ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803//tg-20210603125428

Original scientific paper
Received: 2021-06-03, Accepted: 2021-11-01

Salient Features Selection Techniques for Instruction Detection in Mobile Ad Hoc Networks

Severin Gbetondji Michoagan*, S. M. Mali, Sharad Gore

Abstract: The development of wireless mobile ad hoc networks offers the promise of flexibility, low cost solution for the area where there is difficulties for infrastructure network. A key attraction of this mode of communication is their ease of deployment and operation. However, having a good and robust mobile ad hoc networking will depend entirely on security mechanism system in place. Traditional security mechanisms know as firewalls were used for defensive approach to oppose security obstacle. However, firewalls do not fully or completely defeat intrusions. To cope with this limitation, various intrusions detection systems (IDSs) have been proposed to detect such network intrusion activities. The problem encounter for this particular technique of instruction detections technique is that during network monitoring for data collection for anomaly detection, data that does not contribute to detection must be deleted before detection can be processed or application of learning algorithm for detection of abnormal attacks. In this paper we present a novel feature technique for feature selection before learning technique should be applied. The method has been applied into our own data set, and for the detection purpose we have used most of the well reputed three Machine Learning classifiers with the new selected features for performance evaluation and the experiment shows that higher accuracy results could be achieved with only all the 9 features extracted with our own algorithm with the data set created by using RandomForest classifier.

Keywords: AODV; IDS; Jammer Node; MANET; OPNET; Spoofing attack; Wormhole attack

1 INTRODUCTION

Today, more people use mobile phones than traditional fixed phones. Nevertheless, we are experiencing a huge growth rates in mobile wireless communication. For many countries, mobile wireless communication is the only solution of communication in some location due to the lack of an appropriate fixed communication infrastructure [1]. While traditional communication paradigms deal with fixed networks in which security can be managed, wireless communication raises a new set of questions such as due to the openness of the network appropriate security mechanism are hard to achieve. Especially in the case of Mobile ad hoc networking [2], which is the subject of our research studies, security mechanism is hard to imagine due to some design implementation issues that defined the network such as:

- Dynamic topology
- Limited Bandwidth
- Routing issues
- Lack of central authority
- Lack of association among nodes.

MANETsare more exposed to malicious attacks due to the openness of the network and the autonomous aspect of the connecting nodes. Any node can be able to join and or leave the network at any time.

Attacks in MANETS can be classified as one of the two forms: Horizontal attacks or Vertical Attacks.

The Horizontal attacks are the existing attacks such as: Dos attacks, Blackhole attacks, Malicious attacks, etc, and can be term as going from 1 to *n*. Whereas, vertical attacks means be able to detect news attacks term as going from 0 to 1. Vertical attack is hard to imagine because it require detecting an attack nobody else has ever detect.

To cope with design issues related to MANETs which are the causes of the vulnerabilities of the networks, various IDS (Intrusion Detection Techniques) have been implemented by the researchers in MANET community to have a suitable resources sharing, and communication with less vulnerabilities to malicious attacks.

Intrusion detection is the technique that strives to detect an instructor that attempted into computer system then initiate responses to the intrusion [3].

Various Axes involved in the intrusion detection techniques such as the time at which the detection occurs, the types of inputs examined to detect instructive activities, and the range of responses capabilities as the simple form of alerting an administrator of the potential intrusion. These axes included in the design space for detecting intrusion in the computer systems have yields a wide range of solution known as Intrusion-Detection Systems (IDS). These intrusion detection systems techniques come in two forms: Signature-base detection and anomaly detection.

In **signature-base detection technique**, the system inputs or network traffic are scrutinized for specific behaviour patterns (or signatures) that are known to indicate attacks. In this approach, only known attacks are identified. This issue is well known to virus-detection software vendors, who must release new signatures on a regular basis as new viruses are generated and detected manually. In the case of **anomaly detection**, the attempt or detection is to characterise normal (or non-dangerous) behaviours and detect them when something other than these behaviours occur. Anomaly system activity does not always imply an incursion, but the presumption is that intrusions often induce anomalous behavior in a system. In particular, anomaly detection can detect previous unknown networks of intrusions.

The aim of this research studies are the detection of the salient feature selections techniques used by various researchers to have a best implementation of IDS techniques to remove the node having malicious intent on MANET that constitute security threats. In addition, our feature recommended feature selection approach mechanism is elaborated and the application is performed on our own data generated and the results are applied to existing Machine learning Classifiers algorithms to check the detection accuracy of the proposed feature selection algorithm.

This study is very crucial because feature selection is very important in order to detect malicious behavior during network monitoring to extract various features contributing

to the data collected and the application of the learning algorithm to alert the system of the presence of malicious activities.

The paper's organizational structure is as follows: Section 2 is the description of the previous research work. The proposed method; simulation implementation and data collection is described in Section 3, illustration and results of proposed features extracted are given in Section 4, Experiments and results analysis of applied ML classifiers are given in section 5, and at last Section 6, is the Conclusions and future directions.

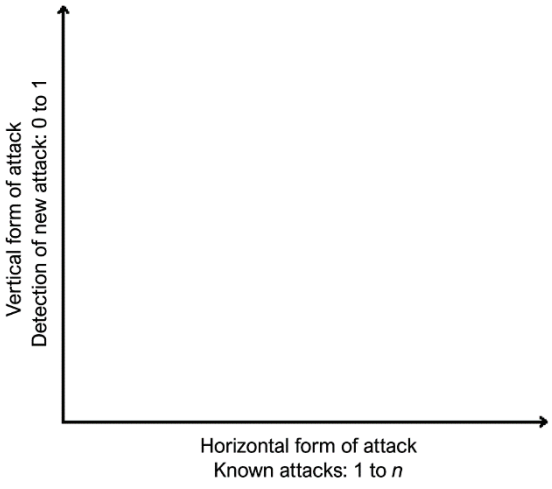

Figure 1 Horizontal and Vertical forms of attack

2 RELATED WORK

In the area of network intrusion detection techniques, relevant features are the main tool with which the detection technique is based on. However, the question arises as what constitute an important feature selection technique. The answer to this question depends entirely upon what constitute the researcher aims and the data generated or data at hand.

Various features solution techniques have been proposed in the literature in order to have a better detection rate in intrusion detection approaches. In [4], for feature selection, a trial-and-error method of deleting one feature at a time is proposed. Neural Network and Support Vector Machines were applied on the selected features for the importance of ranking of the input features. The set of important features was determined to be the reduced feature set that produced the highest detection rate in the experiments. Paper, [5] used a Naive Bayesian classifier to detect network intrusions. They used KDD'99 data with all 31 attributes from the data set and reported overall error rates of 5.1 %. In paper [6,] a feature selection algorithm based on information gain and SVM is developed (Support Vector machine). Its basic principle is to group all data features based on information gain, and then use the SVM algorithm to select the best features subset. In the first stage, Sangkatsance, Watlanapongsakorn, and CharnsriPinyo suggested a real-time intrusion detection system (RT-IDS) and retrieved 12 critical elements from the network packet instructions header. In the second stage, to evaluate the importance of the feature, information gain was used in detecting different forms of attacks. By using (RT-IDS) for detection of different forms of attacks, the rate detection was 98% for probing and denial of service attacks classes. In [8], for feature reduction, the authors developed a clustering conjunct information hybrid technique. Features Clustering was done based on similarity in an unsupervised manner. To increase similarity with response features providing class labels, a supervised learning approach was employed to find important features. Kabiri [9] in his work, DDoS attacks were simulated, and a classifier based on Principle Component Analysis (PCA) was used to select useful attributes from a set of 16 attributes. The three most important attributes he found are routing reply, the number of received packets and total RREP. In paper [8], the authors have elaborated two steps process feature algorithm for intrusion detection system in which redundant features are reduced using mutual information approach with (KDD-99) data set for experimentation. Highest accuracy and in processing speed was achieved by the proposed method. Bayesian networks were use in [10] for data classification as well as to select features with the help of Markov Blanket method on the target variables. Support vector machine and neural network were suggested by the authors in [11] for the classification procedure. In all attack classes, the detection accuracy was outstanding. Barmejo, P. Ossa, L. Gamez, J. A. & Puerta, J. M. [12] proposes a mechanism for dealing with subset selection in datasets with a large number of attributes. The goal of their research was to produce excellent results with a small number of wrapping strategies. To achieve the best results, the suggested approach alternates between filter ranking and wrapper feature subset selection. Furthermore, the approach was tested on 11 high-dimensional data sets using several classifiers.

3 PROPOSED FEATURE SELECTION ALGORITHM

Data is one of the main components in instruction detection techniques analysis. However, large data can occupy more recourse and may result in inefficient of intrusion detection. As a result, data that does not contribute to detection must be removed before processing or using a learning algorithm for atypical attack detection. This necessitates the employment of an appropriate feature reduction technique that cannot only aid minimize training time, but also provide higher detection accuracy and detect anonymous attacks.

Our recommended feature extraction technique is described as two-steps process. In the first step, data pre-pre-processing is elaborated and two algorithms, such as information gain and correlation, perform the second step feature selection or ranking.

3.1 Data Pre-Processing

In reality, due to multiple sources of origins, data used for experimentation are highly unclean, and susceptible to noise [13]. As a result, low-quality data will yield low-quality detection results. Hence, before any feature technique can be implemented it is necessary to check if the data to be use is clean and accurate. Various techniques of data pre-processing have been proposed in the literatures. The inconsistencies in data and removal of noisy data can be achieved using data cleaning technique. To merge data from

multiple sources into a coherent data or the storing of data in data wherehouse can be done with the method of data integration. The reduction of the size of data, or eliminating redundant features or clustering of data, can be achieved with the approach of data reduction. Finally, data transformation is applied to a data scaled within a smaller range related to 0.0 to 0.1. Therefore, the quality of data for experimentation has to satisfy the following requirements such as credibility, accuracy, interpretable, consistency and timeliness.

3.2 Feature Extraction

Data to be analyzedmay contain hundreds of features. Many of them may be unnecessary or redundant to the learning algorithm. Removing relevant features or keeping irrelevant features may be erroneous, and can lower the performance of the learning algorithm to be used. This can lead to the discovery of low-quality patterns. Furthermore, the addiction of an increasing volume of unnecessary or redundant features may gradually affect or slow down the learning process.

3.2.1 Information Gain

The information gain attribute selection technique is a research work approach done by Claude Shannon on information theory [13], by studying the value or "information content". The entropy of each feature is calculated using information gain. The higher the entropy, the more information it contains. The process of identification a given set of features vectors for which attributes is useful for learning process is done using information gain feature selection technique and the selected features will be used for classification in order to identify unknown instances and have a differentiation between types of attacks classes.

Let D be a set of training class-labeled tuples for the partition data. Let assume that the class label attribute has m different values that represent m different classes, for C_i for ($i = 1, \ldots, m$). Let $C_{i,D}$ be theset of tuples of the class C_i in D. Let $|D|$ and $|C_{i,D}|$ denote the number of tuples in D and $C_{i,D}$, respectively. The data required to classify a tuple in D is given by

$$Info(D) = -\sum_{i=1}^{m} p_i \log_2(p_i) \quad (1)$$

Where p_i is the non-zero probability than an arbitrary tuple in D and part of class $C_i$and being estimated by $|C_{i,D}| / |D|$. A log function to the base 2 is used, because the information is encoded in bits. Therefore, the Info (D) is the average information needed to identify the class label of tuple D. If we were suppose to partition the tuples in D on some attributes A having v distinct values, $\{a_1, a_2, \ldots, a_v\}$, as observe from the training data. If A is discrete-value, these correspond directly to the v outcomes of the test on A. Attribute A can be used to partition or subset D into v partitions or subsets, $\{D_1, D_2, \ldots, D_v\}$, where D_j contains those tuples in D that have outcome a_j of A. This split should, in theory, yield a precise classification of the tuples. That is, we want each division to be completely clean. The partitions, on the other hand, are very likely to be impure. How much more data would we require to arrive at a precise classification? This is determined by

$$Info_A(D) = -\sum_{j=1}^{v} \frac{|D_j|}{|D|} \times Info(D_j) \quad (2)$$

The term $\frac{|D_j|}{|D|}$ acts as the weight of the j^{th} partition. The $Info(D)$ is the expected information needed to classify a tuple D based on the partition of A. Hence, the smaller the expected information required, the greatest the cleanness of the partitions. However, information gain is described as the difference between the initial information requirement and the new in requirement, obtain after partition of A. That is,

$$Gain(A) = Info(D) - Info_A(D) \quad (3)$$

3.2.2 Correlation

To uncover features with greater utility values, we use a new mechanism that combines information gain with correlation-based features. Correlation is the second approach for ranking attributes. In a multiclass problem, the lower the correlation of a property in a feature vector, the more powerful it is to distinguish between distinct types of attacks. The pair-wise linear correlation coefficient between each pair of columns is returned by correlation as a matrix. Then, by taking the mean of each column, the correlation coefficient of each feature is computed.

Proposed feature selection Algorithms

Algorithm 1: Data Pre-processing
Input: Dataset with feature subset $f(N) = \{n_1, n_2, n_3, \ldots, n_n\}$
Output: Compact and redundant dataset with feature subset $f(M) = \{m_1, m_2, m_3, \ldots, m_m\}$

1. Removal of duplicate features
2. Replace Infinite values with zero
3. Replace missing values with zero
4. Return $f(M) = \{m_1, m_2, m_3, \ldots, m_m\}$

Algorithm 2: Feature Selection
Input: Pre-processed dataset with feature subset $f(M) = \{m_1, m_2, m_3, \ldots, m_m\}$
Output: Reduced features subset $f(R) = \{r_1, r_2, r_3, \ldots, r_m\}$

1. for $f(M)$
a. Compute weight of IG_{mi}
b. if $IG_{mi} \geq 0.5$ then add into $f(IGFS_1)$
c. if $0.25 \geq IG_{mi} < 0.5$ then add into $f(IGFS_2)$
d. if $0.25 \geq IG_{mi}$ then remove IG_{mi}
2. for $f(M)$
a. Compute weight of CR_{mi}
b. if $CR_{mi} \geq 0.5$ then add into $f(CRFS_1)$

c. if $0.25 \geq CR_{mi} < 0.5$ then add into $f(CRFS_2)$
d. if $0.25 \geq CR_{mi}$ then remove CR_{mi}

End for

3. Compute $f(NMRFS) = f(IGFS_1) \cup f(IGCR_1)$
4. Compute $f(NRFS) = f(IGFS_2) \wedge f(IGCR_2)$
5. Compute $f(R) = f(NMRFS) \cup f(NRFS)$
6. Return $f(R)$

4 SIMULATION IMPLEMENTATION AND DATA COLLECTION

The implementation of our MANET model is design with OPNET modeler 14.5 with AODV as routing protocol and 2 types of attack have been implemented such as selfish nodes attacks and Dos attacks.

4.1 Ad Hoc on Demand Distance Vector Protocol (AODV)

The mechanism to identify routes path if only if there are needed is the functionality of Ad hoc On Demand Distance Vector Protocol [14]. Therefore a route need to be established first, and once a route is being identified, the paths is preserved until there is no needed for it and or once the message desired to the destination is completed the route can be discarded.

4.2 Attacks Implementation

4.2.1 Selfish Node Attack

Selfish node attacks are nodes that are presents in the network and due to lack of energy or in order to preserve their energy consumption for future use do behave maliciously in the network. Selfish node behavior can be categorized as a node that does not perform the packet forwarding after receiving the packet intended to the requested node, or purposely disable its routing protocol to avoid packet forwarding and receiving to preserve it energy, or a node that has power failure or power off during the communication [21].

In the case for our studies, the selfish node implemented is the one that has its routing protocol disabled: Disabling of AODV routing protocol is the configuration of selfishness nodes attacks for our studies.

4.2.2 DOS Attacks

Denial of Service Attack (DOS) floods the network with unnecessary network traffic. The attack traffic consumes network resources, preventing legatine traffic from reaching the destination, wasting nodes energy.

Pulse Jammer attack is simulated in our case. Jammer attack [14-19] floods the network with high wireless radio frequency to disturb the communication in place. Jammer node is different in structure as compared to MANET node. With its radio transmitter, it frequently generates noisy frequencies on wireless channel. Jammer node generates highest bandwidth (in kHz) during the transmission. Jammer transmitter power indicates the transmission power (in Watts) allocated to packets transmitted through the channel. Lastly, the jammer node has a pulse width which point out the length of time (s) a pulse is transferred and a silence width specifies the interval in (s) between pulses [20].

4.2.3 Data Collection and Features Extraction

The recorded data set collected after the simulation was performed contains 15 features plus assigned classes label classifying each record as normal node, Selfish node attacks and dos attacks. The total number of instances that characterize each attack class is distributed in Tab. 3.

Table 1 MANET simulation parameter

Simulation parameters	Value
Network range	500 × 500 m
Routing Protocol	AODV
Number of nodes	80
Number of selfish nodes	5
Pulse jammer node	1
Packet size	512 packet/s
Simulation time	60s

Table 2 List of features extracted

Serial number	Features
1	Total Number of Hops per Route
2	Total Route Discovery Time
3	Total Routing Traffic Received
4	Total Routing Traffic Sent
5	Total Cached Replies Sent
6	Total Packets Dropped
7	Total Replies Sent from Destination
8	Total Route Errors Sent
9	Total Route Replies Sent
10	Total Route Requests Forwarded
11	Total Route Requests Sent
12	Radio transmitter queue size
13	Radio transmitter queuing delay
14	Radio transmitter throughput
15	Radio transmitter utilization

Table 3 Total number of instances for each attack class

Classes	Number of instances	Percentage of class occurrence, %
Normal	7575	92.60
Selfish Attacks	505	6.17
Dos Attacks	101	1.23
Total	8181	100

Table 4 Feature distributing ranking as per information gain

Ranking as per data distribution	Features	IGR
1	Total Number of Hops per Route	0.4291
2	Total Route Discovery Time	0.4291
6	Total Packets Dropped	0.4281
5	Total Cached Replies Sent	0.4281
9	Total Route Replies Sent	0.4281
7	Total Replies Sent from Destination	0.4281
8	Total Route Errors Sent	0.4281
10	Total Route Requests Forwarded	0.4281
11	Total Route Requests Sent	0.4281
3	Total Routing Traffic Received	0.1254
4	Total Routing Traffic Sent	0.1091
13	Radio transmitter queuing delay	0.0960
15	Radio transmitter utilization	0.0343
12	Radio transmitter queue size	0.0335
14	Radio transmitter throughput	0.0300

With the help of WEKA which is one of the powerful data analysis machine learning software, developed at the University of Waikato, New Zealand [15]. The information gain and correlation ratio is calculated with all the 15 features and the results are listed in the Tabs. 4 and 5.

Table 5 Feature distributing ranking as Correlation

Ranking as per data distribution	Features	*CR*
11	Total Route Requests Sent	0.636
10	Total Route Requests Forwarded	0.580
1	Number of Hops per Route	0.561
6	Total Packets Dropped	0.556
8	Total Route Errors Sent	0.536
2	Route Discovery Time	0.487
5	Total Cached Replies Sent	0.433
9	Total Route Replies Sent	0.417
7	Total Replies Sent from Destination	0.390
13	radio transmitter queuing delay	0.370
15	radio transmitter utilization	0.225
3	Routing Traffic Received	0.206
4	Routing Traffic Sent	0.205
12	Radio Transmitter Queue Size	0.205
14	Radio Transmitter Throughput	0.193

Base of the ranking of the information gain and correlation of collected features for our studies, the first stage of our recommended feature selection method results is summarized in Tab. 6.

Table 6 Summarized of First stage of proposed feature extraction algorithm

For Information Gain	
$IGFS_1$	{ θ}
$IGFS_2$	{ 1, 2, 6, 5, 9, 7, 8, 10, 11 }
Removed features	{ 3, 4, 13, 15, 12, 14 }
For correlation	
$CRFS_1$	{ 11, 10, 1, 6, 8 }
$CRFS_2$	{ 2, 5, 9, 7, 13 }
Removed features	{ 15, 3, 4, 12, 14 }

Table 7 Summarized of Second stage of proposed feature extraction algorithm

	Operations	Selected features
NMRFS	$(IGFS_1) \cup (IGCR1)$	{ 11, 10, 1, 6 , 8 }
NRFS	$(IGFS_2) \wedge (IGCR_2)$	{ 2, 5, 9, 7 }
R	(*NMRFS*) ∪ (*NRFS*)	{ 11, 10, 1, 6 , 8, 2 , 5, 9 , 7}

Table 8 Reduced feature with our method

Features
Total Route Requests Sent
Total Route Requests Forwarded
Number of Hops per Route
Total Packets Dropped
Total Route Errors Sent
Route Discovery Time
Total Cached Replies Sent
Total Route Replies Sent
Total Replies Sent from Destination

The second stage of our recommended feature selection method proposed is the computation of the Union operation of $IGFS_1$ and $IGCR_1$ and the Interception operation of $IGFS_2$ and $IGCR_2$, which was store as *NMRFS* and *NRFS* respectively, and the best feature selection is the results of the Union operation of *NMRFS* and *NRFS*.

The best-selected features using our proposed method are {11, 10, 1, 6, 8, 1, 5, 9, and 7}.

The importance of having a feature selection before any IDS methods can be implemented is that there are some features in the data set which can lead to the deterioration of the performance of the classifier learning method considered for anomaly detection. Therefore, any feature F is important if by removing it from the set of features affect the classifier performance. Having feature selection mechanisms in place will contribute to the predictive classifier model to be considered and helping choosing important features that will generate best accuracy and less complexity time when we acquired new data.

5 EXPERIMENTS AND RESULTS ANALYSIS

Experiment in this section used three most existing Machine Learning Classifiers such as NaiveBayes, RandomForest Decision tree and J48. The experiment has two phases: the first phase is the results of the performance using all 15 features applied to the Three ML classifiers. In the second phase, the evaluation is done with the 9 extracted features with all the three ML Classifiers. To measure the performance of the three ML classifiers, precision, recall, and F1 score evaluation measures were used, because they are the most use measurement for performance evaluation in anomaly detection detection techniques. Precision is the percentage of relevant instances found among the retrieved instances. The proportion of important retrieved instances in the total number of important instances is referred to as recall. The harmonic mean of precision and recall is used to calculate the F1 score. These three performance evaluation metrics depend entirely on the confusion matrix in which four possible situations can be defined, as shown in Tab. 9.

The experiment shows that 100% accuracy results is achieved with only all the 9 features extracted with our own algorithm with the data set created by using RandomForest classifier. However, we also have a higher accuracy for normal node and as well as DOS attack in case of NaiveBayes and higher accuracy for DoS Attack for J48 classifier.

Table 9 Evaluation metrics parameters

Measure	Explanation
True positive (*TP*)	Correctly classified instances as an anomaly
True negative (*TN*)	Instances that were accurately labelled as normal
False negative (*FN*)	Instances that were incorrectly classed as normal
False positive (*FP*)	Instances Anomaly incorrectly classified
Precision	*TP* / (*TP* + *FP*)
Recall	*TP*× (*TP* + *FN*)
F1 score	2 ×precision × recall / precision + recall

Table 10 Phase 1: Results for all features (15)

Classifiers	Accuracy, %	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
NaiveBayes	99.99	1	0.002	1	1	1	0.999
RandomForest	100	1	0	1	1	1	1
J48	99.99	1	0.002	1	1	1	0.999

Table 11 Phase 2: Results for selected feature extracted (9)

Classifiers	Accuracy, %	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
NaiveBayes	99.99	1	0.002	1	1	1	0.999
RandomForest	100	1	0	1	1	1	1
J48	99.97	1	0.002	1	1	1	1

Table 12 Accuracy detection for each individual class of attack

Classifiers	Classes	Phase 1, %	Phase 2, %
NaiveBayes	Normal	100	100
	Selfish Attack	99.80	99.80
	Dos Attack	100	100
RandomForest	Normal	100	100
	Selfish Attack	100	100
	Dos Attack	100	100
J48	Normal	100	99.98
	Selfish Attack	99.8	99.88
	Dos Attack	100	100

To sum up our feature selection technique performed well with RandomForest classifier.

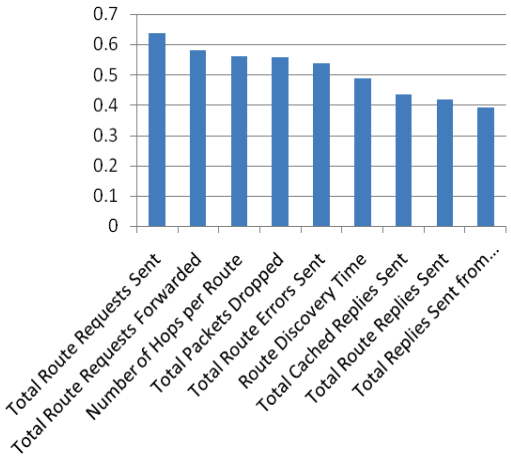

Figure 2 Selected features of dataset generated

Table 13 Comparisons of various feature selection and reduction techniques used in literature with their advantages and drawbacks with proposed feature selections proposed

Papers	Feature selection algorithms used	Types of feature selection categories	Advantages	Drawbacks
Akashdeep, IshfaqManzoor, Neeraj Kumar (2017)	Two algorithms were used: Information Gain and Correlation	Filter Method	The proposed approach detects 98.79 % of probing attacks	Performance for Normal class is low
Mukammala and Sung (2003)	SVM and neural networks were used to classify features.	Wrapper Method	After SVM was compared to neural networks, it was discovered that SVM has higher scalability, but that SVM may be utilized on huge datasets.	SVM can only perform binary classifications, and neural networks required more training time than SVM.
Ahmed Mahfouz, Abdullah Abuhussein, Deepak Venugopal, Sajjan Shiva (2020)	Applied the InfoGainAttributeEval algorithm with Ranker	Filter Method	The precision and recall rates is High, improvement of detection accuracy and improvement of TPR	Decreasing of FPR
Fleuret (2004)	A mutual information-based feature selection method is proposed.	Filter Method	Attain a high level of categorization efficacy.	For all classes, the results were insufficient.
Uguz (2011)	Two-stage feature extraction and selection algorithm has been proposed: Principal Component Analysis and Information Gain	Hybrid Method	Better performance with with naïve bayes than SVM	The focus is more on processing time.
Xiao et al. (2009)	Mutual information was used to eliminate redundant features.	Filter Method	Increased processing speed and accuracy	Just with DoS and Probing attacks only the Experiments showed good results
Karimi et al. (2013)	Information gain and symmetrical uncertainty are used to combine two feature sets.	Hybrid Method	Improvement in Detection rate	The precision of detection in U2R and R2L needs to be increased.
Al-Jarrah et al. (2014)	Random Forest-Forward Selection Ranking (RF-FSR) and Random Forest-Backward Elimination Ranking (RF-BER) are two new rankings that have been proposed (RF-BER)	Wrapper Method	0.01 % increase in detection rate and 0.01% decrease in false alarm.	Only Accurate format was used to generate the results.
Proposed Method	Combination of Information Gain and Correlation	Filter Method	100% accurate detection for Normal node, Selfish attack and Dos Attack With RandomForest algorithm	Not 100% accuracy for all Classifications algorithms used improvement need to be done

6 CONCLUSION AND FUTURE WORK

In this paper, we have presented a novel feature extraction technique to overcome the difficulties encounter during features selection techniques for networks intrusion detection techniques. The dataset generated for our study is completely labeled and 15 network traffic features have been extracted for intrusive flows detection. The feature selection technique proposed has extracted 9 important features. The selected features extracted has been use to compared the

performance of threemost well-known ML classifiers and the experiment shows that the higher accuracy results could be achieved by using RandomForest classifier. Our future work will be the use of Ensemble method for classification to improve the accuracy of detection rate. An Ensemble combines various ML classifiers (Base classifiers) for learning purpose and each base classifier is assigned a unique vote. Based on the votes of the base classifiers, the ensemble returns the prediction class for learning purposes.

7 REFERENCES

[1] Schiller, J. (2009). *Mobile Communication*, 2nd Edition, Published by Darling Kindersley(India) Pvt.Ltd.

[2] Talukder, A. K., Afmed, H., & Yavagal, R. R. (2010). *Mobile Computing: Technology, Application and Service Creation*, Tata Mc Graw Hill Education Private Limited.

[3] Silberschatz, A., Galvin, P. B., & Gagne, G. (2003). *Operating System Concepts*, 6thEdition, Published by John Wiley & Sons, Inc.

[4] Sung, A. H. & Mukkamala, S. (2003). Identifying important features for intrusion detection using support vector machines and neural networks.*Proceedings of International Symposium on Applications and the Internet (SAINT '03)*, 209-216, 2003, IEEE.https://doi.org/10.1109/SAINT.2003.1183050

[5] Panda, M. & Patra, M. (2007). Network Intrusion detection using naïve Bayes.*IJCSNS International Journal of computer science and network security, 7*(12), 258-263.

[6] Zaman, S. & Karray, F. (2009). Features Selection for Intrusion Detection Systems Based on Support Vector Machines. *The 6th IEEE Conference Consumer Communications and Networking Conference (CCNC 2009)*, 1-8. https://doi.org/10.1109/CCNC.2009.4784780

[7] Sangkatsanee, P., Wattanapongsakorn, N., & Charnsripinyo, C. (2011). Practical Real-Time Intrusion Detection Using Machine Learning Approaches. *Computer Communications, 34*, 2227-2235. https://doi.org/10.1016/j.comcom.2011.07.001

[8] Xiao, L. & Liu, Y. (2009). A Two-stepFeature Selection Algorithm Adapting to Intrusion Detection.*2009 International Joint Conference on Artificial Intelligence*, 618-622. https://doi.org/10.1109/JCAI.2009.214

[9] Kabiri, P. & Aghaei, M. (2011). Feature analysis for intrusion detection in Mobile Ad-hoc networks.*InternationalJournal of Network Security, 12*(1), 42-49.

[10] Chebrolu, S., Abraham, A., & Thomas, J. P. (2005). Feature deduction and ensemble design of intrusion detection system. *Computers & Security, 24*(4), 295-307. https://doi.org/10.1016/j.cose.2004.09.008

[11] Sung, A. H. & Mukammala, S. (2003). Feature selection for intrusion detection using neural network and support vector machine. *Transportation Research Record: Journal of the Transportation Research Board*, 1822, 1-11. https://doi.org/10. 3141/1822-05

[12] Barmejo, P., Ossa, L., Gamez, J. A., & Puerta, J. M. (2012). Fast wrapper feature subset selection in high dimensional datasets by means of filter re ranking. *Journal of Knowledge Based Systems*, 25, 35-44. https://doi.org/10.1016/j.knosys.2011.01.015

[13] Han, J., Kamber, M., & Pei, J. (28012). *Data Mining Concepts and Techniques*, 3rd Edition, Published by Morgan Kaufmann.

[14] Krombholz, K., Hobel, H., Huber, M., & Weippl, E. (2015). Advanced Social Engineering Attacks. *Journal of Information Security and Applications*, 22, 113-122. https://doi.org/10.1016/j.jisa.2014.09.005

[15] Lu, Z. & Yang, H. (2012). *Unlocking the Power of OPNET Modeler*, Published by CambridgeUniversity Press.

[16] Salim, S. (2010). Mobile Ad Hoc Network Security Issues, *M.Sc. Thesis*, University of Central Lancashire, 81p.

[17] Sharma, V. & Mittal, S. (2014). Load Balancing in MANETs: A Review.*International Journal of Advanced Research in Computer Science and Software Engineering, 4*(5).

[18] Gagandeep, A. & Kumar, P. (2012). Analysis of Different Security Attacks in MANETs on Protocol Stack A-Review. *International Journal of Engineering and Advanced Technology (IJEAT)*, 1, 269-275.

[19] Biswas, K. & Liaqat, A., Md. (2007). Security Threats in Mobile Ad-Hoc Network, *Master Thesis*, Blekinge Institute of Technology, Blekinge.

[20] Thuente, D. J. & Acharya, M. (2006). Intelligent Jamming in Wireless Networks with Applications to 802.11b and Other Networks. *MILCOM'06: Proceedings of the 2006 IEEE conference on Military communications*, North Carolina State University, 1075-1081.

[21] Michiardi, P. & Molva, R. (2002). Simulation-based Analysis of Security Exposures in Mobile Ad Hoc Networks. Proceedings of European Wireless Conference. *Proceedings of European Wireless Conference*, Florence, 287-292.

Authors' contacts:

Severin Gbetondji Michoagan
(Corresponding author)
Department of Computer Science,
Savitribai Phule Pune University,
Ganeshkhind Road,
Pune - 411 007, Maharashtra State, India
sevegbeton@gmail.com

S. M. Mali
(1) Department of Computer Science,
Savitribai Phule Pune University,
Ganeshkhind Road,
Pune - 411 007, Maharashtra State, India
(2) Dr. Vishwanath Karad,
MIT World Peace University,
Ground Floor, Saraswati Vishwa A, Mitwpu Campus,
S.No.124, Paud Road, Kothrud,
Pune - 411038, Maharashtra State, India
shankarmali007@gmail.com

Sharad Gore
Department of Computer Science,
Savitribai Phule Pune University,
Ganeshkhind Road,
Pune - 411 007, Maharashtra State, India
sharaddgore@gmail.com