Engineering, Communications in Computer and Information Science Vol.148, pp 117-122,2011.

[20] Carlos A. Catania, Facundo Bromberg, Carlos Garcıa Garino,"An Autonomous Labeling approach to Support Vector Machines Algorithms for Network Traffic Anomaly Detection," Preprint submitted to Expert Systems with Applications, Vol.39, no.2, pp.1822-1829, February, 2012.

[21] Yu Guan, Nabil Belacel and Ali A. Ghorbani,"Y-Means: A Clustering Method for Intrusion Detection," Canadian Conference on Electrical and Computer Engineering, vol.2, pp. 1083- 1086, 2003.

[22] K. M. Faraoun and A. Boukelif,"Neural Networks Learning Improvement using the K-Means Clustering Algorithm to Detect Network Intrusions,” International Journal of Computational Intelligence, Vol.3, no.2, 2005.

[23] Shekhar R. Gaddam, Vir V. Phoha and Kiran S. Balagani,"K-Means+ID3: A Novel Method for Supervised Anomaly Detection by Cascading K-Means Clustering and ID3 Decision Tree Learning Methods,"IEEE Transactions On Knowledge And Data Engineering, Vol. 19, No. 3, March 2007.

[24] [24] K.S. Anil Kumar and Dr. V. Nanda Mohan, “Novel Anomaly Intrusion Detection Using Neuro-Fuzzy Inference System, "IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.8, August 2008.

[25] [25] Gang Wang, Jinxing Hao, Jian Ma, Lihua Huang, "A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering, “Expert Systems with Applications, 2010.

[26] Mostafa A. Salama, Heba F. Eid, Rabie A. Ramadan, Ashraf Darwish, and Aboul Ella Hassanien,"Hybrid Intelligent Intrusion Detection Scheme," Soft Computing in Industrial Applications Advances in Intelligent and Soft Computing,Vol.96, pp.293-303,2011.

[27] Iwan Syarif, Ed Zaluska, Adam Prugel-Bennett, Gary Wills, “Application of bagging, boosting and stacking to intrusion detection, “Proceedings of the 8th international

conference on Machine Learning and Data Mining in Pattern Recognition,pp.593-602,2012.

[28] Mrutyunjaya Panda, Ajith Abraham, Manas Ranjan Patra.a “A Hybrid Intelligent Approach for Network Intrusion Detection,” International Conference on Communication Technology and System Design, Procedia Engineering, vol. 30, pp.1-9,2012.

[29] Giorgio Giacinto, Fabio Roli, and Luca Didaci, “Fusion of Multiple Classifiers for Intrusion Detection in Computer Networks,” Journal of Pattern Recognition Letters, vol.24, pp.1795-1803, 2003.

[30] Suseela T. Sarasamma, Qiuming A. Zhu, and Julie Huff “Hierarchical Kohonenen Net for Anomaly Detection in Network Security,” IEEE Transactions on Systems, Man, And Cybernetics-Part B: Cybernetics, Vol. 35, No. 2, April 2005.

[31] Adel Nadjaran Toosi, Mohsen Kahani, “A new approach to intrusion detection based on an evolutionary soft computing model using neuro-fuzzy classifiers,” Computer Communications vol.30, pp.2201–2212, 2007.

[32] H. Gunes Kayacik, A. Nur Zincir-Heywood, Malcolm I. Heywood,” A Hierarchical SOM based Intrusion Detection System, “Engineering Applications of Artificial Intelligence,Vol.20,no.4, pp.439-451,June 2007.

[33] Yang Li, Li Guo, “An active learning based TCM-KNN algorithm for supervised network intrusion detection,” computers & security, vol.26, pp.459-467, 2007.

[34] Saroj Kumar Panigrahy, Jyoti Ranjan Mahapatra, Jignyanshu Mohanty and Sanjay Kumar Jena, “Anomaly Detection in Ethernet Networks using Self Organizing Maps,” Department of Computer Science,2009.

[35] Dewan Md. Farid, Mohammad Zahidur Rahman, “Anomaly Network Intrusion Detection Based on Improved Self Adaptive Bayesian Algorithm,” Journal of Computers, Vol. 5, No. 1, January 2010.

[36] Muna M. Taher Jawhar and Monica Mehrotra, “Anomaly Intrusion Detection System using Hamming Network Approach,” International Journal of Computer Science & Communication,Vol. 1, No. 1, pp. 165-169, January-June 2010.

[37] Mr. Vivek A. Patole,Mr. V. K. Pachghare,Dr. Parag Kulkarni," Self Organizing Maps to Build Intrusion Detection System," International Journal of Computer Applications,pp.0975-8887, Vol.1,No.8,2010.

[38] Hesham Altwaijry, Saeed Algarny,” Bayesian based intrusion detection system,” Journal of King Saud University, Computer and Information Sciences, 2010.

[39] Kapil Kumar Gupta, Baikunth Nath, and Ramamohanarao Kotagiri,"Layered Approach Using Conditional Random Fields for Intrusion Detection,"IEEE Transactions On Dependable And Secure Computing, Vol. 7, No. 1, January-March 2010.

[40] Todd Vollmer, Jim Alves-Foss, Milos Manic,” Autonomous Rule Creation for Intrusion Detection,” IEEE Symposium on Computational Intelligence in Cyber Security (CICS), pp.1-8, 2011.

[41] Kamran Shafi, Hussein A. Abbass, “Evaluation of an Adaptive Genetic-Based Signature Extraction System for Network Intrusion Detection, “Pattern Analysis and Applications, November 2011.

[42] Iwan Syarif, Adam Prugel-Bennett, Gary Wills, “Unsupervised clustering approach for network anomaly detection,” Fourth International Conference on Networked Digital Technologies, 24 - 26 Apr 2012.

[43] Sajal Bhatia, Desmond Schmidt, George Mohay,"Ensemble-based DDoS detection and mitigation model, “Proceedings of the Fifth International Conference on Security of Information and Networks,pp.79-86,2012.

[44] Prasanta Gogoi, Monowar H Bhuyan, D K Bhattacharyya, and J K Kalita,"Packet and Flow Based Network Intrusion Dataset," Contemporary Computing Communications in Computer and Information Science, vol.306, pp.322-334, 2012.