ICAESMT19

Evaluation of Network Intrusion Detection with Features Selection and Machine Learning Algorithms on CICIDS-2017 Dataset

Mr. Shailesh Singh Panwar[a], Dr. Y. P. Raiwani[b], Mr. Lokesh Singh Panwar[c]

[a,b,c]*H.N.B. Garhwal University Srinagar Garhwal, , 246174, Uttarakhand, India*

A R T I C L E I N F O

Article history:
Received 00 December 00
Received in revised form 00 January 00
Accepted 00 February 00

Keywords:
OneR
REPTree
WEKA
CICIDS-2017 Data Set
Feature Selection Algorithms

A B S T R A C T

In the era of network Security, the Intrusion Detection System (IDS) plays an important role in information security. As the usability of the internet among the users in a wide area is increasing day by day so as the importance of security and to keep the system aware of the malicious activities is also increasing. In this paper we have decided to choose four feature selection algorithms i.e. CfsSubset Attribute Evaluator, Classifier Subset Evaluator with Naive Bayes, Classifier Subset Evaluator with J48 and Classifier Subset Evaluator with Decision Tree and the two different machine learning algorithms, namely OneR and REPTree. All these algorithms have been implemented in WEKA machine learning tool to evaluate performance. For experimental work, CICIDS-2017 dataset is used. First we select the features by feature selection algorithms after this each classification algorithm is tested with conducted dataset and finally results have been compared**.**

Introduction

Due to the popularity of Internet and local networks, incidents of intrusion in computer systems are increasing. The rapid spread of computer networks has changed the possibilities of network security. It created the need for a system that could detect not only threats to the network rather rely on intrusion prevention systems. Detecting such hazards not only provides information on the assessment of damage but also helps prevent future attacks. These attacks are commonly detected with Intrusion Detection System.

Researchers have developed intrusion detection systems for different environments based on the security concerns of different networks. The functions of the Intrusion Detection System are gathered to analyze all the possible security violations and to analyze information from different areas within a computer or a network. In the past ten years, intrusion detection and other security technologies such as cryptography, authentication, and firewalls have received its importance rapidly.

Machine learning is a key enabler of Artificial Intelligence. It is about making computers to act without explicitly programming them. The Machine learning out of them can figure out how to perform tasks based on generalizing form data or examples and they can learn to improve themselves from the past data. ML technology has the ability to detect unknown attacks in network traffic sharing facilities with other attacks trained in the general and unusual type of traffic. However, one important problem in ML is to identify and select the most relevant input characteristics, from which to build a specific model based on training data for a particular classification job.

Section II of this paper presents some related work on the basis of intrusion detection, section III gives a brief description with the contents of the CICIDS-2017 dataset, section IV presents an overview of both the feature selection algorithm and the classifier, section V presents the graphical and tabulation analysis report using different classification methods with feature selection algorithm and section VI, relates to the conclusion and future scope.

Nomenclature

LDALinear Discriminate Evaluation
PCA Principal Component Analysis
CICIDS Canadian Institute for Cyber security Intrusion Detection System
IDSsIntrusion Detection Systems and IPSsIntrusion Prevention Systems
CFS Correlation Feature Selection
REPTree Reduce Error Pruning Tree

1. Related Work

Yuxun et al. resolved the problem of decision tree algorithms primarily on the basis of characteristic importance and improved ID3 algorithm properties to gain insight, which have less attributes. After that compared ID3 algorithm with improve ID3 algorithm. An experimental assessment of the facts indicates that the better ID3 sets of algorithms can obtain more realistic and powerful methods.

Tavallaee et al. carried out a statistical analysis on KDDCUP'99 data set and found some important issues that affect the better performance of evaluated system and the outcome of the anomaly detection approaches was very bad. In order to solve these anomalies, they have proposed a new data set, NSL-

KDD **.Olusola et al.** presented the relevance of each feature to detect each class in a dataset that identified the KDD 99 intrusion. To determine the most discriminating characteristics for each class, a certain degree of dependency and dependency ratio of each class was employed.

Weiguo et al. proposed a new customized set of rules for the decision tree. On the idea of ID3, set of rules taken into consideration of the special option in the categories of decision trees and the classification accuracy of the developed set of rules has proved to be better than the ID3.

Ibrahim et al. proposed that accuracy of classification has improved and minimize the high false alarm rate on KDD99 or others. For this some class algorithms have been used such as Linear Discriminate Evaluation (LDA) and Principal Component Analysis (PCA), which minimize the intrusion and anomalies. The experiments of the IDS finished with NSL-KDD information set and tried to improve the methods of classification.

Hora et al. presented a model for the disease prognosis as they brought about proper accuracy and had been used to make predictions using several classification algorithms like J48, Random Forest, etc. Dynamic interfaces can also use built-in models, which suggest that utility works well in each case.

2. CICIDS-2017 Dataset

Intrusion Detection Systems (IDSs) and Intrusion Prevention Systems (IPSs) are the most powerful defense tools against sophisticated and ever-growing network attacks. Due to the lack of reliable test and validation datasets, anomaly-based intrusion detection approaches suffer from the consistent and accurate performance development. The attacks included Brute Force attack, Heartbleed/ Denial-of-service (DoS), Web Attack, Infiltration, Botnet, Port Scan and Distributed Denial-of-service (DDoS). They have been executed in morning and afternoon on Tuesday, Wednesday, Thursday and Friday.

Table 1 - The 79 features of IDS dataset record

Feature no.	Feature label	Feature no.	Feature label	Feature no.	Feature label
1.	Destination Port	28.	Bwd IAT Std	54.	AvgFwd Segment Size
2.	Flow Duration	29.	Bwd IAT Max	55.	AvgBwd Segment Size
3.	Total Fwd Packets	30.	Bwd IAT Min	56.	Fwd Header Length
4.	Total Backward Packets	31.	Fwd PSH Flags	57.	FwdAvg Bytes/Bulk
5.	Total Length of Fwd Packets	32.	Bwd PSH Flags	58.	FwdAvg Packets/Bulk
6.	Total Length of Bwd Packets	33.	Fwd URG Flags	59.	FwdAvg Bulk Rate
7.	Fwd Packet Length Max	34.	Bwd URG Flags	60.	BwdAvg Bytes/Bulk
8.	Fwd Packet Length Min	35.	Fwd Header Len	61.	BwdAvg Packets/Bulk
9.	Fwd Packet Length Mean	35.	Bwd Header Length	62.	BwdAvg Bulk Rate
10.	Fwd Packet Length Std	37.	Fwd Packets/s	63.	SubflowFwd Packets
11.	Bwd Packet Length Max	38.	Bwd Packets/s	64.	SubflowFwd Bytes
12.	Bwd Packet Length Min	39.	Min Packet Length	65.	SubflowBwd Packets
13.	Bwd Packet Length Mean	40.	Max Packet Length	66.	SubflowBwd Bytes
14.	Bwd Packet Length Std	41.	Packet Length Mean	67.	Init_Win_bytes_forward
15.	Flow Bytes/s	42.	Packet Length Std	68.	Init_Win_bytes_backward
16.	Flow Packets/s	43.	Packet Length Variance	69.	act_data_pkt_fwd
17.	Flow IAT Mean	44.	FIN Flag Count	70.	min_seg_size_forward
18.	Flow IAT Std	45.	SYN Flag Count	71.	Active Mean
19.	Flow IAT Max	46.	RST Flag Count	72.	Active Std
20.	Flow IAT Min	47.	PSH Flag Count	73.	Active Max
21.	Fwd IAT Total	48.	ACK Flag Count	74.	Active Min
22.	Fwd IAT Mean	49.	URG Flag Count	75.	Idle Mean
23.	Fwd IAT Std	50.	CWE Flag Count	76.	Idle Std
24.	Fwd IAT Max	51.	ECE Flag Count	77.	Idle Max
25.	Fwd IAT Min	52.	Down/Up Ratio	78.	Idle Min
26.	Bwd IAT Total	53.	Average Packet Size	79.	Label
27.	Bwd IAT Mean				

3. Feature Selection Algorithms and Classifiers

3.1 Correlation Feature Selection (CFS)

Feature selection is a process of choosing a subset of the relevant attribute selected in a large number of basic attribute of a particular dataset by applying unique assessment standards to enhance the pleasant of classifier, while the dimension of the data reduces. It is used to evaluate the subset of features based on the well-suited subsets, which have highly correlated facilities with classification, are still unrelated to each other.

3.2. Classifier Subset Evaluator

- It evaluates the specialty subset on training data or a separate hold-out test set.
- It uses a classification to evaluate the eligibility of a set of features.
- With whom the classification algorithms perform well, it considers subsets of those tasks.

3.3. OneR (One Rule)

OneR is a powerful classifier that produces a one-tier decision tree. It is generally capable of guessing by simple, yet accurate; an example of classification rules. It is capable of handling missing values and numerical characteristics, which is showing optimization capability despite simplicity. OneR algorithm creates a rule for each feature in training data and after that selects the rule with the smallest error rate as "a rule". To create a rule for a feature, the most persistent class for every feature value should be evaluated.

3.4. Reduced Error Pruning Tree (REPTree)

REPTree is considered to be a quick decision-making tree, which uses the benefit of information as the criterion of division to make the decision / regression tree, and prunes it using a low error pruning method. In cases of large volumes of training and test data, the result of reducing the error is a more accurate and simple classification tree.

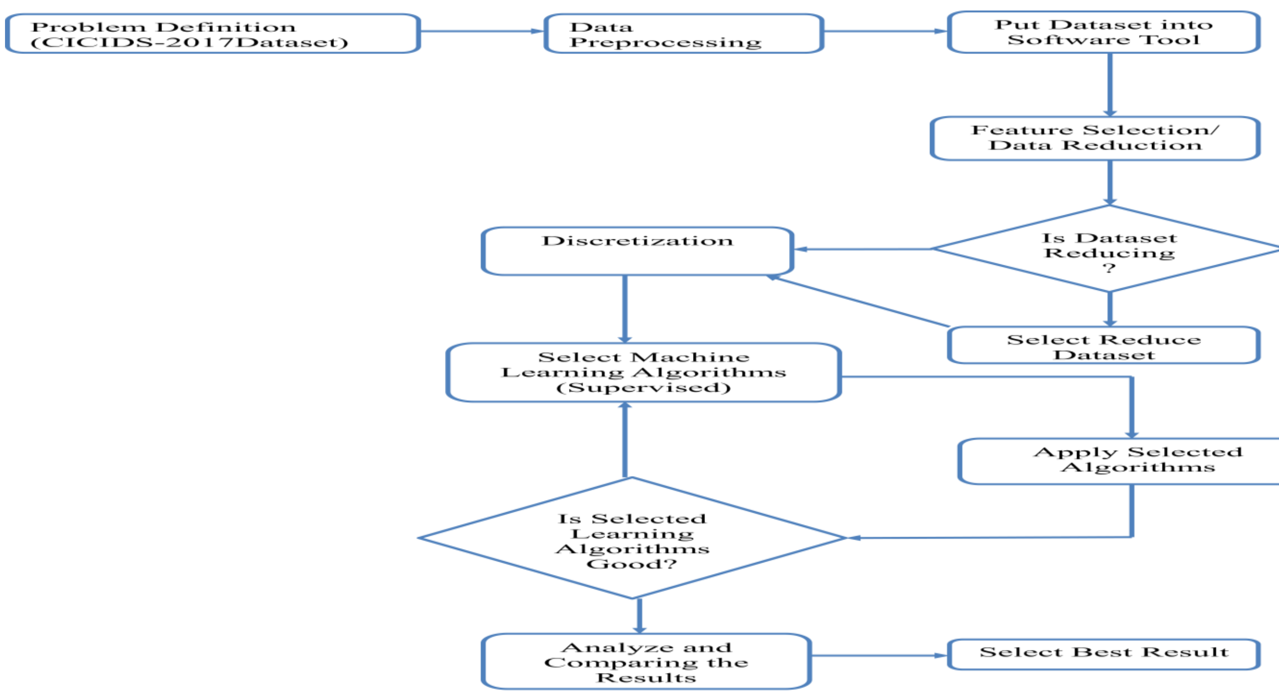

Fig. 1 - Proposed Approach

4.Experiment and Results

To assess the performance of our approach, a sequence of experiments has been performed.

4.1. WEKA Tool

In this paper we have used the WEKA Software tool to investigate and analyze the CICIDS-2017 dataset with the two different machine learning algorithms. WEKA is an open source GUI application which is referred to the Waikato Environment for Knowledge Learning. The University of Waikato in New Zealand developed the WEKA software tool, which identify the data from the lager amount of records that have been collected from the different domain. It helps on several data mining and machine learning applications along with preprocessing, clustering classification, regression, feature selection and visualization.

The essential premise of WEKA software is to use computer software that can be trained machine learning capabilities and useful data can be obtained inside in the form of tendencies and styles.It works on the prediction that the information is available as a document or relationship.For this reason, each data object is described by a variety of characteristics that are usually a special type such as normal alpha-numericor numeric value. WEKA software gives file system informationto novice users with information hidden from the database and easy to implement an alternative system and visual interfaces.

Fig.-2 (Screen Shot) shows the steps of selecting discretization from preprocessing tab.

Table 2-Reduced selected features and number of selected features after applying feature selection algorithms

Attacks	Algorithm	Selected features	No. of selected features
	CfsSubset Attribute Evaluator	1,10,70	3
Brute Force Attack	Classifier Subset Evaluator With Naive Bayes	1,10,18,38,49,67	6
(FTP/ SSH Patator)	Classifier Subset Evaluator With J48	1,35,38,49,67	5
	Classifier Subset Evaluator With Decision Tree	68	1
	CfsSubset Attribute Evaluator	1,6,67,77	4
	Classifier Subset Evaluator With Naive Bayes	1,2,6,14,24,35,41,67,68,70,74	11
	Classifier Subset Evaluator With J48	1,3,6,7,20,24,40,67	8
DoS/ Heartbleed Attack	Classifier Subset Evaluator With Decision Tree	1	1
	CfsSubset Attribute Evaluator	9,25,68	3
	Classifier Subset Evaluator With Naive Bayes	25,67,68	3
	Classifier Subset Evaluator With J48	1,2,4,16,21,25,67,68	8
Web Attack	Classifier Subset Evaluator With Decision Tree	68	1
	CfsSubset Attribute Evaluator	5,72,74,76	4
	Classifier Subset Evaluator With Naive Bayes	1,13,25,67,68,70	6
	Classifier Subset Evaluator With J48	1,8,68	3
Infiltration Attack	Classifier Subset Evaluator With Decision Tree	68	1
	CfsSubset Attribute Evaluator	1, 13, 14, 70	4
	Classifier Subset Evaluator With Naive Bayes	1, 6, 31, 35, 44, 67	6
	Classifier Subset Evaluator With J48	1, 53, 67, 68	4
Botnet Attack	Classifier Subset Evaluator With Decision Tree	1	1
	CfsSubset Attribute Evaluator	13,47,68, 69, 70	5
	Classifier Subset Evaluator With Naive Bayes	5,7,8,10,23,41,47,67,68	9
	Classifier Subset Evaluator With J48	25,35,38,41,68,71	6
Port Scan Attack	Classifier Subset Evaluator With Decision Tree	68	1
	CfsSubset Attribute Evaluator	1, 7, 46, 72	4
	Classifier Subset Evaluator With Naive Bayes	5, 8, 26, 67, 68, 75, 77	7
	Classifier Subset Evaluator With J48	1, 5, 6, 8, 48, 67, 68, 78	8
DDoS Attack	Classifier Subset Evaluator With Decision Tree	1	1

A. Performance Measures

All classifiers are performed on the basis of accuracy, sensitivity, specificity and time. The performance was calculated by True Positive (TP), False Positive (FP), False Negative (FN) and True Negative (TN).All above values are derived from the confusion matrices.

Accuracy gives the possibility that the algorithm can accurately predict positive and negative instances and is calculated:

$$Accuracy= (TP + TN)/ (TP+TN+FP+FN)$$

There is a possibility of sensitivity that the algorithm can accurately predict positive instances and is calculated:

$$Sensitivity= TP/ (TP+FN)$$

There is a possibility of specification that algorithms can accurately predict negative instances and are calculated

$$Specificity= TN/ (TN+FP)$$

WEKA software tool applies in datasets and find out the accuracies by OneR and the REPTree machine learning algorithms with supervised discretization. The performances obtained by this are shown in Fig. 3,4,5,6, 7, 8 and 9.

The Table2 shows reduced selected features and number of selected features after applying feature selection algorithms on the complete data set. Each algorithm does not have a different number of properties based on their evaluation criteria.

Now the main task is to find out which classification algorithms give better results for feature selection on CICIDS-2017 dataset. For this purpose, we have implemented two classification algorithms OneR and REPTree. The feature selection approach of these two algorithms gives good results for performance evaluation.

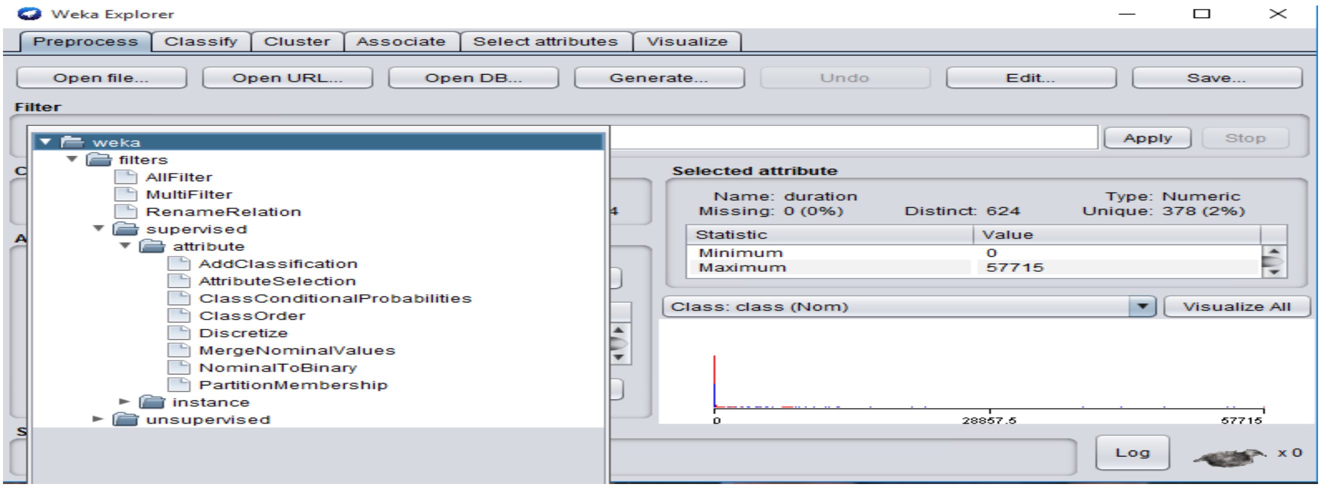

Fig. 2-Selecting Discretization from Preprocessing Tab

Table 3- Performance evaluation for *Brute Force* Attack using classifier with selected features

	Features Selection Algorithm	**Time**	**Accuracy**	**Sensitivity**	**Specificity**
	CfsSubset Attribute Evaluator	0.28	99.2833	99.2605	99.9927
	Classifier Subset Evaluator With Naive Bayes	0.31	99.2833	99.2605	99.9927
OneR	CfsSubset Attribute Evaluator with J48	0.28	99.2833	99.2605	99.9927
	Classifier Subset Evaluator With Decision Tree	0.17	98.867	99.8264	69.0422
	CfsSubset Attribute Evaluator	6.73	99.8695	99.8752	99.6891
	Classifier Subset Evaluator With Naive Bayes	3.77	99.9888	99.9928	99.8626
REPTree	CfsSubset Attribute Evaluator with J48	5.47	99.989	99.9928	99.8698
	Classifier Subset Evaluator With Decision Tree	4.17	98.867	99.8264	69.0422

Table 4- Performance evaluation for *Heartbleed Attack/ DoS Attack* using classifier with selected features

Classifier	**Features selection algorithm**	**Time**	**Accuracy**	**Sensitivity**	**Specificity**
	CfsSubset Attribute Evaluator	0.36	93.7917	97.4647	87.8989
	Classifier Subset Evaluator With Naive Bayes	0.53	93.7917	97.4647	87.8989
OneR	CfsSubset Attribute Evaluator with J48	0.39	93.7917	97.4647	87.8989
	Classifier Subset Evaluator With Decision Tree	0.28	89.8427	88.9187	99.9956
	CfsSubset Attribute Evaluator	3.81	99.4397	99.7988	99.1743
	Classifier Subset Evaluator With Naive Bayes	28.68	99.8284	99.8815	99.8428
REPTree	CfsSubset Attribute Evaluator with J48	12.78	99.8444	99.8884	99.8650
	Classifier Subset Evaluator With Decision Tree	1.41	89.8427	88.9187	99.9956

Table5- Performance evaluation for *Web Attack* using classifier with selected features

Classifier	**Features selection algorithm**	**Time**	**Accuracy**	**Sensitivity**	**Specificity**
	CfsSubset Attribute Evaluator	0.22	99.2469	94.0366	94.0366
	Classifier Subset Evaluator With Naive Bayes	0.20	99.2469	99.6801	94.0366
OneR	CfsSubset Attribute Evaluator with J48	0.22	99.2469	99.6801	94.0366
	Classifier Subset Evaluator With Decision Tree	0.17	99.2469	99.6801	94.0366
	CfsSubset Attribute Evaluator	4.7	99.5791	99.9916	95.0458
	Classifier Subset Evaluator With Naive Bayes	1.7	99.5973	99.2833	97.1559
REPTree	CfsSubset Attribute Evaluator with J48	7.17	99.6173	99.9916	98.2568
	Classifier Subset Evaluator With Decision Tree	0.73	99.2469	99.6801	94.0366

Table 6- Performance evaluation for *Infiltration Attack* using classifier with selected features

Classifier	Features selection algorithm	Time	Accuracy	Sensitivity	Specificity
	CfsSubset Attribute Evaluator	0.23	99.9858	99.9975	5.5555
	Classifier Subset Evaluator With Naive Bayes	0.23	99.9938	99.9982	63.8888
OneR	CfsSubset Attribute Evaluator with J48	0.33	99.9938	99.9982	63.8888
	Classifier Subset Evaluator With Decision Tree	0.17	99.9938	99.9982	63.8888
	CfsSubset Attribute Evaluator	1.08	99.9906	99.9975	44.4444
	Classifier Subset Evaluator With Naive Bayes	1.31	99.9986	99.9989	97.2222
REPTree	CfsSubset Attribute Evaluator with J48	0.91	99.9972	99.9993	83.3333
	Classifier Subset Evaluator With Decision Tree	0.44	99.9938	99.9982	63.8888

Table 7- Performance evaluation for *Botnet* Attack using classifier with selected features

Classifier	Features selection algorithm	Time	Accuracy	Sensitivity	Specificity
	CfsSubset Attribute Evaluator	0.20	99.6304	99.9851	65.5137
	Classifier Subset Evaluator With Naive Bayes	0.23	99.6304	99.9851	65.5137
OneR	CfsSubset Attribute Evaluator with J48	0.19	99.6304	99.9851	65.5137
	Classifier Subset Evaluator With Decision Tree	0.14	99.6304	99.9851	65.5137
	CfsSubset Attribute Evaluator	0.95	99.6393	99.9846	66.4292
	Classifier Subset Evaluator With Naive Bayes	0.98	99.9497	99.9772	97.3041
REPTree	CfsSubset Attribute Evaluator with J48	1.73	99.9686	99.9878	98.1180
	Classifier Subset Evaluator With Decision Tree	0.34	99.6304	99.9851	65.5137

Table 8- Performance evaluation for *PortScan* Attackusing classifier with selected features

Classifier	Features selection algorithm	Time	Accuracy	Sensitivity	Specificity
	CfsSubset Attribute Evaluator	0.23	98.6148	97.8633	99.2178
	Classifier Subset Evaluator With Naive Bayes	0.22	99.5671	99.6495	99.5010
OneR	CfsSubset Attribute Evaluator with J48	0.27	99.5671	99.6495	99.5010
	Classifier Subset Evaluator With Decision Tree	0.16	98.4243	97.0557	99.5224
	CfsSubset Attribute Evaluator	2.5	99.8377	99.6942	99.9528
	Classifier Subset Evaluator With Naive Bayes	4.44	99.9592	99.9310	99.9817
REPTree	CfsSubset Attribute Evaluator with J48	4.44	99.9815	99.9772	99.9848
	Classifier Subset Evaluator With Decision Tree	0.84	98.4243	97.0557	99.5224

Table 9- Performance evaluation for *DDoS Attack*using classifier with selected features

Classifier	Features selection algorithm	Time	Accuracy	Sensitivity	Specificity
	CfsSubset Attribute Evaluator	0.20	96.0442	90.8645	99.9976
	Classifier Subset Evaluator With Naive Bayes	0.22	98.9643	97.9328	99.7516
OneR	CfsSubset Attribute Evaluator with J48	0.27	98.9643	97.9328	99.7516
	Classifier Subset Evaluator With Decision Tree	0.16	96.0442	90.8645	99.9976
	CfsSubset Attribute Evaluator	1.27	98.9125	97.4907	99.9976
	Classifier Subset Evaluator With Naive Bayes	2.5	99.9641	99.9723	99.5782
REPTree	CfsSubset Attribute Evaluator with J48	2.38	99.9805	99.9754	99.9843
	Classifier Subset Evaluator With Decision Tree	0.52	96.0442	90.8645	99.9976

Tables 3,4,5,6,7,8 and 9 shows comparative study using two selected classifiers with features chosen by CSF, Classifier Subset Evaluator with Naive Bayes, Classifier Subset Evaluator with J48 Algorithms and Classifier Subset Evaluator with Decision Tree.

- The REPTree classification algorithm with CfsSubset Attribute Evaluator with J48 features selection technique provides best performance for Brute Force Attack, Heartbleed Attack/ DoS Attack, Web Attack, Botnet Attack, Port Scan Attack and DDoS Attack.
- The REPTree classification algorithm with Classifier Subset Evaluator with Naive Bayes features selection technique provides the best performance for Infiltration Attack.

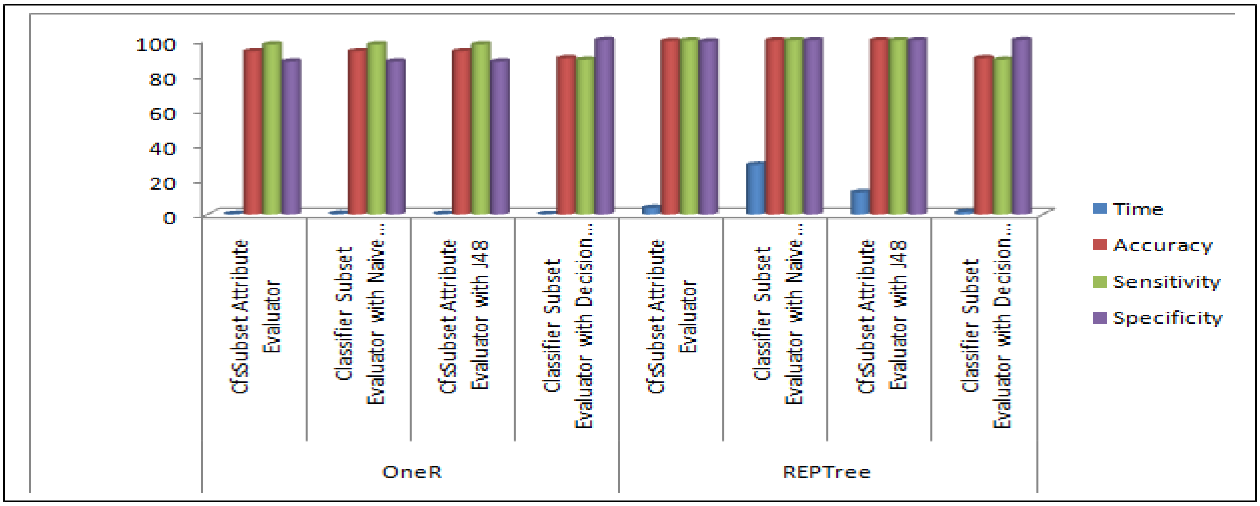

Fig 3- Performance analysis for Brute Force Attack

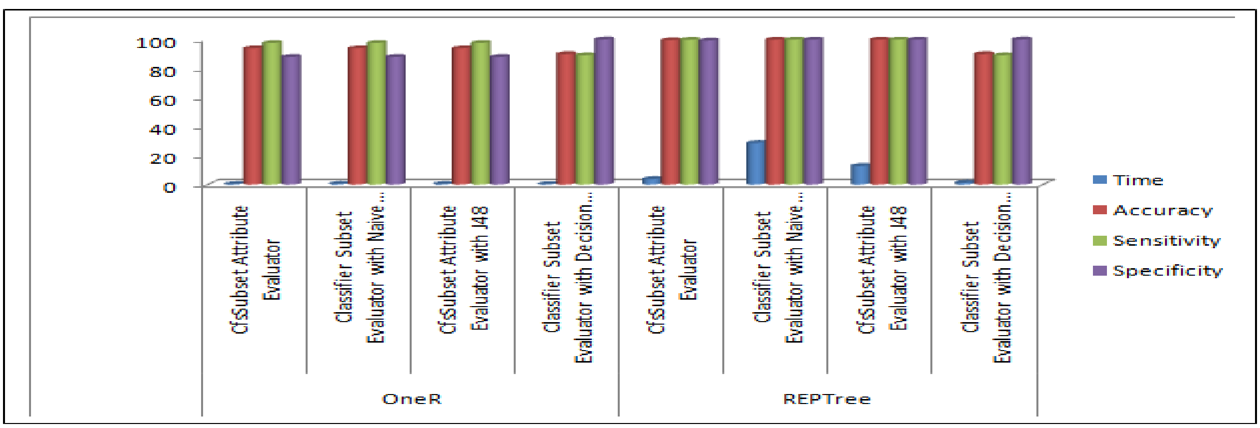

Fig 4- Performance analysis for Heartbleed Attack/ DoS Attack

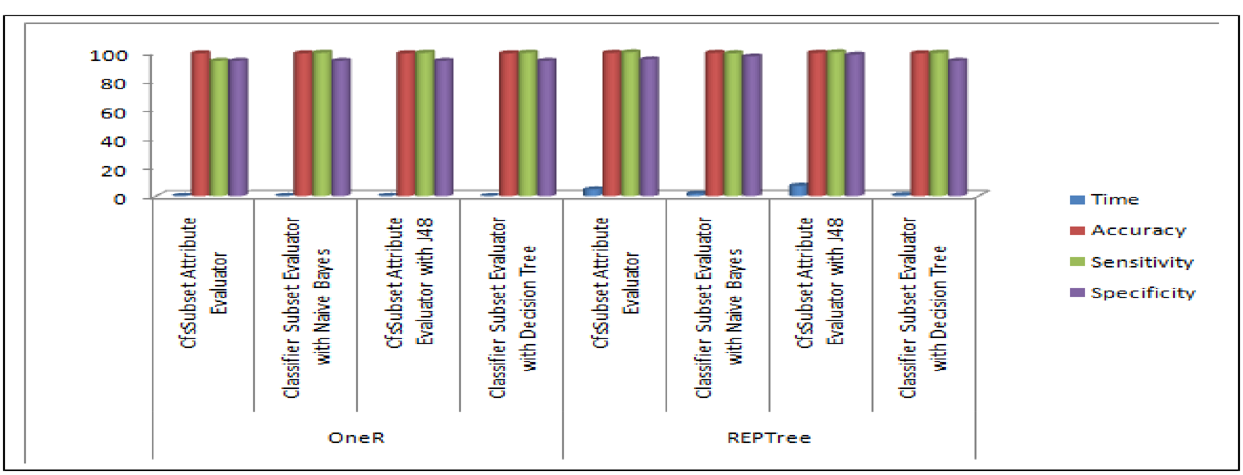

Fig 5- Performance analysis for Web Attack

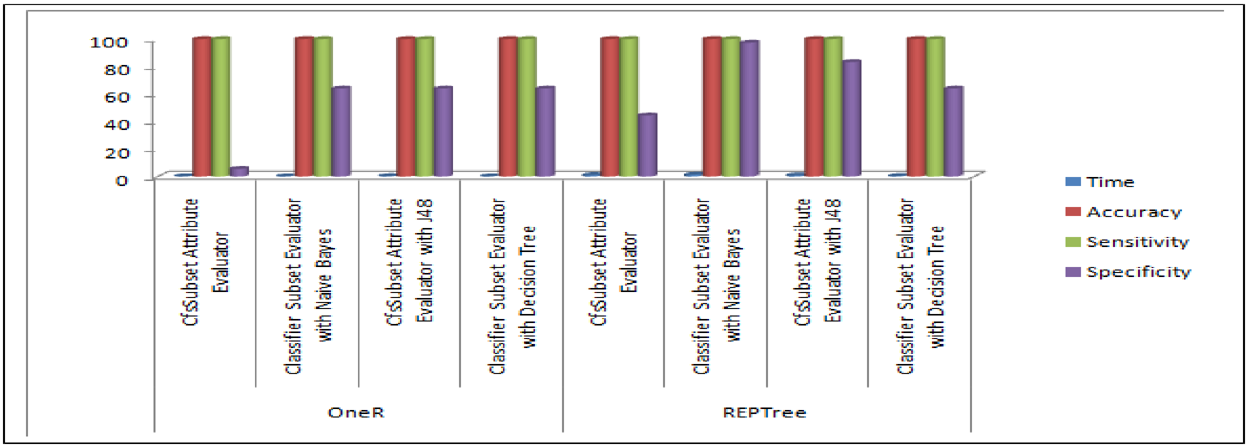

Fig 6- Performance analysis for Infiltration Attack

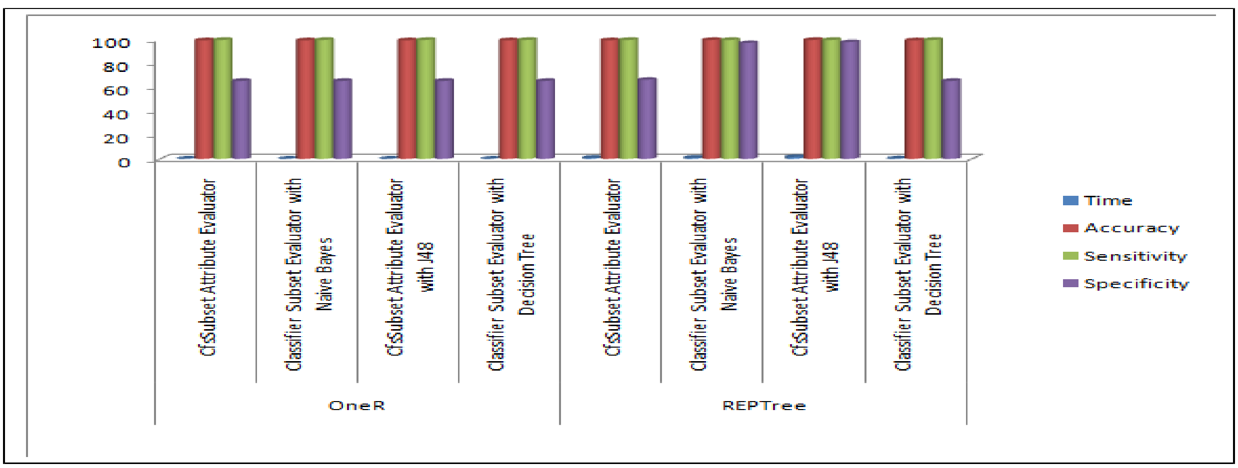

Fig 7- Performance analysis for Botnet Attack

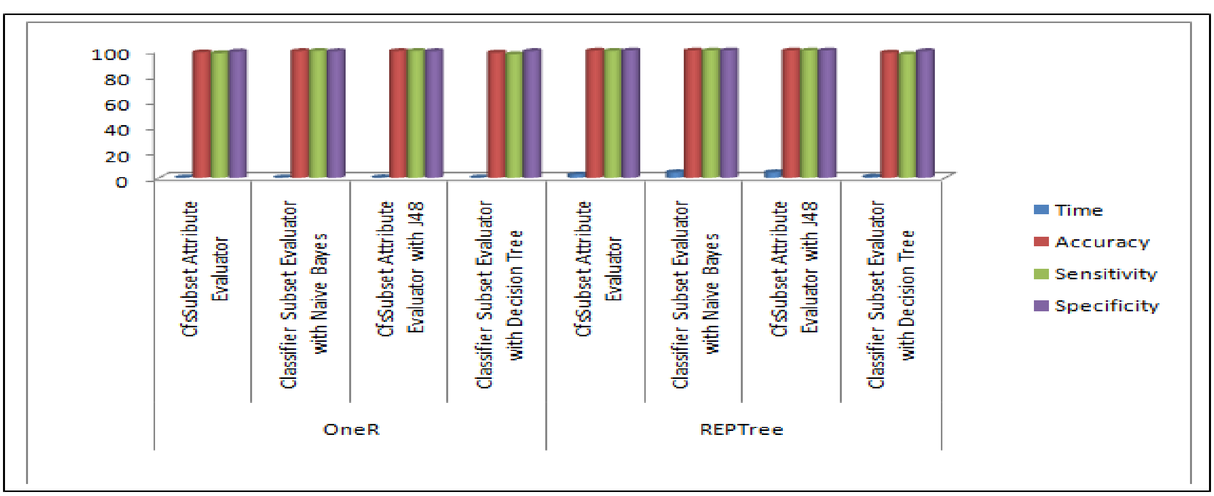

Fig 8 -Performance analysis for Port Scan Attack

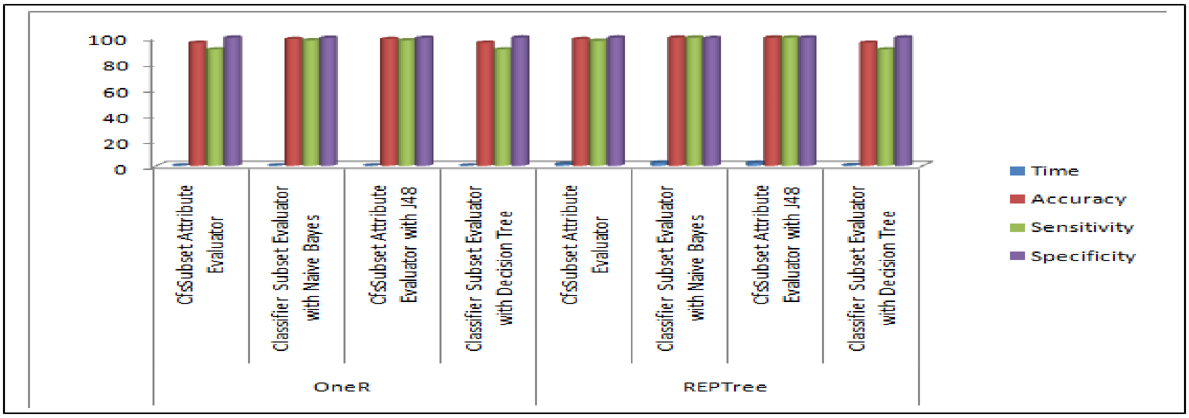

Fig 9 - Performance analysis for DDoS Attack

5. Conclusion and Future Work

In this paper, for evaluation we have used different feature selection algorithms and two classification algorithms with the WEKA tool to detect intrusion. We have used CICIDS-2017 dataset which consists of seven different types of attack. According to results, feature selection reduced the dataset size and time and gives the high performance. The REPTree classification algorithm with CfsSubset Attribute Evaluator with J48 features selection technique provides the best performance for Brute Force Attack, Heartbleed Attack/ DoS Attack, Web Attack, Botnet Attack, Port Scan Attack and DDoS Attack, while the REPTree classification algorithm with Classifier Subset Evaluator with Naive Bayes features selection technique provides the best performance for Infiltration Attack.

We will use other advanced machine learning and deep learning algorithms to detect network intrusion for future work in this field.

REFERENCES

Yao, J. T., Zhao, S. L., & Saxton, L. V. (2005). A study on fuzzy intrusion detection, SPIE: Data Mining, Intrusion Detection, Information Assurance, and Data Networks Security, vol. 58,no. 12, pp. 23-30.

Bakshi, K., & Bakshi, k. (2018). Considerations for Artificial Intelligence and Machine Learning: Approaches and Use Cases, IEEE.

Yuxun, L., & Niuniu, X. (2010). Improved ID3 Algorithm, 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT).

Tavallaee, M., Bagheri, E., Wei, L., & Ghorbani, A. A. (2009). A detailed analysis of the KDD CUP 99 data set, Proceedings of the Second IEEE Symposium on Computational Intelligence for Security and Defence Applications.

The NSL KDD Dataset.[Online]. Available http://nsl.cs.unb.ca/NSL-KDD/,On Dated July 30, 2013.

Olusola, A. A., Oladele, A. S. & Abosede, D. O. (2010). Analysis of NSL KDD'99 Intrusion Detection Dataset for Selection of Relevance Features, Proceedings of the World Congress on Engineering and Computer Science, Vol. 1.

Yi, W., Duan, J., & Lu, M.(2011). Optimization of Decision Tree Based on Variable recision Rough Set, International Conference on Artificial Intelligence and Computational Intelligence.

Ibrahim, K., & Ouaddane, M. (2017). Management of Intrusion Detection Systems based-KDD99: Analysis with LDA and PCA, IEEE.

Robu, R., & Hora, C. (June 2012). Medical data mining with extended WEKA, Intelligent Engineering Systems (INES), IEEE 16th International Conference on, pp.347-350, 13-15.

CICIDS-2017dataset, [Available Online] http://iscx.ca/ CICIDS-2017/.

Aher, S. B., & Lobo, Mr. (2011). Data Mining in Educational System using WEKA, International Conference on Emerging Technology Trends (ICETT'11), pp.20-25.

Kalyani, K., & Lakshmi, G. A. J. (2012). Performance Assessment of Different Classification Techniques for Intrusion Detection, *IOSR Journal of Computer Engineering (IOSRJCE)*, vol. 7(5): 25-29.

Belouch, M., & Hadaj, S. E. Mohamed Idhammad,(2017). A Two-Stage Classifier Approach using RepTree Algorithm for Network Intrusion Detection, *International Journal of Advanced Computer Science and Applications*, Vol. 8, No. 6.

WEKA User Manual, [Available Online] www.gtbit.org/downloads/dwdmsem6/dwd msem6lman.pdf.

www.cs.waikato.ac.nz/ml/weka/.

Panwar, S.S. & Raiwani, Y.P. (2014). Data Reduction Techniques to Analyze NSL-KDD Dataset, *International Journal of Computer Engineering &Technology*. vol 5, issue 10, pp 21-31.

Raiwani, Y. P., & Panwar, S.S. (2014). Research Challenges and Performance of Clustering Techniques to Analyze NSL-KDD Dataset, *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS).* Volume 3, Issue 6,pp 172-177.

Raiwani, Y. P., & Panwar, S.S. (2015). Data Reduction and Neural Networking Algorithms to Improve Intrusion Detection System with NSL-KDD Dataset, *International Journal of Emerging Trends & Technology in Computer Science(IJETTCS)*. Volume 4, Issue 1,pp 219-225.