Research Journal of Applied Sciences 11 (10): 921-932, 2016
ISSN: 1815-932X

Features Evaluation for Anomaly Intrusion Detection System

[1]Adil M. Salman and [2]Safaa O. Al-mamory
[1]College of Info. Technology, University of Babylon, Hillah, Iraq
[2]College of Business Informatics, University of Information Technology and Communications Hillah, Iraq

Abstract: In network security there is an essential field called intrusion detection system it is a method fordetecting abnormal activities in network traffic. Another significant field in these systems is the feature selection methods which reduces the calculation time and tested data.This study introduces an evaluation of the most important features that used in intrusion detection methods of network flow to help the researchers knowing which features are important. Fifty-three different methods are investigated of feature selection and some intrusion detection methods including 39 methods that using different DARPA datasets and 14 methods using other different datasets. We also applied an experiment consists of 96 tests using WEKA 3.8.0 Software for datamining where we utilized 12 combinations of feature selection algorithms, the used datasets were KDD-CUP99 and NSL-KDD datasets. The contribution of this study is the focus on which of the features have the highest selected percentage for both the studied papers and our experiment. We have concluded that the basic features and the features based on the hosts which give the resource of the attacks was the most features that researchers used.

Key words: Feature selection, anomaly detection, network security, network flow, features

INTRODUCTION

IDSs (Intrusion Detection Systems) could be defined as the process of monitoring the computer or network activities to detectany abnormalevent which is considered to be an intrusion. There are three types of IDSs: misuse detection system; anomaly detection system and hybrid detection system. The first system based on building a patterns of the traffic data assumed to be malicious by collecting a set of signatures for the known attacks but it cannot recognize the unknown attacks. The second system, collects information from the flow network, constructinga flow profile, then comparing activities against a (normal) baseline. The hybrid detection system integrates the techniques ofthese two methods (Patcha and Park, 2007; Bhuyan *et al*., 2014).

The problem of the huge size of the originally collected network flows datasets, have a common solution called feature selection technique which decreases the number offeatures. Many papers have been proposed a methods to select the important features through the three types of these techniques: Filter wrapper and hybrid (Chen *et al*., 2006, 2007). This study aims to represent an overview of the most important features used in IDSs. We categorize the studies here into the 3 main types of feature selection methods mentioned above. Then we evaluate the important selected features based on the frequency of use in the research. We have studied 53 methods of features election then we did an experiment of 96 tests using WEKA 3.8.0 software for data-mining then did a comparison between the researcher's methods and our experiment. We conclude that the researcher's trends go toward the direction of traffic detection based on the basic features, then features based on hosts which give the resources of the attacks. Our contribution in this study is to focus on which of the features that have the highest number of been selected by the researchers methods and our experiment.

Literature review

Related work: Chen *et al.* (2006). they surveyed some feature selection algorithms used in IDSs, then made a taxonomy of these algorithms based on three categories: filter, wrapper and hybrid. Then they conclude to identifying the trends and challenges of feature selection methods and development in intrusion detection systems. Garg and Kumar (2014) they represented the comparative performance of compatible classification algorithms using NSL-KDD dataset. Then using WEKA to evaluates these classifiers using 41 attributes. They applied the garret's ranking technique to rank different classifiers according

Corresponding Author: Adil M. Salman,College of Info. Technology University of Babylon, Hillah, Iraq

to their performance. Rotation forest classification approach outperformed the rest. Bjerkestrand *et al.* (2015). they evaluate and compare the usage of various feature selection and reduction algorithms using publicly available datasets. Three feature selection algorithms were used, consisting of an attribute evaluator and a test method. The initial results indicate that the performance of the classifier is unaffected by reducing the number of attributes. Oyebode *et al.* (2011). they examined the accuracy of using data mining techniques in intrusion detection systems using three classification techniques: Naïve bayes, radial basis and rotation forest using KDD99 dataset. Amrita and Ahmed (2012). they survey some feature selection methods for IDS using KDD-CUP'99 benchmark dataset based on three categories which are the filter, wrapper and hybrid approach and different evaluation criteria. Singh Kumar (2015). they represent a review of the three vast kinds of feature selection techniques as the filter, wrapper and hybrid methods and review a numerous feature selection methods for IDS using KDD CUP'99 benchmark dataset with various evaluation criteria. Kumar (2016) they evaluates the performance of data mining classification algorithms namely C4.5, J48, nave bayes, NB-tree and random forest using NSL-KDD dataset and focuses on Correlation Feature Selection (CFS) assess. The results demonstrate that NB-tree and random forest outperforms other two algorithms in terms of predictive accuracy and detection rate.

Intrusion Detection Systems (IDSs): Intrusion detection systemsis defined as a hardware or Software toolused todistinguish between unauthorized and authorized access to a computer system or network. It collects data by monitoring the network traffic. The collected network packets are analyzed for rule violations and when any violation is found the IDS will send an alert to network administrators or to a specific Software in 2007. Intrusion detection is a kind of security management system that deals with computers and networks, the intrusion canbe outlined as a collection of actions aimed to compromise the computer security functions like confidentiality integrity and availability (1990).

MATERIALS AND METHODS

Feature selection methods: In all IDS's learning mechanisms, it is not a good idea to deal with the originally collected dataset because of its large size. Since there are features that are either irrelevant or redundant for the learning algorithm it needs to be optimized. Using data mining in intrusion detection requires a lot of review data to build the proper profile for the datasets. Feature selectionis a preprocessing step to machine learning of selecting a subset of relevant features for building robust learning models. Hence, it is important to determine an optimal set of features that accurately represents the characteristics of the traffic being evaluated. The feature selection methods could be categorized to three types as thefilter wrapper and Hybrid approaches (Chen *et al.*, 2006, 2007; Garg and Kumar, 2014a, b; Lee and Stolfo, 1998; Sheen and Rajesh, 2008; Araujo *et al.*, 2010). We selected some different studies that using various methods for feature selection. These methods are classified into the three categories as we mentioned above (Filter, wrapper and hybrid).

Filter methods: As mentioned above, these methods do not use any learning algorithm but utilize an independent measure and evaluate the selected important features based on this measure. Table 1 is summarizing some studies that use the filter methods for feature selection.

Table 1: Summary for the studies that using filter methods for feature selection

Authors	Feature selection used algorithm	Dataset	Most important features
Lakhina *et al* (2005)	Entropy, multiway subspace method	1	sourc-ip, dest-ip, sourc-port, dest-port
Sheen *et al*(2008)	chi square, information gain, relieff decision tree classifier	2	20 Features selected* (2,3,4,5,12,22,23,24,27,28, 30,31,32,33 34 35,37,38,40, 41)
Zainal and Rajesh (2006)	Using rough set for feature selection and for classification	2	6 Features selected* (3,4,5,24,32,41)
Eid *et al.* (2013)	Analysis of the pearson correlation coefficients	3	17 Features selected* (5,6,18,22,23,25,26,27,28,31,35,36,37 38 39,40,41)
Kayacik	Using information Gain	2	31 Features selected* (1,2,3,4,5,6,12,15,16,17,18,19,23,24, 25,26,27 28,29,30,31,32,33,34,35,36,37,38,39,40,41)
Peng *et al.* (2004)	Sequential Change-point Detection Method based on CUSUM	7	Sourc-IP address
Chen *et al.* (2007)	Change Aggregation Trees(CAT) Distributedchange-point Detection(DCD). secure Infrastructure protocol (ISP)	4	Sourc-IP, dest-ip, sourc-port, dest-port, Applied Protocol
Saad *et al.* (2008)	DHT (distributed hash table) algorithm	8	NodeID, ObjectID, Packet flow frequency, ACK packet ratio, SYN packet ratio,
Rahmani *et al.* (2009)	Joint entropy algorithm	5	Number of received packets, ip-flow
Sengar *et al.* (2009)	Hellinger Distance	1	sourc-ip, dest-ip, sourc-port, dest-port
Tang *et al.* (2014)	X^2 Distance (XD), Mean Deviation (MD)	6	Frequency Distribution of TCP Traffic

*Return to (Stolfo *et al.*, 2000) to see the features name; Abilene and g´eant datasets. KDD-CUP99 dataset. Nsl-kdd dataset. Deter testbed; Caida 2007 dataset. Darpa 1999 dataset. Collected flow records. Connection records

Table 2: Summary for the studies that use wrapper methods for feature selection

Researchers	Feature selection used algorithm	Dataset	Most important features
Moustafa and Slay (2015)	TCP Trace for features selection and NB, DT, ANN, EM Clustering Methods	9	Features are named in (Moustafa and Slay, 2015)
Lee and Stolfo (1998)	Association Rules and Frequent Episodes'Algorithms	1	Start time, duration, participating hosts,ports statistics of the connection, flag, protocol
	Symmetric+Gain Ratio (15)		10 features were selected from 15 features* (2,3,4,5,6,12,23,25,36,37)
	One R+Symmetric (17)		16 features were selected from 17 features* (2,3,4,5,6,12,23,24,25,29,32,33,34,35,36,37)
	One R+ReliefF (21)		21 features were selected from 21 features* (2,3,4,5,6,12,23,24,25,26,29,30,32,33,34,35, 36,37,38,39,40)
	Symmetric+Information Gain (20)		19 features were selected from 20 features* (2,3,4,5,6,12,23,24,25,29,31,32,33,34,35 36,37 ,38,39)
	One R+Symmetric+Gain Ratio (17)		11 features were selected from 17 features* (2,3,4,5,6,12,23,25,33,36,37)
Garg and Kumar (2014a, b)	Symmetric+Information Gain (18)		17 features were selected from 18 features* (2,3,4,5,6,12,23,24,25,29,32,33,34,35,36,37,38)
Using WEKA software	One R+ReliefF (17)	2	14 features were selected from 17 features* (2,3,4,5,6,12,23,24,25,29,32,33,34,36)
	One R+ Symmetric (21)		20 features were selected from 21 features* (2,3,4,5,6,12,23,24,25,26,29,30,31,32,33,34 ,37,39)
	Symmetric+Gain Ratio (24)		18 features were selected from 24 features* (2,3,4,5,6,12,23,24,25,26,31,32,33,34,36,37,38,39)
	One R+ReliefF (20)		17 features were selected from 20 features* (2,3,4,5,6,12,23,24,25,26,29,30,32,33,34,36,37)
	Gain Ratio+Information Gain (15)		9 features were selected from 15 features* (2,3,4,5,6,12,23,36,37)
	Symmetric+Information Gain (17)		16 features were selected from 17 features* (2,3,4,5,6,12,23,24,29,32,33,34,35,36,37,38)
	One R+ReliefF (19)		15 features were selected from 19 features* (2,3,4,5,6,12,23,24,25,26,29,32,33,34,36)
Ghali (2009)	Using rough set to select significant features and neural network algorithm for training	3	7 Features selected* (5,6,23,24,32,33,36)
Sindhu *et al.* (2012)	the proposed method combined GA and neurotree algorithm	3	16 Features selected* (2,3,4,5,6,8,10,11,24,25,29,35,36,37,39,40)
	BestFirst+ConsistencySubsetEval		11 Features selected* (1,3,5,6,23,24,33,35,36,37,38)
	GeneticSearch+CfsSubsetEval		20 Features selected* (1,2,3,4,5,6,8,11,12,23,26,28,29,30,32,33,35,37,38,39)
	GeneticSearch+ConsistencySubsetEval		20 Features selected * (1,3,4,8,10,11,14,16,21,24,25,26,28,29,31,33,36,37 40,41)
	GreedyStepwise+CfsSubsetEval		9 Features selected* (4,6,8,10,19,30,35,36,37)
	Ranker+ChiSquaredAttributeEval		36 Features selected* (1,2,3,4,5,6,7,8,10,11,12,13,14,16,17,19,22,23,24,25, 26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41)
	RankSearch+CfsSubsetEval		22 Features selected* (2,3,4,5,6,7,8,10,11,12,13,25,26,27,29,30,34,35,36,37 38,39)
	RankSearch+ConsistencySubsetEval		34 Features selected* (1,2,3,4,5,6,7,8,10,11,12,13,14,17,22,23,24,25, 26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41)
Suebsing	Use euclidean distance for selecting the robust features, then use C5.0 classifier	3	30 Features selected* (1,2,5,6,7,8,9,10,11,12,13,14, 15,16,17,18,19,22,25,26,27,28,30,31,35,37,38,39,40,41)
Lin *et al.* (2012)	SVM and SA algorithms for feature selection DT for classification	3	28 Features selected* (1,2,3,5,6,7,8,10,11,12,13, 16,22,23,25,28,29,30,32, 33,34,35,36,37,38,39,40,41)
Li *et al.* (2012)	K-means, ANT Colony Optimization, SVM Gradually Feature Removal (GFR) method	3	19 Features selected* (2,4,8,10,14,15,19,25,27,29,31,32,33,34,35,36,37 38,40)
Song *et al.* (2013)	Fuzzy C-means, C4.5 decision tree	3	8 Features selected * (2,3,4,5,6,23,30,36)
mamory and Jassim (2015)	Information Gain for feature selection, Very Fast Decision Tree (VFDT) Algorithm for classification	3	Sourc-IP, Dest-IP, Sourc-Port, Dest-Port, Protocol and 20 Features selected* (1,2,3,4,5,6,12,23,24,25,26,27,28,29,32,33,34,38,39,41)
Gavrilis and Dermatas (2005)	Radial-Basis-Function Neural Network (RBF-NN) Algorithm (RBF-NN) Algorithm	4	(Sourc-Port, SEQ number of client, Window size, Syn, Ack, Fin, Psh, Urg, Rst flags) The most important 3 (Sourc-Port, SEQ number, Syn flag) used in Real-Time
Munz	K-means Clustering Algorithm	5	Num. of bytes, Num. of packets, Num. of different src-despairs, time intervals, service-specific ports
Cheng *et al.* (2009)	IAFV, Support Vector Machin (SVM) Algorithm	6	IP Address Feature Value
Zhong and Yue (2010)	FCM cluster algorithm, Apriori association algorithm	7	Dest-IP, Dest-Port, Flag
Bhaya and Manaa (2014)	Entropy, K-means, Centroid-Based rules	8	Time, Sourc-IP, Dest-IP, Sourc-Port, Dest-Port, Protocol

*Return to see the features name; TCPDUMP DATA.2- NSL-KDD dataset.3- KDD-CUP99 dataset; Simulated Data, Online Data Filtered for www service only; Flow records (Cisco Netflow, IPFIX); ARPA1999 dataset.7- Campus Network; CAIDA2007 and CAIDA2008 datasets; DARPA 2009 dataset

Table 3: Summary for the studies that use Hybrid methods for feature selection

Researchers	Feature selection used algorithm	Dataset	most important features
Srihari and Anitha (2014)	Wavelet, Naïve Bayes, Decision Tree SVMAlgorithms	1	Low-dimensional, High-dimensional
Araújo *et al.* (2010)	Information Gain, K-means	2	14Features selected* (2,3,5,6,9,11,12,14,22,30,31,32,35,37)
Depren *et al.* (2005)	Self-Organizing MAP(SOM), J.48 Decision Tree, Decision Support System (DSS)	2	6 Features selected* (1,2,3,4,5,6)
Zhang *et al.* (2008)	Random Forests Algorithm	2	38 Features selected* (1,2,3,4,5,6,8,9,10,11,12,13,14,15,16,17,18,19 22,23,24,25,26,27,28,29,30,31,32,33,34,35,36 37,38,39,40,41)
Chebrolu *et al.* (2005)	Markov Blanket Model for selection	2	17 Features selected* (1,2,3,5,7,8,11,12,14,17,22,23,24,25,26,30,32)
	Decision Tree		12 Features selected* (3,5,6,12,23,24,25,28,31,32,33,35)
Kim *et al.* (2007)	Random Forest	2	5 Features selected* (3,6,12,13,23)
Chou *et al.* (2008)	Symmetric Uncertainty, Fuzzy belief K-NN Classifier	3	12 Features selected* (1,2,3,4,5,6,12,23,24,31,32,37)
Wang *et al.* (2015)	Information Gain, Bayesian Networks, Decision Trees	2	9 Features selected* (3,5,23,24,31,32,33,37,39)
Ranjbar and Khorsandi (2011)	Danger Theory	4	-
Yao *et al.* (2012)	Entropy, Random Forests Algorithm	5	Sourc-IP, Dest-IP, Sourc-Port, Dest-Port, Time

Interval *Return to to see the features name; NETRESEC, CAIDA, LOIC, HOIC datasets; KDD-CUP 99 dataset; UCI and KDD-CUP99 datasets; Gnutella hybrid peer to peer network.5- DARPA1999 dataset

Wrapper methods: The wrapper methodsis more computationally expensive than the filter model because of using machine learning algorithms. These methods deal with the subsets of features as a search procedure. The wrapper model tends to give superior performance as feature subsets found are better suited to the predeterminedmining algorithm. Table 2 is summarizing some existing work of wrapper method.

Hybrid methods: These methods combinethe wrapper and the filter approaches. They use an independent measure to decide the best feature or subset of features and then use the learning algorithm to select the final best feature or subset of features. Table 3 is summarizing some existing work of hybrid methods.

RESULTS AND DISCUSSION

The experiments: For the purpose of obtaining more information about the features that have been selected we conducted an experiment consist of 96 test using WEKA 3.8.0 software for dataminingby utilizing 12 combinations of algorithms for feature selection using 2 types of KDD datasets which are the KDD-CUP 99 and NSL-KDD datasetsthey are available online for free. Tables 4 and 5 show the number of records and their distribution for 23 and 2 classes, respectively.

Feature selection algorithms in weka: There are some algorithms for feature selection in WEKA Software which help us to select some important features. These algorithms are a combination of two types of algorithms which are attribute evaluators and search methods. The first type we used (Cfs Subset Eval and Consistency Subset Eval). The second type we used (bestFirst, evolutionary search, geneticSearch, greedy stepwise, psosearch, rank search and bvultiobjective evolutionary search), for more information about these algorithms see.

Experiment implementation: We tried a combination of feature selection algorithms using the two datasets mentioned above. For the KDD-CUP 99 dataset we used the test set and the 10% set we didn't use the train set because of its large size and it takesa long time to give one result. For NSL-KDD we used the train set and the 20% set.

Table 6 shows the details of 48 tests of 12 combinations of the algorithms and their selected features using the datasets that are classified into 23 classes as mentioned in Table 4.

Table 7 shows the details of (48) tests of 12 combinations of the algorithms and their selected features using the datasets that classified into 2 classes as mentioned in Table 5. Figure 1 shows a comparison between the feature selection algorithms based on the number of classes and the number of the selected features. It was clearly shown that the datasets with 2 classes have less important features selected than the datasets with 23 classes.

Discussion of important features results: Here we will focus on the features that have been selected from the different types of datasetsto understand the trends in solutions of these issues as proposed by researchers. We have done some work to calculate the number of the

Table 4. Distribution of records for 23 classes in two type of KDD datasets

		KDD-CUP 99			NSL-KDD		
Classes	Specific class types	Train set	Test set	10%	Train set	Test set	20%
Normal	normal	972781	60593	97278	67343	9711	13449
Dos	back	2203	1098	2203	956	359	196
	land	21	9	21	18	7	1
	neptune	1072017	58001	107201	41214	4657	8282
	pod	264	87	264	201	41	38
	smurf	2807886	164091	280790	2646	665	529
	teardrop	979	12	979	892	12	188
Probe	ipsweep	12481	306	1247	3599	141	710
	nmap	2316	84	231	1493	73	301
	portsweep	10413	354	1040	2931	157	587
	satan	15892	1633	1589	3633	735	691
R2L	ftp_write	8	3	8	8	3	1
	guess_passwd	53	4367	53	53	1231	10
	imap	12	1	12	11	1	5
	multihop	7	18	7	7	18	2
	phf	4	2	4	4	2	2
	spy	2	0	2	2	0	1
	warezclient	1020	0	1020	890	0	181
	warezmaster	20	1602	20	20	944	7
U2R	buffer_overflow	30	22	30	30	20	6
	loadmodule	9	2	9	9	2	1
	perl	3	2	3	3	2	0
	rootkit	10	13	10	10	13	4
	Total instances	4898431	292300	494021	125973	18794	25192

Table 5: Distribution of records for 2 classes in two type of kdd 99 datasets

	KDD-CUP 99			NSL-KDD		
Specific class types	Train set	Test set	10%	Train set	Test set	20%
Normal	972781	60593	97278	67343	9711	13449
Anomaly	3925650	231707	396743	58630	9083	11743
Total instances	4898431	292300	494021	125973	18794	25192

Table 6: Numbers offeatures selected by various feature selection algorithms with 23 classes

		Most important features selected			
Feature selection algorithms		KDD-CUP 99		NSL-KDD	
Attribute evaluators	Search methods	Test set	10%	Train set	20%
Best first	Cfs subset eval	13 Features selected* (2,3,4,5,6,7,8,14,	11 Features selected* 21,24,30,33,36)	19 Features selected* (2,3,4,5,6,7,8,	17 Features selected* 14,23,30,36)
	Consistency subset eval	9 Features selected* (1,3,5,6,23,24, 34, 35,40)	10 Features selected* (1,3,5,6,12,33 35,36 37,38)	13 Features selected* (2,3,4,5,6,8,10,12,23 32,33,35,36,37,38,39 40)	11 Features selected* (1,3,4,6,23,24,33, 35,36,37,38)
Genetic search	Cfs subset eval	13 Features selected* (2,3,5,6,7,8,14,21 24,29,30,33,36)	17 Features selected* (2,3,4,5,6,7,8,10,12 19,23,29,30,31,33 36,38)	21 Features selected* (2,3,4,5,6,7,8,10,112 15,23,25,26,29,30,34 35,36,37,38)	17 Features selected* (1,2,3,4,5,6,8,10,14 19 23,25,26,27,29,30 32,35 36,37,38)
	Consistency subset eval	17 Features selected* (3,5,6,7,10,15,23 24,27,28,30,33 34 35,36,40,41)	22 Features selected* (2,4,5,6,8,9,10,12,13 15,20,23,24,25,26,29 32,33,35,37,39,41)	24 Features selected* (3,4,5,6,7,9,10,11,12 13,15 17,20,23,24,26 27,31,32,33,35,36 37,41	19 Features selected* (5,6,9,10,13,14,15,20 21,23,24,25,32,33,35 35,37,38,40,41)
Rank search	Cfs subset eval	28 Features selected* (1,2,3,4,5,6,7,8,10,11 12,13 18,21,22,24,25 27 28,29,30,33,34,35 36,37,40,41)	28 Features selected* (2,3,4,5,6,7,8,9,10,11 12,13,14,22,23,24,25 26,29,30,32,33,34,35 36,37,38,39)	21 Features selected* (2,3,4,5,6,7,8,10,11 12,13 25,26,27,29 30,35,36,37,38,39)	22 Features selected* (2,3,4,5,6,7,8,10,11, 12,13 25,26,27,29,30 34,35,36,37,38,39)
	Consistency subset eval	35 Features selected* (1,2,3,4,5,6,7,8,10,11 12,13,14,16,18,21,22 24,25,26,27,28,29,30 31,32,33,34,35,36,37 38,39,40,41)	26 Features selected* (2,3,4,5,6,7,8,9,10,11 12,13,14,22,23,25,26 29,30,33,34,35,36,37 38,39)	32 Features selected* (2,3,4,5,6,7,8,10,11 1213 14,18,22,23,25 26,27,28,29,30,31 32 33,34,35,36,37 38 ,39,40,41)	34 Features selected* (1,2,3,4,5,6,7,8,10,11 12,13,14,17,22,23,24 25,26,27,28,29 30,31 32,33 34,35,36,37,38 39,40,41)

Table 6: Continue

Feature selection algorithms		Most important features selected: KDD-CUP 99		NSL-KDD	
Attribute evaluators	Search methods	Test set	10%	Train set	20%
PSOSearch	Cfs subset eval	13 Features selected* (2,3,4,5,6,7,8,14,21 23,29,34,36)	19 Features selected* (2,3,4,5,6,7,8,9,10,12 23,24 ,25,29,30,32,34 35,36)	18 Features selected* (2,4,5,6,7,8,12,13,23 25,26,29,30,34,35,36, 37,38)	16 Features selected* (2,4,5,6,8,11,12,14 23,26,29,30,35,36 ,37,38)
	Consistency subset eval	15 Features selected* (3,5,6,7,14,16,17,22 23,24,33,34,35,38,40)	21 Features selected* (2,4,5,6,9,12,13,1416 17,18,20,23,24,26,33 35,37,38,40,41)	20 Features selected* (2,3,4,5,6,7,11,12,13 14,18,22,23,32,33,35 37,38,39,40)	24 Features selected* (2,3,5,6,7,13,14,20,23 24,32,33,35,37,38,39, 40)
Evolutionary search	Cfs subset eval	18 Features selected* (1,2,3,5,8,12,13,24 27 28,29,31,33,34 35,36 37,39)	17 Features selected* (2,3,5,6,7,8,10,23,24 28,29,33,35,36,37,38 39)	21 Features selected* (2,3,4,5,6,7,8,10,12 21,25,26,29,30,32,33 34,35,37,38,40)	24 Features selected* (2,3,4,5,6,8,12,13,17 19,22,23,24,25,26,29 30,32,33,35,37 38,40 41)
	Consistency subset eval	18 Features selected* (1,3,5,6,8,12,17,19 22,23,24,28,31,33,34 35,38,40)	18 Features selected* (2,3,5,6,7,8,10,23,24 28,29,33,35,36,37, 38,39)	23 Features selected* (1,3,5,6,7,10,12,13,19 22,24,26,27,29,30,31 32,33,35,36,37,38,39)	18 Features selected* (1,3,5,8,12,17,19,22 23,24,26,30,32,33,34 24,26,30,32,33,34,35 37,40)
Multiobjective evolutionary search	Cfs subset eval	15 Features selected* (2,3,5,6,7,8,12,18,21 23,29,30,35,36,40)	12 Features selected* (2,3,5,6,7,8,14,23,29, 36,38,40)	22 Features selected* (2,3,4,5,6,7,8,12,13 21,23,25,26,29,30,31 34,35,36,37,39,41)	16 Features selected* (2,3,4,5,8,10,12,25,29 30,31,33,35,36,37,39)
Greedy stepwise	Consistency subset eval	10 Features selected* (1,3,5,6,23,24,33,34 35,40)	11 Features selected* (1,3,5,6,12,23,33,35 36,37,38)	14 Features selected* (1,3,5,6,12,23,32,33 35,36,37,38,38,40)	11 Features selected* (1,3,5,6,23,24,33,35 36,37,38)

*Return to to see the features name

Table 7: Numbers of features selected by various feature selection algorithms with 2 classes

Feature selection algorithms		Most important features selected: KDD-CUP 99		NSL-KDD	
Attribute evaluators	Search methods	Test set	10%	Train set	20%
Best first	Cfs subset eval	6 Features selected* (5,6,12,23,31,37)	5 Features selected* (6,12,23,31,32)	6 Features selected* (4,5,6,12,26,30)	8 Features selected* (4,5,6,12,26,29,30,37)
	Consistency subset eval	9 Features selected* (1,3,5,6,23,24,34,35 40)	7 Features selected* (1,3,5,23,33,34,35)	10 Features selected* (1,3,5,6,23,32,34,35 37,39	10 Features selected* (1,3,4,5,14,23,32,34 35,37)
Genetic search	Cfs subset eval	9 Features selected* (2,5,6,12,15,23,31 36,37)	12 Features selected* (1,2,6,7,8,12,15 23 31,32,36,37)	15 Features selected* (4,5,6,8,10,12,17,23 26,29,30,32,37,38,39)	15 Features selected* (3,4,5,6,12,16,18,25 26,29,30,31,36,37,38)
	Consistency subset eval	17 Features selected* (1,2,3,5,6,11,21,23 24,29,30,33,34,35,37 40,41)	20 Features selected* (1,2,4,5,6,7,8,13,18 21,24,26,28,29,32 33,35,38,39,41)	22 Features selected* (3,5,7,8,9,10,11,13,22 23,24,27,28,29,31,32 33,34,37,38,39,40)	19 Features selected* (1,3,5,6,7,8,9,13,17,18 21,23,27,29,34,36,37 38,40)
Rank search	Cfs subset eval	6 Features selected* (2,6,12,31,32,37)	6 Features selected* (3,6,12,31,32,37)	12 Features selected* (3,4,5,6,12,25,26,29 30,37,38,39)	12 Features selected* (3,4,5,6,12,25,26,29 30,37,38,39)
	Consistency subset eval	23 Features selected* (1,2,3,4,5,6,8,12,15 17,19,23,24,27,28,29 30,31,32,33,35,36,37)	23 Features selected* (1,2,3,5,6,12,15,16,17 18,19,23,24,26,31,32 33,34,35,36,37,38,39)	30 Features selected* (1,2,3,4,5,6,8,10,12,13 15,16,19,23,25,26,27 28,29,30,31,32,33,34 35,36 37,38,39,41)	25 Features selected* (3,4,5,6,8,12,15,16,23 (3,4,5,6,8,12,15,16,23 25 26,27,28,29,30,31,32 33,34,35,36,37,38 39,41)
PSO search	Cfs subset eval	8 Features selected* (2,6,12,23,29,31 32,37)	5 Features selected* (3,6,12,31,32)	9 Features selected* (4,5,6,12,26,29,30, 37,39)	6 Features selected* (4,5,6,12,29,39)
	Consistency subseteval	20 Features selected* (1,2,3,5,6,7,11,14,15 16,23,24,29,30,33,35 38,39,40,41)	16 Features selected* (1,3,5,6,10,13,14,15 18,20,23,31,33,35, 38,41	19 Features selected* (3,5,11,13,14,16,17,20 21,23,31,32,33,34,35 37,38,39,40)	14 Features selected* (1,3,4,5,6,14,24,25,32 35,36,38,39,41)
Evolutionary search	Cfs subset eval	11 Features selected* (5,6,10,11,12,15,24	18 Features selected* (1,2,3,4,5,6,10,11,12	9 Features selected* (5,6,12,25,30,31,36	18 Features selected* (3,4,5,6,8,17,19,23,25

Table 7: Continue

Feature selection algorithms		Most important features selected			
		KDD-CUP 99		NSL-KDD	
Attribute evaluators	Search methods	Test set	10%	Train set	20%
		31 32,36,37)	16,22 23,25,31,32,33 36,37)	37,39)	26,29,30,33,34,37,38 39,41)
	Consistency subset eval	14 Features selected* (1,3,5,6,11,16,18,20 23,24,34,35,36,40)	16 Features selected* (1,2,3,5,10,16,17,24 27,28,31,33,34,35 37,39)	19 Features selected* (1,2,3,5,12,13,20,23, 24 26,29,31,32,33,34 37,38,39,41)	13 Features selected* (1,2,3,5,6,11,17,20 23,34,35,36,37)
Multiobjective Evolutionary search	Cfs subset eval	12 Features selected* (2,3,5,6,8,12,22,23 31,32,35,37)	8 Features selected* (3,5,6,12,23,31,32,37)	19 Features selected* (4,5,6,7,8,11,12,16,21 25,26,29,30,31,34,37 38,39,41)	9 Features selected* (4,5,6,12,15,30,33 ,37,39)
GreedyStepwise	Consistency subset eval	9 Features selected* (1,3,5,6,23,24,34,35, 40)	8 Features selected* (1,3,5,13,23,33,34 35)	11 Features selected* (1,3,5,6,23,32,33,34 35, 37,39)	10 Features selected* (1,3,4,5,14,23,32,34 35,37)

*Return to to see the features name

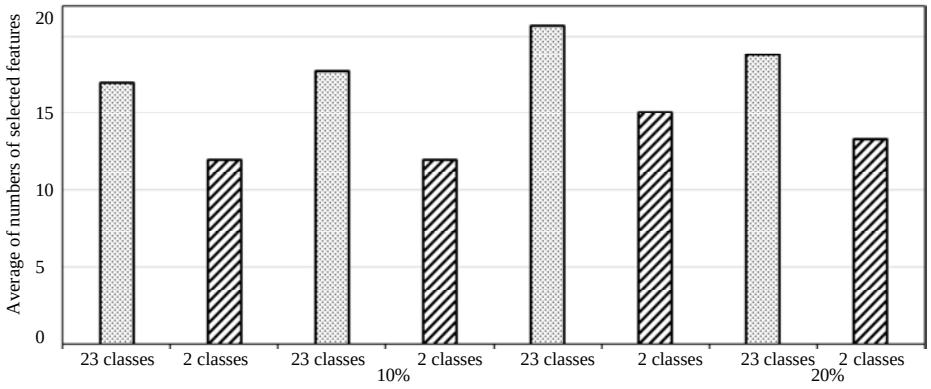

Fig. 1: Comparison between average number of selected features

important features that have been used by the methods we showed previously. At first, we choose 39 different methods of selecting important features from the original 41 features of different DARPA datasets mentioned above. Then we sorted them ascendingly by the smallest number of selection to the highest number as shown in Table 8 below, from this table it is evident that 31 methods have been choosing>21 featuresn with a percentage of 79.5% of the 39 methods which means that the aim of most of the researchers is to get less important features to decrease the time of processing and the utilizing of resources.

We have done a comparison between the features selected by the researcher's methods mentioned previously and the features selected by our experiment, then We suggest ranking these selected features based on the percentage of their selected as shown in Fig. 2. We then take the first 15 selected features of the researcher's methods and do the same of our experiment. These features will be the important features and that is shown in Table 9.

From this table it was obvious that 18 important features have been selected (2, 3, 4, 5, 6, 7, 8, 12, 23, 24, 25, 29, 32, 33, 35, 36, 37, 38). About 12 of the most important features were the same in both the researcher's methods and our experiment and these are (2, 3, 4, 5, 6, 12, 23, 29, 33, 35, 36, 37) and the not repeated were 6 features which are (7, 8, 24, 25, 32, 38). The repeated are shown in bold font and the not repeated are on the grey in Table 10.

These 12 features are considered to be the most important features because we are always looking for fewer features to do the work with good results. In Table 10 shows the features name and the category they belong to and the type by referring to (Stolfo *et al*., 2000). From this table we can see there are five basic features one content feature, two time-based features and four host-Based Features. From Table 10 we count the percentage of the most important selected features classified in their categories and the result is shown in Table 11 which shows that: 5 were most selected of the 9 basic features with a percentage 55.56%;4 were the most

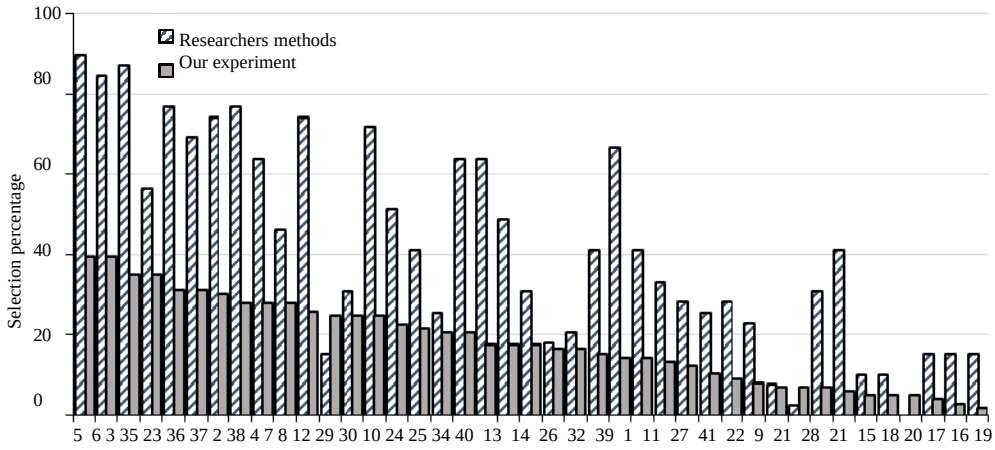

Fig. 2: Descending ranking of the selected features

Table 8: Methods and the number of selected features from DARPA datasets

Reference number	No. of selected features	Percentage selected
Kim	5	12.2
Zainal *et al.* (2006)	6	14.6
Depren *et al.* (2005)	6	14.6
Ghali (2009)	7	17.1
Song *et al.* (2013)	8	19.5
Sindhu *et al.* (2012)	9	22.0
Garg and Kumar (2014a, b)	9	22.0
Wang *et al.* (2015)	9	22.0
Garg and Kumar (2014a, b)	10	24.4
Sindhu *et al.* (2012)	11	26.8
Garg *et al.*(2014)	11	26.8
Chebrolu *et al.* (2005)	12	29.3
Chou *et al.* (2008)	12	29.3
Araújo *et al.* (2010)	14	34.1
Garg and Kumar (2014a, b)	14	34.1
Garg Kumar (2014a, b)	15	36.6
Sindhu *et al.* (2012)	16	39.0
Garg Kumar (2014a)	16	39.0
Garg Kumar (2014b)	16	39.0
Eid *et al.* (2013a, b)	17	41.5
Chebrolu *et al.* (2005)	17	41.5
Garg and Kumar (2014a, b)	17	41.5
Garg Kumar (2014a, b)	17	41.5
Garg and Kumar (2014a, b)	18	43.9
Garg and Kumar (2014a, b)	19	46.3
Li *et al.* (2012)	19	46.3
Sindhu *et al.* (2012)	20	48.8
Sindhu *et al.*(2012)	20	48.8
Sheen and Rajesh (2008)	20	48.8
Garg *et al.* (2014)	20	48.8
Mamory and Jassim (2015)	20	48.8
Garg *et al.* (2014a, b)	21	51.2
Sindhu *et al.* (2012)	22	53.7
Lin *et al.* (2012)	28	68.3
Suebsing	30	73.2
Kayacik and colleagues	31	75.6
Sindhu *et al.* (2012)	34	82.9
Sindhu *et al.* (2012)	36	87.8
Zhang *et al.* (2008)	38	92.7

selected of the 10 host-based features with a percentage 40%; 2 were selected from the 9 time-based features with a percentage 22.22% and 1 was selected from the 13 content-based features with a percentage 7.69%.

Table 9: Important selected features in researchers methods and our experiment

Researcher's methods			The Experiment		
Feature No.	Number times of selections	Selected from 39 methods (%)	Feature No.	Number timeof selections	Selected from 96 tests (%)
5	35	89.74	5	38	39.58
3	34	87.18	6	38	39.58
6	33	84.62	3	34	35.42
2	30	76.92	35	34	35.42
23	30	76.92	23	30	31.25
4	29	74.36	36	30	31.25
37	29	74.36	37	29	30.21
12	28	71.79	2	27	28.13
36	27	69.23	33	27	28.13
32	26	66.67	38	27	28.13
24	25	64.10	4	25	26.04
25	25	64.10	7	24	25.00
33	25	64.10	8	24	25.00
35	22	56.41	12	24	25.00
29	20	51.28	29	22	22.92

Table 10: Important selected features that repeated in researchers methods and our experiment

Feature No.	Feature name	Feature category	Feature type
2	Protocol_type	Basic	Discrete
3	Service	Basic	Discrete
4	Flag	Basic	Discrete
5	Src-bytes	Basic	Continuous
6	Dst-bytes	Basic	Continuous
12	Logged_in	Content	Continuous
23	Count	Time-based	Continuous
29	Same-srv-rate	time-based	Continuous
33	Dst-host-srv-count	Host-based	Continuous
35	Dst-host-diff-srv-rate	Host-based	Continuous
36	Dst-host-same-src-port-rate	Host-based	Continuous
37	Dst-host-srv-diff-host-rate	Host-based	Continuous

Table 11: Most important selectedfeatures classified on their category

Feature category	No. of all features	No. of features selected with highest percentage from table 6	Selected (%)
Basic features	9	5	55.56
Host-based features	10	4	40.00
Time-based features	9	2	22.22
Content-based features	13	1	7.690

Table 12: Number of methods with the number of selected features

Reference No.	No. of selected	FeaturesFeatures name
Peng *et al.* (2004)	1	Sourc-IP Address
Lakhina *et al.*(2005)	4	Sourc-IP, Dest-IP, Sourc-Port, Dest-Port
Chen *et al.* (2007)	5	Sourc-IP, Dest-IP, Sourc-Port, Dest-Port, Protocol Applied
Saad *et al.* (2008)	5	NodeID, ObjectID, Packet flow frequency, SYN packet ratio, ACK packet ratio
Rahmani *et al.* (2009)	2	Number of Received Packets, IP-Flow
Sengar *et al.* (2009)	4	Sourc-IP, Dest-IP, Sourc-Port, Dest-Port
Tang *et al.* (2014)	1	Frequency Distribution of TCP Traffic
Lee and Stolfo (1998)	6	start time, duration, participating hosts, ports,statistics of the connection, flag, protocol
Gavrilis and Dermatas (2005)	3	Sourc-Port, SEQ number, Syn flag
Munz *et al.* (2007)	5	Number of bytes, Number of packets, Number of different source-destination pairs time intervals, service-specific ports
Cheng *et al.* (2008, 2009)	1	IP Address Feature Value
Zhong and Yue (2010)	3	Dest-IP, Dest-Port, Flag
Bhaya *et al.* (2014)	6	Sourc-IP, Dest-IP, Sourc-Port, Dest-Port, Protocol, Time
Yao *et al.* (2012)	5	Sourc-IP, Dest-IP, Sourc-Port, Dest-Port, Time

Some other features were used by other researchers who were using other data sets. Table 12 summaries the 14 different methods usingvarious datasets and the features they selected. Table 13 shows the percentage of the selecting of these features which shows that the 5 tuples of any flow which are source IP, destination IP, sourceport, destinationport, protocol are the most used in these studies.

Table 13: Another features usedby researchers using another datasets

Feature name	No. of used	Used (%)
Sourc-ip address	7	50.00
Sourc-port	7	50.00
Dest-port	7	50.00
Dest-ip address	6	50.00
Protocol applied	3	28.57
Time intervals	3	21.43
Flag	2	14.29
Number of packets	1	7.140
Start time	1	7.140
Duration	1	7.140
Participating hosts	1	7.140
Statistics of connection	1	7.140
Syn flag	1	7.140
Number of bytes	1	7.140
Number of different Source-destination pairs	1	7.140
Service-specific ports	1	7.140
Nodeid	1	7.140
Objectid	1	7.140
Packet flow frequency	1	7.140
Syn packet ratio	1	7.140
Ack packet ratio	1	7.140
Frequency distribution of TCP traffic	1	7.140

CONCLUSION

For all the above we conclude that the researcher's trends go in the direction of the traffic detection based on the basic features, then the features based on the hosts which give the resources of the attacks. This means that the researchers get interested about the hosts that are sending the traffic data which is appropriate because when an attack happens we need to know it's source in order to try to stop or mitigate it. In addition, the content-based features were no use and this is also right because most of the anomaly detection uses edge routers and routers cannot read the content of packets theyjust need the information of the packet's header. The final conclusion is to recommend using the basic, host-based and time-based features for training offline and then to use a random select for testing online. This will reduce the number of features which in turn will reduce the calculation time that we need for testing the traffic flow online.

The future work will be to evaluate our experiment with multi algorithms of classification to focus on improved methods for selecting features and to propose a new method for selecting features in less time and with more accuracy.

REFERENCES

Amrita and P. Ahmed, 2012. A study of feature selection methods in intrusion detection system: A survey. Int. J. Comput. Sci. Eng. Inf. Technol. Res., 2: 1-25.

Araujo, N., D.R. Oliveira, A.A. Shinoda and B. Bhargava, 2010. Identifying important characteristics in the KDD99 intrusion detection dataset by feature selection using a hybrid approach. Proceedings of the IEEE 17th International Conference on Telecommunications (ICT), April 4-7, 2010, IEEE, Cuiaba, Brazil, ISBN:978-1-4244-5246-0, pp: 552-558.

Bhaya, W. and M.E. Manaa, 2014. A proactive DDoS attack detection approach using data mining cluster analysis. J. Next Gener. Inf. Technol., 5: 36-47.

Bhuyan, M.H., D.K. Bhattacharyya and J.K. Kalita, 2014. Network anomaly detection: methods, systems and tools. IEEE. Commun. Surv. Tutorials, 16: 303-336.

Bjerkestrand, T., D. Tsaptsinos and E. Pfluegel, 2015. An evaluation of feature selection and reduction algorithms for network IDS data. Proceedings of the International Conference on Cyber Situational Awareness Data Analytics and Assessment (CyberSA), June 8-9, 2015, IEEE, London, UK., ISBN:978-0-9932-3380-7, pp: 1-2.

Chebrolu, S., A. Abraham and J.P. Thomas, 2005. Feature deduction and ensemble design of intrusion detection systems. Comput. Secur., 24: 295-307.

Chen, Y., K. Hwang and W.S. Ku, 2007. Collaborative detection of DDoS attacks over multiple network domains. IEEE Trans. Parallel Distrib. Syst., 18: 1649-1662.

Chen, Y., Y. Li, X.Q. Cheng and L. Guo, 2006. Survey and taxonomy of feature selection algorithms in intrusion detection system. Proceedings of the International Conference on Information Security and Cryptology, November 29- December 1, 2006, Springer, Berlin, Germany, ISBN:978-3-540-49608-3, pp: 153-167.

Cheng, J., J. Yin, Y. Liu, Z. Cai and M. Li, 2009. DDoS attack detection algorithm using IP address features. Proceedings of the International Workshop on Frontiers in Algorithmics, June 20-23, 2009, Springer, Berlin, Germany, ISBN:978-3-642-02269-2, pp: 207-215.

Chou, T.S., K.K. Yen and J. Lou, 2008. Network intrusion detection design using feature selection of soft computing paradigms. Int. J. Comput. Intell., 4: 196-200.

Depren, O., M. Topallar, E. Anarim and M.K. Ciliz, 2005. An intelligent Intrusion Detection System (IDS) for anomaly and misuse detection in computer networks. Expert Syst. Appl., 29: 713-722.

Eid, H.F., A.E. Hassanien, T.H. Kim and S. Banerjee, 2013. Linear Correlation-Based Feature Selection for Network Intrusion Detection Model. In: Advances in Security of Information and Communication Networks, Ismail, A.A., E.H. Aboul and B. Kensuke (Eds.). Springer, Berlin, Germany, ISBN:978-3-642-40596-9, pp: 240-248.

Garg, T. and S.S. Khurana, 2014a. Comparison of classification techniques for intrusion detection dataset using WEKA. Proceedings of the IEEE Conference on Recent Advances and Innovations in Engineering (ICRAIE), May 9-11, 2014, IEEE, Bathinda, India, ISBN:978-1-4799-4040-0, pp: 1-5.

Garg, T. and Y. Kumar, 2014b. Combinational feature selection approach for network intrusion detection system. Proceedings of the 2014 International Conference on Parallel Distributed and Grid Computing (PDGC), December 11-13, 2014, IEEE, Kapurthala, India, ISBN:978-1-4799-7682-9, pp: 82-87.

Gavrilis, D. and E. Dermatas, 2005. Real-time detection of distributed denial-of-service attacks using RBF networks and statistical features. Comput. Networks, 48: 235-245.

Ghali, N.I., 2009. Feature selection for effective anomaly-based intrusion detection. Int. J. Comput. Sci. Network Security, 9: 285-289.

Kumar, M.S., 2016. A survey on improving classification performance using data pre processing and machine learning methods on NSL-KDD data. Int. J. Eng. Comput. Sci., 5: 16156-16161.

Lakhina, A., M. Crovella and C. Diot, 2005. Mining anomalies using traffic feature distributions. ACM Sigcomm Comput. Commun. Rev., 35: 217-228.

Lee, W. and S. Stolfo, 1998. Data mining approaches for intrusion detection. Proceedings of the 7th USENIX Security Symposium, January 26-29, 1998, USENIX Association, Berkeley, CA., USA., pp: 79-94.

Li, Y., J. Xia, S. Zhang, J. Yan, X. Ai and K. Dai, 2012. An efficient intrusion detection system based on support vector machines and gradually feature removal method. Expert Syst. Appl., 39: 424-430.

Lin, S.W., K.C. Ying, C.Y. Lee and Z.J. Lee, 2012. An intelligent algorithm with feature selection and decision rules applied to anomaly intrusion detection. Applied Soft Comput. J., 12: 3285-3290.

Mamory, S.O.A. and F.S. Jassim, 2015. On the designing of two grains levels network intrusion detection system. Karbala Int. J. Mod. Sci., 1: 15-25.

Moustafa, N. and J. Slay, 2015. Creating novel features to anomaly network detection using DARPA-2009 data set. Proceedings of the 14th European Conference on Cyber Warfare and Security, July 2-3, 2015, ACPI Publisher, Hatfield, England, UK., ISBN:978-1-910810-28-6, pp: 204-207.

Oyebode, E.O., S.G. Fashoto, O.A. Ojesanmi, O.E. Makinde and O. State, 2011. Intrusion detection system for computer network security 1. Aust. J. Basic Appl. Sci., 5: 1317-1320.

Patcha, A. and J.M. Park, 2007. An overview of anomaly detection techniques: Existing solutions and latest technological trends. Comput. Networks, 51: 3448-3470.

Peng, T., C. Leckie and K. Ramamohanarao, 2004. Proactively detecting distributed denial of service attacks using source IP address monitoring. Proceedings of the 3rd International IFIP-TC6 Networking Conference on Networking Technologies, Services and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications, May 9-14, 2004, Athens, Greece, pp: 771-782.

Rahmani, H., N. Sahli and F. Kammoun, 2009. Joint entropy analysis model for DDoS attack detection. Proceedings of the 5th International Conference on Information Assurance and Security, Volume 2, August 18-20, 2009, Xian, China, pp: 267-271.

Ranjbar, L. and S. Khorsandi, 2011. A collaborative intrusion detection system against ddos attack in peer to peer network. Proceedings of the International Conference on Software Engineering and Computer Systems, June 27-29, 2011, Springer, Berlin, Germany, ISBN:978-3-642-22202-3, pp: 353-367.

Saad, R., N.F. Abdesselam and A. Serhrouchni, 2008. A collaborative peer-to-peer architecture to defend against DDoS attacks. Proceedings of the 2008 33rd IEEE Conference on Local Computer Networks (LCN), October 14-17, 2008, IEEE, Lille, France, ISBN:978-1-4244-2412-2, pp: 427-434.

Sengar, H., X. Wang, H. Wang, D. Wijesekera and S. Jajodia, 2009. Online detection of network traffic anomalies using behavioral distance. Proceedings of the 17th International Workshop on Quality of Service (IWQoS), July 13-15, 2009, IEEE, Fairfax, Virginia, ISBN: 978-1-4244-3875-4, pp: 1-9.

Sheen, S. and R. Rajesh, 2008. Network intrusion detection using feature selection and decision tree classifier. Proceedings of the IEEE Region 10 Conference TENCON, November 19-21, 2008, Hyderabad, pp: 1-4.

Sindhu, S.S.S., S. Geetha and A. Kannan, 2012. Decision tree based light weight intrusion detection using a wrapper approach. Expert Syst. Applic., 39: 129-141.

Singh, H. and D. Kumar, 2015. A study on performance analysis of various feature selection techniques in intrusion detection system. Int. J., 3: 50-54.

Song, J., Z. Zhu, P. Scully and C. Price, 2013. Selecting Features for Anomaly Intrusion Detection: A Novel Method using Fuzzy C Means and Decision Tree Classification. In: Cyberspace Safety and Security, Guojun, W. I. Ray, D. Feng and M. Rajarajan (Eds.). Springer, Berlin, Germany, ISBN:978-3-319-03583-3, pp: 299-307.

Srihari, V. and R. Anitha, 2014. DDoS detection system using wavelet features and semi-supervised learning. Proceedings of the International Symposium on Security in Computing and Communication, September 24-27, 2014, Springer, Berlin, Germany, ISBN:978-3-662-44965-3, pp: 291-303.

Tang, D., K. Chen, X. Chen, H. Liu and X. Li, 2014. A new collaborative detection method for LDoS attacks. J. Networks, 9: 2674-2681.

Wang, W., Y. He, J. Liu and S. Gombault, 2015. Constructing important features from massive network traffic for lightweight intrusion detection. IET. Inf. Secur., 9: 374-379.

Yao, D., M. Yin, J. Luo and S. Zhang, 2012. Network anomaly detection using random forests and entropy of traffic features. Proceedings of the 2012 Fourth International Conference on Multimedia Information Networking and Security, November 2-4, 2012, IEEE, Zhengzhou, China, ISBN:978-1-4673-3093-0, pp: 926-929.

Zaina, A., M.A. Maarof and S.M. Shamsuddin, 2006. Feature selection using rough set in intrusion detection. Proceedings of the TENCON 2006 IEEE Region of 10 Conference, November 14-17, 2006, Teknologi Malaysia, Johor, pp: 1-4.

Zhang, J., M. Zulkernine and A. Haque, 2008. Random-forests-based network intrusion detection systems. IEEE. Trans. Syst. Man Cybernetics Part C Appl. Rev., 38: 649-659.

Zhong, R. and G. Yue, 2010. DDoS detection system based on data mining. Proceedings of the 2nd International Symposium on Networking and Network Security, April 2-4, 2010, Jinggangshan, China, pp: 62-65.