Australian Journal of Basic and Applied Sciences, 7(7): 799-813, 2013
ISSN 1991-8178

Evolutionof Intrusion Detection Systems Based on Machine Learning Methods

Abdulla Amin Aburomman, Mamun Bin Ibne Reaz

Department of Electrical, Electronic and Systems Engineering, Universiti Kebangsaan Malaysia, 43600, UKM, Bangi, Selangor, Malaysia.

Abstract: In this immense data communication and accumulation age, it is crucial to protect the valuable information. In this context an intrusion detection system (IDS) which applied artificial intelligence techniques comes into the helping hand. Generally, intrusion detection system scrutinizes all incoming and outgoing network activities and distinguishes abnormal (intrusions) and normal patterns. A lot of fees able IDSs are available in the market today, its detection for intrusions (attacks) are depending on signature based, yet their efficiency in terms of identifying new unseen before intrusions (attacks) is a big question mark. Anomaly-based intrusion detection systems comes to solve this issue by applying different machine learning techniques and offering the technical platform for data mining, it is applied for data extractionfrom the raw in databases, and the comprehensive can be used for a variety of purposes. Employing data mining for intrusion detection is a novel idea. This paper has reviewed machine learning techniques and provided quite a few studies associated with single, hybrid, and ensemble classifiers. Furthermore, this paper reviewed information relative to classifiers design, employed datasets, feature extraction/selection, clustering techniques, baseline classifiers and/or anchor paper used for comparisons, dataset validation, accuracy detection measures, and other test configurations, these researches were compared. Ultimately, suggestions and guidelines have also been presented for designing good network intrusion detection system.

Key words: Classifiers, Data mining, feature selection, IDS, Support vector Machines (SVM).

INTRODUCTION

In today's global economy, Internet communication plays a very crucial role in accomplishment of business; economical and technological objectives. However, like the other side of a coin, protecting the priceless information from falling into the hands of intruders is the biggest challenge. Hackers and other forms of cyber criminals pose a critical threat to intrusion detection system (IDS). Despite these threats, the intrusion detection systems strive very hard to combat the cyber-attacks. The fundamental objective of IDS is to shield computer systems from cyber criminals. Generally, IDSs detect various forms of detrimental incoming network connection and application of computer systems, compared to the typical firewall. Misuse and anomaly detection are two well-known types of IDS. IDS compares the incoming network pattern and signatures with the saved ones in its profile, if an attack accrues and its signatures harmonized with the saved one in the IDS profile, IDS can detect that kind of attack (Catania, Bromberg *et al*. 2012).Based on various kinds of machine learning techniques, some of them employed single classifier, such as decision tree, K-nearest neighbor, and SVM, etc. In contrast, some systems combine various classifiers, such as ensemble or hybrid classifications techniques, which are employed to categorize the incoming Internet access as regular access or an invasion. This study evaluate the relevant researches and systems for the purpose of recognizing their strategies employed, experiments carried out, and understanding the directions of future work in the context of machine learning. This paper has reviewed the data sets used in classification tasks, data preprocessing steps, Validation, Performance measure and baseline classifier, Dimensionality reduction techniques, machine learning techniques, and briefs some of the intrusion detection techniques, and comparative analysis of associated work regarding different classification method. Finally, conclusion and discussion for future research is presented.

Data Set for Classification Tasks:

There are available online data sets which can be used for classification tasks; such as DARPA 1998, 1998, NSL-KDD99, and KDD99. The most used data set for benchmark is KDD 99 data set which can be found on:http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html. McHugh (2000) criticized DARPA data set, and indicated that the results has been subjective since there are some doubts in the evaluation of the methodologies applied. For instance, he claimed that the samples of normal and attack data were not realistic. Also, the inadequacy of the datasets in the training phase has been figured. Moreover, the false alarm behavior of IDSs which was under investigation has not been crucially validated to show the variation between the real data and the synthetic ones. McHugh's results has been approved by Malhony and Chan (Mahoney and Chan 2003) who finds out that several features are different from the real network traffics as in real traffic they are large and

Corresponidng Author: Abdulla Amin Aburomman, Department of Electrical, Electronic and Systems Engineering, Universiti Kebangsaan Malaysia, 43600, UKM, Bangi, Selangor, Malaysia.
E-mail: reoroman@yahoo.com mamun.reaz@gmail.com

always growing, but in the computer simulation they have fixed ranges.KDD99 datas*et al*so share the same limitation because it is inherited from the DARPA dataset. Also the research results of Sabhnani *et al* (Sabhnani and Serpen 2004) stated that the U2R and R2L attacks in both training and testing datasets has differentcharacteristic. They stated that U2R and R2L attack has novel attacks in the test dataset which form 80% for 4 attacks belong to U2R and 60% for 7 attacks belong to R2L, and that results in poor detection for those attacks in all IDSs cases.With all the critics has been made toward KDD99 and DARPA datasets, but it stays the most biggest online available data which better suits researches for evaluation oftheirIDS and machine learning. Table 1 represents data sets used for experiments, in which some of them have been generated by specific tool, nevertheless these data can't be guaranteed that it does not contain attacks among normal group. Also these self-produced data might be inefficient in the training mode, and that will degrade the detectionquality.The Kdd99 data set has 41different features for every TCP/IP connection, 34 are numerical and 7 are characters. These 41 features distributed into 3 parts of features which are: intrinsic, traffic and content (Chunhua and Xueqin 2009). Otherscan divide it into 4 categories of features which are: basic, content, time-base and host-based. The basic feature has features from 1 to 9; those are deduced from packet header with no attention to payload. The second one is content feature where attention paid to the payload such as number of failed login tries. The third one is time-based where attention goes for connection time to the same host if the connection exceeds 2-s. the fourth one is host-based traffic which is projected to evaluate the attack, where span interval duration is more than 2s (Altwaijry and Algarny 2012). Network features label with its type and description can are clearly shown in(Amiri, Rezaei Yousefi *et al*. 2011). The task is to build up a model of network intrusion detectorcandifferentiate between attacksand normal traffic. It is very important to study the data and preprocess it before proceeding to classification part. There are 22 attacks in the training file (kddcup.data_10_percent.gz A 10% subset), and new 17 attacks in the testing file (corrected KDD). A total of 39 attacksavailable in testing set and classified into four classes which are: DOS, PROB, U2R, R2L. NSL-KDD (http://nsl.cs.unb.ca/NSL-KDD/),is another dataset used for the classification task suggested to work out problems located beneath the KDD'99 data set which is mentioned in (Tavallaee, Bagheri *et al*. 2009).KDD99 data ismeasured as a binary problem. Researchers first concern is to classify between normal Vs attack traffic, then classify the other 4 types of attacks (Bolón-Canedo, Sánchez-Maroño *et al*. 2011).

Table 1 illustrates the datasets used for experiments based on year wise distribution. It is noteworthy that, some works have divided the same data into more than one part, so that one can be used in training and other for testing purposes. Latterly, due to the lack of famous Knowledge Discovery and Data Mining data set, (KDD99) and DARPA Intrusion Detection Data Sets 1998 and 1999 datasets, these datasets have been considered by a lot of studies for analyses. However, other researchers have used their own datasets. The KDD and DARPA indicates it publicity for use as a typical dataset.

Table 1: Year wise for experimental dataset.

datasets	Years					
	2008	2009	2010	2011	2012	2013
DARPA 1998					1	
DARPA 1999	2	3	3	2	1	
KDD99	5	10	12	12	19	1
NSL-KDD			2		4	
Network tcpdump	1	1		1	1	
Windows-SYS		1				
University/campus		1		3		
Generated data	1		1	1	1	
IDS_Bag,					1	
kdd99,Kyoto honeypot data				1		

Data Preprocessing:

Symbolic to numeric: In KDD99 data should be first preprocessed, such as symbolic to numeric (replacing characters attributes by numerical one). Since some algorithms supports only numerical data, so replacement of character by numerical value ismust. This means a decimalnumber will replace the character value.

Normalization or scaling of data: here the target is make the data values fall between same ranges, example, to scale the date between [0,1]. This makes sure that there is nooverpoweringfor some input vectors in training mode. One of the famous formula applied for data normalization is: $xi = (xi - x\ min)\ (x\ max - x\ min)$(Chunhua and Xueqin 2009). In (li and Liu 2011) an example of normalization has been experimented on 10000 records of KDD99 dataset using Min-Max normalization method based on SVM classifier with RBF kernel , and the results showed that classification is better with normalization applied than without it, which hints that normalization can provide more speed and difference reduction of data.

Discretization (continues to discrete conversion, which of course depends on the classifier). In (Chunhua and Xueqin 2009), three methods used to discretize the value of attribution, which are IEA, EFA and BRA, and

that led to accuracy improvement of the classifier.This information intersect with earlier study which emphasized on discretization to improve the classification results (Koc, Mazzuchi *et al*. 2012).

Redundant removal:redundant records are one of the weak points in KDD dataset, because their attendancemakes the classifier to be biased toward repeated values and not toward rare one (Eid, Darwish *et al*. 2010; Sivatha Sindhu, Geetha *et al*. 2012). Take for example the R2L which has very less records comparing to DOS.Some researchers remove these redundant records to enhancedetection accuracy.

Validation, Performance Measure and Baseline Classifier:

Recognition Rates are calculated as the ratio between the amount of properly detected attacks and also the final amount of attacks, while False Alarm (false positive) Rates are calculated as the ratio relating to the amounts of normal connections which are improperly misclassified as attacks. They are good indications of performance, given that they measure what number of attacks the machine has the capacity to identify and just how many incorrect classifications come in the procedure(Somwang and Lilakiatsakun 2011).

You will find many techniques for evaluation of predictive precision, for example: K-fold mix validation, Holdout, Re-substitution and then leave-one-out(li and Liu 2011). K-fold cross validation is the ideal one particular. It's really a method of evaluating the efficiency of the classifier. Primary, the main records are in random portioned into 10 subsets. Next, one particular subset is designated to be the testing data-set as well as the leftover 9 subsets are handled as training data. Later on, the cross validation process repeat 10 times as well as evaluation accuracy of the classifier could be examined through the average accuracy from the ten estimations. The 10-fold cross validation is much more well-known within the conditions of enormous data set, in contrast to the Leave-one-out cross validation which is very time costly based on the great difficulty of training times (Li, Xia *et al*. 2012; Sivatha Sindhu, Geetha *et al*. 2012).

Table 2: Articles which used k-fold cross validation, overall accuracy, false positive and false negative measure.

Reference	Year	k-fold C.V	Overall acc.	F.P	F.N
(Panda, Abraham *et al*. 2012)	2012	10-fold C.V	Yes	Yes	No
(Koc, Mazzuchi *et al*. 2012)	2012	10-fold C.V	Yes	No	Yes
(Muntean, Valean *et al*. 2010)	2010	10-fold C.V	Yes	No	No
(Tsai and Lin 2010)	2009	10-fold C.V	Yes	Yes	Yes
(Li, Xia *et al*. 2012)	2012	10-fold C.V	Yes	No	No
(Lin, Ying *et al*. 2012)	2012	10-fold C.V	Yes	No	No
(Pereira, Nakamura *et al*. 2012)	2012	10-fold C.V	Yes	No	No
(Sivatha Sindhu, Geetha *et al*. 2012)	2012	10-fold C.V	Yes	Yes	No
(Mukherjee and Sharma 2012)	2012	10-fold C.V	Yes	No	No
(Govindarajan and Chandrasekaran 2011)	2012	10-fold C.V	Yes	No	No
(Chi, Wee-Peng *et al*. 2012)	2012	10-fold C.V	Yes	No	No
(Horng, Fan *et al*. 2008)	2008	3-fold C.V	Yes	Yes	Yes
(Su 2011)	2011	4-fold C.V	Yes	Yes	No
(li and Liu 2011)	2011	5-fold C.V	Yes	No	No
(Kavitha, Karthikeyan *et al*. 2012)	2012	5-fold C.V	Yes	Yes	No
(Lee, Kim *et al*. 2012)	2012	5-fold C.V	Yes	No	No
(Winter, Hermann *et al*. 2011)	2011	8-fold C.V	Yes	Yes	No
(Chi, Wee-Peng *et al*. 2012)	2012	10-fold C.V	Yes	Yes	No
(Sharma and Mukherjee 2012)	2012	10-fold C.V	Yes	Yes	No
(Pingjie, Rong-an *et al*. 2010)	2010	10-fold C.V	Yes	Yes	No
(Arau, x *et al*. 2010)	2010	10-fold C.V	Yes	Yes	No

Throughout a 5-class problem in IDS, the dataset is broken into five classes and every sample faces 5 various prospects. The next proportions can be used to assess the effectiveness of the classifier(Li, Xia *et al*. 2012):

- True negative (TNi): The amount of outer samples that is certainly the right way classified;
- False negative (FNi): The amount of ith class records which is incorrectly classified tothe other classes;
- Accuracy= $\frac{\sum(TP_i+TN_i)}{(TP_i+TN_i+FP_i+FN_i)}$;
- (MCC) Matthews correlation coefficient, which usually functions properly even just in the out ofbalance classes. $MCCi = \frac{TP_i\ X\ TN_i - FP_i XFN_i}{\sqrt{(TP_i+ FN_i)\ (TP_i+ FP_i)(TN_i+ FP_i)(TN_i+ FN_i)}}$
- $MCC_{avg} = \frac{\sum MCCi}{5}$

For validation of new classification methods, a baseline classifier (anchor paper) should be provided as a proof paper to compare the work and results with it. Generally each work selects diverse classification technique to authorize their new system, some authors compared their work to more than one classifier. Table 3 shows different articles which considered different classifiers or other articles as a comparison for their work. In fact, SVM is most popular baseline technique. Furthermore, of late, it has also been widely used for model comparisons.

Table 3: Baseline approaches.

Reference	Comparedwith
(Gengming and Junguo 2008),(Huike and Daquan 2009),(Liu, Kang *et al*. 2010),(Muntean, Valean *et al*. 2010),(Yongli and Yanwei 2010),(Lei, Zhi-ping *et al*. 2010),(Catania, Bromberg *et al*. 2012),(Xie and Zhang 2012),(Sujatha, Priya *et al*. 2012),(Chi, Wee-Peng *et al*. 2012),(Agarwal and Mittal 2012),(li and Liu 2011),(Chunhua and Xueqin 2009),(Zhenguo and Guanghua 2009),(Lin Li, Zhang Ya *et al*. 2010),(Shirazi 2009),(Xuejun, Guiling *et al*. 2008),(Amiri, Rezaei Yousefi *et al*. 2011),(Lin, Ying *et al*. 2012),(Farid and Rahman 2010),(Pereira, Nakamura *et al*. 2012),(Mok, Sohn *et al*. 2010),(Wei and Wu 2008),(Altwaijry and Algarny 2012),(Bolón-Canedo, Sánchez-Maroño *et al*. 2011),(Muniyandi, Rajeswari *et al*. 2012),(Tsai and Lin 2010),(Kai-mei, Xu *et al*. 2009),(Wei, Shaohua *et al*. 2010),(Kuang, Xu *et al*. 2012),(Horng, Su *et al*. 2011)	SVM
(Lin Li, Zhang Ya *et al*. 2010),(Wang, Hao *et al*. 2010),(Altwaijry and Algarny 2012),	BBNN
(Shirazi 2009),(Wei and Wu 2008),(Xiang, Yong *et al*. 2008),(Dartigue, Hyun Ik *et al*. 2009),(Bolón-Canedo, Sánchez-Maroño *et al*. 2011),	KDD cup winner, KDD cup runner.
(Amiri, Rezaei Yousefi *et al*. 2011),(Pereira, Nakamura *et al*. 2012),(Sangkatsanee, Wattanapongsakorn *et al*. 2011),	Bayesian networks
(Lin, Ying *et al*. 2012),(Mok, Sohn *et al*. 2010),(Sivatha Sindhu, Geetha *et al*. 2012),(Arau, x *et al*. 2010),(Wang, Hao *et al*. 2010),(Altwaijry and Algarny 2012),(Sangkatsanee, Wattanapongsakorn *et al*. 2011),	Decision tree
(Farid and Rahman 2010),	NN
(Farid and Rahman 2010),(Abadeh, Mohamadi *et al*. 2011),	GA
(Farid and Rahman 2010),(Sivatha Sindhu, Geetha *et al*. 2012),(Wang, Hao *et al*. 2010),(Muniyandi, Rajeswari *et al*. 2012),(Om and Kundu 2012),(Koc, Mazzuchi *et al*. 2012),(Sharma and Mukherjee 2012),	Naïve Bayes
(Pereira, Nakamura *et al*. 2012),(Altwaijry and Algarny 2012),	SOM
(Raj Kumar and Selvakumar 2011),	bagging, boosting, adaboost
(Sivatha Sindhu, Geetha *et al*. 2012),	Random forest
(Altwaijry and Algarny 2012),(Muniyandi, Rajeswari *et al*. 2012),(Tsai and Lin 2010),	K-means
(Muniyandi, Rajeswari *et al*. 2012),(Giacinto, Perdisci *et al*. 2008),(Om and Kundu 2012),(Tsai and Lin 2010)	K-NN

Dimensionality Reduction: Feature Selection and Clustering:

Feature selection helps to improve the performance of a classifier by selecting a subset of relevant features and eliminating most irrelevant and redundant features. It also named attribute selection, variable subset selection or variable selection.

Features are statistical attributes produced from the accumulated dataset. Real time features subset selection are important for classification of live (online) traffic, but science more features means more accuracy, the computation time may take long and make overhead and time wasting. (Raj Kumar and Selvakumar 2011). Feature selection is essential for an additional reasons (i). To relieve the effect from the curse of dimensionality, (ii). To boost generalization capacity, (iii). To accelerate learning method and (iv). To enhance model interpretability. It is essential to select the appropriate features for any classifier. Utilization of more features may produce the problem of lack of generalization whereas utilization of less features sometimes causes degradation in the level of classification quality. Feature selection is very important for any classifier. Use of more features may create the problem of loss of generalization whereas use of fewer features sometimes causes degradation in classification quality.Feature selection likewise helps individuals to acquire better understanding regarding their data by letting them know what are essential features and just how they're related to one another (Saha, Sairam *et al*. 2012). Furthermore experimental results demonstrate that an IDS with feature selection works much better than that with no feature selection in computational cost and recognition precision (Amiri, Rezaei Yousefi *et al*. 2011). From research completed with feature selection, it has noticed that feature selection led to enhance overall accuracy, reduced the amount of false positives, and enhanced the recognition of samples with low rate within the training data. This is the main reason feature selection was introduced in several

proposed model(Dartigue, Hyun Ik *et al*. 2009).Feature selection is usually classified into wrapper and filter methods. While wrapper methods attempt to optimize several predetermined conditions with regards to the feature set included in the selection procedure, filter methods depend on the overall features of the training data to pick features which can be separate from one another and therefore are extremely determined by the end result. Feature selection methods make use of a search algorithm to find information about the entire feature space and assess feasible subsets. To judge these subsets, they might need a feature benefits measure which scores any subset of features. Generally speaking, an attribute is a plus when it is tightly related to the result, however is not repetitive along with other related features. An attribute benefits evaluate could possibly be the reliance in between 2 features. A couple of the most significant benefits measurements to decide on the characteristics are mutual information and correlation coefficient(Amiri, Rezaei Yousefi *et al*. 2011). Attribute reduction decide the minimum subset attribute whose classification quality is similar towards the original attribute of data set(Chunhua and Xueqin 2009).These types of reduced features is going to be used in training mode to learn the classifier the possible patterns of intrusions and normal traffic, after which detect them in cross validation and testing stage.The system is going to be monitored regards to false positive (FP), false negative (FN), true positive (TP) and true negative (TN) . reduced subset of features should help in increasing the detection rate and reducing the false alarms, therefore decreasing the classifier overhead on the time of training and testing(Kausar, Samir *et al*. 2012).In general sometimes features could have false correlations, which impede the actual process of learning task. Furthermore, a number of features could possibly be unimportant or redundant. These additional features can have an impact on computation time and the classification accuracy. For this reason, feature selection methods are suitable for classification domains because it is effectively describe the problem without the need of degrading performance.Also Feature selection can have other motives, such as data reduction, set of features reduction, overall performance enhancement and much better data understanding.Wrapper type have a tendency to acquire much better performances than filters since it utilize a classifier along with a search method to rank subsets of attributes based on their own predictive strength, regardless the higher computational cost. Filter type depend on the typical features of the training data to decide on the most effective features without any dependency of classifier.Alternatively, when confronted with huge datasets, filter type would be the best option.In feature selection, it is not easy to locate techniques that could handle multiple class problems, as hardly any studies have already been completed in this part yet. The multiple class problems are provided to be affected by a term called accumulative impact, which usually grows more noticeable once the number of classes increases.

Some consideration should be taken in multiple classes' problem:

- Unbalanced classes: in which more than one class has a greater number of samples when compared with other classes.
- It is hard to Identify the best features for every class, for the reason that feature selection provides a number of attributes that could stand for only the vast majority classes(Bolón-Canedo, Sánchez-Maroño *et al*. 2011).

The clustering technique could very well generate excellent datas*et al*ong with much less cases which adequately symbolize every one of the cases with in the original dataset. The two main kinds of cluster algorithms are; hierarchical and partitioning. Partitioning technique is unacceptable for IDS case considering that the amount of clusters ought to be pre-determined in partitioning, despite the fact that simply no sufficient details about it. Consequently, investigators follow a hierarchical approach. This approach is usually accustomed to classify animals and plants, and is particularly likely to be sufficient for classifying the stages in the DDoS attack by means of their own attributes (Lee, Kim *et al*. 2008; Horng, Su *et al*. 2011).

Techniques Employed for Machine Learning:

Recognition of Pattern:

Classifications of the patterns is the process of using computers and related equipment to understand desired patterns from their environment, and build rational assessments regarding fundamental characters of patterns. Herbert Simon the eminent Nobel award recipient has accentuated that pattern recognition was his central finding, which is vital for the human decision-making process. As a matter of fact, a number of data mining processes which can be considered as pattern recognition are involved in many aspects such as, bank evaluation for customer credits, hospital diagnostics for patient medical records… etc. However, this study focuses on more conventional pattern recognition problem, which is network intrusion detection. Basically, there are three steps of pattern recognition process, as follows: the first step is data acquisition, during which data are collected with computers or other devices; this is followed by the second step, which is data processing, during which the data collected from the previous step are processed by feature selection, extraction, and reconstruction processes, and third and last step in pattern recognition process is pattern analysis and identification, in which depending on the processed data with the help of classification techniques, such as, decision trees and neural networks, the decision support models are constructed.Later on these models can be used for patterns prediction. Each pattern recognition data possess some unique features and properties. Pattern recognition is associated with some challenges, first, the remote sensing and image scanning devices, used for generating and collecting data, can without doubt produce terabytes and petabytes of data, hence, it is essential to construct model for decision support to recognize the patterns with the capability of handling huge volume of

data; secondly, due to the numeric measurements, in most of the pattern-recognition issues, the properties and capabilities in the initial and refined data will be of primarily numeric, hence, it is mandatory to have techniques for building decision support models, to effectively deal with numeric attributes; third, a lot of studies have observed that, binomial or normal (statistical distributions)can't depicted the pattern recognition behavior, which means the traditional strategies of parametric statistical may not help. Lastly, fundamentally pattern identification problems involveother types of classifications such as intrusion detection.Traditionally, firewalls are used as protective shield to prevent the computers from intruders. Basically, the differences between the behaviors of intruders and normal users are the key of intrusion detection techniques; furthermore, IDS can be learned on the intrusions patterns for easy detection. Due to significance of IDS, the study has escalated of late. Detection of intrusions and patterns recognition processes are identical to each other (Li 2005).

Table 4: Dimension reduction methods.

Reference	Feature selection/extraction	Reference	Clustering	Reference	Feature selection/extraction and clustering
(Panda, Abraham *et al*. 2012),(Kausar, Samir *et al*. 2012),(Guiling, Yongzhen *et al*. 2010),(Somwang and Lilakiatsakun 2011),(al 2010),(Eid, Darwish *et al*. 2010),(Tsai and Lin 2010)	PCA	(Shirazi 2009),(Guanghui, Jiankang *et al*. 2011),(Giacinto, Perdisci *et al*. 2008),(Muniyandi, Rajeswari *et al*. 2012)	k-means	(Wei and Wu 2008)	(KFDA)+ Fuzzy c-means
(Liu, Kang *et al*. 2010),(Chunhua and Xueqin 2009),(Chen, Cheng *et al*. 2009),(Lei and Ke-nan 2011),	RST (Rough Set Theory)	(Wang, Hao *et al*. 2010),(Xiaozhao, Wei *et al*. 2010),(Ganapathy, Kulothungan *et al*. 2012)	soft clustering (Fuzzy c-means)	(Om and Kundu 2012)	Entropy+ K-means
(Dartigue, Hyun Ik *et al*. 2009),(Sangkatsanee, Wattanapongsakorn *et al*. 2011),(Cohen, Avrahami *et al*. 2008),(Arau, x *et al*. 2010),(Xiang, Yong *et al*. 2008)	information gain			(Pingjie, Rong-an *et al*. 2010)	Information gain+ K-means Based TASVM
(Saha, Sairam *et al*. 2012),(Su 2011),(Zhenguo and Guanghua 2009),(Meijuan, Jingwen *et al*. 2009),(Sivatha Sindhu, Geetha *et al*. 2012)	Genetic Algorithm			(Ashok, Lakshmi *et al*. 2011)	Information Measure + k-means Cluster
(Raj Kumar and Selvakumar 2011),(Lee, Kim *et al*. 2011)	decision tree			(Ashok, Lakshmi *et al*. 2011)	Optimized Feature Selection with k-Means
(Agarwal and Mittal 2012),(Qazanfari, Mirpouryan *et al*. 2012)	Entropy			(Zhao, Yu *et al*. 2009)	extension clustering+ PCA
(Zaman and Karray 2009)	Enhanced Support Vector Decision Function (ESVDF)			(Li, Xia *et al*. 2012)	GFR, gradually feature removal method)+Kmeans clustering, ACO method
(Kavitha, Karthikeyan *et al*. 2012)	best first search			(Tjhai, Furnell *et al*. 2010)	association-rule+K-means
(Bolón-Canedo, Sánchez-Maroño *et al*. 2011)	Correlation-based (CFS),INTERACT, and Consistency-based filters				
(Shafi and Abbass 2009)	Dixon's rule reduction algorithm				
(Kuang, Xu *et al*. 2012)	feature reduction (KPCA)				
(Visumathi and Shunmuganathan 2012)	FFS algorithm				
(Mukherjee and Sharma 2012)	FVBRM,Correlation-based Information Gain, Gain Ratio				
(Sujatha, Priya *et al*. 2012)	SVM				

The Single Classifiers:

Based on this literature review, single classifier has been employed in IDS, (e.g. ANN, SVM, KNN etc.) to answer the problem of classification.

A. Fuzzy Logic:

Fuzzy logic is a very effectual and potential technique, which deals with human reasoning and decision-making processes. Zadeh (1965) has proposed the Fuzzy set theory, which offers a common means to obtain IF-THEN rules in the linguistic form. Fuzzy logic has been widely used in a lot of engineering applications due to its efficiency in solving complicated non-linear problems and providing linguistic portrayal (Guo and Li 2011).

B. ANN Classifier:

Artificial neural networks are most existing effectual classification techniques. Flexibility and the inherent speed are the benefits of employing NN in the detection of instances. NN is also capable of analyzing the non-linear data sets with multi-variables (Wu and Huang 2010).

C. K-NN:

K-NN is popular classification scheme, which employs distance measures. The K- nearest neighbor considers the whole collection of sampling incorporatesthe ideal classification for every single object, apart from the data in the collection. It is essential to compute the distance of each item in the sampling set for classifying a new item. The k-closest items in the collection of sampling are only deemedfurther. The novel item is subsequently categorized, to the category, which consists most of the items from this collection of k closest items (Jiaqi, Ru *et al*. 2011).

D. SVM:

Support vector machines technique is projected by Vapnik (1998); it maximizes the margin to increase the efficiency of classification. SVM classify the data into different groups by constricting a hyper plane, basically it divides data into two groups. Support vectors are the data points which are close the class boundary, and it can be described as a boundary function.In SVM, quadratic programming problem computes the vector of boundary function to solve its margin-maximizing. After all SVM locate a separating hyper plane that has a maximal margin in higher dimensional space when all data are mapped there. Kernels trick are used in SVM to classify non liner classification, an example of those kernels can be RBF kernel, Polynomial Kernel.(Horng, Su *et al*. 2011).

E. Naïve Bayes Networks:

NB is widely used in classifications technique. It depends on directed acyclic graph, where the attributes are represented by nodes and attribute dependencies are represented by arcs. This classifier is quite popular due to its convenience, and `effectiveness of calculations, which are learnt from its aspect of conditional independence assumption. Although this classifier has great efficiency, it is not suitable for large datasets (Koc, Mazzuchi *et al*.).

F. Self-Organizing Maps:

Kohonen (1995) has introduced the fundamental concept, architecture and implementation of Self Organizing Maps. It is kind of neural network with unsupervised learning technique. Based on dataresemblance, SOM sustain the datainput topology by generating feature maps.The normal neural networks must be trained with their desired outputs, whereas, during training, SOM can self-categorize all input data types.Self-organizing maps providesstraightforward methods for data clustering.According to Labib and Vemuri(2002),the fast speed results and fast rates conversion of SOM has been pragmatically established it as an appropriate data classification technique, Furthermore, SOM is expected to surpass other approaches due to capability of preserving topological mappings among the input data. The initiative of the SOM algorithm is to represent or map a high dimensional data in a simple visual 2-dimensional array with the help of data compression technique(Tjhai, Furnell *et al*. 2010).

G. Genetic Algorithms:

Of late, the genetic algorithm is getting very popular due to its potential in the intrusion detection field. GA also evolves solutions for the purpose of adapting to the needs of a problem and emulates the method of normal assess, by replicating concepts of normal choice and duplication.A group of hypotheses, example (individuals) which is called population is indiscriminately produced when GA initially searches the huge hypothesis space, and within this population GA repeatedly picks the superior individuals, the next generation will be reproduced by changing the individuals and crossing it over. Choosing top individuals is stochastically worked based on few preferred performance measure, which is known as fitness of the individual (Shafi and Abbass 2009).

H. Decision Trees:

The other popular classification algorithms in data mining are decision trees. Initially it is composed of a collection of pre-classified data, where ideals of the Features attributes are describing the data points,however, it is hard to select the attributeswhich capably classifies the data.Attributes with largest information gain, are effective in classifications.Decision trees comprisearcs (edges), nodes and leaves, Features attribute signified by node where data has to be segmented. Number of edges belongs to each node; according to common possible values between edges and parent nodes, the edges will be labeled. The edge will be connected to either node and leaf or two nodes. Leaves will be tagged with a decision value for classifying data points. In DT, the classification of an unseen target is initiated at the root and followed in the branch, which is suggested by the consequence of each and every examination, until reaching the leaf node, where class namethere is the consequential classification(Peddabachigari, Abraham *et al*. 2007).

In terms of the works depending on developing single classifiers, table 5 illustrates several papers employed them.

Table 5: Single classification methods.

Reference	Classifiers
(Mohammed and Sulaiman 2012),(li and Liu 2011),(Muntean, Valean *et al*. 2010),(Horng, Su *et al*. 2011),(Mohammad, Sulaiman *et al*. 2011),(Kausar, Samir *et al*. 2012),(Li, Xia *et al*. 2012),(Zhao, Yu *et al*. 2009),(Somwang and Lilakiatsakun 2011),(Zaman and Karray 2009),(Saha, Sairam *et al*. 2012),(Agarwal and Mittal 2012),(Ashok, Lakshmi *et al*. 2011),(Eid, Darwish *et al*. 2010),(Chen, Cheng *et al*. 2009),(Chunhua and Xueqin 2009),(Gengming and Junguo 2008),(Catania, Bromberg *et al*. 2012),(Xie and Zhang 2012),(Yu, Lee *et al*. 2008),(Xuejun, Guiling *et al*. 2008),(Winter, Hermann *et al*. 2011),(Jingbo, Haixiao *et al*. 2010),(Kai-mei, Xu *et al*. 2009)	SVM
(Lin Li, Zhang Ya *et al*. 2010),(Yongli and Yanwei 2010),(Amiri, Rezaei Yousefi *et al*. 2011)	Least squares support vector machine
(Sangkatsanee, Wattanapongsakorn *et al*. 2011),(Muniyandi, Rajeswari *et al*. 2012)	C4.5 decision tree
(Horng, Fan *et al*. 2008)	back-propagation neural network
(Mukherjee and Sharma 2012),(Sharma and Mukherjee 2012),(Koc, Mazzuchi *et al*. 2012)	Naïve Bayes
(Cohen, Avrahami *et al*. 2008)	Info-Fuzzy Network (IFN)
(Altwaijry and Algarny 2012),(Feng, Wang *et al*. 2009)	Bayesian algorithm
(Sujatha, Priya *et al*. 2012),(Lee, Kim *et al*. 2012)	GA
(Bolón-Canedo, Sánchez-Maroño *et al*. 2011)	C4.5, naive Bayes, one-layer (FNN), (PSVM), Multilayer Feed forward Neural Network (FNN)
(Jiaqi, Ru *et al*. 2011)	(CSWC-SVM)
(Farid and Rahman 2010)	improved self-adaptive Bayesian algorithm (ISABA)
(Arau, x *et al*. 2010)	K-means
(Shirazi 2009)	KNN
(Pereira, Nakamura *et al*. 2012)	optimum-path forest
(Mok, Sohn *et al*. 2010)	random effect logistic regression model
(Yi, Wu *et al*. 2011)	RS-ISVM
(Suthaharan and Panchagnula 2012)	RST
(Chi, Wee-Peng *et al*. 2012)	xtremelearningmachines (ELMs)

The Hybrid Classifiers:

The employment of multiple and hybrid classifiers, enhances the accuracy of classification and facilitates understanding difficult problems. Hybrid classifiers have drawn the attention of a lot of scholars, especially in machine learning and statistics. The objective hybrid classifier is to merge quite a few machine learning techniques, to significantly enhance the effectiveness from the hybrid system. Class accuracy is the process of assessing the classification accuracy of the hybrid system. Based on extensive review of various learning algorithms, it is evident that each algorithm has its some selected supremacy; each algorithm is best suitable for some specific problems, but definitely not all. M.Govindarajan *et al* (2011) have proposed the hybrid architecture. It has proved that, the performance is better for distinct classification methods. The hybrid approach facilitates anomaly and misuse detection. Combination of network and host IDS is another hybrid IDS, which merges anomaly misuse detection and was proposed by Duanyang Zhao *et al* (2010) (Sujatha, Priya *et al*. 2012).

In terms of the works depending on developing hybrid classifiers, table 6 illustrates several papers employed them.

Table 6: Hybrid classification method approaches.

Reference	Classifiers
(Powers and He 2008)	artificial immune system (AIS) (SOM)Kohonen Self Organising Map
(Panda, Abraham *et al*. 2012)	combo of DT, PCA, SPegasos, END, RF and Grading
(Liu, Kang *et al*. 2010)	RS-SVM
(Om and Kundu 2012)	KNN+NB
(Wang, Hao *et al*. 2010)	FC-ANN
(Lei and Ke-nan 2011)	(RS) + (FSVM)
(Kuang, Xu *et al*. 2012)	SVM with GA
(Huike and Daquan 2009)	SVM with fuzzy algorithm
(Ganapathy, Kulothungan *et al*. 2012)	IGA(immune genetic algorithm)
(Tjhai, Furnell *et al*. 2010)	neural network SOM + K means
(Hoang, Hu *et al*. 2009)	fuzzy based HMM (Hidden Markov Model)
(Xiaozhao, Wei *et al*. 2010)	FCM clustering algorithm and heuristic PSO algorithm +SVM
(Tong, Wang *et al*. 2009)	hybrid RBF/Elman neural network
(Tsai and Lin 2010)	K-NN (TAAN)
(Shafi and Abbass 2009)	UCSM
(Visumathi and Shunmuganathan 2012)	EDTSVM (enhanced decition tree based support vector machines)
(Guiling, Yongzhen *et al*. 2010)	Noise reduced Payload based fuzzy support vector Machine(PAYL-FSVM)
(Lin, Ying *et al*. 2012)	SA+SVM+DT
(Gan, Duanmu *et al*. 2013)	PLS–CVM
(Mulay, Devale *et al*. 2010)	Combination of HM-SVM and TSM-SVM
(Xiang, Yong *et al*. 2008)	DT+Bayesian clustering
(al 2010)	PCANNA
(Pingjie, Rong-an *et al*. 2010)	TASVM
(Lei, Zhi-ping *et al*. 2010)	fuzzy support vector machine
(Wei, Shaohua *et al*. 2010)	Fuzzy-SVM
(Srinivasu and Avadhani 2012)	GA-NN
(Meijuan, Jingwen *et al*. 2009)	SVM + RBFNN
(Devarakonda, Pamidi *et al*. 2012)	Bayesian Network+HMM
(Wei and Wu 2008)	kernel fisher discriminant analysis (KFDA), and Multiclass SVM
(Pachghare and Kulkarni 2011)	decision tree algorithms +SVM
(Su 2011)	a genetic algorithm combined with KNN (GA/KNN hybrid)
(Lee, Kim *et al*. 2011)	SOM+K-means
(Zhenguo and Guanghua 2009)	Improved support vector machines using artificial immunization algorithm
(Su 2011)	genetic weighted KNN
(Qazanfari, Mirpouryan *et al*. 2012)	MLP and SVM
(Abadeh, Mohamadi *et al*. 2011)	Genetic fuzzy systems (GFSs) based on (Pittsburg approach)

Ensemble Classifiers:

A lot of applications have successfully employed Ensemble based methods. It is noteworthy that, augmenting the differences among ensembles methods, devoid of augmenting their particular test error, inevitably decreases the ensemble test error. In contrast, if all ensemble members are similar in classifications, then the cumulative classification will be same as any single one of them, so no decreasing to the ensemble test error will happen. Different factors for different classifiers can be employed in machine learning methods. The table below illustrates that employing ensemble based techniques produces better results in terms of single classifier technique(Majidi, Mirzaei *et al*. 2008). Ensemble of classifiers-based methods offers a new and well accepted solution for a wide range of applications. These methods constitute uncomplicated combination formats, such as max/min rules, majority vote, averaged Bayes classifier and threshold voting. At first, an collection of classifiers is trained for every specific feature set, and later bundled, due to the fact the collection of classifiers provide superior classification accuracy as againstsingle classifiers (Parikh and Tsuhan 2008).

Table 7: Ensemble classification method approaches.

Reference	Classifiers
(Dartigue, Hyun Ik *et al*. 2009)	C4.5 decision tree
(Kavitha, Karthikeyan *et al*. 2012)	Neutrosophic Logic Classifiers and improvised genetic algorithm
(Li, Wang *et al*. 2009)	DT whose nodes consist linear SVM
(Sivatha Sindhu, Geetha *et al*. 2012)	neurotree
(Govindarajan and Chandrasekaran 2011)	neural: MLP, RBF
(Raj Kumar and Selvakumar 2011)	Resilient Back Propagation (RBP) Boost
(Giacinto, Perdisci *et al*. 2008)	Parzen density estimation+m-SVC+k-means
(Guanghui, Jiankang *et al*. 2011)	MK-SVM

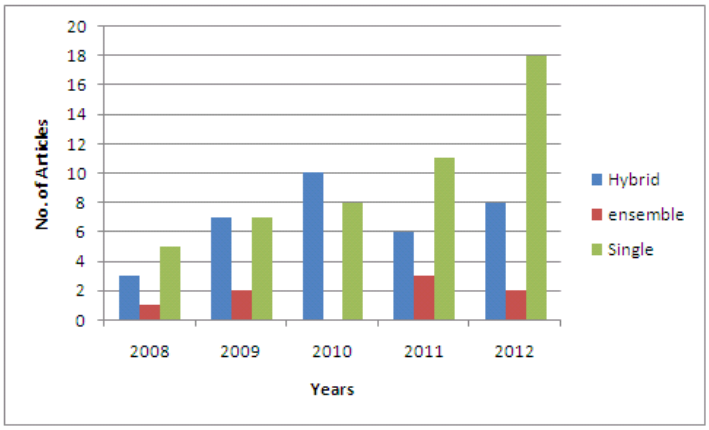

Fig. 1: Classifiers type use, based on different years.

The Conclusions and Discussion Section:

Up to now, significant amounts of resources and time are already dedicated to IDS and other machine learning Techniques for example: SVM, Bayesian belief networking,artificial neural network, data mining methods, and hybrid intelligent system usually are researched to design IDS. Nevertheless, it would appear that not one of them has the capacity to identify all type of invasion tries effectively when it comes to recognition and false alarm rate. Therefore, the necessity is to blend various classifiers like a hybrid data mining technique to boost the recognition precision with the model integrated to create successful smart conclusions in determining the intrusions (Panda, Abraham *et al*. 2012). This paper had evaluated a number of existing researches related to intrusion detection system and classifications techniques such as single classification techniques, hybrid classification techniques, and ensemble classification techniques. With regards to the outcomes of the evaluation, it is evident that, still intensive investigations have to be conducted, for building intrusion detection systems, with machine learning techniques. Because of this it may be contended the fact that mixture of classifiers trained on several feature sets may possibly have better Functionality compared to every individual classifier. Additionally, performances also needs to be greater than that surrounding classifiers with different single feature vector that contains all of the obtainable features(Giacinto and Roli 2002).

The following aspects could facilitate future research.

- Baseline paper (anchor paper). It is not adequate to perform classification results comparison with baseline classification techniques as the results of one classifier in most of cases is not good in accuracy and prediction; it is worth to try comparisons with combination of more than one classifier to show the new good results.
- Multiple classification techniques: It is essential to design more classy classifiers by merging ensemble and hybrid classifiers. The concept of hybridizing multiple classifiers is aimed at combining each other rather than competing with one another.
- Feature selection/extraction: feature selection could result in best detection. Every feature selection algorithm has its advantages and disadvantages for selecting the best features which will affectthe classification output, so combining more than feature selection could lead to better accuracy, instead of them competing to each other, they will work together. Empowering them together could prove better classification

REFERENCES

2000. "Testing Intrusion detection systems: a critique of the 1998 and 1999 DARPA intrusion detection system evaluations as performed by Lincoln Laboratory." ACM Trans. Inf. Syst. Secur., 3(4): 262-294.

Abadeh, M.S., H. Mohamadi, *et al*., 2011. "Design and analysis of genetic fuzzy systems for intrusion detection in computer networks." Expert Systems with Applications, 38(6): 7067-7075.

Agarwal, B. and N. Mittal, 2012. "Hybrid Approach for Detection of Anomaly Network Traffic using Data Mining Techniques." Procedia Technology, 6(0): 996-1003.

al, S.L. e., 2010. "Feature Reduction using Principal Component Analysis for Effective Anomaly-Based Intrusion Detection on NSL-KDD." International Journal of Engineering Science and Technology, 2(6): 1790-1799.

Altwaijry, H. and S. Algarny, 2012. "Bayesian based intrusion detection system." Journal of King Saud University - Computer and Information Sciences, 24(1): 1-6.

Amiri, F., M. Rezaei Yousefi, *et al*., 2011. "Mutual information-based feature selection for intrusion detection systems." Journal of Network and Computer Applications, 34(4): 1184-1199.

Arau, X., *et al*., 2010. Identifying important characteristics in the KDD99 intrusion detection dataset by feature selection using a hybrid approach. Telecommunications (ICT), 2010 IEEE 17th International Conference on.

Ashok, R., A.J. Lakshmi, *et al*., 2011. Optimized feature selection with k-means clustered triangle SVM for Intrusion Detection. Advanced Computing (ICoAC), 2011 Third International Conference on.

Bolón-Canedo, V., N. Sánchez-Maroño, *et al*., 2011. "Feature selection and classification in multiple class datasets: An application to KDD Cup 99 dataset." Expert Systems with Applications 38(5): 5947-5957.

Catania, C.A., F. Bromberg, *et al*., 2012. "An autonomous labeling approach to support vector machines algorithms for network traffic anomaly detection." Expert Systems with Applications, 39(2): 1822-1829.

Chen, R.C., K.F. Cheng, *et al*., 2009. Using Rough Set and Support Vector Machine for Network Intrusion Detection System. Proceedings of the 2009 First Asian Conference on Intelligent Information and Database Systems, IEEE Computer Society, 465-470.

Chi, C., T. Wee-Peng, *et al*., 2012. Extreme learning machines for intrusion detection. Neural Networks (IJCNN), The 2012 International Joint Conference on.

Chunhua, G. and Z. Xueqin, 2009. A Rough Set and SVM Based Intrusion Detection Classifier. Computer Science and Engineering, 2009. WCSE '09. Second International Workshop on.

Cohen, L., G. Avrahami, *et al*., 2008. "Info-fuzzy algorithms for mining dynamic data streams." Applied Soft Computing, 8(4): 1283-1294.

Dartigue, C., J. Hyun Ik, *et al*., 2009. A New Data-Mining Based Approach for Network Intrusion Detection. Communication Networks and Services Research Conference, 2009. CNSR '09. Seventh Annual.

Devarakonda, N., S. Pamidi, *et al*., 2012. "Intrusion Detection System using Bayesian Network and Hidden Markov Model." Procedia Technology, 4(0): 506-514.

Eid, H.F., A. Darwish, *et al*., 2010. Principle components analysis and Support Vector Machine based Intrusion Detection System. Intelligent Systems Design and Applications (ISDA), 2010 10th International Conference on.

Farid, D.M. and M.Z. Rahman, 2010. Anomaly Network Intrusion Detection Based on Improved Self Adaptive Bayesian Algorithm.

Feng, L., W. Wang, *et al*., 2009. "Predicting intrusion goal using dynamic Bayesian network with transfer probability estimation." Journal of Network and Computer Applications, 32(3): 721-732.

Gan, X.S., J.S. Duanmu, *et al*., 2013. "Anomaly intrusion detection based on PLS feature extraction and core vector machine." Knowledge-Based Systems, 40(0): 1-6.

Ganapathy, S., K. Kulothungan, *et al*., 2012. "A Novel Weighted Fuzzy C –Means Clustering Based on Immune Genetic Algorithm for Intrusion Detection." Procedia Engineering, 38(0): 1750-1757.

Gengming, Z. and L. Junguo, 2008. Research of Intrusion Detection Based on Support Vector Machine. Advanced Computer Theory and Engineering, 2008. ICACTE '08. International Conference on.

Giacinto, G., R. Perdisci, *et al*., 2008. "Intrusion detection in computer networks by a modular ensemble of one-class classifiers." Information Fusion, 9(1): 69-82.

Giacinto, G. and F. Roli, 2002. Intrusion detection in computer networks by multiple classifier systems. Pattern Recognition, 2002. Proceedings. 16th International Conference on.

Govindarajan, M. and R.M. Chandrasekaran, 2011. "Intrusion detection using neural based hybrid classification methods." Computer Networks, 55(8): 1662-1671.

Guanghui, S., G. Jiankang, *et al*., 2011. An Intrusion Detection Method Based on Multiple Kernel Support Vector Machine. Network Computing and Information Security (NCIS), 2011 International Conference on.

Guiling, Z., K. Yongzhen, *et al*., 2010. An Improvement of Payload-Based Intrusion Detection Using Fuzzy Support Vector Machine. Intelligent Systems and Applications (ISA), 2010 2nd International Workshop on.

Guo, N.R. and T.H.S. Li, 2011. "Construction of a neuron-fuzzy classification model based on feature-extraction approach." Expert Systems with Applications, 38(1): 682-691.

Hoang, X.D., J. Hu, *et al*., 2009. "A program-based anomaly intrusion detection scheme using multiple detection engines and fuzzy inference." Journal of Network and Computer Applications 32(6): 1219-1228.

Horng, S.J., P. Fan, *et al*., 2008. "A feasible intrusion detector for recognizing IIS attacks based on neural networks." Computers & Security, 27(3-4): 84-100.

Horng, S.J., M.Y. Su, *et al*., 2011. "A novel intrusion detection system based on hierarchical clustering and support vector machines." Expert Systems with Applications, 38(1): 306-313.

Huike, L. and G. Daquan, 2009. A Novel Intrusion Detection Scheme Using Support Vector Machine Fuzzy Network for Mobile Ad Hoc Networks. Web Mining and Web-based Application, 2009. WMWA '09. Second Pacific-Asia Conference on.

Jiaqi, J., L. Ru, *et al*., 2011. A New Intrusion Detection System Using Class and Sample Weighted C-support Vector Machine. Communications and Mobile Computing (CMC), 2011 Third International Conference on.

Jingbo, Y., L. Haixiao, *et al*., 2010. Intrusion Detection Model Based on Improved Support Vector Machine. Intelligent Information Technology and Security Informatics (IITSI), 2010 Third International Symposium on.

Kai-mei, Z., Q. Xu, *et al*., 2009. Intrusion Detection Using Isomap and Support Vector Machine. Artificial Intelligence and Computational Intelligence, 2009. AICI '09. International Conference on.

Kausar, N., B.B. Samir, *et al*., 2012. An approach towards intrusion detection using PCA feature subsets and SVM. Computer & Information Science (ICCIS), 2012 International Conference on.

Kavitha, B., D.S. Karthikeyan, *et al*., 2012. "An ensemble design of intrusion detection system for handling uncertainty using Neutrosophic Logic Classifier." Knowledge-Based Systems, 28(0): 88-96.

Koc, L., T.A. Mazzuchi, *et al*., "A network intrusion detection system based on a Hidden Naïve Bayes multiclass classifier." Expert Systems with Applications(0).

Koc, L., T.A. Mazzuchi, *et al*., 2012. "A network intrusion detection system based on a Hidden Naïve Bayes multiclass classifier." Expert Systems with Applications, 39(18): 13492-13500.

Kuang, F., W. Xu, *et al*., 2012. "A novel approach of KPCA and SVM for intrusion detection." Journal of Computational Information Systems, 8(8): 3237-3244.

Lee, K., J. Kim, *et al*., 2008. "DDoS attack detection method using cluster analysis." Expert Systems with Applications, 34(3): 1659-1665.

Lee, S., G. Kim, *et al*., 2011. "Self-adaptive and dynamic clustering for online anomaly detection." Expert Systems with Applications, 38(12): 14891-14898.

Lee, S.M., D.S. Kim, *et al*., 2012. "Detection of DDoS attacks using optimized traffic matrix." Computers & Mathematics with Applications, 63(2): 501-510.

Lei, L. and Z. Ke-nan, 2011. A New Intrusion Detection System Based on Rough Set Theory and Fuzzy Support Vector Machine. Intelligent Systems and Applications (ISA), 2011 3rd International Workshop on.

Lei, L., G. Zhi-ping, *et al*., 2010. Fuzzy Multi-class Support Vector Machine Based on Binary Tree in Network Intrusion Detection. Electrical and Control Engineering (ICECE), 2010 International Conference on.

li, W. and Z. Liu, 2011. "A method of SVM with Normalization in Intrusion Detection." Procedia Environmental Sciences 11, Part A(0): 256-262.

Li, X.B., 2005. "A scalable decision tree system and its application in pattern recognition and intrusion detection." Decision Support Systems, 41(1): 112-130.

Li, Y., J.L. Wang, *et al*., 2009. "Building lightweight intrusion detection system using wrapper-based feature selection mechanisms." Computers & Security, 28(6): 466-475.

Li, Y., J. Xia, *et al*., 2012. "An efficient intrusion detection system based on support vector machines and gradually feature removal method." Expert Systems with Applications, 39(1): 424-430.

Lin Li, Z., M. Zhang Ya, *et al*., 2010. Network intrusion detection method by least squares support vector machine classifier. Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on.

Lin, S.W., K.C. Ying, *et al*., 2012. "An intelligent algorithm with feature selection and decision rules applied to anomaly intrusion detection." Applied Soft Computing, 12(10): 3285-3290.

Liu, Z., J. Kang, *et al*., 2010. A hybrid method of rough set and support vector machine in network intrusion detection. Signal Processing Systems (ICSPS), 2010 2nd International Conference on.

Mahoney, M. and P. Chan, 2003. An Analysis of the 1999 DARPA/Lincoln Laboratory Evaluation Data for Network Anomaly Detection. Recent Advances in Intrusion Detection. G. Vigna, C. Kruegel and E. Jonsson, Springer Berlin Heidelberg, 2820: 220-237.

Majidi, F., H. Mirzaei, *et al*., 2008. A diversity creation method for ensemble based classification: Application in intrusion detection. Cybernetic Intelligent Systems, 2008. CIS 2008. 7th IEEE International Conference on.

Meijuan, G., T. Jingwen, *et al*., 2009. Intrusion Detection Method Based on Classify Support Vector Machine. Intelligent Computation Technology and Automation, 2009. ICICTA '09. Second International Conference on.

Mohammad, M.N., N. Sulaiman, *et al*., 2011. "A novel local network intrusion detection system based on support vector machine." Journal of Computer Science, 7(10): 1560-1564.

Mohammed, M.N. and N. Sulaiman, 2012. "Intrusion Detection System Based on SVM for WLAN." Procedia Technology, 1(0): 313-317.

Mok, M.S., S.Y. Sohn, *et al*., 2010. "Random effects logistic regression model for anomaly detection." Expert Systems with Applications, 37(10): 7162-7166.

Mukherjee, S. and N. Sharma, 2012. "Intrusion Detection using Naive Bayes Classifier with Feature Reduction." Procedia Technology, 4(0): 119-128.

Mulay, S.A., P.R. Devale, *et al*., 2010. Decision tree based Support Vector Machine for Intrusion Detection. Networking and Information Technology (ICNIT), 2010 International Conference on.

Muniyandi, A.P., R. Rajeswari, *et al*., 2012. "Network Anomaly Detection by Cascading K-Means Clustering and C4.5 Decision Tree algorithm." Procedia Engineering, 30(0): 174-182.

Muntean, M., H. Valean, *et al*., 2010. A novel intrusion detection method based on support vector machines. Computational Intelligence and Informatics (CINTI), 2010 11th International Symposium on.

Om, H. and A. Kundu, 2012. A hybrid system for reducing the false alarm rate of anomaly intrusion detection system. Recent Advances in Information Technology (RAIT), 2012 1st International Conference on.

Pachghare, V.K. and P. Kulkarni, 2011. Pattern based network security using decision trees and support vector machine. Electronics Computer Technology (ICECT), 2011 3rd International Conference on.

Panda, M., A. Abraham, *et al*., 2012. "A Hybrid Intelligent Approach for Network Intrusion Detection." Procedia Engineering, 30(0): 1-9.

Parikh, D. and C. Tsuhan, 2008. "Data Fusion and Cost Minimization for Intrusion Detection." Information Forensics and Security, IEEE Transactions on, 3(3): 381-389.

Peddabachigari, S., A. Abraham, *et al*., 2007. "Modeling intrusion detection system using hybrid intelligent systems." Journal of Network and Computer Applications, 30(1): 114-132.

Pereira, C.R., R.Y.M. Nakamura, *et al*., 2012. "An Optimum-Path Forest framework for intrusion detection in computer networks." Engineering Applications of Artificial Intelligence, 25(6): 1226-1234.

Pingjie, T., J. Rong-an, *et al*., 2010. Feature Selection and Design of Intrusion Detection System Based on k-Means and Triangle Area Support Vector Machine. Future Networks, 2010. ICFN '10. Second International Conference on.

Powers, S.T. and J. He, 2008. "A hybrid artificial immune system and Self Organising Map for network intrusion detection." Information Sciences, 178(15): 3024-3042.

Qazanfari, K., M.S. Mirpouryan, *et al*., 2012. A novel hybrid anomaly based intrusion detection method. Telecommunications (IST), 2012 Sixth International Symposium on.

Raj Kumar, P.A. and S. Selvakumar, 2011. "Distributed denial of service attack detection using an ensemble of neural classifier." Computer Communications, 34(11): 1328-1341.

Sabhnani, M. and G. Serpen, 2004. "Why machine learning algorithms fail in misuse detection on KDD intrusion detection data set." Intell. Data Anal., 8(4): 403-415.

Saha, S., A.S. Sairam, *et al*., 2012. Genetic algorithm combined with support vector machine for building an intrusion detection system. Proceedings of the International Conference on Advances in Computing, Communications and Informatics. Chennai, India, ACM: 566-572.

Sangkatsanee, P., N. Wattanapongsakorn, *et al*., 2011. "Practical real-time intrusion detection using machine learning approaches." Computer Communications, 34(18): 2227-2235.

Shafi, K. and H.A. Abbass, 2009. "An adaptive genetic-based signature learning system for intrusion detection." Expert Systems with Applications, 36(10): 12036-12043.

Sharma, N. and S. Mukherjee, 2012. "A Novel Multi-Classifier Layered Approach to Improve Minority Attack Detection in IDS." Procedia Technology, 6(0): 913-921.

Shirazi, H.M., 2009. "Anomaly Intrusion Detection System Using Information Theory, K-NN and KMC algorithms." Australian Journal of Basic & Applied Sciences, 3(3): 251-2597.

Sivatha Sindhu, S.S., S. Geetha, *et al*., 2012. "Decision tree based light weight intrusion detection using a wrapper approach." Expert Systems with Applications, 39(1): 129-141.

Somwang, P. and W. Lilakiatsakun, 2011. Computer network security based on Support Vector Machine approach. Control, Automation and Systems (ICCAS), 2011 11th International Conference on.

Srinivasu, P. and P.S. Avadhani, 2012. "Genetic Algorithm based Weight Extraction Algorithm for Artificial Neural Network Classifier in Intrusion Detection." Procedia Engineering, 38(0): 144-153.

Su, M.Y., 2011. "Real-time anomaly detection systems for Denial-of-Service attacks by weighted k-nearest-neighbor classifiers." Expert Systems with Applications, 38(4): 3492-3498.

Su, M.Y., 2011. "Using clustering to improve the KNN-based classifiers for online anomaly network traffic identification." Journal of Network and Computer Applications, 34(2): 722-730.

Sujatha, P.K., C.S. Priya, *et al*., 2012. Network intrusion detection system using genetic network programming with support vector machine. Proceedings of the International Conference on Advances in Computing, Communications and Informatics. Chennai, India, ACM: 645-649.

Suthaharan, S. and T. Panchagnula, 2012. Relevance feature selection with data cleaning for intrusion detection system. Southeastcon, 2012 Proceedings of IEEE.

Tavallaee, M., E. Bagheri, *et al*., 2009. A detailed analysis of the KDD CUP 99 data set. Computational Intelligence for Security and Defense Applications, 2009. CISDA 2009. IEEE Symposium on.

Tjhai, G.C., S.M. Furnell, *et al*., 2010. "A preliminary two-stage alarm correlation and filtering system using SOM neural network and K-means algorithm." Computers & Security, 29(6): 712-723.

Tong, X., Z. Wang, *et al*., 2009. "A research using hybrid RBF/Elman neural networks for intrusion detection system secure model." Computer Physics Communications, 180(10): 1795-1801.

Tsai, C.F. and C.Y. Lin, 2010. "A triangle area based nearest neighbors approach to intrusion detection." Pattern Recognition, 43(1): 222-229.

Visumathi, J. and K.L. Shunmuganathan, 2012. "An Effective IDS for MANET Using Forward Feature Selection and Classification Algorithms." Procedia Engineering, 38(0): 2816-2823.

Wang, G., J. Hao, *et al*., 2010. "A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering." Expert Systems with Applications, 37(9): 6225-6232.

Wei, Y.X. and M.Q. Wu, 2008. "KFDA and clustering based multiclass SVM for intrusion detection." The Journal of China Universities of Posts and Telecommunications, 15(1): 123-128.

Wei, Z., T. Shaohua, *et al*., 2010. Fuzzy Multi-Class Support Vector Machines for cooperative network intrusion detection. Cognitive Informatics (ICCI), 2010 9th IEEE International Conference on.

Winter, P., E. Hermann, *et al*., 2011. Inductive Intrusion Detection in Flow-Based Network Data Using One-Class Support Vector Machines. New Technologies, Mobility and Security (NTMS), 2011 4th IFIP International Conference on.

Wu, H.C. and S.H.S. Huang, 2010. "Neural networks-based detection of stepping-stone intrusion." Expert Systems with Applications, 37(2): 1431-1437.

Xiang, C., P.C. Yong, *et al*., 2008. "Design of multiple-level hybrid classifier for intrusion detection system using Bayesian clustering and decision trees." Pattern Recognition Letters, 29(7): 918-924.

Xiaozhao, F., Z. Wei, *et al*., 2010. A Research on Intrusion Detection Based on Support Vector Machines. Communications and Intelligence Information Security (ICCIIS), 2010 International Conference on.

Xie, Y. and Y. Zhang, 2012. An intelligent anomaly analysis for intrusion detection based on SVM. Computer Science and Information Processing (CSIP), 2012 International Conference on.

Xuejun, D., Z. Guiling, *et al*., 2008. High Efficient Intrusion Detection Methodology with Twin Support Vector Machines. Information Science and Engineering, 2008. ISISE '08. International Symposium on.

Yi, Y., J. Wu, *et al*., 2011. "Incremental SVM based on reserved set for network intrusion detection." Expert Systems with Applications, 38(6): 7698-7707.

Yongli, Z. and Z. Yanwei, 2010. Application of Improved Support Vector Machines in Intrusion Detection. e-Business and Information System Security (EBISS), 2010 2nd International Conference on.

Yu, J., H. Lee, *et al*., 2008. "Traffic flooding attack detection with SNMP MIB using SVM." Computer Communications, 31(17): 4212-4219.

Zaman, S. and F. Karray, 2009. Features Selection for Intrusion Detection Systems Based on Support Vector Machines. Consumer Communications and Networking Conference, 2009. CCNC 2009. 6th IEEE.

Zhao, R., Y. Yu, *et al*., 2009. An Intrusion Detection Algorithm Model Based on Extension Clustering Support Vector Machine. Artificial Intelligence and Computational Intelligence, 2009. AICI '09. International Conference on.

Zhenguo, C. and Z. Guanghua, 2009. Support Vector Machines Improved by Artificial Immunisation Algorithm for Intrusion Detection. Information Engineering and Computer Science, 2009. ICIECS 2009. International Conference on.