Cloud Computing and Data Science
http://ojs.wiserpub.com/index.php/CCDS/

UNIVERSAL WISER
PUBLISHER

Research Article

Machine Learning for Intrusion Detection in Ad-hoc Networks: Wormhole and Blackhole Attacks Case

Aurelle Tchagna Kouanou[1,2], Theophile Fozin Fonzin[2,3], Franck Mani Zanga[2], Adèle Ngo Mouelas[2,4], Gerad Nzebop Ndenoka[5], Michael Sone Ekonde[1]

[1]Department of Computer Engineering, College of Technology, University of Buea, Cameroon
[2]Department of Training, Research, Development and Innovation, InchTech's Solutions, Yaounde, Cameroon
[3]Department of Electrical and Electronic Engineering, Faculty of Engineering, University of Buea, Cameroon
[4]National Advanced School of Engineering Yaounde 1, Yaounde, Cameroon
[5]Department of Computer Science, University of Yaounde 1, Cameroon
Email: tkaurelle@gmail.com

Received: 10 August 2023; **Revised:** 20 September 2023; **Accepted:** 20 September 2023

Abstract: This paper addresses the security concerns associated with Mobile Ad-hoc Networks (MANET) and proposes a new method for detecting and preventing attacks using machine learning. The study involved the creation of a MANET with 26 nodes in NetSim (Network Simulator) software, followed by the implementation of wormhole and blackhole attacks. A dataset was generated from the network traffic obtained during the simulations, and a machine-learning model was designed to predict and detect these attacks. The model achieved high sensitivity, accuracy and f1 scores of 99%. The effectiveness of the model was tested by developing a real-time application. This method can be applied to any wireless network and is particularly relevant for companies that use Ad-hoc networks for communication.

***Keywords*:** network security, Mobile Ad-hoc Networks (MANET), wormhole and blackhole, machine learning

1. Introduction

A Mobile Ad-hoc Network (MANET) is a dynamic wireless network that transmits information about neighboring nodes through a temporary configuration [1-3]. MANET is also a set of mobile, self-organizing, and decentralized nodes used in special cases such as military [3-5]. Each node in the MANET is equipped with a wireless receiver and transmitter, allowing it to communicate with other nodes within the wireless transmission range [5-7]. Nowadays, the use of MANETs is highly appealing and widespread in a lot of applications such as space communication, disaster relief, mission-critical battlefield communication, road or accident guidance, trade fairs, sports stadiums, shopping malls, and avoiding vehicle crashes [8]. However, MANET properties make the network's environment vulnerable to various types of attacks, including wormholes, black holes, Grey Holes, Rushing, and flood-based attacks [2-4, 9-11]. Many techniques and methods have been developed to deal with these attacks. For example, various intrusion detection methods have been developed to detect common network attacks, focusing on other routing attacks such as black hole attacks, Sybil attacks, identity replication attacks, selective forwarding attacks, wormhole attacks, and hello flood attacks [12]. Prasad et al. investigated the detection method that can classify benign and malicious information in the MANET

DOI: https://doi.org/10.37256/ccds.5120243516

networks based only on routing attacks [13]. Ezhilarasi et al. Introduced in 2022, a new intrusion detection system that uses fuzzy and feed-forward neural networks to detect only routing attacks in wireless sensor networks [12]. The two previously cited papers work on routing attacks that include a set of attacks according to [12]. In this paper, our work is based on Black Hole and Wormhole attacks because they are the major attacks in a MANET [14-16]. Many researchers applied ML algorithms to develop a detection and prediction model for these two major attacks in MANET. It's the case of Prasad et al. who used ML along with Naïve Bayes and Stochastic Gradient Descent (SGD) for Wormhole detection in an Ad hoc Network [17]. Shams et al. worked in Vehicular Ad Hoc Networks (a specialized type of MANET) based on a Support Vector Machine (ML algorithm) to identify any signs of bad nodes that may be impacting system performance [18]. However, to the best of our knowledge, many works using ML conducted on blackhole and wormhole attacks in MANET have shown accuracy rates ranging from 59% to 98%, and any application developed for real-time detection.

The main problem addressed in this paper is the lack of effective security mechanisms for Mobile Ad-hoc networks. Works done in the literature to protect Mobile Ad-hoc networks from attacks such as blackhole and wormhole attacks are not very accurate. Therefore, there is a need for new methods that can detect and prevent these attacks in real time with good accuracy and precision. The objective of this paper is to address the security concerns associated with MANET and propose an ML approach for detecting and preventing attacks.

In this paper, we designed an ML model to detect and predict Black Hole and Wormhole attacks in MANET. We first simulated our network by using NetSim, secondly, we performed attacks and registered the log file simulation. Thirdly, after pre-processing the dataset, we applied data analysis and ML to construct a Model. To get a good model, we test various ML algorithms and, in the end, we choose Random Forest as our best model. The model is implemented in an application that allows us to detect and predict blackhole and wormhole attacks in real time.

The rest of our work is designed as follow: Section 2 presents the state of the art where the different definition and related work are given. The methodology of our work is presented in section 3. Here, the workflow used algorithms, and performance metrics are presented. The results and discussions with related work are given in section 4. This work ends with a conclusion and future work in section 5.

2. Related works

MANET represents an independent system of porTable nodes to form a self-organizing, infrastructure-less, and quickly deployable wireless network [8, 11, 19]. MANET is a promising technology that can provide important facilities for up-to-date transportation systems [1]. Due to its inherent nature, MANET is strongly vulnerable to miscellaneous security attacks [13, 17, 20]. Security attacks against MANET are divided into two categories according to their nature: active attacks and passive attacks [13, 17, 20].

• *Active attacks* mainly target the confidentiality and integrity of data. Active attacks involve modifying, dropping, manufacturing, duplicating, or blocking the exchange of packets on the network. These attacks are usually launched from authorized nodes on the network. They use various functions of the network to launch attacks.

• *Passive attacks* mainly target data confidentiality. In a passive attack, malicious nodes attempt to compromise the system based solely on monitoring transmissions on the channel, without directly harming the network. They extract valuable information and use it for future attacks. These attacks are hard to spot because they don't cause direct damage.

In addition, security attacks on MANET are also divided into two categories by domain: insider attacks and outsider attacks [20-22].

• *Outsider attacks* are carried out by unauthorized external nodes to cause congestion, disrupt the normal operation of the network, or spread incorrect routing information.

• *Insider attacks* are caused by internal compromised/malicious nodes to disrupt the normal operation of the network.

In this work, we focused on Blackhole and wormhole attacks. Indeed, in a wormhole attack, an attacker inserts fake nodes to broadcast data and transmit packets from one location on the network to another. On the other hand, an attacker records packets from one place and tunnels them to another place in the MANET, where those packets are sent back to the network. Black hole attack is one of the known security threats in wireless MANET. An intruder exploits this vulnerability for malicious behavior because the process of route discovery is necessary and unavoidable. This attack is known as a node dropping all packets it should forward, claiming it has the shortest path to the destination. Black hole attack in MANET refers also to the attack of malicious nodes, which force the route from source to destination by falsely advertising the shortest hops to reach the destination node [23].

In the literature, some authors worked to find an optimal solution for MANET attacks. Alhaidari and Alrehan conducted an extensive literature review in 2021 and found many limitations in the datasets that can be used for DDoS attacks on vehicular ad hoc networks [1]. However, they only focus their research on DDOS attacks. Hassan et al. introduced an intelligent black hole attack detection scheme tailored to autonomous and connected vehicles [24]. But their scheme wasn't based on machine learning. Meddeb et al. in 2019 proposed just an approach to integrate an IDS able to detect the majority and not all security attacks occur in MANET [25]. Their model wasn't based on Machine Learning and based only on behavioral databases. Another's researchers like Shukla et al. based on cryptographic methods to deal with wormhole and blackhole attacks [14]. Subba et al. in 2016 proposed an Intrusion Detection Systems (IDS) scheme using a novel process for cluster leader election and based on an information Bayesian game for Modeling the intrusion detection process [6]. Abdan and Seno in 2022 investigated on classification of wormhole attacks in the MANET with several ML methods [26]. They reached an accuracy of 98.9% with Decision Trees (DT) higher than other proposed ML models. However, they based on the unbalanced dataset to reach this result. Joon and Chopra in 2021 based on deep learning (DL) and proposed a wireless network with a Hybrid DL Prediction (HDLP) model that used Auto Encoding for Key Management and cluster-based network [27]. Table 1 presents a summary of the existing solution.

Table 1. Summarize of existing research

Authors	Approach	Work done/Limitations	ML-Based
Alhaidari and Alrehan [1]	Extensive literature review on network attacks	Limited datasets for DDoS attacks on VANETs	No
Hassan et al. [24]	Intelligent black hole attack detection	Not based on machine learning	No
Meddeb et al. [25]	IDS integration	Detects majority, not all security attacks in MANET	No
Shukla et al. [14]	Cryptographic methods	Deals with wormhole and blackhole attacks	No
Subba et al. [6]	IDS scheme using Bayesian game	Novel process for cluster leader election	No
Abdan and Seno [26]	Classification of wormhole attacks	Based on unbalanced dataset	Yes (Decision Trees)
Joon and Chopra [27]	Hybrid DL Prediction model	Uses Auto Encoding for Key Management and cluster-based network	Yes (Deep Learning)

Based on the literature, no proposed work has designed a real-time application to detect and predict blackhole and wormhole attacks in MANET. Also, the proposed ML models were not very accurate. Based on this drawback, we design in this paper an optimal ML model that runs to a real-time application in MANET. The next section describes each step of our proposed method.

3. Methodology

In this section, we give a detailed presentation of the different methods, techniques, and tools used to carry out the work. We first present our proposed pipeline. Afterward, we present the methods used to generate our dataset, preprocess the dataset, perform exploratory data analysis (EDA), and, the ML used to build our prediction model. At the end of this section, we present the metric evaluation of the ML model and tools used to build our real time blackhole and wormhole attack detection and prediction application. Figure 1a presents our proposed pipeline that contains steps involved in the realization of our ML model. Figure 1b presents the pipeline to detect intrusion using machine learning.

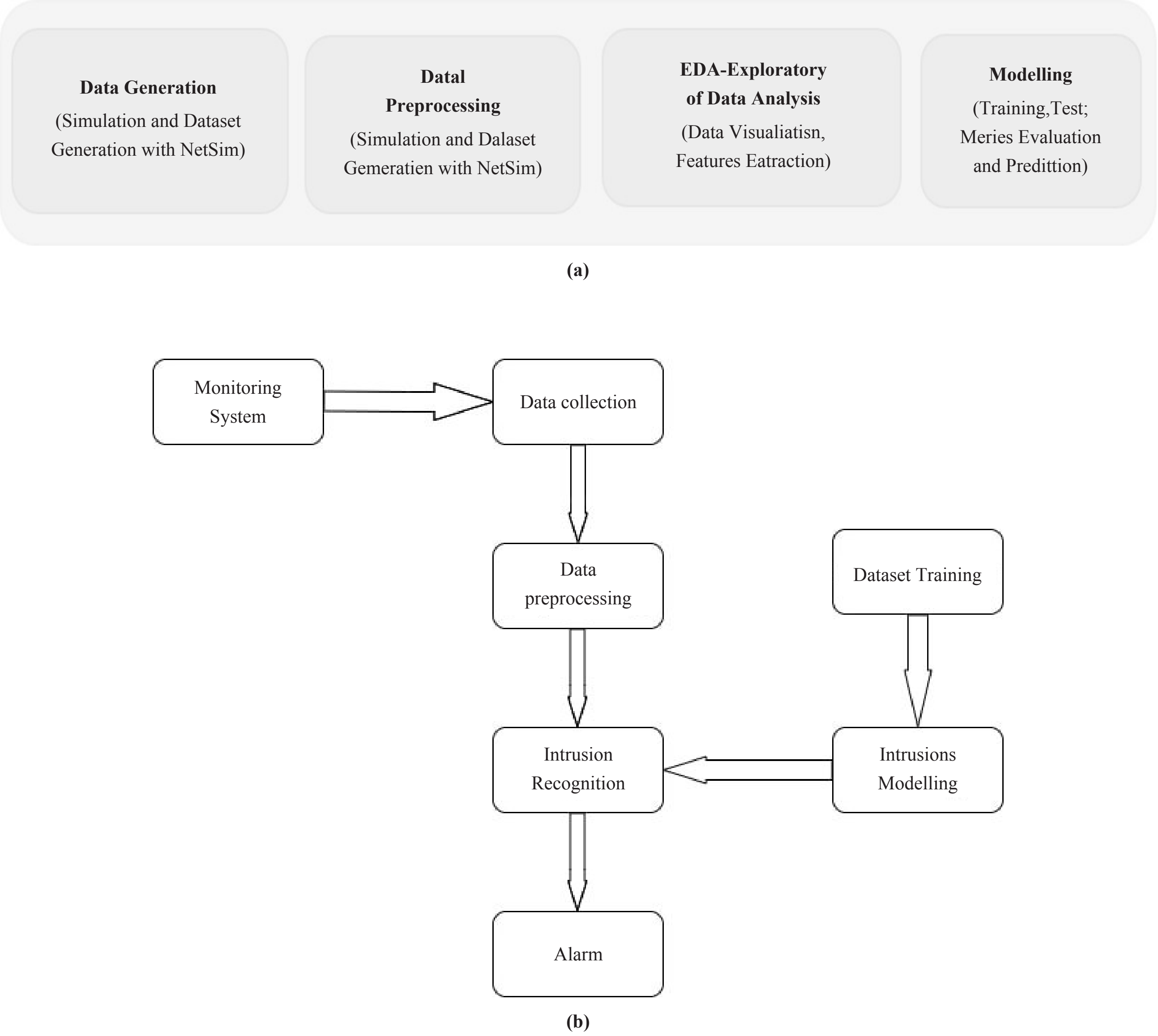

Data Generation
(Simulation and Dataset Generation with NetSim)
Datal Preprocessing
(Simulation and Dalaset Gemeratien with NetSim)
EDA-Exploratory of Data Analysis
(Data Visualiatisn, Features Eatraction)
Modelling
(Training,Test; Meries Evaluation and Predittion)
(a)
Monitoring System
Data collection
Data preprocessing
Dataset Training
Intrusion Recognition
Intrusions Modelling
Alarm
(b)

Figure 1. (a) Pipeline of our Proposed Solution (b) Pipeline of the machine learning technique used for intrusion detection

3.1 *Data generation*

In this subsection, we simulate our MANET and also simulate Blackhole and wormhole attacks in this network and, save the log file as our dataset. To carry out our simulation, we based on NetSim software. Indeed, NetSim is a tool that can be licensed for research, professional, or teaching use [28]. NetSim provides native parsing support, a packet animation, a user-friendly tool to support miscellaneous activities and allows both emulation and simulation [28-30]. So, build a network with 26 nodes and conFigure the routing protocol on AODV (Ad-Hoc On-Demand Vector) between them. Blackhole and Wormhole processes have been installed as applications. Figure 2 presents the Initial position radio characteristics of each node.

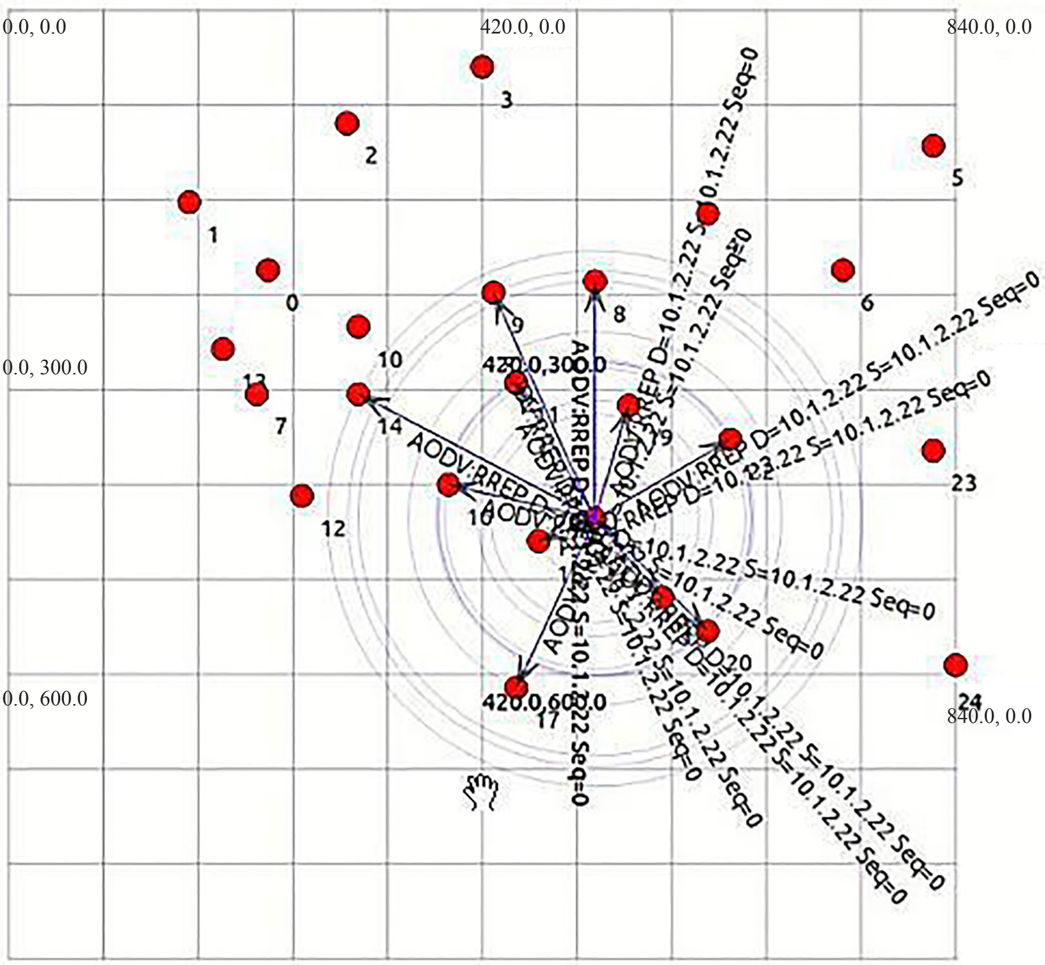


Figure 2. Initial position radio characteristics of each node

If the configurations are ended for all the nodes, we can now perform general properties configuration. Afterwards, we create the simulation network by specifying the simulation time and tick the record animation and then launch the simulation. Figure 3b shows the initial position of each node and their radio characteristics. At the end of the simulation, we can export all the features as a csv file. Table 2 presents the form of our csv file. Our data set contains 21 columns. In Table 2 we present just the head of 15 columns.

Table 2. Head of the generated dataset

duration	protocol	psize	flag	dsn	msn	si	land	mode	neighbor	lflow	avghopcount	nfc	frate	label
0	AODV	84	0	0	0	0	0	0	1	101	0.078049	206	100	normal
0.028136	AODV	84	0	0	1	0	0	0	1	101	0.078049	206	100	normal
0.972864	AODV	84	0	0	0	1	0	0	1	101	0.078049	206	100	normal
0.028136	AODV	84	0	0	1	1	0	0	1	101	0.078049	206	100	normal
0.967864	AODV	84	0	0	0	2	0	0	1	101	0.078049	206	100	normal
0.024136	AODV	84	0	0	1	2	0	0	1	101	0.078049	206	100	normal
0.978864	AODV	84	0	0	0	3	0	0	1	101	0.078049	206	100	normal
0.028136	AODV	84	0	0	1	3	0	0	1	101	0.078049	206	100	normal
0.971864	AODV	84	0	0	0	4	0	0	1	101	0.078049	206	100	normal
0.026136	AODV	84	0	0	1	4	0	0	1	101	0.078049	206	100	normal
0.975864	AODV	84	0	0	0	5	0	0	1	101	0.078049	206	100	normal
0.018136	AODV	84	0	0	1	5	0	0	1	101	0.078049	206	100	normal

3.2 *Data pre-processing*

The aim here is to clean, encode, impute, and standardize our dataset. These steps are classic and simple to deal with.

• **Cleaning:** It consists of deleting variables that have at least 90% of missing values. Our new data set has the dimension (13,480.21 for the blackhole data set and 37,862.21 for the wormhole dataset) and contains respectively 40% attack cases and 60% normal cases for one, and 60% attack cases and 40% normal cases for the other.

• **Encoding:** Here the target and other discrete features are to associate each qualitative value to a numerical value.

• **Imputation:** It consists of deleting or replacing missing values with other values in order to facilitate future operations. In this paper, we replaced the missing values by the mean of elements because of 7% of missed values.

• **Standardization:** It consists of putting all the variables (features and target) under the same scale by making them follow the same law of probability.

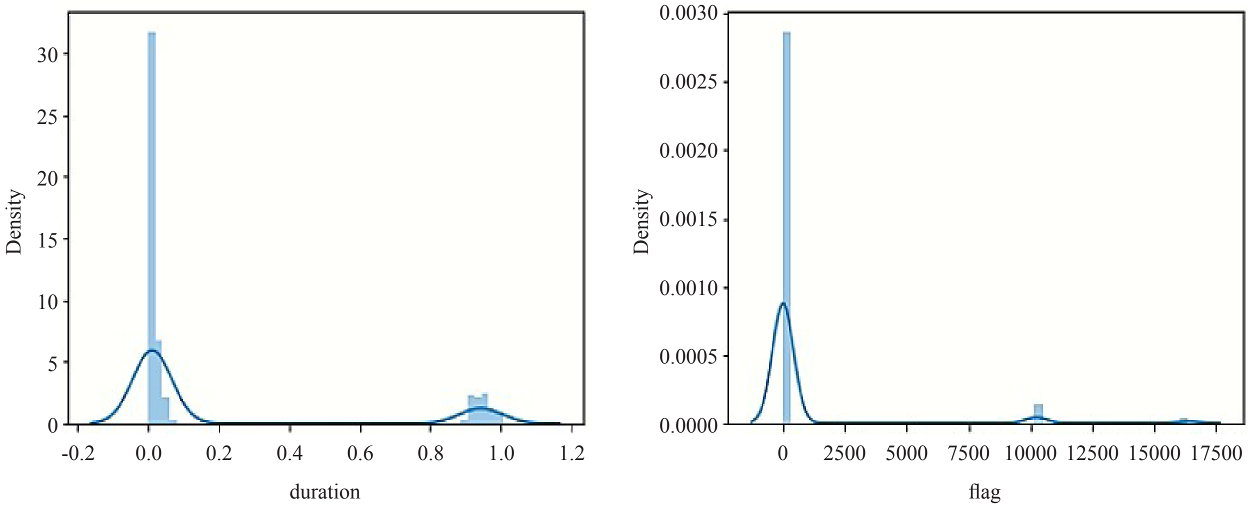
3.3 *Exploratory of Data Analysis (EDA)*

Our analysis is based on the generated dataset. The dataset contains respectively 13,480 and 37,862 entries for blackhole and wormhole network activities. The features of this dataset are obtained after simulations in the AdHoc Network using NetSim software. In total, we have 21 features and the target is represented by the variable Label, which contains attack in case of malignant activity and normal in case of benign activity.

In Table 3, *duration* indicates the transferring time of the packet from source to destination, *flag* shows the status of packets and *hopcount* shows the intermediate nodes. The *Size of packets* is defined in a packet size that includes *header length* in themselves. *Messages* are divided into many categories which are mainly *Route Request, Route Reply, Route Acknowledgment, etc.* A *Neighbor node* is a number of nodes surrounding the node in the communication range. When the sender and originator of the message are the same, then *land* is indicated by Zero Otherwise One. *Unicast* and *broadcast* are two different types of message-transferring modes. *Message sequence number*, originator sequence number, and stream index are generated sender or receiver for uniquely identified packets. The *flow* of the message through the nodes can define the highest flow, lowest flow, and average flow. The *Number of failed connections* and *failure rate* can be computed using the Route Error message [17].

Table 3. Features names and type

No	Feature name	Type	No	Feature name	Type
1	Duration	Continuous	12	Land	Discrete
2	Protocol	Discrete	13	Message sequence number	Continuous
3	Packet size	Continuous	14	Stream index	Continuous
4	Flag	Discrete	15	Highest flow	Continuous
5	Header length	Continuous	16	Average flow	Continuous
6	Hop count	Continuous	17	Lowest flow	Continuous
7	Life time	Continuous	18	Average hop count	Continuous
8	Message type	Discrete	19	Number of failed connections	Continuous
9	Destination sequence number	Continuous	20	Failed connection rate	Continuous
10	Message transfer mode	Discrete	21	Label	Discrete
11	Number of neighbors	Continuous			



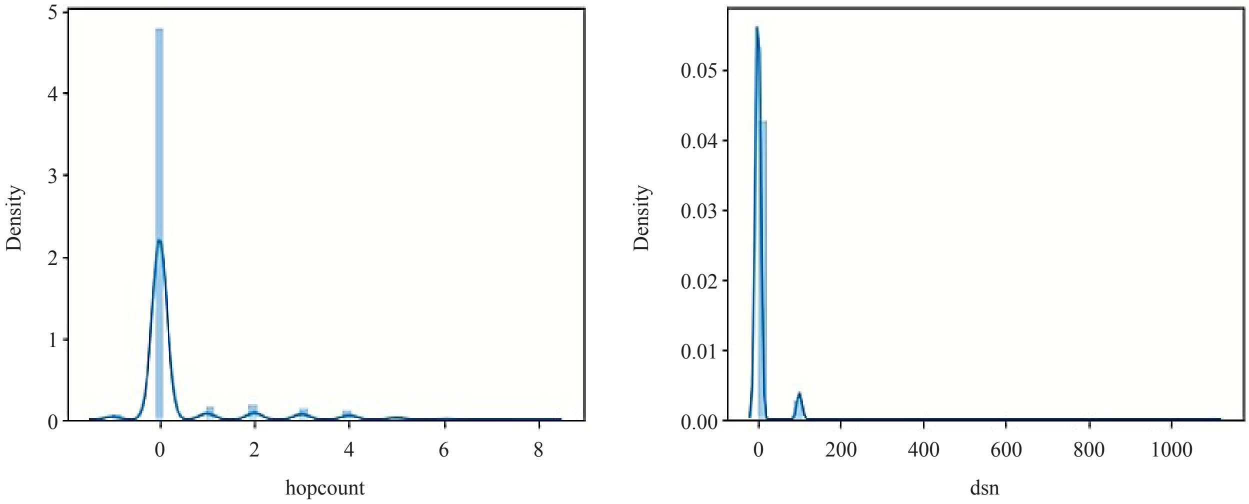


Figure 3. Data Distribution of features Duration, Flag, HCPcount, DSN

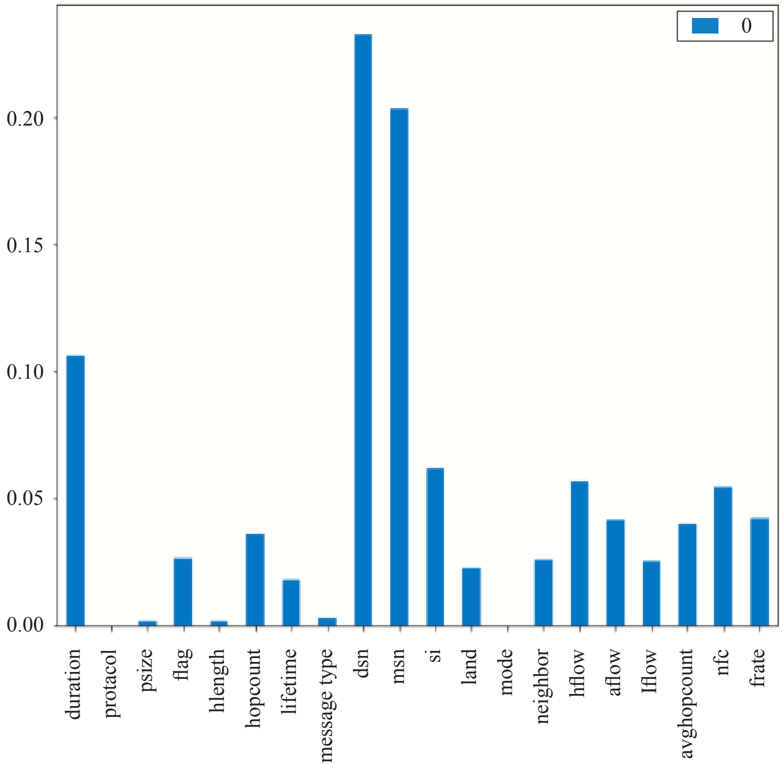


Figure 4. Features Importance with ANOVA F-test

ML needs features that follow the probabilistic normal law [31]; so, we have plotted the feature distribution in order to ensure that all features follow a normal distribution. Figure 3 presents us with some feature distribution of our dataset.

Our dataset contains 20 features plus the target. We used the ANOVA with F-test to see the importance of each feature. Indeed, Analysis of variance (ANOVA) is a statistical technique used to test whether the means of two or more groups are significantly different [32]. ANOVA tests the effect of one or more factors by comparing the means of different samples. ANOVA tests the equality of means using the F-test statistic [31-32]. Based on [31], we can see how to evaluate the feature importance by using ANOVA with the F test. Figure 4 shows us the most important features in our dataset. In this Figure, the value is not different so far because the variation range is between 0.00 to 0.25 which is negligible. ANOVA test tells us that we will use all our features.

3.4 *Modelling*

In this part, we discuss the ML algorithms used to develop our model. We start with five ML algorithms and next, we choose the best one that gives good metrics to our dataset. In this part, 80% of the dataset is used as training data, and, 20% constitutes the test set or data for evaluation or validation. The evaluation criteria used here are accuracy, precision, and recall given in the eq. (1), eq. (2), and eq. (3) [31, 33-35].

$$accuracy(y, y_{pred}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}^{-1}} 1(y_{pred} = y_i) \tag{1}$$

$$recall = \frac{\sum True\_Positive}{\sum True\_Positive + \sum False\_Negative} \tag{2}$$

$$specificity = \frac{\sum True\_Negative}{\sum True\_Negative + \sum False\_Positive} \tag{3}$$

With: $n_{samples}$: The number of samples; y_{pred} : The predicted value of the *i-th* sample; y_i : The corresponding true value. *True_Positive* result indicates a correct identification of a threat, while a *True_Negative* result indicates a correct determination that no threat exists. A *False_Positive* result is an incorrect identification of a threat, and a *False_Negative* result is a failure to identify a threat.

In this paper, we implement Random Forests (RF). Indeed, we implement five ML algorithms (Support Vector Machine, Logistic Regression, K-Nearest Neighbors, Random Forests and Decision Trees). But, our best algorithm among them is RF.

RF classifier is an ensemble method that trains multiple decision trees in parallel with bootstrapping and subsequent aggregation, collectively known as bagging [36-38]. RF merges the decisions of multiple decision trees in order to find an answer, which represents the average of all these decision trees (predictions from all trees are pooled to make the final prediction) [38]. We based on the *Gini index* to perform RF on the classification dataset. Its formula is given by the eqs. (4) and (5) [39].

$$GiniIndex = 1 - \sum^{C} (P_i)^2 \tag{4}$$

$$= 1 - \left[(P_+)^2 + (P_-)^2 \right] \tag{5}$$

Eq. (5) uses the class and probability to determine the *Gini* of each branch on a node, determining which of the branches is more likely to occur. Here, P_i represents the relative frequency of the class you are observing in the dataset, C represents the number of classes, P_+ represents the probability of a positive class and P_- represents the probability of a negative class [39].

Eq. (6) shows us that we can also use entropy to determine how nodes branch in a decision tree [37-39].

$$Entropy = \sum_{i=1}^{C} -p_i * \log_2(p_i) \tag{6}$$

Entropy uses the probability of a given outcome to decide how a node should branch. Unlike the Gini index, it is more mathematically intensive due to the use of a logarithmic function in the calculation [37-39]. In this paper, we implement the RF Model with the eqs. (5) and (6).

4. Results and discussions

In this section, we present the obtained results in a simple and realistic way. Interpretation of the results is discussed to put them in context and, explain why they are important.

4.1 *Results*

As explained in the methodology section, we train our model by using five algorithms. Table 4 presents us with the evaluation metric of all five algorithms. In Table 4, we can easily see why we chose Random Forest to build our proposed application.

Table 4. comparative results of our fives used algorithms

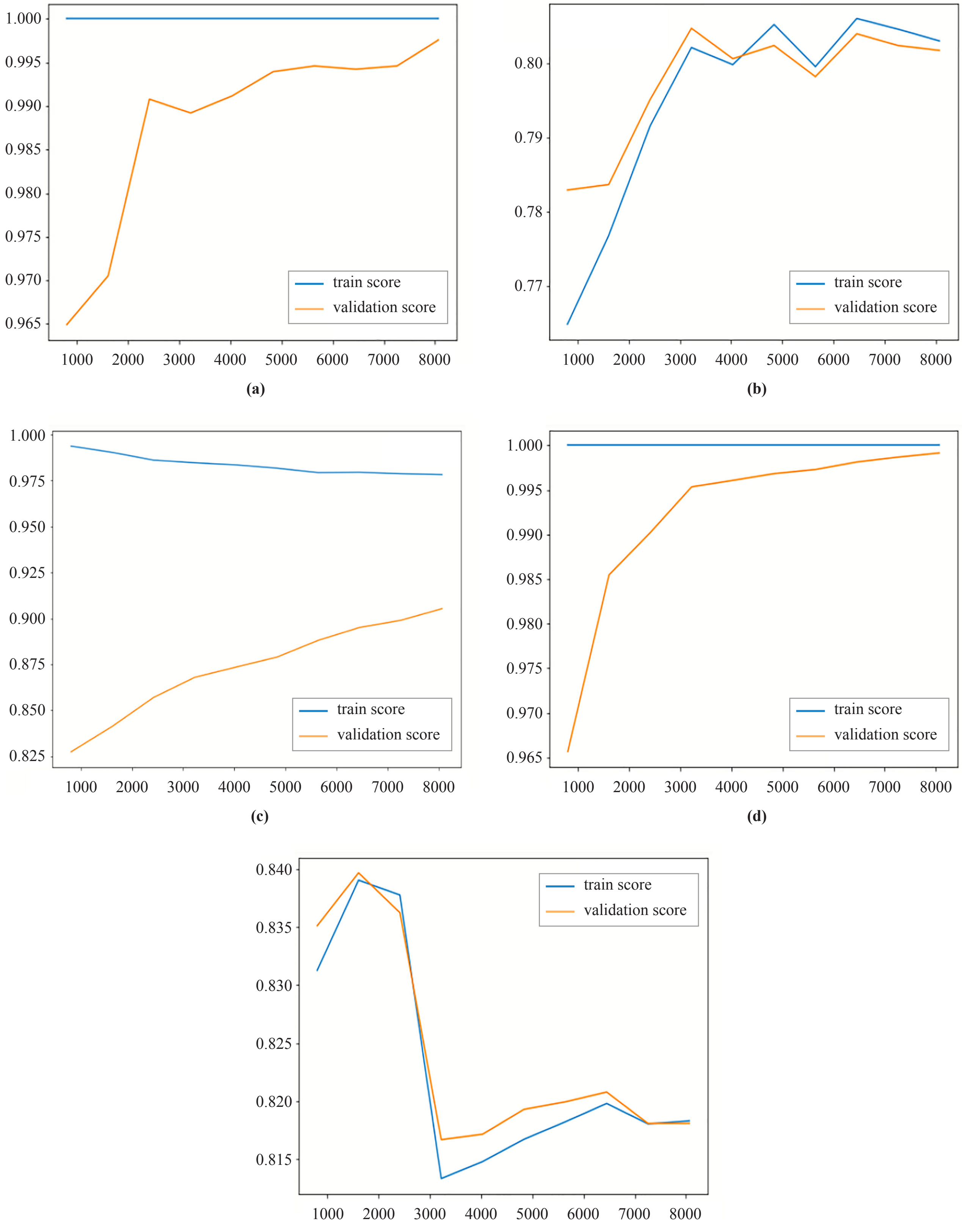
Model	True Negative	False Negative	False Positive	True Positive	Precision
Logistic Regression	2106	552	5	23	0.80
Support Vector Machine	1991	401	120	174	0.71
K-Nearest-Neighbors	1781	456	22	427	0.75
Decision Tree	2111	3	0	572	0.98
Random Forest	2107	1	4	574	0.99

The learning curves of the different algorithm on the model are presented on Figure 5.

Figure 5 permits us to notice that the Decision Tree (DT) and Random Forest (RF) are almost the same. These results can be explained because an RF is a set of many DTs. Also, an RF is more stage than a decision tree. To ensure that we are not faced with overfitting, we apply the cross-validation method to our RF model. By applying the GridSearchCV method, we obtained the same learning curves presented in Figure 5a. The RF model is used to develop our application. We based on Flask and developed our application namely *APP-IPS.* To launch our application, we need to start the Flask service by running in prompt command the following commands:

- *set FLASK_APP = main.py*
- *set FLASK_DEBUG = 1*
- *Flask run*

When the service is run, we need to copy and paste the displayed address to the navigator. After logging successfully, we have the main page that gives us the eventuality to train or re-train the model before the prediction and, the prevention. Figure 6 presents the interfaces for these operations.



(a)

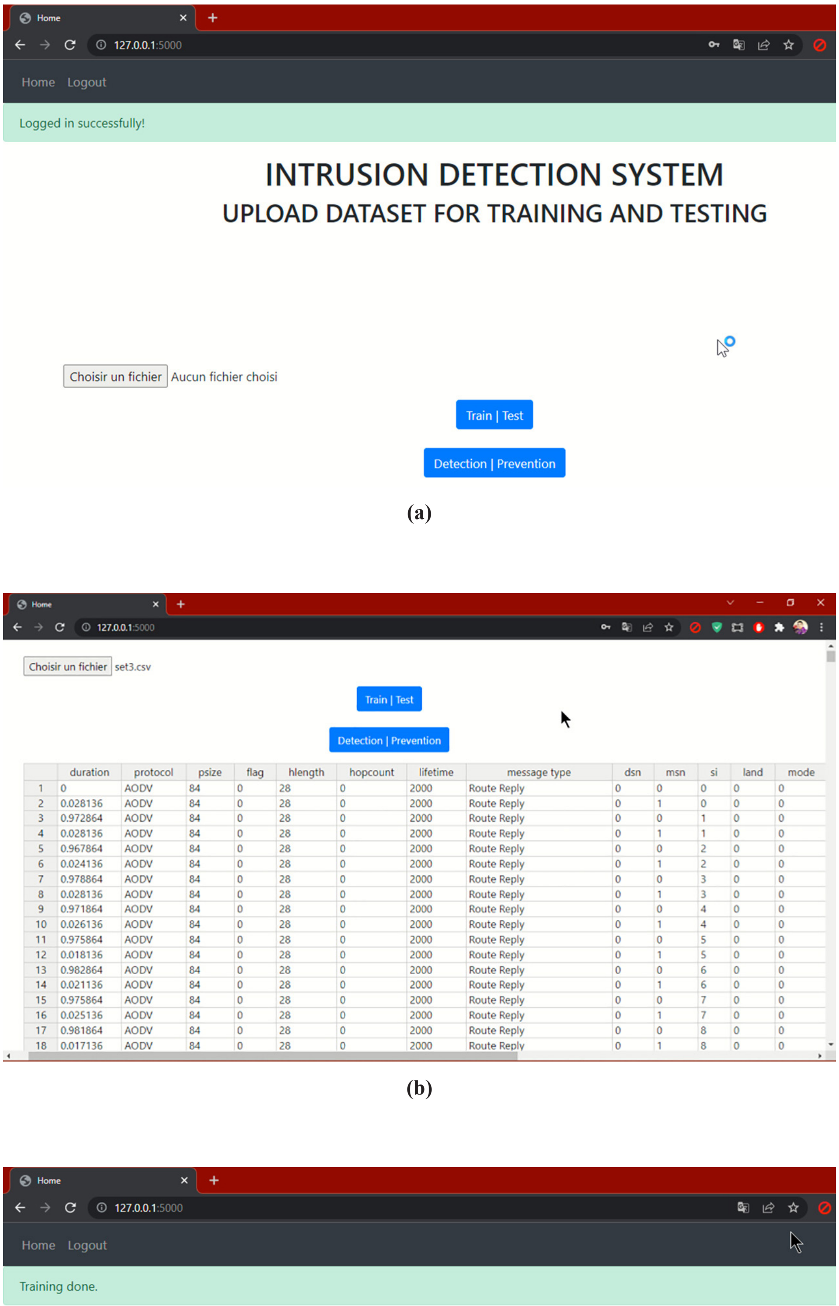
(b)

(c)

(d)

(e)

Figure 5. Learning curves of each algorithm, (a): Random Forest, (b): Logistic Regression, (c): Support Vector Machine, (d): Decision Tree, (e): K-Nearest Neighbor



Logged in successfully!

INTRUSION DETECTION SYSTEM

UPLOAD DATASET FOR TRAINING AND TESTING

(a)

	duration	protocol	psize	flag	hlength	hopcount	lifetime	message type	dsn	msn	si	land	mode
1	0	AODV	84	0	28	0	2000	Route Reply	0	0	0	0	0
2	0.028136	AODV	84	0	28	0	2000	Route Reply	0	1	0	0	0
3	0.972864	AODV	84	0	28	0	2000	Route Reply	0	0	1	0	0
4	0.028136	AODV	84	0	28	0	2000	Route Reply	0	1	1	0	0
5	0.967864	AODV	84	0	28	0	2000	Route Reply	0	0	2	0	0
6	0.024136	AODV	84	0	28	0	2000	Route Reply	0	1	2	0	0
7	0.978864	AODV	84	0	28	0	2000	Route Reply	0	0	3	0	0
8	0.028136	AODV	84	0	28	0	2000	Route Reply	0	1	3	0	0
9	0.971864	AODV	84	0	28	0	2000	Route Reply	0	0	4	0	0
10	0.026136	AODV	84	0	28	0	2000	Route Reply	0	1	4	0	0
11	0.975864	AODV	84	0	28	0	2000	Route Reply	0	0	5	0	0
12	0.018136	AODV	84	0	28	0	2000	Route Reply	0	1	5	0	0
13	0.982864	AODV	84	0	28	0	2000	Route Reply	0	0	6	0	0
14	0.021136	AODV	84	0	28	0	2000	Route Reply	0	1	6	0	0
15	0.975864	AODV	84	0	28	0	2000	Route Reply	0	0	7	0	0
16	0.025136	AODV	84	0	28	0	2000	Route Reply	0	1	7	0	0
17	0.981864	AODV	84	0	28	0	2000	Route Reply	0	0	8	0	0
18	0.017136	AODV	84	0	28	0	2000	Route Reply	0	1	8	0	0

(b)

Training done.

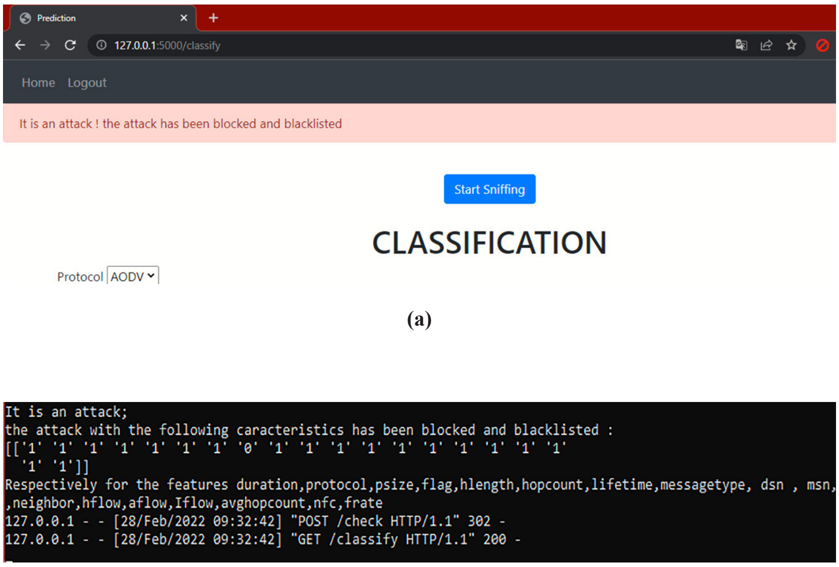
INTRUSION DETECTION SYSTEM

UPLOAD DATASET FOR TRAINING AND TESTING

(c)

Figure 6. APP-IPS (a): Main page, (b): Overview of dataset, (c): Training and Testing done, (d): Classification

In Figure 6, we can see that, after selecting the dataset (Figure 6a), we can previsualization the dataset before moving to the train and test part (Figure 6b). when the train and test were done, we received a notification (Figure 6c). We can use the model to perform the classification of data (Figure 6d). The system takes all the values in the fields of any packet that transits and, creates an instance of the prediction function which prints either attack in case of attack or normal in case of normal activity (Figure 7).



It is an attack ! the attack has been blocked and blacklisted

Protocol AODV

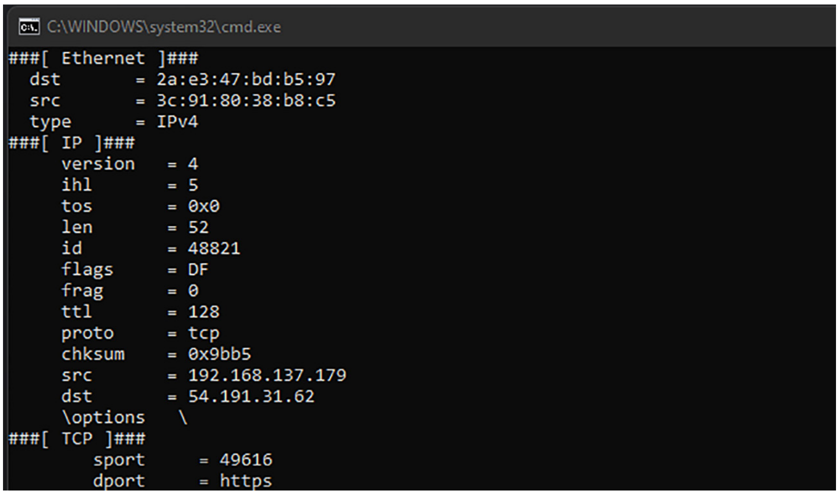
(a)

```
It is an attack;
the attack with the following caracteristics has been blocked and blacklisted :
[['1' '1' '1' '1' '1' '1' '1' '0' '1' '1' '1' '1' '1' '1' '1' '1' '1' '1'
  '1' '1']]
Respectively for the features duration,protocol,psize,flag,hlength,hopcount,lifetime,messagetype, dsn , msn,
,neighbor,hflow,aflow,Iflow,avghopcount,nfc,frate
127.0.0.1 - - [28/Feb/2022 09:32:42] "POST /check HTTP/1.1" 302 -
127.0.0.1 - - [28/Feb/2022 09:32:42] "GET /classify HTTP/1.1" 200 -
```

(b)

Figure 7. Classification and IPS action to anomaly detection. (a): Classifying before IPS action, (b): IPS reaction

Figure 7b presents the reaction of the application background. We can notice that the attack has been blocked and blacklisted. Figure 8 presents an example of capturing on a server.



```
###[ Ethernet ]###
  dst       = 2a:e3:47:bd:b5:97
  src       = 3c:91:80:38:b8:c5
  type      = IPv4
###[ IP ]###
     version   = 4
     ihl       = 5
     tos       = 0x0
     len       = 52
     id        = 48821
     flags     = DF
     frag      = 0
     ttl       = 128
     proto     = tcp
     chksum    = 0x9bb5
     src       = 192.168.137.179
     dst       = 54.191.31.62
     \options   \
###[ TCP ]###
        sport     = 49616
        dport     = https
```

Figure 8. Capturing packets from our proposed application (*APP-IPS*)

4.2 *Discussions*

In this paper, we proposed a method based on ML to detect and predict Blackhole and wormhole attacks to a

MANET. We first design our MANET network in NETSIM software using 26 nodes. We did an attack in this network and generated a dataset. This dataset has been used to build our ML model. We constructed an application that allowed us to predict and block blackhole and wormhole attacks on our network. Although the good performance of our results, we need to compare them to see the effectiveness of our proposed method. However, we perform only subjective comparisons because we haven't used the same database as those in the literature. Indeed, most of the projects done in this domain have been performed by using KDD or IRIS datasets. It's the case of Prasad et al. in [17] that based on their generated dataset got after simulation, and, obtained a precision of their proposed model of 80%. Also, Sebopelo et al., based on the IRIS dataset in MANET obtained a model accuracy equal to 100% [40]. It's important to notice that Sebopelo et al., only detect malicious nodes in MANET and cannot identify the type of attack. Sebopelo et al. performed binary classification. They classified packet data in MANET as either abnormal or normal and, for that reason, they can reach a high precision and accuracy. Gad et al. worked on a variant of MANET, VANET, and based on the KDD dataset with a multiclass classification, they reached an accuracy of 98.3% and a precision of 98.3% using XGboost as the best ML algorithm [41]. Meddeb et al. in [42] and [43] worked also in MANET and performed multiclass classification on their own generated dataset and obtained encouraging results. However, other authors proposed security by using blockchain and encryption methods in communication networks [44-46]. Table 3 summarizes and compares our results with those in the literature. Figure 9 represents the graphical version of Table 3 for a good comparison. Based on Table 5, and, to the best of our knowledge, no literature work did not propose a real-time application to detect an attack in the MANET network. This is a good advantage of our proposed paper.

Table 5. Comparison Table of proposed work based on Data based, algorithms, Precision and accuracy

Authors	Data Base	ML Algorithms	Precision (%)	Accuracy (%)
Prasad et al. [17]	Their Generated dataset	Naive Bayes (NB)	94	93.06
Sebopelo et al. [40]	IRIS	Logistic Regression (LR)	100	100
Gad et al. [41]	KDD	XGBoost	98.3	98.3
Meddeb et al. [42]	Their Generated dataset	Fuzzy-KNN	-	94
Meddeb et al. [43]	Their Generated dataset	SVM	-	96.2
US	Our Generated Dataset	RF	99.8	99.8

In terms of future works, there is a need to implement real-time intrusion detection systems for MANETs using ML algorithms in a local company. This requires developing a pipeline of our solution and describing all dependencies. Furthermore, researchers should focus on developing methods that can detect all the attacks and sophisticated attacks, such as those that use advanced evasion techniques. This requires that we will need very large data and use deep learning instead of machine learning for a robust model.

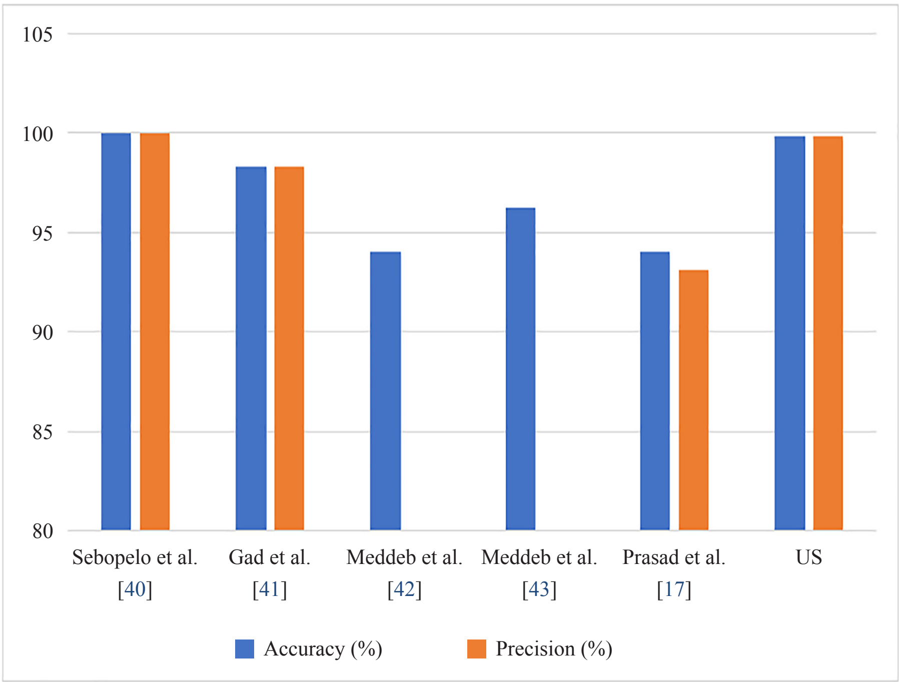


Figure 9. Comparison graph of proposed work based on Precision and accuracy

5. Conclusion

Decentralized wireless networks are one of the options of the future in terms of cost (without infrastructure) and connectivity. In contrast to these advantages, these networks are subject to vulnerability to various attacks. This paper proposed a method for modeling blackhole and wormhole attacks using machine learning methods in MANET. We described different attacks in MANET and, based on the literature, chose to work on the two famous attacks blackhole and Wormhole. After modeling of this MANET and the configuration of all the nodes, we launched attacks and recorded data. The obtained dataset is used. We tested five ML algorithms and chose RF as our best one. RF gave us 99.8% of precision and accuracy and, is used to construct our application to detect and predict Blackhole and wormhole attacks in MANET. The proposed method introduced in this paper can be applied to any wireless network, but it is particularly relevant for companies that use Ad-hoc networks for communication. This paper focuses on two types of attacks (blackhole and wormhole), which may not cover all possible attacks that can occur in MANETs. This can be a limitation of the paper. As the number of data continues to grow in communication networks, we propose in the future to use Deep Learning along with autoencoder to develop a robust model for our real-time application in MANET.

Acknowledgements including declarations

Acknowledgements: The authors would like to acknowledge and thank InchTech's team (www.inchtechs.com) for their support and assistance during the conception of that work.

Funding and competing interests: We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Ethical approval: This article does not contain any studies with human participants and/or animals performed by any of the authors.

Conflict of interest

The authors declare no competing financial interest.

References

[1] Alhaidari FA, Alrehan AM. A simulation work for generating a novel dataset to detect distributed denial of service attacks on Vehicular Ad hoc Network systems. *International Journal of Distributed Sensor Networks*. 2021; 17(3): 1-25. Available from: https://doi.org/10.1177/15501477211000287.

[2] Aluvala S, Sekhar R, Vodnala D. An empirical study of routing attacks in mobile Ad-hoc networks. *Procedia Computer Science*. 2016; 92: 554-561. Available from: https://doi.org/10.1016/j.procs.2016.07.382.

[3] Tseng FH, Chou LD, Chao HC. A survey of black hole attacks in wireless mobile ad hoc networks. *Human-Centric Computing and Information Sciences*. 2011; 1(1): 4.

[4] Abdelhaq M, Alsaqour R, Abdelhaq S. Securing mobile ad hoc networks using danger theory-based artificial immune algorithm. *PLoS ONE*. 2015; 10(5): e0120715. Available from: https://doi.org/10.1371/journal.pone.0120715.

[5] Anusha K, Sathiyamoorthy E. A new trust-based mechanism for detecting intrusions in MANET. *Information Security Journal: A Global Perspective*. 2017; 26(4): 153-165. Available from: https://doi.org/10.1080/19393555.2017.1328544.

[6] Subba B, Biswas S, Karmakar S. Intrusion detection in mobile Ad-hoc networks: Bayesian game formulation. *Engineering Science and Technology, an International Journal*. 2016; 19: 782-799.

[7] Amiri E, Keshavarz H, Heidari H, Mohamadi, E, Moradzadeh H. Intrusion detectionsystems in MANET: A review. *Procedia-Social and Behavioral Sciences*. 2014; 129: 453-459. Available from: https://doi.org/10.1016/j.sbspro.2014.03.700.

[8] Thanuja R, Umamakeswari A. Unethical network attack detection and prevention using fuzzy based decision system in mobile Ad-hoc networks. *Journal of Electrical Engineering and Technology*. 2018; 13(5): 2086-2098. Available from: https://doi.org/10.5370/JEET.2018.13.5.2086.

[9] Alheeti KMA, Gruebler A, McDonald-Maier K. Intelligent intrusion detection of grey hole and rushing attacks in self-driving vehicular networks. *Computers*. 2016; 5(3): 16. Available from: https://doi.org/10.3390/computers5030016.

[10] Popli R, Sethi M, Kansal I, Garg A, Goyal N. Machine learning based security solutions in MANETs: State of the art approaches. *Journal of Physics: Conference Series*. 2021; 1950: 012070. Available from: https://doi.org/10.1088/1742-6596/1950/1/012070.

[11] Imran M, Khan FA, Jamal T, Durad MH. Analysis of detection features for wormhole attacks in MANETs. *Procedia Computer Science*. 2015; 56: 384-390. Available from: https://doi.org/10.1016/j.procs.2015.07.224.

[12] Ezhilarasi M, Gnanaprasanambikai L, Kousalya A, Shanmugapriya M. A novel implementation of routing attack detection scheme by using fuzzy and feed-forward neural networks. *Soft Computing*. 2022; 27: 4157-4168. Available from: https://doi.org/10.1007/s00500-022-06915-1.

[13] Prasad M, Tripathi S, Dahal K. An enhanced detection system against routing attacks in mobile Ad-hoc network. *Wireless Networks*. 2022; 28: 1411-1428. Available from: https://doi.org/10.1007/s11276-022-02913-1.

[14] Shukla M, Joshi K, Singh U. Mitigate wormhole attack and blackhole attack using elliptic curve cryptography in MANET. *Wireless Personal Communications*. 2021; 121: 503-526. Available from: https://doi.org/10.1007/s11277-021-08647-1.

[15] Malik TS, Siddiqui MN, Mateen M, Malik KR, Sun S, Wen J. Comparison of blackhole and wormhole attacks in cloud MANET enabled IoT for agricultural field monitoring. *Security and Communication Networks*. 2022; 2022: 4943218. Available from: https://doi.org/10.1155/2022/4943218.

[16] Siddiqui MN, Malik KR, Malik TS. Performance analysis of blackhole and wormhole attack in MANET based IoT. *2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2)*. Islamabad, Pakistan: IEEE; 2021. p.1-8. Available from: https://doi.org/10.1109/ICoDT252288.2021.9441515.

[17] Prasad M, Tripathi S, Dahal K. *Wormhole attack detection in ad hoc network using machine learning technique*. *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*. Kanpur, India: IEEE; 2019. p.1-7. Available from: https://doi.org/10.1109/ICCCNT45670.2019.8944634.

[18] Shams EA, Rizaner A, Ulusoy AH. Trust aware support vector machine intrusion detection and prevention system

in vehicular ad hoc networks. *Computers & Security*. 2018; 78: 245-254. Available from: https://doi.org/10.1016/j.cose.2018.06.008.

[19] Alghamdi SA. Novel trust-aware intrusion detection and prevention system for 5G MANET-Cloud. *International Journal of Information Security*. 2022; 21: 469-488. Available from: https://doi.org/10.1007/s10207-020-00531-6.

[20] Kumar S, Dutta K. Intrusion detection in mobile ad hoc networks: techniques, systems, and future challenges. *Security and Communication Networks*. 2016; 9(14): 2484-2556. Available from: https://doi.org/10.1002/sec.1484.

[21] Khraisat A, Alazab A. A critical review of intrusion detection systems in the internet of things: Techniques, deployment strategy, validation strategy, attacks, public datasets and challenges. *Cybersecurity*. 2021; 4(1): 18. Available from: https://doi.org/10.1186/s42400-021-00077-7.

[22] Sivanesh S, Dhulipala S. Accurate and Cognitive Intrusion Detection System (ACIDS): A novel black hole detection mechanism in mobile ad hoc networks. *Mobile Networks and Applications*. 2021; 26: 1696-1704. Available from: https://doi.org/10.1007/s11036-019-01505-2.

[23] Mohanapriya M, Krishnamurthi I. Modified DSR protocol for detection and removal of selective black hole attack in MANET. *Computers and Electrical Engineering*. 2014; 40: 530-538. Available from: https://dx.doi.org/10.1016/j.compeleceng.2013.06.001.

[24] Hassan Z, Mehmood A, Maple C, Khan MA, Aldegheishem A. Intelligent detection of black hole attacks for secure communication in autonomous and connected vehicles. *IEEE Access*. 2020; 8: 199618-199628. Available from: https://doi.org/0.1109/ACCESS.2020.3034327.

[25] Meddeb R, Jemili F, Triki B, Korbaa O. Anomaly-based behavioral detection in mobile Ad-hoc networks. *Procedia Computer Science*. 2019; 159: 77-86. Available from: https://doi.org/10.1016/j.procs.2019.09.162.

[26] Abdan M, Seno SAH. Machine learning methods for intrusive detection of wormhole attack in Mobile Ad-hoc Network (MANET). *Wireless Communications and Mobile Computing*. 2022; 2022: 2375702. Available from: https://doi.org/10.1155/2022/2375702.

[27] Joon D, Chopra K. Hybrid deep learning prediction model for blackhole attack protection in wireless communication. *Natural Volatile and Essential Oils*. 2021; 8(4): 10228-10243.

[28] Campanile L, Gribaudo M, Iacono M, Marulli F, Mastroianni M. Computer network simulation with ns-3: A systematic literature review. *Electronics*. 2020; 9(2): 272. Available from: https://doi.org/10.3390/electronics9020272.

[29] Patel KN, Jhaveri RH. A survey on emulation testbeds for mobile Ad-hoc networks. *Procedia Computer Science*. 2015; 45: 581-591. Available from: https://doi.org/10.1016/j.procs.2015.03.111.

[30] Dorathy I, Chandrasekaran M. Simulation tools for mobile ad hoc networks: A survey. *Journal of Applied Research and Technology*. 2018; 16(5): 437-445. Available from: https://doi.org/10.22201/icat.16656423.2018.16.5.739.

[31] Kouanou AT, Attia TM, Djeumo AF, Mouelas AN, Nzogang MP, Tchapga CT, et al. An overview of data analysis and machine learning for Covid-19 detection. *Journal of Healthcare Engineering*. 2021; 2021: 4733167. Available from: https://doi.org/10.1155/2021/4733167.

[32] Cleophas TJ, Zwinderman AH. Analysis of Variance (Anova). *Regression Analysis in Medical Research*. Springer, Cham; 2021. Available from: https://doi.org/10.1007/978-3-030-61394-5_7.

[33] Tchapga CT, Mih TA, Kouanou AT, Fonzin TF, Fogang PK, Mezatio BA, et al. Biomedical image classification in a big data architecture using machine learning algorithms. *Journal of Healthcare Engineering*. 2021; 2021: 9998819. Available from: https://doi.org/10.1155/2021/9998819.

[34] Kouanou AT, Tchiotsop D, Kengne R, Zephirin DT, Armele NMA, Tchinda R. An optimal big data workflow for biomedical image analysis. *Informatics in Medicine Unlocked*. 2018; 11: 68-74. Available from: https://doi.org/10.1016/j.imu.2018.05.001.

[35] Alla Takam C, Samba O, Tchagna Kouanou A, Tchiotsop D. Spark architecture for deep learning-based dose optimization in medical imaging. *Informatics in Medicine Unlocked*. 2020; 29: 1-13. Available from: https://doi.org/10.1016/j.imu.2020.100335.

[36] Misra S, Li H. Noninvasive fracture characterization based on the classification of sonic wave travel times. *Machine Learning for Subsurface Characterization*. 2020; 243-287. Available from: https://doi.org/10.1016/B978-0-12-817736-5.00009-0.

[37] Best K, Gilligan J, Baroud H, Carrico A, Donato K, Mallick B. Applying machine learning to social datasets: A study of migration in southwestern Bangladesh using random forests. *Regional Environmental Change*. 2022; 22: 52. Available from: https://doi.org/10.1007/s10113-022-01915-1.

[38] Wang S, Aggarwal C, Liu H. Random-forest-inspired neural networks. *ACM Transactions on Intelligent Systems and Technology*. 2018; 9(6): a69. Available from: https://doi.org/10.1145/3232230.

[39] Algehyne EA, Jibril ML, Algehainy NA, Alamri OA, Alzahrani AK. Fuzzy neural network expert system with

an improved gini index random forest-based feature importance measure algorithm for early diagnosis of breast cancer in saudi arabia. *Big Data and Cognitive Computing*. 2022; 6: 13. Available from: https://doi.org/10.3390/bdcc6010013.

[40] Sebopelo R, Isong B, Gasela N. Identification of compromised nodes in MANETs using machine learning technique. *International Journal of Computer Network and Information Security*. 2019; 1: 1-10. Available from: https://doi.org/10.5815/ijcnis.2019.01.01.

[41] Gad A, Nashat A, Barkat T. Intrusion detection system using machine learning for vehicular ad hoc networks based on ToN-IoT dataset. *IEEE Access*. 2021; 9: 142206-142217. Available from: https://doi.org/10.1109/ACCESS.2021.3120626.

[42] Meddeb R, Triki B, Jemili F, Korbaa O. Dataset for intrusion detection in mobile Ad-hoc networks. In: Abraham A, Siarry P, Ma K, Kaklauskas A. (eds.) *Intelligent Systems Design and Applications. ISDA 2019. Advances in Intelligent Systems and Computing, vol 1181*. Springer, Cham; 2021. Available from: https://doi.org/10.1007/978-3-030-49342-4_3.

[43] Meddeb R, Jemili F, Triki B, Korbaa O. Anomaly-based behavioral detection in mobile Ad-hoc network. *Procedia Computer Science*. 2019; 159: 77-86. Available from: https://doi.org/10.1016/j.procs.2019.09.162.

[44] Ploder C, Spiess T, Bernsteiner R, Dilger T, Weichelt R. A risk analysis on blockchain technology usage for electronic health records. *Cloud Computing and Data Science*. 2021; 2(2): 1-16. Available from: https://doi.org/10.37256/ccds.222021777.

[45] Khasim S, Basha SS. An improved fast and secure CAMEL based authenticated key in smart health care system. *Cloud Computing and Data Science*. 2022; 3(2): 77-91. Available from: https://doi.org/10.37256/ccds.3220221423.

[46] Tchagna Kouanou A, Tchapga CT, Sone Ekonde M, Monthe V, Mezatio BA, Manga J, et al. Securing data in an internet of things network using blockchain technology: Smart home case. *SN Computer Science*. 2022; 3: 167. Available from: https://doi.org/10.1007/s42979-022-01065-5.