Network Intrusion Classification Employing Machine Learning: A Survey

Afroja Bhuiyan Akhi (ID: 011133049)
Esrat Jahan Kanon (ID: 011131040)
Arpita Kabir (ID: 011131014)
Atika Banu (ID: 011132127)
Department of Computer Science and Engineering
United International University

A thesis submitted for the degree of
BSc in Computer Science & Engineering

January 2019

Abstract

In this modern era computer network security is a vital issue. Network security is developed by an efficient Intrusion Detection System (IDS). It is used to identify unauthorized access, malicious attacks and give an alert when monitors any kind of unusual activity. Over the past 30 years, there have been lots of work on intrusion detection system using machine learning algorithms. Basically, realizing the present status of application of machine learning algorithms for solving intrusion classification task, this review work gives a proper guideline. This survey work selected 84 papers based on highest citations number from the years of 2009-2018. This thesis work gives an overview of a different intrusion detection systems, a statistical comparison based on different classifier like single, hybrid and ensemble learning. In addition, we have discussed best machine learning classifiers, best datasets and some feature selections process in this thesis work.

This work is devoted to our mother and father.

Acknowledgements

First of all, we would like to thank our thesis supervisor Dr. Dewan Md. Farid, Associate Professor, Department of Computer Science and Engineering, United International University (UIU). Without his help it was impossible for us to complete this thesis work. His positive efforts towards us and pushing us to our limits were the only reason that we have made it till here. We would like to thank Dr. Swakkhar Swatabda, Associate professor, Department of Computer Science and Engineering, United International University (UIU). We would also want to thank our loving United International University (UIU).

Contents

List of Figures **vii**

List of Tables **viii**

1 Introduction **1**

1.1 Motivation 1

1.2 Objectives of the Thesis 2

1.3 Organization of the Thesis 2

2 Intrusion Detection Systems **3**

2.1 Types of Intrusion Detection Systems 3

2.1.1 Host-based IDS 4

2.1.2 Network-based IDS 4

2.1.3 Misuse-based IDS 5

2.1.4 Anomaly-based IDS 5

2.1.5 Hybrid-based IDS 6

2.2 Intrusion Detection Tools 6

2.2.1 SNORT 6

2.2.2 OSSEC-HIDS 6

2.2.3 Bro IDS 7

3 Machine Learning for Intrusion Classification **8**

3.1 Learning Types 10

3.1.1 Supervised Learning 10

3.1.2 Unsupervised Learning 10

3.1.3 Reinforcement Learning 11

3.2 Learning Algorithms 11
3.2.1 Single Classifiers 11
3.2.2 Hybrid Classifiers 14
3.2.3 Ensemble Classifiers 14
3.3 Papers Collection 15
3.4 Single Classifiers 17
3.5 Hybrid Classifiers 19
3.6 Ensemble Classifiers 19

4 Data Set and Comparison of Performance **33**
4.1 Data sets 33
4.2 Comparison of Performance 37

5 Conclusions and Future Work **43**
5.1 Conclusions 43
5.2 Future Work 43

Bibliography **44**

List of Figures

2.1 Types of Intrusion Detection System. 4
2.2 Host-based Intrusion Detection System. 4
2.3 Network-based Intrusion Detection System. 5

3.1 Distribution of papers from year (2009-2018). 15
3.2 Year (2009-2018) of study of the type of classifier design. 16
3.3 Shows the annual work from 2009 to 2018 for single classifiers with different algorithms. 18

4.1 The relation between DARPA, KDD99 and NSL-KDD. 33
4.2 Shows the annual operation of the data set from 2009 to 2018. 36

List of Tables

3.1 Total number of papers for each classifier. 16
3.2 Number of papers collected from several sources. 17
3.3 Algorithms used in studies based on a single classifier. 18
3.4 Paper details using single classifier from 2009-2010. 20
3.5 Paper details using single classifier from 2011-2012. 21
3.6 Paper details using single classifier from 2013, 2014 and 2016. 22
3.7 Paper details using single classifier from 2016-2017. 23
3.8 Paper details using single classifier from 2017-2018. 24
3.9 Paper details using hybrid classifier from 2009-2010. 25
3.10 Paper details using hybrid classifier from 2011. 26
3.11 Paper details using hybrid classifier from 2012-2013. 27
3.12 Paper details using hybrid classifier from 2013-2015. 28
3.13 Paper details using hybrid classifier from 2015-2016. 29
3.14 Paper details using hybrid classifier from 2017-2018. 30
3.15 Paper details using ensemble learning from 2009, 2011-2015. 31
3.16 Paper details using ensemble learning from 2015-2018. 32

4.1 DARPA, KDD99, NSL-KDD data sets information. 34
4.2 Different types of attacks on KDD 99 data set and NSLKDD data set. . 35
4.3 The KDD 99 data set features. 36
4.4 Shows the distribution of single classifier by year in terms of results for each paper (2009-2012). 37
4.5 Shows the distribution of single classifier by year in terms of results for each paper (2013-2018). 38

4.6 Shows the distribution of hybrid classifier by year in terms of results for each paper (2009-2013). 39
4.7 Shows the distribution of hybrid classifier by year in terms of results for each paper (2013-2018). 40
4.8 Shows the distribution of ensemble classifier by year in terms of results for each paper (2009-2018). 41

Chapter 1

Introduction

Intrusion Detection System (IDS) is an application that effectively monitors dangerous alerts and activities in a computer network. There are several types of intrusion detection systems that gather information and notify unusual actions intrusions. In addition, some intrusion detection systems actually detect and prevent intrusions. In this thesis, we have reviewed several IDS that use machine learning algorithms from the year of 2009-2018.

1.1 Motivation

In recent times, the number of Internet users is increasing rapidly as a result number of attacks are also increasing repeatedly, which make a concern about the safety of the users and also it turns into system vulnerabilities of current network systems. So ensuring the safety of the users it is essential to make intrusion detection system (IDS) that is efficient enough to detect an attack. So with the flow of time, there have been a lot of change in IDS to ensure efficiency. Basically, this is the reason behind doing review paper because our motive is to find out the best classifier which will be very helpful for the development and improvement of IDS. Intrusion detection system became an important tool in security management. Nowadays IDS can block malicious attacks to retain the performance normal and also perform security analysis which is very urgent. Different research papers use different data sets. Over the past 30 years, automated learning research has increased the number of algorithms to detect intrusions, increasing recognition speed and meeting requirements. The purpose of this

work is to review recent papers machine learning for intrusion classification task. The survey should be consistent and capture useful information about each machine learning algorithm. It should also be extensible and benchmarking new algorithms should be straightforward and easily allow for more comparisons. As mentioned earlier, one of the biggest obstacles to comparing algorithms is the contradiction in data. The survey should be consistent useful information and provide a statistical comparison.

1.2 Objectives of the Thesis

- In this thesis, we have reviewed several intrusion detection systems have been proposed in the past 10 years. We have reviewed 84 research papers.
- This thesis describes several intrusion classification technology and it's architecture using machine learning algorithms.

1.3 Organization of the Thesis

Chapter 2 The chapter provides a general description of Intrusion Detection Systems.

Chapter 3 The chapter provides review of the papers. Explain the chapter about commonly used machine learning algorithms in intrusion detection system with experimental analysis and evaluation of the models.

Chapter 4 The chapter discusses comparison of machine learning algorithms performance and datasets used.

Chapter 5 Present conclusions and future work.

Chapter 2

Intrusion Detection Systems

Intrusion Detection System (IDS) is an important and widely used tool for ensuring network security. In recent years, intruder detection based on statistical recognition technology has attracted widespread attention in response to the growing need for highly reliable and intelligent IDS [1]. We cannot conceive our life without security, as it is one of the fundamental needs. Therefore, we need to be worried about having safe and secure networking channel. Without any authorized access to these networks means a lot of problems and basically, which means losses either financially or other. Before we think about the future, we need to think about securing our Internet or any other network that begins with detecting intrusion in the existing network. Several IDS approaches have been proposed in based on machine learning, each approach has their own advantages and disadvantages. IDS is a device or application that actively monitors network alerts when malicious actions and attacks are detected. There are several types of IDS that can actually deal with attacks and defend themselves. Intrusions are attacks that can come from outdoor or inside the network.

2.1 Types of Intrusion Detection Systems

The IDS can be extended based on two parameters based on the data source used and the detection method. Based on the data source used, the IDS is divided into two types of IDS: host-based IDS and network-based IDS.

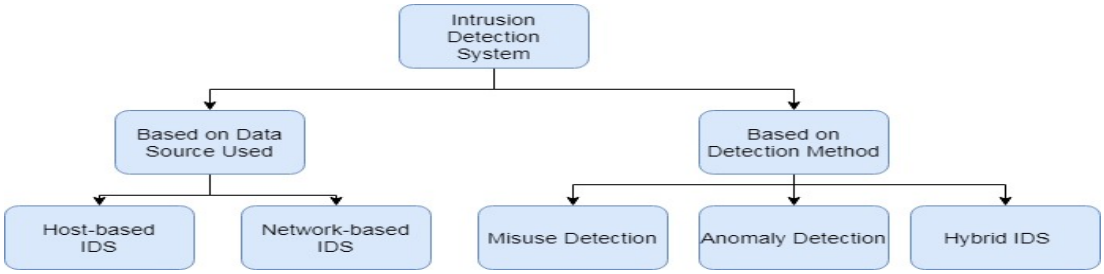

Figure 2.1: Types of Intrusion Detection System.

2.1.1 Host-based IDS

Host-based works like software application or agent that is deployed on individual hosts. When it's on an individual workstation, it can detect any threat and report it without delay to the user or administrator. The advantage of a host-based IDS is that it is capable of processing highly specific data, which is usually very useful. However, host-based IDS can significantly affect the performance of a running computer depending on the processing method. The general structure for generating host-based identifiers is shown in Figure 2.2.

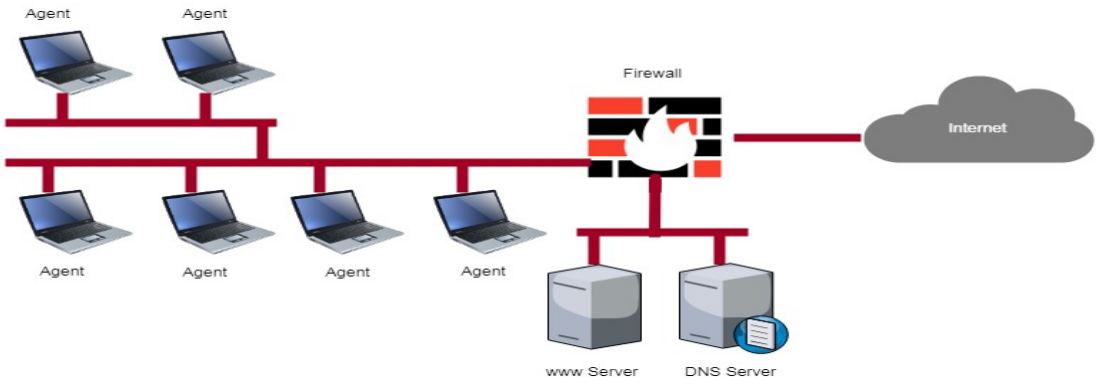

Figure 2.2: Host-based Intrusion Detection System.

2.1.2 Network-based IDS

As the popularity of the Internet grows, attacks on the network itself, which cannot only be detected using the host's audit trail, are increasing. For these reasons, special tools for network attacks are being developed. Network-based IDS are used to monitor and analyze network traffic and to protect the system from network threats. They are used

to monitor the transport layer and the IP layer header. The network identifier analyzes various types of information obtained by monitoring the network infrastructure. Figure 2.3 shows the general structure for creating network-based network IDSs.

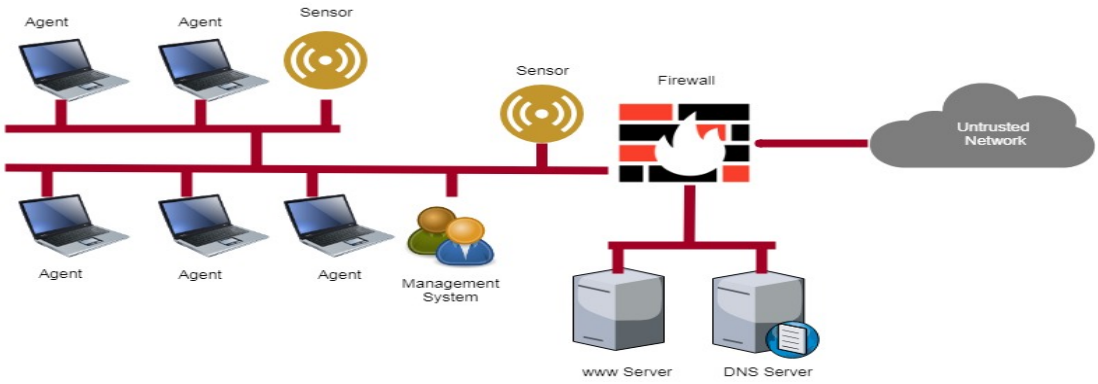

Figure 2.3: Network-based Intrusion Detection System.

Based on the detection method- misuse-based IDS, anomaly-based IDS and hybrid-based IDS are the main approaches to detecting attacks.

2.1.3 Misuse-based IDS

Detection of misuse is based on a broad knowledge of experts about known attacks and system vulnerabilities. A misuse-based approach executes these attacks and searches for intruders trying to exploit known vulnerabilities. Detection of ill-treatment can detect known attacks, but approaches to abuse cannot detect unknown or outgoing cyber threats. Fraud detection is the most common approach used in modern IDSs (IDS). The abuse detection approach can be divided into four main categories: a signature method, a rule technology, a state transition analysis method and a data method.

2.1.4 Anomaly-based IDS

In the meantime, the detection of anomalies is based on the analysis of profiles that describe the normal operation of users, hosts or network connections. Anomalous detector separates the normal operation of a computer with different technologies and uses different sizes to determine deviations from a given normal operation as a dynamic anomaly. The main advantage of anomaly detection algorithms is the ability to dynamically detect unexpected attacks. The anomaly detection method can be divided

into 5 groups: statistical method, rule method, distance method, profiling method and model method.

2.1.5 Hybrid-based IDS

The current trend in IDSs is the development of hybrid systems by combining both misuse IDS and anomalous IDS types. In such a hybrid system, signature capture technology recognizes known attacks and anomaly detection technology detects new or unknown attacks [2]. Hybrid IDS technology reduces the number of false alarms generated by anomaly detection methods.

2.2 Intrusion Detection Tools

Snort is one of the best intrusion detection tools, but there are many other tools. This section describes some security tools.

2.2.1 SNORT

A popular software for both host-based and network-based configuration is SNORT [114]. This compact open source intrusion detection software allows real-time traffic analysis and packet logging. The SNORT warning is displayed as soon as malicious code is detected. Regardless of the brand, intrusion detection software has the flexibility. IDS functions as a sensor, it identifies any anomalies within one specific segment and sends an alert if something malicious happens. IDS software offers two types of protection: active and passive. Active IDS software tries to prevent hacker from gaining access to system. Passive IDS software logs and reports potential attacks, but it doesn't actively respond to threats.

2.2.2 OSSEC-HIDS

OSSEC [115] is an open source IDS as well as intrusion detection. OSSEC can do a lot. Since this is a very good analysis, we can detect a rootkit. OSSEC runs on the host operating system and uses a client / server based architecture. OSSEC is an open source software and it is properly free so that it becomes the most popular tools for intrusion detection system. It can also detect anomaly, give alerts.

2.2.3 Bro IDS

Bro IDS is an IDS based on abnormality [116]. It is commonly used in combination with Snort. Bro IDS is a domain-specific language for network applications written by Bro IDS. This method is especially effective when almost all signatures have the correct arrows.

Chapter 3

Machine Learning for Intrusion Classification

Over the past 30 years, several research projects related to intrusion detection systems have been implemented. The concept of intrusion detection was launched in 1980 by James P. Anderson, presenting a threat classification model that developed a security monitoring system based on detection of user behavior anomalies [3]. In the second half of 1986, Dorothy E. Denning proposed various commercial IDS models based on statistics, the Markov chain and time series. This model is independent of the specific system, application environment, system vulnerability or type of attack and provides the basis for a common intrusion detection system [4]. Haystack et al. in 1988, was developed using the statistically abnormal IDS database, which uses user and group based anomaly detection strategies. In this system, the threshold was considered normal for each attribute. An alarm has been activated if the attribute has exceeded the normal limit. It is designed to detect six types of intrusions: unauthorized users invasion, denial of service, intrusion into the security management system, masked rider attacks [5]. Forrest et al. in 1996, an analogy was proposed between the human immune system and intrusion detection, which analyzed the sequence of system calls for a program to create a normal profile [6]. In 1999, Wenke Li reviewed data to calculate abuses and anomalies, observed the behavior of the generated data and proposed a data mining structure for creating an intrusion detection model [7]. There are variant surveys that are topical to this paper. Chih Phong Tsai et al. in 2009, the use of different classification methods in intrusion detection systems is another study of automatic learning. Studies related

to workbook design, data sets and other laboratory equipment are compared. They take into account the various machine learning methods used in the intrusion detection domain to test single, hybrid and ensemble classifiers [27]. In 2011, Shu-Xian Liao et al. review and explains that the development of data mining technologies that focus on expert knowledge and the development of data mining applications are more concerned with problems. The ability to constantly change and provide new insights is the main advantage of data mining methods and will become the basis of future applications for data mining technologies [28]. Michał Wozniak et al. research focuses on the 2014, sample classification and is performed by combining different classification systems that create the same model or different models or data sets. This document presents the latest overview of several classification systems from a hybrid system perspective [29]. A study by Dewan Farid et al in 2015, used various learning classifier in IDS is an external study in ML (Machine Learning) and AI (Artificial Intelligence). However, the focus of the study is on identifying research papers for relevant applications that use different classifiers for intrusion detection systems and show reliable statistical comparisons of work in this area [23]. In 2016, Anna L. Buchak and Erhan Guven will focus on a review of the literature on machine learning methods (ML) and data mining (DM) used in cyberspace. Consider the understanding of accuracy, time of classification, the complexity of unknown instances with trained patterns and the final solution of each ML or DM method. Methods ML and DM do not work if the data has not been delegated. Obtaining such records is difficult and time consuming [24]. In 2017, Bruno Bogaz Zarpelaoa et al. an IDS survey was announced based on the detection method, IDS deployment strategy, security threat and verification strategy. Efforts on IoT. IoT is very much expected, because it has the ability to convert the physical objects of different application domains into Internet hosts [30]. In 2018, Nicholas J. Miller proposed a benchmark that measures both accuracy and performance to generate purpose metrics that can be used in the evaluation of each algorithm implementation. In this survey paper, the benchmark will be used to compare variant ML algorithms (Support Vector Machine (SVM), Naïve Bayes (NB), Neural Network (NN) and K-means Clustering) with NSL-KDD dataset. Experimental results show the change in accuracy and performance between these algorithms in the data set [25].

3.1 Learning Types

Machine learning (ML) is part of artificial intelligence. Machine learning is used to make computers seem intelligent and by definition machine learning is like teaching computers to learn so rather than programming every single possible outcome what we do is we take a computer and we feed it a bunch of data and it's going to learn from that data and be able to improve over time. The reason machine learning is becoming so popular now is that we have so much data at our hands. So we are turning to computers to do the analyzing for us and this is really super popular in something called data analytics. Our data well we can add some discipline by having predictive data analytics. Machine learning algorithms fall into three categories: supervised learning, unsupervised learning and reinforcement learning.

3.1.1 Supervised Learning

Training is supervised, as a name indicating that training in the presence of a supervisor can be a teacher or instructor. In this case, you must specify or train the data in which the data label is indicated, that is the device in which some data has already been marked with the correct answer. Learning monitoring is a general concept applicable to real-world scenarios. The controlled learning algorithm analyzes the learning data and generates a classifier function (if discrete output) or a regression function (if output continuous). There are many supervised learning algorithms. Artificial Neural Network (ANN), Support Vector Machine (SVM), K-Nearest-Neighbors (KNN), Bayesian Statistics, Bayesian Networks, Boosting, Ensembles classifiers (Bagging, Boosting), Decision Trees(C4.5, ID3, CART, Random Forrest), Linear Classifiers (Naïve Bayes (NB), Logistic Regression) classifiers are popular supervised learning algorithms.

3.1.2 Unsupervised Learning

Uncontrolled training is the study of tasks that are performed without supervision by supervisors. With Unsupervised learning, tasks are performed based on past experience. Training includes learning by data, but not goals and predictions. Unsupervised learning refers to the problem of finding a secret structure in unmarked data. Unlike supervised learning, there is no clear goal or environmental assessment for each entry. Rather, an impartial apprentice gives rise to previous prejudices, according to which

aspects of the input composition should be included in the output. There are many unsupervised learning algorithms. Clustering and Deductive algorithm.

3.1.3 Reinforcement Learning

Reinforcement learning is a very simple way to teach agents and computers the best possible behavior in a certain sense, trying to earn rewards of various kinds. The extension approach may prompt the user to identify instances [23]. There are many reinforcement learning algorithms. Q-Learning, Temporal Difference (TD), Deep Adversarial Networks and Markov Decision Process algorithm.

3.2 Learning Algorithms

Although various machine learning algorithms are used for intrusion detection systems, the IDS method can consist of three categories: single, hybrid and ensemble. Methods for developing IDS or machine learning algorithms can be used as one classifier. There are DT (Decision Tree), NB (Naïve Bayes), KNN (K-Nearest Neighbors), ANN (Artificial Neural Network), FL (Fuzzy Logic) and SVM (Support Vector Machines). Hybrid classifiers can combine several machine learning algorithms to improve the performance of intrusion detection systems. Preprocessing samples into training data using several clustering-based algorithms (fuzzy clustering, K-clustering) to remove unrepresentative learning patterns. Ensemble method & machine learning is a type of meta-algorithm that studies predictor sets using other machine learning algorithms as components. This classifier is slightly better than a weak learners. Bagging and Boosting are ensemble classifiers.

3.2.1 Single Classifiers

The following section provides single classifier algorithms description.

Naïve Bayesian

A simple naïve Bayes (NB) classifier is a simple stochastic method that can predict the probability of a class belonging [18]. This classifier can easily process missing attribute values by simply omitting the cor-responding probability of these attributes when calculating the membership probability of each class. The result of the attributes of a given class doesn't depend on the consequences of some attributes [19]. Naïve bayes is

among most effective algorithms to do this there are many more algorithms that are more efficient than naive bayes but naive bayes is a very is a very good baseline and it's pretty effective algorithm. Valdez et al. describe anomaly-based IDS to detect impulses of pulsed traffic using a naïve Bayesian network in 2000 [8]. In 2003, the method of multisensory synthesis by Krugel et al. proposed with the Bayesian classifier false alarms are classified and suppressed [10]. Amor et al. using the KDD 99 dataset in 2004, compare the Naïve Bayesian (NB) classifier and the Decision Tree (DT) classifier and developed an experimental analysis [11]. In 2007, Panda and Patra proposed a way to sign a specific attack using naïve bayes. They use the KDD99 data set for experiments and the authors conclude that the NB classifier resets the neural network classifier [14].

Decision Tree

It is a popular algorithm. It represents decisions and decision making. There are three main elements in the decision tree: node, leaf and edge [17]. Root node: the topmost node of the tree structure. There is no incoming edge. Internal node: Internal node displays a test attribute. It has an outgoing edge. Branch Node: Each branch node displays the result of the test. Leaf nodes: Let class tag leaves be called end nodes or decision nodes. This means a target value. Decision tree to reach the goal. When a decision tree is used for the classification work, it is more specifically called a classification tree. The decision tree algorithm is an ID3 (Iterative Dichotomiser 3), C4.5, a CART (Classification and Regression Tree), or a decision tree algorithm. In 2006, Bouzida et al. used to detect intrusions based on anomalies by using decision tree (DT), which assigns a test instance to standard classes that are not obfuscated by the tree. Then investigate the unknown attack analysis, examining the standard class [13]. In 2010, Dewan Farid et al. combine naïve Bayes for adaptive intrusion detection with a decision tree to achieve high recognition rates and reduce false positives in various types of network solutions [16].

K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) is one of the most fundamental and important classification algorithms in machine learning for pattern recognition, data mining and intrusion detection. KNN is a non parametric method used for classification and regression. KNN is a classification algorithm based on the use of distance measures [20]. Adjacent classifier K can choose the appropriate value for K based on mutual verification. KNN

is a simple classifier that uses the distance learning feature widely used in machines. In 2011, G. Davantseo et al. experimentally determined the values of false negative rate and false positive rate to be low enough to achieve an average false positive rate of less than 1%, but can be recognized correctly. K-nearest, GaussianParzen, LOF, seems to be different from the corresponding performance [41].

Artificial Neural Network

Artificial Neural Networks (ANNs) originate from the human brain or human nervous system, consisting of large parallel connections of numerous neurons. Numerous neurons or neurons perform the treatment. They are usually related to each other in a very complex way. In addition, since there is a connection from one neuron to another neuron, the network is realized. A typical neural network is built with layers of several interconnected nodes with activation functions. Templates are presented on the network through the input level. The input level is associated with one or more hidden layers, where the weighted join system completes the initial processing. A hidden level is then added to the output level and the identification result is generated as a result [23]. In 2017, Akashdeep et al. proposed a combining the obtained ranks is associated with a decrease in the gain and correlation of information using a new method for determining useful and useless functions. These reduced functions are provided to the neural direct communication network for training and testing the KDD99 data set [21].

Support Vector Machine

Support Vector Machine (SVM) is supervised learning using related learning algorithms used for classification and regression analysis. One of the most efficient classifiers with a certain linear classification. Supporting vector machines that can handle specific cases of nonlinearities consisting of using nonlinear basis functions, or call these basic functions. Supporting vector machines are one of the most common machine learning algorithms because of their good generalization properties and their ability to overcome sacrifice sizes. Some researchers still pay special attention to some aspects of the SVM-based IDS health. Another problem is that SVM processes every data function without distinction. In a real attack detection record, many functions are excessive or insufficient. It is best to take into account weights during SVM training. In 2001, Wun-Hwa Chen et al. suggested that the use of the network and the Internet should increase interest in the security of information systems that detect intruders individually. An artificial neural network (ANN) and a support vector machine (SVM) are two

mining methods and use two coding methods to detect possible system failures [9]. In 2007, L. Khan et al. evaluated a combination of a dynamically growing self-organizing tree algorithm and SVM for the clustering algorithm of the DARPA98 dataset. They compare approaches to random selection with loss of accuracy using the Rocchio binding method and one actual set of control data and it also improves accuracy [12].

Fuzzy Logic

Fuzzy logic (FL) has been around a long time. It's first described in a paper around 1965 by a man called Lotfi Aliasker Zadeh. It is been in widespread use for decades. Fuzzy logic is a way of dealing with uncertainties which is something that computers do not do naturally which human do very well. So normally fuzzy logic something can be represented by either having a value of true which is 1 or false which is 0. Fuzzy logic can be anywhere between 0 & 1. In 2000, J. E. Dickerson and J. A. Dickerson used a fuzzy analysis mechanism to evaluate fuzzy input recognition engine (FIRE) mechanism that triggered a security warning level. It can detect general assaults [22].

3.2.2 Hybrid Classifiers

A combination of hybrid classifiers of multiple machine learning algorithms to increase the effectiveness of intrusion detection systems. By using multiple cluster-based algorithms to preprocess example learning data to remove non-representative learning patterns, the result of clustering is used as a template for learning pattern recognition to design classifiers [23].

3.2.3 Ensemble Classifiers

The following section provides ensemble classifier algorithms description.

Bagging

Bagging is a way to get to know the ensemble learning. Bagging carriers are provided for summarizing trunk straps. A learner ensemble is used and a balanced classification work allows learners to combine to compare individual base classifiers for better accuracy. The classification is to perform a weighted vote, calculating the weighted average in numerical prediction.

Boosting

Another ensemble methods that enhance to collect the increase in the collection of prophecies, to adapt the information of primary education in a simple model, then

analyzing data errors, are an ensemble of learners procedures for consistent learning. These defects reveal problems and problems that are difficult to understand. In particular, in these examples, to try to correct them, there is an end to the mass of all samples, the statement is associated with some of the prediction. Acceleration is a way to change the sequence of weak learners with very complex predictions. AdaBoost algorithm called adaptive enhancement algorithm, it was first introduced in 1997 by Fround and Shapier. They were classified as an example by adjusting the weight gain predicted on the basis of the basic classifier to achieve a round series [26]. In 2008, Waming Hu proposed a network IDS algorithm based on the AdaBoost algorithm. This indicates that the algorithm has lower computational overhead and error rate compared with higher computational complexity, as was verified using data from control samples [15]. There are many papers about different algorithms and their success in intrusion detection systems. Different papers will use their own IDS dataset. However, each paper trains and tests their algorithm differently with their own IDS dataset.

3.3 Papers Collection

In the thesis, we have reviewed 84 research papers from 2009 to 2018. Eight papers were watched from the years 2009, 2010, 2012 and 2015. Since 2011 there are 11 papers. 10 papers were collected from in 2013, 2016 and 2017. In 2014, four papers and in 2018 seven papers were collected. Table 3.1 shows total number of papers for each classifier. Table 3.2 shows number of papers collected from several sources- 77 journals and 7 conferences we review for our paper. And paper search key word is intrusion detection.

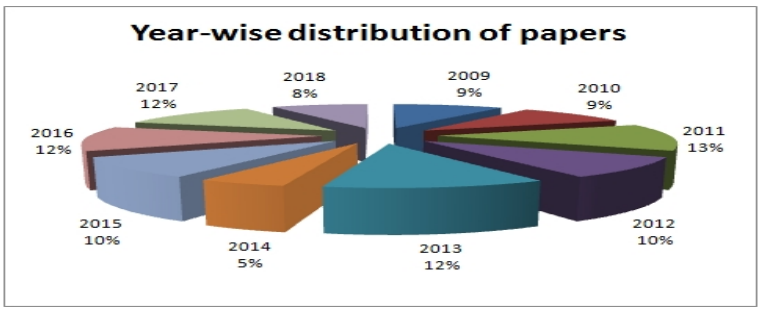

Figure 3.1: Distribution of papers from year (2009-2018).

Intrusion detection methods can be divided into three categories: Single, Hybrid, Ensemble [27]. Figure 3.2. determining the number of experiments in each year used to classify single, hybrid and ensembles. According to a statistical comparison of published papers, single are in research activities studied in 33 papers. Hybrid have the highest number of coordinates with the largest number of 35 coordinates for a given time interval. Ensemble, we researched 16 empirically papers.

No.	Classifier design type	No. of research paper	References
1	Single	33	[31], [32], [33], [34], [35], [36], [37], [38],[39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [84], [85], [86], [87], [104], [105], [106], [107]
2	Hybrid	35	[55], [56], [57], [58], [59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [76], [77] , [78], [79], [80], [81], [82], [83], [88], [108], [109], [110], [111], [112]
3	Ensemble	16	[89], [90], [91], [92], [93], [94], [95], [96], [97], [98], [99], [100], [101], [102], [103], [113]
	Total	84	

Table 3.1: Total number of papers for each classifier.

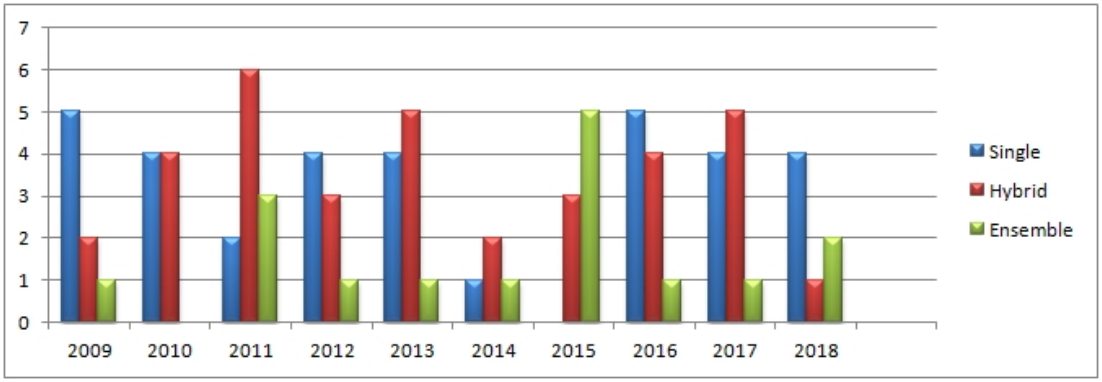

Figure 3.2: Year (2009-2018) of study of the type of classifier design.

Table 3.2: Number of papers collected from several sources.

Year	IEEE Xplore	ScienceDirect	Springer	Others	Journals	Conference
2009	-	6	-	2	8	-
2010	-	7	-	1	8	-
2011	-	10	-	1	11	-
2012	-	5	-	3	8	-
2013	-	6	-	4	9	1
2014	-	2	-	2	4	-
2015	6	2	-	-	6	2
2016	1	3	1	5	10	-
2017	1	5	1	3	10	-
2018	6	-	-	1	3	4
				Total	77	7

3.4 Single Classifiers

The number of studies written to refer to single type of classification with different categories and publications from 2009 to 2018. Figure 3.3 shows the annual work from 2009 to 2018 for single classifiers with different categories. The number of elements in this group will be used using different classification algorithms such as Baizian, Vector Algorithm, Tree Algorithm, Artificial Algorithm, Fuzzy Algorithm etc. For all the products described here. Table 3.3 shows algorithms used in studies based on a single classifier. Table 3.4-3.8 shows the distribution of individual categories by year in terms of algorithms and data sets for each element.

SVM and DT are the most well-known algorithms in classifiers with single learning algorithm. Although we have graphed 84 related papers, the number of comparative samples is small, but the result of the comparison is that the support vector machine is the most common and is still considered the only classification method that means. Fuzzy logic the most less-known algorithms in classifiers with single learning algorithm.

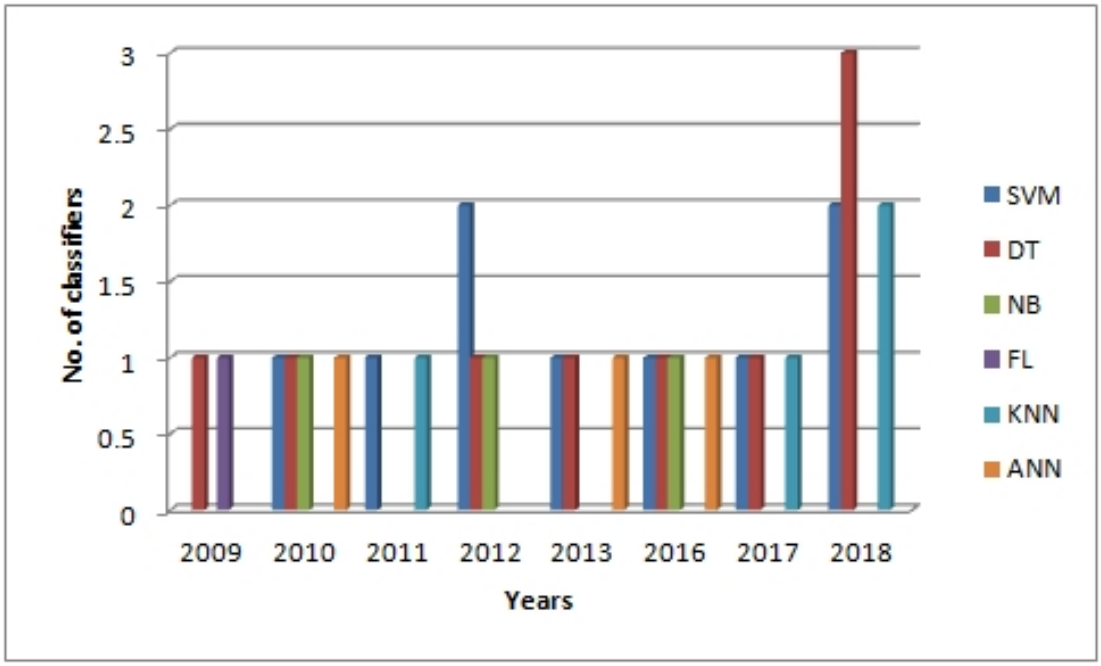

Figure 3.3: Shows the annual work from 2009 to 2018 for single classifiers with different algorithms.

Algorithm	**Reference**
Naïve Bayes (NB)	(2012)[42], (2010)[36]
Support Vector Machine (SVM)	(2010)[36], (2011)[41], (2012)[43], (2012)[44], (2013)[48], (2016)[52], (2017)[107], (2018)[85], (2018)[87]
Decision Tree	(2009)[32], (2010)[36], (2012)[45], (2013)[46], (2016)[54], (2017)[106], (2018)[84], (2018)[85], (2018)[87]
Artificial Neural Network (ANN)	(2010)[39], (2013)[47], (2016)[54]
Fuzzy Logic (FL)	(2009)[32]
K-Nearest Neighbour (KNN)	(2011)[41], (2017)[106], (2018)[86], (2018)[87]

Table 3.3: Algorithms used in studies based on a single classifier.

3.5 Hybrid Classifiers

In most cases, this work is done to create a superior system, leading to the success of hybrid classifiers for intrusion detection systems. Hybrid classifiers have machine learning methods for improving viable systems such as SVM, DT, K-NN, GA. There were two aspects to this hybrid approach. The first data captures the raw data to get a quick result and the second information gets quick results and gives the final result. Table 3.9-3.14 shows the annual distribution of hybrid classifiers in terms of data sets and algorithms of each thesis. Because of the accuracy of today's results, hybrid classifiers for intrusion detection systems have been created in basic research. According to statistics, the hybrid classifier has the largest number of articles in 2011.

3.6 Ensemble Classifiers

Some weak learners are classified as ensemble classifiers. According to statistics, AdaBoost is the most commonly used learning algorithm with the most votes. Table 3.15-3.16 also includes each classifier and information about each article for the entire period. According to statistics, the ensemble classifier has the largest number of articles in 2015.

Table 3.4: Paper details using single classifier from 2009-2010.

Year	Title of The Paper	Author(s)	Algorithm(s)	Data set(s)	Find Out Attacks	Cited by	Ref
2009	Association rules applied to credit card fraud detection	D. Sanchez, M.A. Vila, L. Cerda, J.M. Serrano	Association rule methodology	Credit approval data set	credit card fraud detection	175	[31]
2009	Data Mining based intrusion detectors	Su-Yun Wu, Ester Yen	C4.5	KDD99	Probe, Dos and U2R	131	[32]
2009	Some issues about out-lier detection in roughset theory	Feng Jiang, Yuefei Sui and Cungen Cao	Outlier Detection algorithm	KDD99	handle uncertain and/or incomplete data.	49	[33]
2009	Project out-lier detection in high-dimensional mixedattributes data set	Mao Ye, Xue Li and Maria E. Orlowska	PODM algorithm	Credit approval data set	fraud detection, network intrusion detection, monitoring criminal activities etc.	43	[34]
2009	Information inconsistencies detection use a rule-map	Jun Ma, Jie Lu, Guangquan Zhang	RMDID algorithm	database data		4	[35]
2010	Malicious web-content detection	Yung-Tsung Hou, Yimeng Chang, Tsuhan Chen, Chi-Sung Laih and Chia-Mei Chen	Naïve Bayes, DT, SVM and AdaBoost	benign and malicious DHTML	malicious web pages detection and present relevant features	96	[36]
2010	Random effect logistic regression model for anomaly detection	Min Seok Mok, So Young Sohn, Yong Han Ju	Logistic regression model	KDD99	normal and anomaly	23	[37]
2010	The Decision Model for Responding to an Intrusion Based on a Hierarchical Network Task Plan	Chengpo Mu and Yingjiu Li	Hierarchical task network planning	Roc curve : excellent		54	[38]
2010	Neural Networks based detection of stepping stone intrusion	Han-Ching Wu, Shou-Hsuan Stephen Huang	Neural Network		long connection chain by estimating the number of stepping-stones.	23	[39]

Table 3.5: Paper details using single classifier from 2011-2012.

Year	Title of The Paper	Author(s)	Algorithm(s)	Data set	Find Out Attacks	Cited by	Reference
2011	Evolving boundary detectors for anomaly detection	Wang Dawei, Zhang Fengbin, Xi Liang	Detector Generation algorithm	KDD99	R2L, U2R, DoS and Probing	15	[40]
2011	Network Failure Service Anomaly Detection Technology	G. Davanzo, E. Medvet and A. Bartoli	KNN and SVM	Small–Short, Large–Long	web defacement monitoring service	25	[41]
2012	Hidden Naïve Bayes Network Intrusion Detection System Based on Multi-Class Classifier	Levent Koc, Thomas A. Mazzuchi and Shahram Sarkani	Hidden Naïve Bayes	KDD99	Dos	178	[42]
2012	An autonomous labeling approach to support vector machines algorithms for network traffic anomaly detection	Carlos A. Catania, Facundo Bromberg, Carlos García Garino	Support Vector machine	1998 DARPA		47	[43]
2012	A differentiated one-class classification method with applications to intrusion detection	Inho Kang, Myong K. Jeong, Dongjoon Kong	Support Vector machine	simulated data set and the DARPA 1998 BSM data set	DARPA: U2R	44	[44]
2012	Madame ID For Intrusion Detection Using Data Mining	Prabhjeet Kaur, Amit Kumar Sharma and Sudesh Kumar Prajapat	J48	given data set	Dos	7	[45]

Table 3.6: Paper details using single classifier from 2013, 2014 and 2016.

Year	Title of The Paper	Author(s)	Algorithm(s)	Data set(s)	Find Out Attacks	Cited by	Reference
2013	A Cost - Conscious Decision Tree Approach to Detecting Fraud	Yusuf Sahin, Serol Bulkan and Ekrem Duman	Decision Tree	real world credit card data set	fraud detection	97	[46]
2013	Accuracy Detection for an Intrusion Detection System Using a Neural Network Classifier.	S. Devaraju and S. Ramakrishnan	Neural Network	KDD99	Dos, Probe, U2R, R2U and Normal	25	[47]
2013	Abrupt change detection with one class time adaptive Support Vector Machine	Guillermo L. Grinblat, Lucas C. Uzal, Pablo M. Granitto	Support Vector Machine			17	[48]
2013	Application of feature detection rule to the negative selection algorithm	Mario Poggiolini, Andries Engelbrecht	Negative Selection algorithm	Glass, Mushroom		17	[49]
2014	A New Approach to the Selection of Functions, Based on the Optimization Algorithm of Cuttlefish for an Intrusion Detection System	Adel Sabry Eesa, Zeynep Orman, Adnan Mohsin Abdulazeez Brifcani	Cuttlefish algorithm	KDD99		93	[50]
2016	A multi-level intrusion detection method for abnormal network behaviors	Soo-Yeon Ji, Bong-Keun Jeong, Seonho Choi, Dong Hyun Jeong	SVM, Naïve Byes and NN	NSL-KDD	Dos, Probe, R2L	36	[52]
2016	An Approach for Intrusion Detection Using Novel Gaussian Based Kernel Function	Gunupudi Rajesh Kumar, Nimmala Mangathayaru, Gugulothu Narsimha	K-means, Gaussian	DARPA, KDD99		30	[53]

Table 3.7: Paper details using single classifier from 2016-2017.

Year	Title of The Paper	Author(s)	Algorithm(s)	Data set(s)	Find Out Attacks	Cited by	Reference
2016	Feature Selection for Intrusion Detection System Using Ant Colony Optimization	Mehdi Hosseinzadeh Aghdam and Peyman Kabiri	Proposed method uses ACO algorithm and a simple classifier (nearest neighbor classifier)	KDD99, NSL-KDD	Normal,DoS, Probe, U2R and R2L	34	[51]
2016	Assessment of the Network Anomaly Detection System: Statistical Analysis of the UNSW-NB15 Data Set and Comparison With the KDD99 Data Set	Nour Moustafa & Jill Slay	DT (Decision Tree), LR (Logistic Regression), NB (Naïve Bayes), ANN and Expectation Maxi-misation	UNSW-NB15, KDD99		57	[54]
2017	Fuzziness based semi-supervised learning approach for intrusion detection system	Rana Aamir Raza Ashfaq, Xi-Zhao Wang, Joshua Zhexue Huang, Haider Abbas and Yu-Lin He	semi-supervised learning using privileged information	NSL-KDD	Deniel of service, User to root, Remote to local, Probing	157	[104]
2017	Intrusion Detection Model Using fusion of Chi-square feature selection and multi class SVM.	I. Sumaiya Thaseen, Ch. Aswani Kumar	multi class support vector machine (SVM)	NSL-KDD	Normal, DoS, R2L, Probe, U2R	39	[107]
2017	Intrusion Detection System Based On Flows Using Machine Learning Algorithms	E.M. Kakihata, H.M. Sapia, R.T. Oikawa, D.R. Pereira, J.P. Papa, V.H.C.Alburquerque and F.A. Silva	Optimum-Path Forest (OPF)	Random	Normal, Port Scanning, Força Bruta, DoS	3	[105]

Table 3.8: Paper details using single classifier from 2017-2018.

Year	Title of The Paper	Author(s)	Algorithm(s)	Data set(s)	Find Out Attacks	Cited by	Reference
2017	Enhancing Collaborative Intrusion Detection Networks Against Insider Attacks Using Supervised Intrusion Sensitivity-Based Trust Management Model	Wenjuan Li, Weizhi Meng, Lam-For Kwok, Horace H. S. IP	KNN, BPNN and DT	Knowledge Database	DoS	16	[106]
2018	Detection of Intrusion Into Computer Network Using Various Classifiers and Ensemble Learning	Ali H.Mirza	Neural Network Logistic Regression Decision Trees	KDD99		not yet	[84]
2018	Machine Learning and deep learning for cyber security	Yang xin, Zhi Liu, Yuling Chen	Svm Decision Tree DBM CNN	KDD99		3	[85]
2018	An intrusion detection using model based on Fuzzy c means clustering	Liyo Duan, Youan Xiao	Fuzzy c means k-means kNN	KDD99		not yet	[86]
2018	Data mining techniques in intrusion detection	Abdallah Shami Fadi Salo	DT KNN BN SVM Ad boost			not yet	[87]

Table 3.9: Paper details using hybrid classifier from 2009-2010.

Year	Title of The Paper	Author(s)	Algorithm(s)	Data set(s)	Find Out Attacks	Cited by	Reference
2009	Abnormality Detection Method Using Fusion of Unsupervised Control Neural Network	M.Bahrololum, E.Salahi and M.Khaleghi	NN	KDD99	Dos, Probe and R2L	38	[55]
2009	Adaptive Genetic Intrusion Detection Signature Training System	Kamran Shafi and Hussein A.Abbass	GA	10% KDD	Normal, Probe, DOS, U2R and R2L	62	[56]
2010	Novel Approach Using Fuzzy Clustering and Artificial Neural Networks to Intrusion Detection	Gang Wang, Jinxing Hao, Jian Mab and Lihua Huang	ANN and Fuzzy clustering	KDD99	Probe, DOS, U2R and R2L	348	[57]
2010	A distributed sinkhole detection method using cluster analysis	Woochul Shim, Gisung Kim, Sehun Kim	Hierarchical cluster analysis			19	[58]
2010	Design Network Intrusion Detection System using hybrid Fuzzy-Neural Network	M.M.T.Jawhar and M.Mehrotra	Fuzzy clustering, and NN	KDD CUP 1999	Probe, DOS, U2R, R2L and unknown	43	[59]
2010	Histological Algorithm of Negative Selection Based on Chaotic Approach and Its Application to Detection of Abnormalities and Anomalies	Ilhan Aydin, Mehmet Karakose and Erhan Akin	Negative selection, Clonal selection and KNN	Random		77	[60]

Table 3.10: Paper details using hybrid classifier from 2011.

Year	Title of The Paper	Author(s)	Algorithm(s)	Data set(s)	Find Out Attacks	Cited by	Reference
2011	Detect Online Games and Lottery Fraud	I.T.Christou, M.Bakopoulos, T.Dimitriou, E.Amolochitis, S.Tsekeridou and C.Dimitriadis	Local Outlying Factor, K-means clustering and EXAMCE	Random	games and lottery fraud in online	13	[61]
2011	Fast Detection of Outliers for Very Large Log Data	S.Kim, N.W.Cho, K.Bokyoung, Suk-Ho-Kang	KNN and LOF	IonoSphere, Musk and Synthetic Control Chart		37	[62]
2011	An Integrated Intrusion Detection System for Cluster-based Wireless Sensor Networks	Shun-Sheng Wang, Kuo-Qin Yan, Shu-Ching Wang, Chia-Wei Liu	BPN. ART. Rule based method	KDD CUP 1999	Normal, Probe, DOS, U2R and R2L	84	[64]
2011	Development and Analysis of Genetic Fuzzy System for Intrusion Detection in Computer Network	Mohammad S.Abadeh, H.Mohamadi and Jafar Habibi	Michigan, Pitsburg, IRL	KDD CUP 1999	normal and abnormal behaviors	72	[63]
2011	RealTime Anomaly Detection for DoS Attacks Using KNN Weighted Classifiers	Ming-Yang Su	GA and KNN	KDD CUP 1999	DoS/DDoS	69	[65]
2011	Online Anomaly Detection for Self-Adaptive and Dynamic Clustering	S.Lee, G.Kim and S.Kim	SOM. K-means clustering	KDD CUP 1999	new attacks to be clearly detected	49	[66]

Table 3.11: Paper details using hybrid classifier from 2012-2013.

Year	Title of The Paper	Author(s)	Algorithm(s)	Data set(s)	Find Out Attacks	Cited by	Ref
2012	An efficient intrusion detection system based on support vector machines and gradually feature removal method	Yinhui Li, Jingbo Xi, Silan Zhang, Jiakai Yan, Xiaochuan Ai, Kuobin Dai	K-means clustering. SVM. Ant colony	KDDcup99	Normal, Probe, DOS, U2R and R2L	183	[67]
2012	The combined approach for anomaly detection using neural networks & clustering techniques	A.S. Aneetha and Dr. S. Bose	SOM. K-means clustering	10% KDD	normal and attack	19	[68]
2012	Anomaly Based Intrusion Detection System Using Artificial Neural Network and Fuzzy Clustering	Prof. D.P. Gaikwad, Sonali Jagtap, Kunal Thakare, Vaishali Budhawant	ANN. Fuzzy clustering	KDDcup99	R2L and U2R	6	[69]
2013	By Using Neural Network and SVM Building a Enhancement Hybrid Intrusion Detection System	A.M.Chandrashekhar and K.Raghuveer	Fuzzy C-means clustering. Fuzzy neural network. SVM	KDDcup99	Probe, DOS, U2R and R2L	16	[70]
2013	Advanced probabilistic approach for network intrusion forecasting and detection	Seongjun Shin, Seungmin Lee, Hyunwoo Kim, Sehun Kim	Markov chain, K-means clustering, APAN	DARPA 2000	DDoS	42	[71]

Table 3.12: Paper details using hybrid classifier from 2013-2015.

Year	Title of The Paper	Author(s)	Algorithm(s)	Data set(s)	Find Out Attacks	Cited by	Ref
2013	Intrusion Detection System Based on NSLKDD Data Set by Fuzzy Clustering and Neural Network in Hybrid	D.A.Ahmad Zainaddin and Z.M.Hanapi	Fuzzy clustering	KDDCup 1999, NSL-KDD	Probe, DOS, U2R and R2L	16	[72]
2013	Hybrid structure based on neural network MLP & K-means clustering implies for intrusion detection system	M.Mohammadi Lisehroodi, Zaiton Muda and Warusia Yassin	K-means clustering and MLP	KDDcup99	Probe, DOS, U2R and R2L	5	[73]
2013	A novel hybrid intrusion detection method integrating anomaly detection with misuse detection	Gisung Kim, Seungmin Lee, Sehun Kim	C4.5, 1-class SVM	NSL-KDD		203	[74]
2014	New Clustering Method for Abnormality Detection	R.Ranjan and G.Sahoo	C4.5, SVM, K-means clustering	kDD99	Probe, DOS, U2R and R2L	45	[76]
2014	Mining Network Data for Intrusion Detection through combining SVMs with ant colony networks	Wenying Feng, Qinglei Zhang, Gongzhu Hu, Jimmy Xiangji Huang	CSOACN (self organized ant colony network) and SVMCSVAC (combining support vectors with ant colony)	KDD99		150	[75]
2015	Detecting Wireless Anomalies Depends On IEEE802.11 Behavior Analysis	H.Alipour, Youssif B.Al-Nashif, P.Satam and Salim Hariri	supervised learning and n-gram		triggered by wireless network attacks	28	[77]

Table 3.13: Paper details using hybrid classifier from 2015-2016.

Year	Title of The Paper	Author(s)	Algorithm(s)	Data set(s)	Find Out Attacks	Cited by	Ref
2015	Intrusion Detection System (IDS): Anomaly Detection Using Outlier Detection Approach	Jabej J, Dr.B. Muthukumar	Neighborhood outlier factor	KDD99		42	[79]
2015	Developing a Hybrid Intrusion Detection System Using Data Mining for Power Systems	Shengyi Pan, Thomas Morris and Uttam Adhikari	The common paths mining-based IDS	Random	power system scenarios including disturbances, normal control operations and cyber-attacks	76	[78]
2016	A Hybrid Data Mining Approach for Intrusion Detection on Imbalanced NSL-KDD Data set	Mohammad Reza Parsaei, Samaneh Miri Rostami, Reza Javidan	Synthetic minority oversampling technique (SMOTE), cluster center and nearest neighbor (CANN), leave one out method (LOO).	NSL-KDD	R2L and U2R	12	[80]
2016	Hybrid mode consists of SVM and GA for Intrusion detection system	B.M.Aslahi-Shahri, R.Rahmani, M.Chizari, A.Maralani, M.Eslami, M.J.Golkar and A.Ebrahimi	GA and SVM	KDD99		32	[81]
2016	Building an intrusion detection system using a filter-based feature selection algorithm	Mohammed A. Ambusaidi, Xiangjian He, Priyadarsi Nanda and Zhiyuan Tan	Flexible Mutual Information Feature Selection (FMIFS), Least Square Support Vector Machine based IDS (LSSVM-IDS)	KDD99, NSL-KDD, Kyoto 2006+	Normal, Probe, DOS, U2R and R2L	87	[82]
2016	An effective MCLP / SVM based intrusion detection system optimized by optimizing particle steepness in the time zone	S.M.H.Bamaka, Huadong Wang, Tian Yingjie and Yong Shia	timevarying chaos particle swarm optimization (TVCPSO), multiple criteria linear programming (MCLP), SVM	NSL-KDD	R2L and U2R	47	[83]

Table 3.14: Paper details using hybrid classifier from 2017-2018.

Year	Title of The Paper	Author(s)	Algorithm(s)	Data set(s)	Find Out Attacks	Cited by	Ref
2017	A Hypergraph and Arithmetic Residue-based Probabilistic Neural Network for Classification in Intrusion Detection Systems	Gauthama Raman M R, Nivethitha Somu, Kannan Kirthivasan, Shankar Sriram V S	Hypergraph and Arithmetic Residue-based Probabilistic Neural Network (HG AR-PNN)	KDD99	Normal, DoS, Probe, R2L, U2R	16	[108]
2017	A novel technique for intrusion detection system for network security using hybrid SVM-CART	Aastha Puri, Nidhi Sharma	SVM-CART	KDD99	Normal, DoS, U2R, R2L and Probe	4	[109]
2017	Distributed Host-based Collaborative Detection for False Data Injection Attacks in Smart Grid Cyber-Physical System	Beibei Li, Rongxing Lu, Wei Wang, Kim-Kwang Raymond Choo	Adaptive reputation updating (ARU), Phasor measurement units (PMUs)		smart grid cyber-physical system	25	[110]
2017	Security Enrichment in Intrusion Detection System Using Classifier Ensemble	Uma R. Salunkhe, SureshN. Mali	J48, Logistic Regression and Naive Bayes	KDD99	Normal, Phf, Teardrop, Loadmodule, Smurf	2	[111]
2018	On Cloud Environment Using Hybrid Classification and OTS Generation by Enhanced Intrusion Detection and Prevention System	V.Balamurugan and R.Saravanan	K-means clustering, neural network, PSA, NK-RNN	SCADA, KDD	DDOS, flood attacks, U2R attacks, zero day attacks and R2L	2	[112]
2018	Signature anomaly intrusion detection system	Roshan Sharma Deepak Sharma	k-means, neural network	KDD99		not yet	[88]

Table 3.15: Paper details using ensemble learning from 2009, 2011-2015.

Year	Title of The Paper	Author(s)	Algorithm(s)	Data set(s)	Find Out Attacks	Cited by	Reference
2009	Novel intrusion detection using probabilistic neural network & adaptive boosting	Tich Phuoc Tran, Longbing Cao, Dat Tran and Cuong Duc Nguyen	NN, AdaBoost, BSPNN	KDD-99	Normal, Probe, DoS, U2R, R2L	30	[89]
2011	A novel artificial immune system for fault behavior detection	C.A. Laurentys, R.M. Palhares and W.M. Caminhas	GA, Majority Vote	not use	fault behavior detection	45	[90]
2011	Adaptive intrusion detection based on boosting & naive Bayesian classifier	Dewan Md. Farid, Mohammad Zahidur Rahman and Chowdhury Mofizur Rahman	NB, AdaBoost	KDD9	Normal, Probe, DoS, U2R, R2L	41	[91]
2011	Incremental SVM based on reversed set for network intrusion detection	Yang Yi, Jiansheng Wu and Wei Xu	SVM, ISVM	KDD99	U2R and R2L	65	[92]
2012	Decision tree based light weight intrusion detection using a wrapper approach	Siva S. Sivatha Sindhu, S. Geetha and A. Kannan	Neural ensemble decision tree	KDD99	DoS	208	[93]
2013	An adaptive ensemble classifier for mining concept drifting data streams	Dewan Md. Farid, Li Zhang, Alamgir Hossain, C.M.Rahman, R.Strachan, G.Sexton and K.Dahal	NB, C4.5 and AdaBoost	NSLKDD		94	[94]
2014	Compilation model to classify attacks by choosing a function based on KDD-99 and NSL-KDD datasets	A.K.Shrivas and A.Kumar Dewangan	ANN, Bayesian Network, Gain ratio FS	KDD99, NSLKDD	DoS and U2R	22	[95]
2015	A Collaborative Intrusion Detection Mechanism Against False Data Injection Attack in Advanced Metering Infrastructure	Xiaoxue Liu, Peidong Zhu, Yan Zhang	Smart Meter Architecture Based On Petri NET, colored Petri net			40	[96]

Table 3.16: Paper details using ensemble learning from 2015-2018.

Year	Title of The Paper	Author(s)	Algorithm(s)	Data set(s)	Find Out Attacks	Cited by	Reference
2015	Development and analysis of multi-level intrusion detection system to automate production processes	C Zhou, S Huang, N Xiong, S-Hua Yang, H.Li, Y.Qin and X.Li	Hidden Markov model, Multi model based anomaly ids		spoofing, tampering and dos	39	[97]
2015	APT malware detection based on malicious DNS and traffic analysis	G Zhao, K.Xu, L.Xu and B.Wu	J48, C4.5 and 10-fold cross validation			34	[98]
2015	CANN: An intrusion detection system based on combining cluster centers and nearest neighbors	Wei-Chao Lin, Shih-Wen Ke , Chih-Fong Tsai	K-NN, SVM, CANN	KDD99	U2L and R2L	180	[99]
2015	Intrusion Detection for Applying Data Mining Techniques	Jonathon Ng, D Joshi and Shankar M.Banik	Clustering		dos	15	[100]
2016	A novel SVM-kNN-PSO ensemble method for intrusion detection system	Abdulla Amin Aburomman, Mamun Bin Ibne Reaz	SVN, KNN, PSO, LUS, WMA	KDD99	Normal,DoS, Probe, U2R and R2L.	100	[101]
2017	Intrusion detection system for improvement of ensemble technology based on vector support machine and neural network	A.K.Siddiqui and T Farooqui	SVM, NN, KNN	KDD99	Normal, DoS, Probe, R2L, U2R	not yet	[113]
2018	Network intrusion detection using clustering and Gradient boosting	Parag Verma,Shayan Anwar, Shadap khan, Dr. Punil B Mane	DT Adaptive Boosting K-means clustering Gradient boosting	NSL-KDD		not yet	[102]
2018	Network intrusion detection mechanisms using outlier detection	Ch. Nagamani , Dr. Suneetha Chittineni	Random Forest	NSL-KDD		not yet	[103]

Chapter 4

Data Set and Comparison of Performance

4.1 Data sets

In 1998, the DARPA data set was created by the Cyber System and the technology group of the Lincoln Institute at the Massachusetts Institute of Technology. DARPA is a raw data set. KDD99 is the DARPA data extraction version. The KDD99 data set was used at the 3rd International Competition for Knowledge Discovery and Data Development held at the 5th International Conference on Intellectual Discovery and Data Development KDD99. This is a very old record. In 1999, the KDD99 data set also creates the context for data analysis, that year they provided real traffic for 7 weeks. A record KDD99 Cup amounts to 5 billion and the function is also very excellent. In addition, they proposed a new NSLKDD data set, consisting of individual records of the entire KDD record. NSLKDD is a remote and compact version of the KDD99 data set.

Figure 4.1: The relation between DARPA, KDD99 and NSL-KDD.

Data set name	Training Size	Testing Size	Information
DARPA	6.591.458 kb (6.2gb)	3.853.522 kb (3.67gb)	Raw data set
KDD99	4898431	311029	Features extracted and preprocessed for machine learning
NSL-KDD	125973	22544	Duplicates removed, size reduced

Table 4.1: DARPA, KDD99, NSL-KDD data sets information.

Characteristics of the Data Set

There are 41 features, including not only normal traffic, but also DoS, U2R, R2L, Probe.

1. Normal: Normal connections are performed in that it is not classified as an attack.

2. Denial of Service (DoS): In denial of service attacks, hackers will deny service to authorized users. In addition to clear rejected, this computer, on the server or on the network, does not respond to user requests and also includes the delay time of critical operations. Hackers, all available bandwidth controlled from resources, consuming or wasting the capacity of hard disk or disk storage and create these delays. They either crash into a car, find a way to be a bad operation, it's impossible to work as intended. Denial of service attacks are not intentionally malicious. In the early stages of network computing, this can lead to chaos to individuals and the website was quite simple. In the past, one of the hackers was using TCP / IP. Hacker, it takes a large load on the server, without pouring the request multiplier ping on the server, while there is no response, the server was almost closed. Early hackers developed a method to exploit a vulnerability in the implementation of the TCP / IP packet, ICMP. This is because in the previous version of the operating system it was found that it was not able to cope with these errors. Hacker effectively protect the server, he rejected all requests for legitimate services. In addition to TCP / IP vulnerabilities, hackers can attack a poorly designed website.

3. User to Root (U2R): These attack intruders somehow become a root user by privileges using the system. The ideal way to detect u2r is through a host-based

detector which may use inputs such as the changes to the filesystem, syscall trace and process table.

4. Remote to Local (R2L): R2L is an attack with which an attacker can gain access to the local machine through a network connection. For most R2L attacks, an attacker penetrates the computer system via the Internet. A typical example of an R2L attack is the use of software vulnerabilities for guessing passwords and accessing local computers.

5. Probing (Probe): Probing attacks involve some sort of method to gather information or find some known debilitation. Most of that action is noisy and detectable by Snort and given the correct rules. Overall, sixteen rules detected over 20% of the probe connections.

In the KDD99 data set and the NSLKDD data set, these four attacks are divided into 22 different attacks-

Sr. No.	Four Attacks	22 Attacks Class
1	DOS	neptune, land, back, smurt, teardrop, pod
2	U2R	perl, loadmodule, buffer_overflow, rootkit
3	R2L	guess_passwd, ftp_write, imap, warezmaster, multihop, phf,spy,warezclient
4	Probe	ipsweep, portsweep, satan, nmap

Table 4.2: Different types of attacks on KDD 99 data set and NSLKDD data set.

Features

The KDD 99 record has 41 features. It includes-

-nine main features of individual TCP connections

-13 features of content in the connection, offered by domain knowledge and

-nine traffic features calculated with a second time window.

The table is as follows-

Index	Label	Data Type	Index	Label	Data Type
1	duration	conti.	22	is-guest-login	discr.
2	protocol-type	discr.	23	count	conti.
3	service	discr.	24	srv-count	conti.
4	flag	discr.	25	s-error-rate	conti.
5	src-bytes	conti.	26	srv-serror-rate	conti.
6	dst-bytes	conti.	27	r-error-rate	conti.
7	land	discr.	28	srv-rerror-rate	conti.
8	wrong-fragment	conti.	29	same-srv-rate	conti.
9	urgent	conti.	30	diff-srv-rate	conti.
10	hot	conti.	31	srv-diff-host-rate	conti.
11	num-failed-logins	conti.	32	dst-host-count	conti.
12	logged-in	discr.	33	dst-host-srv-count	conti.
13	num-compromised	conti.	34	dst-host-same-srv-rate	conti.
14	root-shell	conti.	35	dst-host-diff-srv-rate	conti.
15	su-attempted	conti.	36	dst-host-same-sre-port-rate	conti.
16	num-root	conti.	37	dst-host-srv-diff-host-rate	conti.
17	num-file-creations	conti.	38	dst-host-serror-rate	conti.
18	num-shells	conti.	39	dst-host-srv-serror-rate	conti.
19	num-access-files	conti.	40	dst-host-rerror-rate	conti.
20	num-outbound-cmds	conti.	41	dst-host-srv-rerror-rate	conti.
21	is-host-login	discr.	42	attack-type	discr.

Table 4.3: The KDD 99 data set features.

DARPA, KDD99, NSL-KDD, Credit Approval, DHTML and UNSW-NB15 are records used for classification tasks. KDD99 is the most commonly used entry. Figure 4.2 shows the annual operation of the data set from 2009 to 2018. These data sets are published and recognized as standard intrusion detection system (IDS) records.

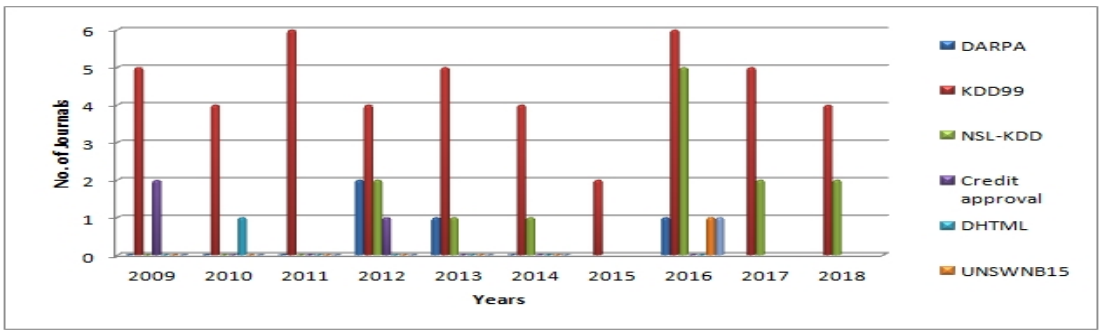

Figure 4.2: Shows the annual operation of the data set from 2009 to 2018.

4.2 Comparison of Performance

Table 4.4: Shows the distribution of single classifier by year in terms of results for each paper (2009-2012).

Year	Title of The Paper	Result Percentage(%)	Ref
2009	Association rules applied to credit card fraud detection	Certainty factor: 80.08	[31]
2009	Data Mining based intrusion detectors	Decision Rate: 70.62, FAR: 1.44	[32]
2009	Some issues about out-lier detection in rough-set theory	Decision Rate: SEQ: 90, DIS:92	[33]
2009	Project out-lier detection in high-dimensional mixedattributes data set	Decision Rate: Credit approval data: 70	[34]
2009	Information inconsistencies detection use a rule-map	Error=5.0%, Train Set:5, Test Set:4	[35]
2010	Malicious web-content detection	Accuracy: Naïve Bayes: 58.28, DT: 94.74, SVM: 93.51, Boosted DT: 96.14	[36]
2010	Random effect logistic regression model for anomaly detection	Classification accuracy on training data set: Normal: 79.43, Attack: 20.57	[37]
2010	The Decision Model for Responding to an Intrusion Based on a Hierarchical Network Task Plan		[38]
2010	Neural Networks based detection of stepping stone intrusion	Accuracy : 99.0	[39]
2011	Evolving boundary detectors for anomaly detection	Decision Rate: KDD99 data set: DOS: 94.5, U2R: 78.85, Probe: 93.64, R2L: 50.69	[40]
2011	Network Failure Service Anomaly Detection Technology	false positive rate: KNN: 19.43, and SVM: 6.45	[41]
2012	Hidden Naïve Bayes Network Intrusion Detection System Based on Multi-Class Classifier	93.73	[42]
2012	An autonomous labeling approach to support vector machines algorithms for network traffic anomaly detection	Decision Rate: 88.64 (80% attack), 98.37 (1% attack)	[43]
2012	A differentiated one-class classification method with applications to intrusion detection	Decision Rate: M=200*, Targeted attack: 96.9	[44]
2012	Madame ID For Intrusion Detection Using Data Mining	false positive rate: 75.00	[45]

Table 4.5: Shows the distribution of single classifier by year in terms of results for each paper (2013-2018).

Year	Title of The Paper	Result Percentage(%)	Ref
2013	A Cost - Conscious Decision Tree Approach to Detecting Fraud	Gini: 92.8	[46]
2013	Accuracy Detection for an Intrusion Detection System Using a Neural Network Classifier.	Accuracy : FFNN: 79.49, ENN: 78.1, GRNN: 58.74, PNN:85.50 and RBNN: 83.51	[47]
2013	Abrupt change detection with one class time adaptive Support Vector Machine	495.9 sequences correctly classified within 500 sequences.	[48]
2013	Application of feature detection rule to the negative selection algorithm	Feature Detection rule: 0.9375	[49]
2014	A New Approach to the Selection of Functions, Based on the Optimization Algorithm of Cuttlefish for an Intrusion Detection System	Decision Rate: 71.067 and false positive rate: 17.685	[50]
2016	An Approach for Intrusion Detection Using Novel Gaussian Based Kernel Function		[53]
2016	Feature Selection for Intrusion Detection System Using Ant Colony Optimization	Accuracy: 98.9, 2.59 false positive rate, Normal: 97.41, DoS: 99.78, Probe: 74.65, U2R: 93.52, R2L: 99.17	[51]
2016	A multi-level intrusion detection method for abnormal network behaviors	average accuracy: SVM: 95.5471, Naïve Bayes: 89.024 and NN: 96.67	[52]
2016	Assessment of the Network Anomaly Detection System: Statistical Analysis of the UNSW-NB15 Data Set and Comparison With the KDD99 Data Set	KDD99 data set accuracy: DT: 92.30, LR: 92.75, NB: 95.00, ANN: 97.04, EM: 78.06 UNSW-NB15 data set accuracy: DT: 85.56, LR: 83.15, NB: 82.07, ANN: 81.34, EM: 78.47	[54]
2017	Fuzziness based semi-supervised learning approach for intrusion detection system	accuracy: KDDTest^{+} 84.12 and KDDTest^{21} 68.82	[104]
2017	Intrusion Detection Model Using fusion of Chi-square feature selection and multi class SVM	accuracy: 98%	[107]
2017	Intrusion Detection System Based On Flows Using Machine Learning Algorithms	accuracy: 93.57	[105]
2017	Enhancing Collaborative Intrusion Detection Networks Against Insider Attacks Using Supervised Intrusion Sensitivity-Based Trust Management Model	accuracy: increased from 0.88 to 0.96	[106]
2018	Detection of Intrusion Into Computer Network Using Various Classifiers and Ensemble Learning	Accuracy: NN: 91.73%, DT: 88.00%, Logistic regression: 80.2%	[84]
2018	Machine Learning and deep learning for cyber security		[85]
2018	An intrusion detection using model based on Fuzzy c means clustering	FCM accuracy: 51.08, Detect rate: 50.92%, False alarm: 37.48%	[86]
2018	Data mining techniques in intrusion detection		[87]

Table 4.6: Shows the distribution of hybrid classifier by year in terms of results for each paper (2009-2013).

Year	Title of The Paper	Result Percentage(%)	Ref
2009	Abnormality Detection Method Using Fusion of Unsupervised Control Neural Network	True Positive rate: 97.00 (Dos), 71.65 (Probe), 26.69 (R2L)	[55]
2009	Adaptive Genetic Intrusion Detection Signature Training System	Accuracy: 92	[56]
2010	Novel Approach Using Fuzzy Clustering and Artificial Neural Networks to Intrusion Detection	Accuracy: 96.71	[57]
2010	A distributed sinkhole detection method using cluster analysis	Decision Rate: 96.61	[58]
2010	Design Network Intrusion Detection System using hybrid Fuzzy-Neural Network	Accuracy: 100 (Dos), 100 (U2R), 99.8 (Probe), 40 (R2L) and 68.6 (Unknown)	[59]
2010	Histological Algorithm of Negative Selection Based on Chaotic Approach and Its Application to Detection of Abnormalities and Anomalies	Accuracy: 97.65	[60]
2011	Detect Online Games and Lottery Fraud	Decision Rate: 98	[61]
2011	Fast Detection of Outliers for Very Large Log Data		[62]
2011	An Integrated Intrusion Detection System for Cluster-based Wireless Sensor Networks	Accuracy: 95.13	[64]
2011	Development and Analysis of Genetic Fuzzy System for Intrusion Detection in Computer Network	Decision Rate: 88.13 (Mitchigan), 99.53 (Pitsburg), 93.2 (IRL)	[63]
2011	RealTime Anomaly Detection for DoS Attacks Using KNN Weighted Classifiers	Accuracy: 97.42 (known) and 78 (unknown)	[65]
2011	Online Anomaly Detection for Self-Adaptive and Dynamic Clustering	Decision Rate: 83.4 (offline), 84.4 (online)	[66]
2012	An efficient intrusion detection system based on support vector machines and gradually feature removal method	Decision Rate: 98.6249	[67]
2012	The combined approach for anomaly detection using neural networks & clustering techniques	Decision Rate: 98.5 (Dos)	[68]
2012	Anomaly Based Intrusion Detection System Using Artificial Neural Network and Fuzzy Clustering		[69]
2013	By Using Neural Network and SVM Building a Enhancement Hybrid Intrusion Detection System	Accuracy: 98.94 (Dos), 97.11 (Probe), 97.80 (U2R), 97.78 (R2L)	[70]
2013	Advanced probabilistic approach for network intrusion forecasting and detection	Decision Rate: 90	[71]
2013	Intrusion Detection System Based on NSLKDD Data Set by Fuzzy Clustering and Neural Network in Hybrid	Recall: 99.1 (Dos), 94.1 (Prob), 78 (U2R), 89 (R2L)	[72]

Table 4.7: Shows the distribution of hybrid classifier by year in terms of results for each paper (2013-2018).

Year	Title of The Paper	Result Percentage(%)	Ref
2013	Hybrid structure based on neural network MLP & K-means clustering implies for intrusion detection system	Decision Rate: 99.99 (Dos), 99.97 (Probe), 99.99 (U2R) and 99.98 (R2L)	[73]
2013	A novel hybrid intrusion detection method integrating anomaly detection with misuse detection	Decision Rate: 99.98 (known attack) and 97.4 (unknown attack)	[74]
2014	New Clustering Method for Abnormality Detection	Decision Rate: 96.12 (Dos), 90.10(Probe), 70.51(U2R) and 70.13 (R2L) Accuracy: 96.38	[76]
2014	Mining Network Data for Intrusion Detection through combining SVMs with ant colony networks	Decision Rate: 94.86	[75]
2015	Detecting Wireless Anomalies Depends On IEEE802.11 Behavior Analysis	Decision Rate: 99.00	[77]
2015	Intrusion Detection System (IDS): Anomaly Detection Using Outlier Detection Approach		[79]
2015	Developing a Hybrid Intrusion Detection System Using Data Mining for Power Systems	average detection accuracy: 73.43	[78]
2016	A Hybrid Data Mining Approach for Intrusion Detection on Imbalanced NSL-KDD Data set	Accuracy: R2L:92.97, U2R:55.91	[80]
2016	Hybrid mode consists of SVM and GA for Intrusion detection system	True-positive: 0.973, false-positive: 0.017	[81]
2016	Building an intrusion detection system using a filter-based feature selection algorithm	99.94 accuracy, 98.93 detection rate and 0.28 false positive rate	[82]
2016	An effective MCLP / SVM based intrusion detection system optimized by optimizing particle steepness in the time zone	TVCPSO-MCLP accuracy: 94.69% and TVCPSO-SVM accuracy: 95.75%	[83]
2017	A Hypergraph and Arithmetic Residue-based Probabilistic Neural Network for Classification in Intrusion Detection Systems	Accuracy: 97.63 (Normal), 98.86 (DoS), 88.12 (Probe), 82.41 (R2L), 82.88 (U2R)	[108]
2017	A novel technique for intrusion detection system for network security using hybrid SVM-CART	accuracy: 95%	[109]
2017	Distributed Host-based Collaborative Detection for False Data Injection Attacks in Smart Grid Cyber-Physical System		[110]
2017	Security Enrichment in Intrusion Detection System Using Classifier Ensemble		[111]
2018	On Cloud Environment Using Hybrid Classification and OTS Generation by Enhanced Intrusion Detection and Prevention System	false alarm rate: 4%, Detection rate: 85%	[112]
2018	Signature anomaly intrusion detection system		[88]

Table 4.8: Shows the distribution of ensemble classifier by year in terms of results for each paper (2009-2018).

Year	Title of The Paper	Result Percentage(%)	Ref
2009	Novel intrusion detection using probabilistic neural network & adaptive boosting	Decision Rate: 94.31	[89]
2011	A novel artificial immune system for fault behavior detection	Decision Rate: 97.85	[90]
2011	Adaptive intrusion detection based on boosting & naive Bayesian classifier	Decision Rate : 99.75	[91]
2011	Incremental SVM based on reversed set for network intrusion detection	Decision Rate: 81.377	[92]
2012	Decision tree based light weight intrusion detection using a wrapper approach	Decision Rate: 98.38	[93]
2013	An adaptive ensemble classifier for mining concept drifting data streams	Decision Rate: 92.65	[94]
2014	Compilation model to classify attacks by choosing a function based on KDD-99 and NSL-KDD datasets	Decision Rate: 97.53 (NSLKDD) and 99.41 (KDD99)	[95]
2015	A Collaborative Intrusion Detection Mechanism Against False Data Injection Attack in Advanced Metering Infrastructure	average intrusion detection rate 69.6%	[96]
2015	Development and analysis of multi-level intrusion detection system to automate production processes	Detection Accuracy: 98.61%	[97]
2015	APT malware detection based on malicious DNS and traffic analysis	10-fold cross validation: 96%	[98]
2015	CANN: An intrusion detection system based on combining cluster centers and nearest neighbors	CANN (k = 1) 99.46%, SVM (degree = 2) 95.37%, k-NN (k = 1) 99.89%	[99]
2015	Intrusion Detection for Applying Data Mining Techniques		[100]
2016	A novel SVM-kNN-PSO ensemble method for intrusion detection system		[101]
2017	Intrusion detection system for improvement of ensemble technology based on vector support machine and neural network	Detection Rate: 95.8 (Normal), 93.83 (Dos), 93.83 (Probe), 96.5 (R2L), 95.5 (U2R)	[113]
2018	Network intrusion detection using clustering and Gradient boosting	Accuracy: Gradient boost and adaboost with k-means clustering: 84.00% and 82.00%	[102]
2018	Network intrusion detection mechanisms using outlier detection		[103]

Decision:

Single-

Considering the table, single classifier used Neural Network algorithm and achieve 99.0% accuracy.

[Neural Networks based detection of stepping stone intrusion - Han-Ching Wu, Shou-Hsuan and Stephen Huang] [39]

Hybrid-

Considering the table, hybrid classifier proposed Flexible Mutual Information Feature Selection (FMIFS), and Least Square Support Vector Machine based IDS (LSSVM-IDS). Accuracy: 99.94%.

[Building an intrusion detection system using a filter-based feature selection algorithm - Mohammed A. Ambusaidi, Xiangjian He, Priyadarsi Nanda and Zhiyuan Tan] [82]

Ensemble-

Considering the table, ensemble classifier used Naïve Bayesian and AdaBoost algorithms and achieve 99.75% decision rate with KDD99 data set. **[Adaptive intrusion detection based on boosting & Naïve Bayesian classifier - Dewan Md. Farid, Mohammad Zahidur Rahman and Chowdhury Mofizur Rahman] [91]**

Chapter 5

Conclusions and Future Work

5.1 Conclusions

Various activities are performed to detect and prevent intrusions. A number of machine learning methods are used for intrusion detection systems, including single, hybrid and ensemble classifiers. Although these methods are very useful for IDS, it is well known that there is no single method for detecting all types of attacks. Therefore, further efforts are needed to improve the ability to detect all types of attacks using machine learning methods and reduce false positives. The use of several classification methods in an intrusion detection system is an increase in machine learning and research in the field of artificial intelligence. This attracted the attention of researchers for a long time. In this thesis, we have reviewed 84 research papers related to the use of various classifiers for the detection of intruders declared in the period from 2009 to 2018.

5.2 Future Work

In future, we will work on adaptive ensemble classifiers to deal with novel intrusions in concept drifting environment, where feature importance may change over the time.

Bibliography

[1] Srilatha Chebrolu, Ajith Abraham, and Johnson P. Thomas, "Feature Deduction and Ensemble Design of Intrusion Detection Systems," *International Journal of Computers and Security, vol. 24, pp. 295–307, 2005.* 3

[2] Rajni Tewatia, and Asha Mishra, "Introduction To Intrusion Detection System: Review," *International journal of scientific technology research volume 4, issue 05,pp. 219-223(5), may 2015.* 6

[3] James P. Anderson, "Computer security threat monitoring and surveillance," *Technical Report, James P. Anderson Company, Fort Washington, USA, 1980.* 8

[4] Dorothy E. Denning, "An Intrusion Detection Model," *IEEE Transaction on Software Engineering, 1987, SE-13(2), pp. 222-232.* 8

[5] Smaha and Stephen E., "Haystack: An Intrusion Detection System," *The 4th Aerospace Computer Security Applications Conference, Orlando, FL, pp. 37-44, December 1988.* 8

[6] S. Forrest, S.A. Hofmeyr, A. Somayaji, T.A. Longstaff, "A Sense of Self for Unix Processes," *Proc. of the IEEE Symposium on Research in Security and Privacy, Oakland, CA, USA, pp. 120-128, 1996.* 8

[7] Wenke Lee, S.J. Stolfo, K.W. Mok, "A Data Mining Framework for Building Intrusion Detection Models," *Proceedings of the 1999 IEEE Symposium on Security and Privacy, Oakland, CA, USA, , pp. 120-132, 14 May 1999.* 8

[8] Alfonso Valdes and Keith Skinner, "Adaptive Model-Based Monitoring for Cyber Attack Detection," *in Recent Advances in Intrusion Detection Toulouse, France, 11 November 2000, pp. 80-93.* 12

[9] Wun-Hwa Chen, Sheng-Hsun Hsu, Hwang-Pin Shen, "Application of SVM and ANN for Intrusion Detection," *Computers Operations Research, vol. 32, pp. 2617–2634, 2005, ELSEVIER.* 14

[10] Christopher Kruegel, Darren Mutz, William Robertson, Fredrik Valeur, "Bayesian Event Classification for Intrusion Detection," *Proc. of the 19th Annual Computer Security Applications Conference, Las Vegas, NV, pp. 14, 2003.* 12

[11] Nahla Ben Amor, Salem Benferhat, and Zied Elouedi, "Naïve Bayes vs. Decision Trees in Intrusion Detection Systems," *2004 ACM Symposium on Applied Computing, New York, pp. 420-424, 2004.* 12

[12] Latifur Khan, Mamoun Awad and Bhavani Thuraisingham, "A New Intrusion Detection System Using Support Vector Machines and Hierarchical Clustering," *The VLDB Journal, vol. 16, pp. 507–521. ISSN 1066-8888, 2007.* 14

[13] Yacine Bouzida and Frédéric Cuppens, "Detecting Known and Novel Network Intrusions," *Security and Privacy in Dynamic Environments, pp. 258-270, 2006.* 12

[14] Mrutyunjaya Panda, and Manas Ranjan Patra, "Network Intrusion Deteciton Using Naïve Bayes," *International Journal of Computer Science and Network Security(IJCSNS), Vol. 7, pp. 258-263, December 2007.* 12

[15] Weiming Hu, Wei Hu, and Steve Maybank, "Adaboost-Based Algorithm for Network Intrusion Detection," *IEEE Transactions on Systems, Man, and Cybernetics, Part B(Cybernetics), Volume: 38, Issue: 2, pp. 577-583, April 2008.* 15

[16] Dewan Md. Farid, Nouria Harbi and Mohammad Zahidur Rahman, "Combining Naive Bayes and Decision Tree for Adaptive Intrusion Detection," *International Journal of Network Security Its Applications (IJNSA), pp. 12-25, April 2010.* 12

[17] Dewan Md. Farid, Nouria Harbi, Emna Bahri, Mohammad Zahidur Rahman, Chowdhury Mofizur Rahman, "Attacks Classification in Adaptive Intrusion Detection Using Decision Tree," *International Conference on Computer Science (ICCS), pp. 86-90, 2010.* 12

[18] Dewan Md. Farid and Mohammad Zahidur Rahman , "Anomaly Network Intrusion Detection Based on Improved Self Adaptive Bayesian Algorithm," *Journal of Computers, vol. 5, pp. 23–31, 2010.* 11

[19] Dewan Md. Farid, Li Zhang, Chowdhury Mofizur Rahman, M. A. Hossain, and Rebecca Strachan, "Hybrid Decision Tree and Naive Bayes Classifiers for Multi-Class Classification Tasks," *Expert Systems with Applications, vol. 41, no. 4, pp. 1937–1946, 2014.* 11

[20] Richard O. Duda, Peter E. Hart, David G. Stork, "Pattern Classification," *International Journal of Computational Intelligence and Applications, vol. 1, pp. 335–339, 2001.* 12

[21] Akashdeep, Ishfaq Manzoor, Neeraj Kumar, "A Feature Reduced Intrusion Detection System Using ANN Classifier," *Expert Systems With Applications, vol. 88, pp. 249-257, 2017, ELSEVIER.* 13

[22] John E. Dickerson and Julie A. Dickerson, "Fuzzy Network Profiling for Intrusion Detection," *Annual Conference of The North American Fuzzy Information Processing Society-nafips, pp. 301-306, February 2000* 14

[23] Nutan Farah Haq, Musharrat Rafni, Abdur Rahman Onik, Faisal Muhammad Shah, Md. Avishek Khan Hridoy and Dewan Md. Farid, "Application Of Machine Learning Approaches in Intrusion Detection System: A Survey," *(IJARAI) International Journal of Advanced Research in Artificial Intelligence, Vol. 4, No.3, pp. 9-18, 2015.* 9, 11, 13, 14

[24] Anna L. Buczak and Erhan Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection," *IEEE communications surveys tutorials, vol. 18, no. 2, pp. 1153 - 1176, 2016.* 9

[25] Nicholas J. Miller, "Benchmarks for Evaluating Anomaly-Based Intrusion Detection Solutions," *Presented to the Department of Computer Engineering and Computer Science California State University, Long Beach, 2018.* 9

[26] A. J. M. Abu Afza, Dewan Md. Farid, and Chowdhury Mofizur Rahman, "A Hybrid Classifier using Boosting, Clustering, and Naïve Bayesian Classifier," *World*

of Computer Science and Information Technology Journal (WCSIT),Vol. 1, No. 3, pp. 105-109, 2011 15

[27] Chih-Fong Tsai, Yu-Feng Hsu, Chia-Ying Lin, Wei-Yang Lin (2009), "Intrusion Detection by Machine Learning: A Review," *expert systems with applications, Vol. 36, Issue 10, pp. 11994-12000, December 2009, ELSEVIER.* 9, 16

[28] Shu-Hsien Liao, Pei-Hui Chu, Pei-Yuan Hsiao (2011), "Data Mining Techniques and Applications – a Decade Review From 2000 to 2011," *expert systems with applications, Volume 39, Issue 12, 15 September 2012, Pages 11303-11311, ELSEVIER.* 9

[29] Michał Wozniak, Manuel Graña, and Emilio Corchado (2014), "A Survey of Multiple Classifier Systems as Hybrid Systems," *expert systems with applications, Volume 16, March 2014, pp. 3-17, ELSEVIER.* 9

[30] Bruno Bogaz Zarpelão, Rodrigo Sanches Miani, Cláudio Toshio Kawakani, Sean Carlisto de Alvarenga (2017), "A Survey of Intrusion Detection in Internet of Things," *expert systems with applications, Volume 84, 15 April 2017, pp. 25-37, ELSEVIER.* 9

[31] D. Sanchez, M.A. Vila, L. Cerda, J.M. Serrano (2009), "Association Rules Applied to Credit Card Fraud Detection," *expert systems with applications, Volume 36, Issue 2, Part 2, March 2009, pp. 3630-3640, ELSEVIER.* 16, 20, 37

[32] Su-Yun Wu, Ester Yen (2009). "Data Mining-Based Intrusion Detectors," *expert systems with applications, Volume 36, Issue 3, Part 1, April 2009, pp. 5605-5612, ELSEVIER.* 16, 18, 20, 37

[33] Feng Jiang, Yuefei Sui and Cungen Cao (2009). "Some Issues About Outlier Detection in Rough Set Theory," *expert systems with applications, Volume 36, Issue 3, Part 1, April 2009, pp. 4680-4687, ELSEVIER.* 16, 20, 37

[34] Mao Ye, Xue Li, Maria E. Orlowska (2009). "Projected Outlier Detection in High Dimensional Mixed-Attributes Data Set," *expert systems with applications, Volume 36, Issue 3, Part 2, April 2009, Pp. 7104-7113, ELSEVIER.* 16, 20, 37

[35] Jun Ma, Jie Lu, Guangquan (2009). "Information Inconsistencies Detection Using a Rule-Map Technique," *expert systems with applications, Volume 36, Issue 10, December 2009, pp. 12510-12519, ELSEVIER.* 16, 20, 37

[36] Yung-Tsung Hou, Yimeng Chang, Tsuhan Chen, Chi-Sung Laih, Chia-Mei Chen (2010), "Malicious Web Content Detection by Machine Learning," *expert systems with applications, Volume 37, Issue 1, January 2010, pp. 55-60, ELSEVIER.* 16, 18, 20, 37

[37] Min Seok Mok, So Young Sohn, Yong Han Ju (2010), "Random Effects Logistic Regression Model for Anomaly Detection," *expert systems with applications, Volume 37, Issue 10, October 2010, Pages 7162-7166, ELSEVIER.* 16, 20, 37

[38] Chengpo Mua, Yingjiu Li (2010), "An Intrusion Response Decision Making Model Based on Hierarchical," *expert systems with applications, Volume 37, Issue 3, 15 March 2010, Pages 2465-2472, ELSEVIER.* 16, 20, 37

[39] Han-Ching Wu, Shou-Hsuan Stephen Huang (2010), "Neural Networks-Based Detection of Stepping-Stone Intrusion," *expert systems with applications, Volume 37, Issue 2, March 2010, pp. 1431-1437, ELSEVIER.* 16, 18, 20, 37, 42

[40] Wang Dawei, Zhang Fengbin, Xi Liang (2011), "Evolving Boundary Detector for Anomaly Detection," *expert systems with applications, Volume 38, Issue 3, March 2011, Pages 2412-2420, ELSEVIER.* 16, 21, 37

[41] G. Davanzo, E. Medvet, A. Bartoli (2011), "Anomaly Detection Techniques for a Web Defacement Monitoring Service," *expert systems with applications, Volume 38, Issue 10, 15 September 2011, Pages 12521-12530, ELSEVIER.* 13, 16, 18, 21, 37

[42] Levent Koc, Thomas A. Mazzuchi, Shahram Sarkani (2012), "A Network Intrusion Detection System Based on a Hidden Naïve Bayes Multiclass Classifier," *expert systems with applications, Volume 39, Issue 18, 15 December 2012, Pages 13492-13500, ELSEVIER.* 16, 18, 21, 37

[43] Carlos A. Catania, Facundo Bromberg, Carlos García Garino (2012), "An Autonomous Labeling Approach to Support Vector Machines Algorithms for Network

Traffic Anomaly Detection," *expert systems with applications, Volume 39, Issue 2, 1 February 2012, Pages 1822-1829, ELSEVIER.* 16, 18, 21, 37

[44] Inho Kang, Myong K. Jeong, Dongjoon Kong (2012), "A Differentiated One-Class Classification Method With Applications to Intrusion Detection," *expert systems with applications, Volume 39, Issue 4, March 2012, Pages 3899-3905, ELSEVIER.* 16, 18, 21, 37

[45] Prabhjeet Kaur, Amit Kumar Sharma, Sudesh Kumar Prajapat (2012), "Madam Id for Intrusion Detection Using Data Mining," *International Journal of Research in IT Management, Volume 2, Issue 2, ISSN 2231-4334, pp. 256-263, February 2012, IJRIM.* 16, 18, 21, 37

[46] Yusuf Sahin, Serol Bulkan, Ekrem Duman (2013), "A Cost-Sensitive Decision Tree Approach for Fraud Detection," *expert systems with applications, Volume 40, Issue 15, 1 November 2013, Pages 5916-5923, ELSEVIER.* 16, 18, 22, 38

[47] S. Devaraju, S. Ramakrishnan (2013), "Detection of Accuracy for Intrusion Detection System Using Neural Network Classifier," *International Journal of Emerging Technology and Advanced Engineering, ISSN 2250-2459, Volume 3, Issue 1, pp. 338-345, January 2013.* 16, 18, 22, 38

[48] Guillermo L. Grinblat, Lucas C. Uzal, Pablo M. Granitto (2013), "Abrupt Change Detection With One-Class Time-Adaptive Support Vector Machines," *expert systems with applications, Volume 40, Issue 18, 15 December 2013, Pages 7242-7249, ELSEVIER.* 16, 18, 22, 38

[49] Mario Poggiolini, Andries Engelbrecht (2013), "Application Of the Feature Detection Rule to The Negative Selection Algorithm," *expert systems with applications, Volume 40, Issue 8, 15 June 2013, Pages 3001-3014, ELSEVIER.* 16, 22, 38

[50] Adel Sabry Eesa, Zeynep Orman, Adnan Mohsin Abdulazeez Brifcani (2014), "A Novel Feature-Selection Approach Based on the Cuttlefish Optimization Algorithm for Intrusion Detection Systems," *expert systems with applications, Volume 42, Issue 5, 1 April 2015, Pages 2670-2679, ELSEVIER.* 16, 22, 38

[51] Mehdi Hosseinzadeh Aghdam and Peyman Kabiri (2014), "Feature Selection for Intrusion Detection System Using Ant Colony Optimization," *International Journal of Network Security, Vol.18, No.3, PP.420-432, May 2016.* 16, 23, 38

[52] Soo-Yeon Ji, Bong-Keun Jeong, Seonho Choi, Dong Hyun Jeong (2016), "A Multi-Level Intrusion Detection Method for Abnormal Network Behaviors," *Journal of Network and Computer Applications, Volume 62, February 2016, Pages 9-17.* 16, 18, 22, 38

[53] Gunupudi Rajesh Kumar, Nimmala Mangathayaru, Gugulothu Narsimha (2016), "An Approach for Intrusion Detection Using Novel Gaussian Based Kernel Function," *Journal of Universal Computer Science, vol. 22, no. 4 (2016), pp. 589-604.* 16, 22, 38

[54] Nour Moustafa and Jill Slay (2016), "The Evaluation of Network Anomaly Detection Systems: Statistical Analysis of The UNSW-NB15 Data Set and the Comparison With The KDD99 Dataset," *Information security journal: A global perspective, ISSN: 1939-3555, pp. 1-14, 13 January 2016.* 16, 18, 23, 38

[55] M. Bahrololum, E. Salahi and M. Khaleghi (2009), "Anomaly Intrusion Detection Design Using Hybrid of Unsupervised and Supervised Neural Network," *International Journal of Computer Networks Communications (IJCNC), Vol.1, No.2, pp. 26-33, July 2009.* 16, 25, 39

[56] Kamran Shafi, Hussein A. Abbass (2009), "An adaptive genetic-based signature learning system for intrusion detection," *Expert Systems with Applications, Volume 36, Issue 10, December 2009, pp. 12036-12043, ELSEVIER.* 16, 25, 39

[57] Gang Wang, Jinxing Hao, Jian Mab, Lihua Huang (2010), "A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering," *Expert Systems with Applications, Volume 37, Issue 9, September 2010, Pages 6225-6232, ELSEVIER.* 16, 25, 39

[58] Woochul Shim, Gisung Kim, Sehun Kim (2010), "A distributed sinkhole detection method using cluster analysis," *Expert Systems with Applications, Volume 37, Issue 12, December 2010, pp. 8486-8491, ELSEVIER.* 16, 25, 39

[59] Muna Mhammad T. Jawhar, M. Mehrotra (2010), "Design Network Intrusion Detection System using hybrid," *International Journal of Computer Science and Security, Volume 4, Issue 3, pp. 285-294, 2010.* 16, 25, 39

[60] Ilhan Aydin, Mehmet Karakose, Erhan Akin (2010), "Chaotic-based hybrid negative selection algorithm and its applications in fault," *Expert Systems with Applications, Volume 37, Issue 7, July 2010, Pages 5285-5294, ELSEVIER.* 16, 25, 39

[61] I.T. Christou, M. Bakopoulos, T. Dimitriou, E. Amolochitis, S. Tsekeridou, C. Dimitriadis (2011), "Detecting fraud in online games of chance and lotteries," *Expert Systems with Applications, Volume 38, Issue 10, 15 September 2011, Pages 13158-13169, ELSEVIER.* 16, 26, 39

[62] Seung Kim, Nam Wook Cho, Bokyoung Kang, Suk-Ho Kang (2011), "Fast outlier detection for very large log data," *Expert Systems with Applications, Volume 38, Issue 8, August 2011, Pages 9587-9596, ELSEVIER.* 16, 26, 39

[63] Mohammad Saniee Abadeh, Hamid Mohamadi, Jafar Habibi (2011), "Design and analysis of genetic fuzzy systems for intrusion detection in computer networks," *Expert Systems with Applications, Volume 38, Issue 6, June 2011, Pages 7067-7075, ELSEVIER.* 16, 26, 39

[64] Shun-Sheng Wang, Kuo-Qin Yan, Shu-Ching Wang, Chia-Wei Liu (2011), "An Integrated Intrusion Detection System for Cluster-based Wireless," *Expert Systems with Applications, Volume 38, Issue 12, November–December 2011, pp. 15234-15243, ELSEVIER.* 16, 26, 39

[65] Ming-Yang Su (2011), "Real-time anomaly detection systems for Denial-of-Service attacks by weighted," *Expert Systems with Applications, Volume 38, Issue 4, April 2011, Pages 3492-3498, ELSEVIER.* 16, 26, 39

[66] Seungmin Lee, Gisung Kim, Sehun Kim (2011), "Self-adaptive and dynamic clustering for online anomaly detection," *Expert Systems with Applications, Volume 38, Issue 12, December 2011, Pages 14891-14898, ELSEVIER.* 16, 26, 39

[67] Yinhui Li, Jingbo Xia, Silan Zhang, Jiakai Yan, Xiaochuan Ai, Kuobin Dai (2012), "An efficient intrusion detection system based on support vector machines and gradually feature removal method," *Expert Systems with Applications, Volume 39, Issue 1, January 2012, Pages 424-430, ELSEVIER.* 16, 27, 39

[68] A.S. Aneetha and Dr. S. Bose SOM (2012), "The combined approach for anomaly detection using neural networks and clustering techniques," *Computer Science Engineering: An International Journal (CSEIJ), Vol.2, No.4, pp. 37-46, August 2012.* 16, 27, 39

[69] Prof. D.P. Gaikwad, Sonali Jagtap, Kunal Thakare, Vaishali Budhawant (2012), "Anomaly Based Intrusion Detection System Using Artificial Neural Network and Fuzzy clustering," *International Journal of Engineering Research Technology (IJERT), Vol. 1 Issue 9, ISSN: 2278-0181, pp. 1-6, November 2012.* 16, 27, 39

[70] A. M. Chandrashekhar and K. Raghuveer (2013), "Fortification of hybrid intrusion detection system using variants of neural networks support vector machines," *International Journal of Network Security Its Applications (IJNSA), Vol.5, No.1, pp. 71-90, January 2013.* 16, 27, 39

[71] Seongjun Shin, Seungmin Lee, Hyunwoo Kim, Sehun Kim (2013), "Advanced probabilistic approach for network intrusion forecasting and detection," *Expert Systems with Applications, Volume 40, Issue 1, January 2013, pp. 315-322, ELSEVIER.* 16, 27, 39

[72] Dahlia Asyiqin Ahmad Zainaddin and Zurina Mohd Hanapi(2013), "Hybrid of fuzzy clustering neural network over nsl-kdd dataset for intrusion detection system," *Journal of Computer Science, ISSN 1549-3636, pp. 391-403, 2013.* 16, 28, 39

[73] Mazyar Mohammadi Lisehroodi, Zaiton Muda and Warusia Yassin (2013), "A hybrid framework based on neural network mlp and k-means clustering for intrusion detection system," *Proceedings of the 4th International Conference on Computing and Informatics, ICOCI, pp. 305-311, 2013.* 16, 28, 40

[74] Gisung Kim, Seungmin Lee, Sehun Kim (2013), "A novel hybrid intrusion detection method integrating anomaly detection with misuse detection," *Expert Systems with Applications, Volume 41, Issue 4, Part 2, March 2014, pp. 1690-1700, ELSEVIER.* 16, 28, 40

[75] Wenying Feng, Qinglei Zhang, Gongzhu Hu, Jimmy Xiangji Huang (2014), "Mining network data for intrusion detection through combining SVMs with ant colony networks," *Future Generation Computer Systems, Volume 37, July 2014, pp. 127-140, ELSEVIER.* 16, 28, 40

[76] Ravi Ranjan and G. Sahoo (2014), "A new clustering approach for anomaly intrusion detection," *International Journal of Data Mining Knowledge Management Process (IJDKP), Vol. 4, No. 2, March 2014, pp. 29-38.* 16, 28, 40

[77] Hamid Alipour, Youssif B. Al-Nashif, Pratik Satam, Salim Hariri (2015), "Wireless Anomaly Detection Based on IEEE 802.11 Behavior Analysis," *IEEE Transactions on Information Forensics and Security, vol. 10, pp. 2158-2170, 2015.* 16, 28, 40

[78] Shengyi Pan, Thomas Morris and Uttam Adhikari (2015), "Developing a Hybrid Intrusion Detection System Using Data Mining for Power Systems," *IEEE Transactions on Smart Grid, vol. 6, no. 6, pp. 3104-3113, Nov. 2015.* 16, 29, 40

[79] J.Jabez and B.Muthukumar Dr. (2015), "Intrusion Detection System (IDS): Anomaly Detection Using Outlier Detection Approach," *International Conference on Intelligent Computing, Communication Convergence (ICCC-2015), Vol. 48, 2015, pp. 338-346.* 16, 29, 40

[80] Mohammad Reza Parsaei, Samaneh Miri Rostami and Reza Javidan (2016), "A Hybrid Data Mining Approach for Intrusion Detection on Imbalanced NSL-KDD Dataset," *(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 6, pp. 20-25, 2016.* 16, 29, 40

[81] B. M. Aslahi-Shahri, R. Rahmani, M. Chizari, A. Maralani, M. Eslami, M. J. Golkar and A. Ebrahimi (2016), "A hybrid method consisting of GA and SVM for intrusion detection system," *Neural Computing and Applications, 2016, Volume 27, Number 6, pp. 1669–1676.* 16, 29, 40

[82] Mohammed A. Ambusaidi, Xiangjian He, Priyadarsi Nanda and Zhiyuan Tan (2016), "Building an Intrusion Detection System Using a Filter-Based Feature Selection Algorithm," *IEEE Transactions on Computers, Volume: 65 , Issue: 10 , Oct. 1, pp. 1-13, 2016.* 16, 29, 40, 42

[83] Seyed Mojtaba Hosseini Bamakana, Huadong Wang, Tian Yingjie, and Yong Shia (2016), "An effective intrusion detection framework based on MCLP/SVM optimized by time-varying chaos particle swarm optimization," *Neuro computing, Volume 199, 26 July 2016, pp. 90-102, ELSEVIER.* 16, 29, 40

[84] Ali H. Mirza (2018), "Computer network intrusion detection using various classifiers and ensemble learning," *26th Signal Processing and Communications Applications Conference (SIU), pp. 1-4, 2018.* 16, 18, 24, 38

[85] Y. Xin, L. Kong, Z. Liu, Y. Chen, Y. Li, H. Zhu, M. Gao, H. Hou and C. Wang (2018), "Machine learning and deep learning methods for cybersecurity," *IEEE Access, pp. 35365-35381, vol. 6, 2018.* 16, 18, 24, 38

[86] Liyo Duan, Youan Xiao (2018), "An intrusion detection using model based on Fuzzy c means clustering," *8th international conference Electronics information and Emergency communication (ICEIEC).* 16, 18, 24, 38

[87] Fadi Salo, Mohammadnoor Injadat, Ali Bou Nassif, Abdallah Shami, and Aleksander Essex (2018), "Data mining techniques in intrusion detection," *IEEE Access, Vol. 4, pp. 1-14, 2018.* 16, 18, 24, 38

[88] Roshan Kumar and Deepak Sharma (2018), "Signature-Anomaly Intrusion Detection System," *9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1-7, 12 July 2018.* 16, 30, 40

[89] Tich Phuoc Tran, Longbing Cao, Dat Tran and Cuong Duc Nguyen (2009), "Novel Intrusion Detection using Probabilistic Neural Network and Adaptive Boosting," *(IJCSIS) International Journal of Computer Science and Information Security, Vol. 6, No. 1, pp. 083-091, October 2009, USA.* 16, 31, 41

[90] C.A. Laurentys, R.M. Palhares and W.M. Caminhas (2011), "A novel artificial immune system for fault behavior detection," *Expert Systems with Applications, Vol. 38, Issue 6, June 2011, pp. 6957-6966, ELSEVIER.* 16, 31, 41

[91] Dewan Md. Farid, Mohammad Zahidur Rahman and Chowdhury Mofizur Rahman (2011), "Adaptive intrusion detection based on boosting naive Bayesian classifier," *International Journal of Computer Applications, pp. 0975–8887, Vol. 3, June 2011.* 16, 31, 41, 42

[92] Yang Yi, Jiansheng Wu and Wei Xu (2011), "Incremental SVM based on reversed set for network intrusion detection," *Expert Systems with Applications, Volume 38, June 2011, pp. 7698-7707, ELSEVIER.* 16, 31, 41

[93] Siva S. Sivatha Sindhu, S. Geetha and A. Kannan (2012), "Decision tree based light weight intrusion detection using a wrapper approach," *Expert Systems with Applications, Volume 39, January 2012, pp. 129-141, ELSEVIER.* 16, 31, 41

[94] Dewan Md. Farid, Li Zhang, Alamgir Hossain, Chowdhury Mofizur Rahman, Rebecca Strachan, Graham Sexton, and Keshav Dahal (2013), "An adaptive ensemble classifier for mining concept drifting data streams," *Expert Systems with Applications, Volume 40, Issue 15, 1 November 2013, pp. 5895-5906, ELSEVIER.* 16, 31, 41

[95] Akhilesh Kumar Shrivas and Amit Kumar Dewangan (2014), "An ensemble model for classification of attacks with feature selection based on KDD-99 and NSL-KDD data set," *International Journal of Computer Applications (0975 – 8887), Volume 99, pp. 8-13, August 2014.* 16, 31, 41

[96] Xiaoxue Liu, Peidong Zhu, Yan Zhang and Kan Chen (2015), "A Collaborative Intrusion Detection Mechanism Against False Data Injection Attack in Advanced Metering Infrastructure," *IEEE Transactions on Smart Grid, vol. 6, no. 5, pp. 2435-2443, Sept. 2015.* 16, 31, 41

[97] Chunjie Zhou, Shuang Huang, Naixue Xiong, Shuang-Hua Yang, Huiyun Li, Yuanqing Qin and Xuan Li (2015), "Design and Analysis of Multimodel-Based Anomaly Intrusion Detection Systems in Industrial Process Automation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems (Volume: 45, Issue: 10, pp. 1345-1360, Oct. 2015).* 16, 32, 41

[98] Guodong Zhao, Ke Xu , Lei Xu, and Bo Wu (2015), "Detecting APT Malware Infections Based on Malicious DNS and Traffic Analysis," *IEEE Access, vol. 3, pp. 1132-1142, 2015.* 16, 32, 41

[99] Wei-Chao Lin, Shih-Wen Ke and Chih-Fong Tsai (2015), "CANN: An intrusion detection system based on combining cluster centers and nearest neighbors," *Knowledge-Based Systems, Volume 78, April 2015, pp. 13-21, ELSEVIER.* 16, 32, 41

[100] Jonathon Ng, Deepti Joshi and Shankar M. Banik (2015), "Applying Data Mining Techniques to Intrusion Detection," *2015 12th International Conference on Information Technology - New Generations (ITNG), Las Vegas, NV, USA, 2015, pp. 800-801.* 16, 32, 41

[101] Abdulla Amin Aburomman and Mamun Bin Ibne Reaz (2016), "A novel SVM-kNN-PSO ensemble method for intrusion detection system," *Applied Soft Computing, Volume 38, January 2016, Pages 360-372, ELSEVIER.* 16, 32, 41

[102] Parag Verma, Shayan Anwar, Shadab Khan and Sunil B Mane (2018), "Network Intrusion Detection Using Clustering and Gradient Boosting," *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Bangalore, 2018, pp. 1-7.* 16, 32, 41

[103] Ch. Nagamani and Suneetha Chittineni (2018), "Network Intrusion Detection Mechanisms Using Outlier Detection," *Second International Conference on Inventive Communication and Computational Technologies (ICICCT), pp. 1468-1473, 2018.* 16, 32, 41

[104] Rana Aamir Raza Ashfaq, Xi-Zhao Wang, Joshua Zhexue Huang, Haider Abbas, Yu-Lin He (2017), "Fuzziness based semi-supervised learning approach for intrusion detection system," *Information Sciences, Volume 378, 1 February 2017, pp. 484-497, ELSEVIER.* 16, 23, 38

[105] Eduardo Massato Kakihata, Helton Molina Sapia, Ronaldo Toshiaki Oiakawa, Danillo Roberto Pereira, Joao Paulo Papa, Victor Hugo Costa de Albuquerque and Francisco Assis da Silva (2017), "Intrusion Detection System Based On Flows

Using Machine Learning Algorithms," *IEEE Latin America Transactions, Volume: 15, Issue: 10, pp. 1988 - 1993, Oct. 2017.* 16, 23, 38

[106] Wenjuan Li, Weizhi Meng, Lam-For Kwok, and Horace H. S. IP (2017), "Enhancing Collaborative Intrusion Detection Networks Against Insider Attacks Using Supervised Intrusion Sensitivity-Based Trust Management Model," *Journal of Network and Computer Applications, Volume 77, 1 January 2017, pp. 135-145, ELSEVIER.* 16, 18, 24, 38

[107] Ikram Sumaiya Thaseen, and Cherukuri Aswani Kumar (2017), "Intrusion Detection Model Using fusion of Chi-square feature selection and multi class SVM," *Journal of King Saud University - Computer and Information Sciences, Volume 29, Issue 4, October 2017, pp. 462-472.* 16, 18, 23, 38

[108] Gauthama Raman M R, Nivethitha Somu, Kannan Kirthivasan, Shankar Sriram V S (2017), "A Hypergraph and Arithmetic Residue-based Probabilistic Neural Network for Classification in Intrusion Detection Systems," *Neural Networks, Volume 92, August 2017, pp. 89-97, ELSEVIER.* 16, 30, 40

[109] Aastha Puri and Nidhi Sharma(2017), "A Novel Technique for Intrusion Detection System for Network Security Using Hybrid SVM-CART," *International Journal of Engineering Development and Research, Volume 5, Issue 2, ISSN: 2321-9939, pp. 155-161.* 16, 30, 40

[110] Beibei Li, Rongxing Lu, Wei Wang, Kim-Kwang Raymond Choo(2017), "Distributed Host-based Collaborative Detection for False Data Injection Attacks in Smart Grid Cyber-Physical System," *Journal of Parallel and Distributed Computing, Volume 103, May 2017, pp. 32-41, ELSEVIER.* 16, 30, 40

[111] Uma R. Salunkhe, SureshN. Mali (2017), "Security Enrichment in Intrusion Detection System Using Classifier Ensemble," *Journal of Electrical and Computer Engineering, pp. 1-6, 2017.* 16, 30, 40

[112] V. Balamurugan, and R. Saravanan (2017), "Enhanced Intrusion Detection and Prevention System on Cloud Environment Using Hybrid Classification and OTS Generation," *Cluster Computing, pp 1–13, 2017.* 16, 30, 40

[113] Aafreen K. Siddiqui and Tanveer Farooqui (2017), "Improved Ensemble Technique based on Support Vector Machine and Neural Network for Intrusion Detection System," *International Journal Online of Science, Vol. 3, No. 12, pp. 1-8, 2017.* 16, 32, 41

[114] "Snort(software)," https://www.snort.org/, Online accessed: 24-Dec-2018, Time: 4:30 PM. 6

[115] "OSSEC(software)," https://www.ossec.net/, Online accessed: 24-Dec-2018, Time: 4:38 PM. 6

[116] "The Bro Network Security Monitor," https://www.bro.org/, Online accessed: 24-Dec-2018, Time: 4:49 PM.

7