EAI Endorsed Transactions
on Scalable Information Systems
Research Article EAI.EU

Design of Intrusion Detection and Prevention Model Using COOT Optimization and Hybrid LSTM-KNN Classifier for MANET

G. Madhu*

Asst. Professor, Computer science Department, MVSR Engineering College, Affiliated to Osmania University, Hyderabad, Telangana 501510, India.

Abstract

INTRODUCTION: MANET is an emerging technology that has gained traction in a variety of applications due to its ability to analyze large amounts of data in a short period of time. Thus, these systems are facing a variety of security vulnerabilities and malware assaults. Therefore, it is essential to design an effective, proactive and accurate Intrusion Detection System (IDS) to mitigate these attacks present in the network. Most previous IDS faced challenges such as low detection accuracy, decreased efficiency in sensing novel forms of attacks, and a high false alarm rate.
OBJECTIVES: To mitigate these concerns, the proposed model designed an efficient intrusion detection and prevention model using COOT optimization and a hybrid LSTM-KNN classifier for MANET to improve network security.
METHODS: The proposed intrusion detection and prevention approach consist of four phases such as classifying normal node from attack node, predicting different types of attacks, finding the frequency of attack, and intrusion prevention mechanism. The initial phases are done through COOT optimization to find the optimal trust value for identifying attack nodes from normal nodes. In the second stage, a hybrid LSTM-KNN model is introduced for the detection of different kinds of attacks in the network. The third stage performs to classify the occurrence of attacks.
RESULTS: The final stage is intended to limit the number of attack nodes present in the system. The proposed method's effectiveness is validated by some metrics, which achieved 96 per cent accuracy, 98 per cent specificity, and 35 seconds of execution time.
CONCLUSION: This experimental analysis reveals that the proposed security approach effectively mitigates the malicious attack in MANET.

Keywords: Intrusion detection & prevention; Ad hoc network security; MANETs; COOT optimization; hybrid KSTM-KNN; FIS; two-factor authentication; DNA cryptography.

Received on 20 August 2022, accepted on 17 December 2022, published on 27 December 2022

doi: 10.4108/eetsis.v10i3.2574

1. Introduction

MANETs are a network of self-contained gadgets with no infrastructure that also serves as intermediary routers. MANETS contain large applications in numerous fields recently. Reactive and proactive protocols are the two routing protocols in MANET [1]. For that reason, they have reduced routing overhead. Further, DSR and AODV are regarded as more effective and scalable than proactive counterparts like OLSR. DSR and AODV have predicated on the premise that all nodes in the network are trustworthy of one another and that no intruder attacker nodes are present. As a result, the occurrence of any nodes poses a risk to its security [2]. Malicious nodes can cause severe disruption by conducting a series of threats, such as routing and data dissemination assaults [3]. The two most common types of attacks are passive and active assaults. The attacker does not

* Corresponding author. Email: madhug.mvsrec@gmail.com

interfere with the network's functionality but instead seeks out important information in passive attacks. Active attacks produced varying degrees of network damage depending on the sort of attack [4]. As a result, several researchers focused on providing security agencies in MANETs, even though security is the most important factor in MANET application acceptance. Intruders can affect the network's operation via attacking any of the layers like the network layer, physical layer or MAC layer, making MANETs vulnerable in their functionality [5]. Because traditional data security techniques like encryption and authentication cannot give complete protection, Intrusion Detection and Prevention (IDP) technologies are widely utilized for safeguarding MANETs [6].

The IDP system can be widely used to provide more security in MANETs and to defend nodes from routing threats. Intrusion detection (ID) mechanisms are categorized into two categories: anomaly-based intrusion detection (ABID) and knowledge-based intrusion detection (KBID) [7]. KBID is a level of expertise that only contains known attack patterns or signatures, and it searches for these patterns in an attempt to detect them. Testing and training are the two stages of ABID. This method not only detects potential intrusions but can also detect efforts to utilize known vulnerabilities. Nevertheless, it is more likely to produce false positives than KBID [8]. Because of the shift in networking technology from fixed to wireless networks over the last decade, intrusion identification and monitoring have become one of the primary levels of defence. Because of the difficulty in meeting the needs of IDS, ID in MANETs is more sophisticated as well as demanding than in fixed networks, and some MANET characteristics create operational and implementation complexities [9]. Because of the mobility of the nodes, the network topology is dynamic and unpredictable, making intrusion detection difficult. In addition, due to the limited processing capabilities of most nodes in MANETs, intrusion detection systems (IDSs) are more sophisticated.

The primary goal of this current study is to develop an effective intrusion detection technique for predicting various types of attacks in MANET in order to improve network security. Machine Learning (ML) processes have been extensively researched in order to develop IDSs capable of performing optimally in order to meet network security requirements [10]. Machine learning (ML) is a growing field of computational algorithms that mimic human intelligence. ML approaches make use of statistical models to find patterns in massive amounts of data [11]. In the context of IDS research and development, the following ML techniques have been extensively used such as Artificial Neural Networks (ANNs), Decision Tree (DT), k-Nearest Neighbors (k-NN), Naive Bayes (NB), Support Vector Machine (SVM) and Random Forest (RF) [12]. Emergence of wireless communication technologies, wireless handheld devices, Wireless Computer Networks (WCN) has allowed for processing large volumes of data at a given moment. As a result, these systems are subjected to a myriad of malicious attacks and security vulnerabilities. It is therefore imperative to develop Intrusion Detection Systems (IDS) that are proactive, efficient and accurate in mitigating these attacks. An IDS is the first line of defence in a computer network. At the highest level, IDSs are categorized as: Host-based IDS (H-IDS) and Network-based IDS (N-IDS). The N-IDS operates in a distributed manner within a network system, and the H-IDS runs on a host computer or system. Additionally, both types of IDSs operate using two methods, namely, anomaly-based detection and signature-based recognition. Traditional IDSs suffer from shortcomings such as a high false alarm rate, a reduced efficiency in detecting novel forms of attacks and a low detection accuracy. It is therefore imperative to design robust IDSs that can improve the detection accuracy, decrease the false alarm rate and increase the efficiency in the discovery of new types of attacks. The present research work focuses on Deep Learning (DL) to tackle the issues that arise in the prior methods. DL is a subset of machine learning that deals with algorithms inspired by the structure, function, and operation of biological neurons in the brain. In this current research, a hybrid classifier is designed by coupling deep learning with machine learning for attaining effective prediction. The main contribution of the present work is organized as follows,

- An effective intrusion detection technique is designed for predicting different kinds of attacks in MANET to improve network security.
- The normal and attack nodes are identified based on the trust value. It is selected based on the COOT optimization algorithm.
- A hybrid LSTM-KNN classification approach is designed for attaining an effective prediction of different kind’s attacks in the MANET.
- The fuzzy interference system (FIS) is utilized in order to classify the occurrence of attacks, whether it is frequent or rare.
- Through a two-factor authentication technique, the intrusion protection process is designed to limit the direct exposure of attack nodes in the system.

The following sections of this manuscript are structured as below, and section 2 illustrates certain research articles related to existing methods used for intrusion detection and prevention. Section 3 briefly explains the description of the proposed intrusion detection and prevention methodology. The obtained results and the performance metrics values attained in this proposed work are illustrated in section 4. Finally, the complete research work is concluded in section 5.

2. Related Work

Various researches had developed by many authors to overcome the security threats in MANETS. Most of the existing IDS system is designed based on utilizing different machine learning and deep learning approaches. A lot of intrusion prevention mechanisms are designed based on encryption algorithms and single authentication schemes.

Among those prior existing security mechanisms, a few of them are reviewed below.

Alghamdi *et al.* [13] presented a trust-aware intrusion detection and protection system (TA-IDPS) for network defence. This proposed work comprises a cloud service layer, cloudlet and MANET. The first process of this method uses an ultra-lightweight symmetric cryptographic approach to register and authenticate mobile nodes, which is ideal for resource-constrained contexts. Authentication, High energy consumption and scalability are all critical challenges in MANETs, and the presented moth flame optimization technique addresses these issues. A deep belief network classifies packet data obtained by the cluster head (CH) as normal, malicious, or suspicious. Cloudlets are utilized in the cloudlet layer to collect packets from the CH and verify their legitimacy before forwarding them to the cloud service layer.

Arul *et al.* [14] introduced a novel detection mechanism to mitigate black hole attacks and enhance MANET security and performance. Multi-hop communication is used to carry out the engagement. In any case, a few nodes in a comprehensive or integrated environment will provide truthful data to determine the best route to the destination network. This is caused by the greedy nature of those nodes that require network traffic over time. Nodes that provide wrong information will not forward data packets and will instead drop them, a technique known as a black hole attack. In this regard, reducing the consequences of the Black hole through Identification and Protection (ABIP), a unique intrusion detection method, had developed. The ABIP is composed of a non-static receiver succession number threshold level, which produces a high receiver succession number for nodes that supply incorrect information.

Kondaiah *et al.* [15] presented a trust factor as well as a fuzzy-firefly integrated particle swarm optimization-based intrusion detection and prevention scheme for safe MANET routing. As a result, a trust factor, as well as a fuzzy based intrusion detection and protection approach, was developed to protect MANET, which is a primary responsibility. The intruder is identified by the fuzzy system that relies on the true value of the nodes. In this way, the secure path can be produced in the MANET. Furthermore, by combining the Firefly Technique (FA) and Particle Swarm Optimization (PSO), an optimization algorithm called Fuzzy integrated Particle Swarm Optimization (FuzzyFPSO) is developed for the best path selection in order to ensure secure routing. Doss *et al.* [16] presented a novel technique for accurately preventing and detecting jellyfish attacks (APD-JFAD). It's a hybrid of an authenticated routing-based attack detection system and a support vector machine (SVM). For learning packet forwarding behaviour, SVM is used. On the basis of the hierarchical trust evaluation property of nodes, the designed technique selects trusted nodes in the network for packet routing.

Verma *et al.* [17] developed a new vampire attack detection and prevention technique for MANET based on the hash value and timestamp concepts. A MANET can be employed in unusual situations where the installation structure is exceedingly complex, such as a disaster region or a fighting zone. The life span of a wireless ad hoc network is determined by the battery power of nodes. In most situations, recharging or changing the battery appears to be impossible. As a result, data packets must be transferred from one node to the next using the least amount of energy possible. However, in the case of vampire attacks, malicious nodes consume more energy than normal, resulting in node power drooping and network failure. The vampire assault depletes the battery's resources, causing the battery to drain prematurely. The key idea is that there are honest and dishonest nodes. If a dishonest node is situated in the middle of a network of honest nodes, it will use more energy than usual to transport data packets to the surrounding node.

For secure data fusion analysis, Zou *et al.* [26] introduced a certificateless short signature system based on integrated neural networks and elliptic curve encryption. The Inv-CDH problem is the foundation for the solution's security. According to the stochastic predictor model, the full security proof is provided. It has been demonstrated that the new model can withstand an adaptive selective message attack from a new enemy. Nouman *et al.* [27] presented construct a new blockchain technology that has the ability to form a decentralised application in education With the aid of several components, including Meta Mask, IPFS, Ganache, additional test networks like Rinke by and Rostand, Web3 JS, and Ethereum cryptocurrency. Analysis of the top three blockchain coins is the first step (Bitcoin, Litecoin, and Ethereum). The second is to research the characteristics and problems related to the Bitcoin cryptocurrency. The third step entails using Web3 JS and Smart Contracts to develop a graphical interface for the IPFS bandwidth analysis for network file storage. Yin *et al.* [28] introduced a Modality-Aware Graph Convolutional Network (MAGCN) module that integrates topological graph connectivity data and multi-modality entity properties into a single lower-dimensional feature space in order to improve link prediction performance. A Graph Knowledge Transfer Learning (GKTL) strategy is also designed using this method to transfer knowledge between subgraphs taken from the same knowledge graph. A technique to build an access control knowledge network using user and resource attributes was given by coworkers *et al.* [29]. On the basis of the created knowledge graph, a proposed online learning framework for access control decision-making is also built. This method considered topological features from the framework to express resource and user attributes with high cardinality. Yin *et al.* [30] was presented an algorithm called Adaptive Sliding Window Weighted Learning (ASWWL). To address the dynamic multiclass imbalance problem that arises in many industrial applications, including the prediction of exploitation time.

Intrusion is described as any type of unwelcome or unexpected network activity that compromises the integrity, confidentiality, or access to a network resource without permission. The "Intrusion Detection System (IDS) is a system that is used to detect anomalous behaviours in this network." In MANET, there is a variety of Intrusion Detection techniques utilized. There are a number of concerns with the aforementioned connected works' intrusion detection and prevention procedures, based on the aforementioned

connected works. MANETs lack concentration locations where data collection and monitoring can be done. MANET routing systems necessitate nodes cooperating and acting as routers, which opens the door for attacks. The network topology is dynamic and unpredictable due to the movement of the nodes, making intrusion detection difficult. IDSs in MANETs are more complicated because most nodes in MANETs have limited processing capabilities. To mitigate these burdens, the proposed method designed an effective intrusion detection technique for predicting different kinds of attacks in M ANET to improve network security.

3. Proposed Methodology of the Intrusion Detection and Prevention Technique in Manet

An effective intrusion detection and prevention technique is presented for improving the MANET security to mitigate attackers in MANET utilizing artificial intelligence strategies. MANET has gained popularity in various applications because of its various supporting characteristics such as easy node configuration, dynamic topology and distributed administration. This MANET's primary function is to route packets from source to destination. The nodes in MANET are said to be mobile because they move randomly and change their location often. So, these mobile nodes are often subjected to security vulnerability due to their existing supporting characteristic. To secure MANET from various kinds of attacks and data breaches designing intrusion detection and prevention techniques is considered mandatory. The primary goal of this current study is to develop an effective intrusion detection and prevention technique for predicting various types of attacks that cause MANET security vulnerabilities. The intrusion prevention technique is designed based on an artificial intelligence system. The architecture of the proposed intrusion detection and prevention technique related to MANET is given in figure 1.

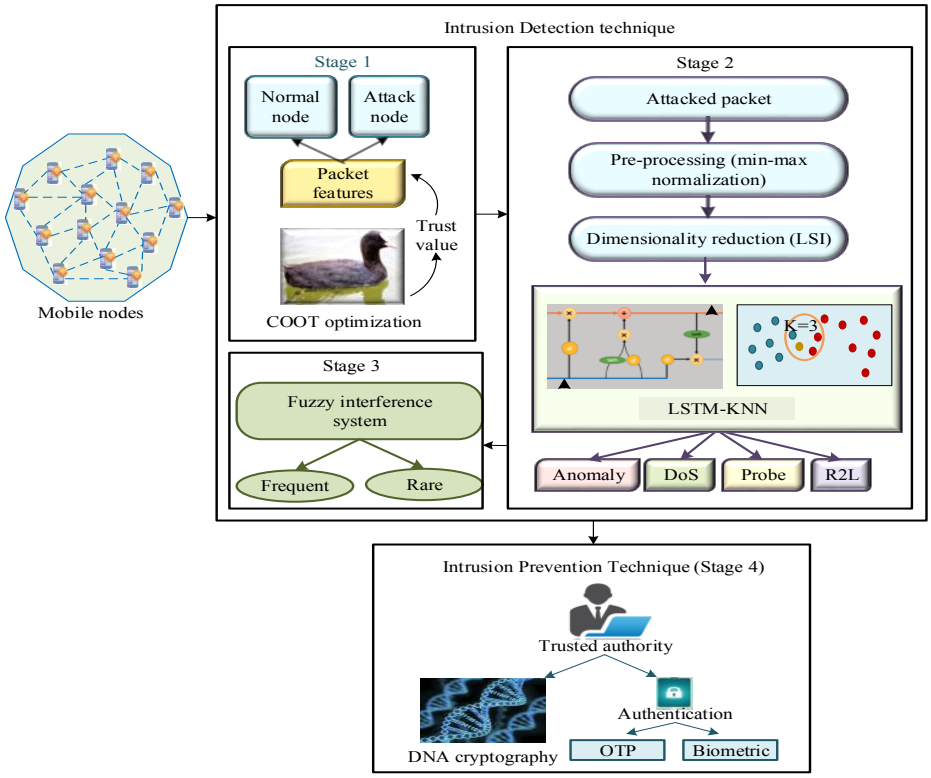

Figure 1. Architecture of the proposed intrusion detection and prevention technique

The suggested intrusion and detection model consist of four stages such as classifying normal node from attack node, predicting different types of attacks, finding the frequency of attack and intrusion prevention mechanism. At the beginning of the process, the nodes are deployed within the specified area. These nodes send information in the form of packets from the sender to the receiver. In this way, the attack node is identified from the normal node based on the packet features and the trust value (threshold). The trust value is selected based on utilizing the COOT optimization algorithm. After finding the trust value, the attack node is categorized from the normal node. Then the next stage, the packets corresponding to the attack node are considered and which is subjected to preprocessing utilizing min-max normalization. After that,

the dimension of the data is reduced using the Latent Semantic Index (LSI) to reduce the process complexity. The dimension-reduced data is given to the input of the hybrid classifier LSTM-KNN. The classifier predicts four classes such as anomaly, DoS, Probe and R2L. These packet features are given as input into the fuzzy interference system in order to classify the occurrence of attack, whether it is frequent or rare. Then, the intrusion prevention mechanism is designed in order to restrict the access of malicious nodes present in the network. Initially, the mobile nodes will register with the trusted authority by means of using a two-factor authentication scheme. Thus, the proposed intrusion prevention and detection model provides mode security for each mobile user and can completely protect the system from attacker nodes.

3.1. Process Involved in the Proposed Model

There are two phases involved in this proposed intrusion prevention and detection model. The primary goal of this current study is to develop an effective intrusion detection and prevention technique for predicting various kinds of attacks that causes security vulnerability to MANET.

Node Deployment: At the beginning of the process is nodes deployment within the specified area. The deployed nodes are considered to be mobile. These nodes capture the information from that specified location and transmit it to the destination in the form of packets.

Intrusion Detection Technique

The intrusion detection technique consists of three stages such as,

- First stage: Classifying normal node from attack node.
- Second stage: Prediction of different types of attacks
- Third stage**:** Finding the frequency of attack

First stage: Classifying Normal Node from Attack Node

There is the possibility of various types of attacks intruding all through packet transmission from the sender to the receiver. So, the attack node is categorized from the normal node based on packet features. Packets from different locations of mobile nodes in a deployed network are attained to gather packet features. Since the packet's threshold value is constant, it may vary, which is inefficient and might lead to inaccurate results. Thus, it must be flexible and dynamic in order to classify assault patterns. In order to deal with these difficulties, the threshold function is applied as well as it is estimated and when each packet arrives at the intrusion detection engine, it is upgraded and updated. In addition to basic packet features derived from packet header information, the node i trust value is calculated utilizing the below methods.

$$T_i = \frac{Number\ of\ packets\ sent\ sucessfully}{Number\ of\ packets\ totally\ sent} \times 100\% \quad (1)$$

Where, T_i denotes the trust value of node, i, which is expressed as a percentage. If the total number of packets delivered by node i is 0, T_i becomes zero, indicating that node i has dropped all incoming packets from the attacker node, and this information has been broadcast to node i 's neighbours. Table 1. tabulates Feature Name and Packet Features.

Table 1. Feature Name and Packet Features

Feature Name	Packet Features
Packet arrival time	PF1
Num. of packets per flow	PF2
Packet counts	PF3
Packet size	PF4
Packet type	PF5
Inter-packet interval	PF6
Flow direction	PF7

The attacking node is categorized from a normal node using the trust value (threshold). In this model, the threshold value is set as 10 and based on the threshold value, and the attack node is categorized in the mobile node. The trust value selection is performed based on utilizing the COOT optimization algorithm. The fitness function considered for finding the trust value is to minimize the packet travel time.

(a) Background of Coot Optimization

Optimization is the procedure of determining the best or global optimal solution to an issue. When solving issues, the ideal global answer is the function's smallest value. This method, which is based on the actions of this bird on the surface of the water, employs the coot optimization technique to provide a better answer. Coots appear to be well within what surf scoters consider a repulsion zone as they travel at an angle to their own momentum. The swarming behaviour of coots on the water is comprised of three actions. On the water's surface, the coots form a chain, with each coot moving behind its front coot [18]. The process involved in the coot algorithm is given as follows.

Step1: Initialization

The initial random population can be specified as $\vec{x} = \{\vec{x}_1, \vec{x}_1, \ldots \vec{x}_n\}$. The target function evaluates the random population repeatedly and determines a target value. It is also aided by a set of principles that form the basis of an optimization method. As the number of sample alternatives and optimization stages (iteration) increases, so does the possibility of obtaining the global optimal. Using the formula, the population is generated at random in a small space (2).

$$CootPos(i) = rand(1, d) * (ub - lb) + lb \quad (2)$$

Where, the coot position is represented as $CootPos(i)$, the number of variables or problem dimensions is denoted as d, ub represents the upper bound of the search space, and lb denotes the lower bound of the search space, and that is defined as the formula (3).

$$lb = [lb_1, lb_2, .., lb_d], ub = [ub_1, ub_2, .., ub_d] \quad (3)$$

Step 2: Random movement to one side or the other

Consider a random place in the search space according to formula (4) and move the coot towards this random position to implement this movement.

$$Q = rand(1, d) * (ul - lb) + lb \quad (4)$$

This coot movement investigates several aspects of the search space. If the algorithm becomes stuck in the local optimal, this movement will cause it to exit. The new position of Coot is calculated using equation (5).

$$CootPos\ (i) = CootPos(i) + A \times R2 \times (Q - CootPos(i)) \quad (5)$$

Where R2 indicate a random number which is a range between 0 and 1, A is estimated by using the following expression (6).

$$A = 1 - L \times (\tfrac{1}{Iter}) \quad (6)$$

Where, $Iter$ represents the max iteration, and the current iteration is denoted as L.

Step 3: Chain Movement

The ordinary position of two coots can be utilized to implement performance. Another way to implement a chain movement is to first calculate the distance vector between the two coots and then move the Coot toward the other Coot by half the distance vector. According to the following expression (7), the new position is estimated,

$$CootPos(i) = 0.5 \times (CootPos(i-1) + CootPos(i)) \quad (7)$$

Where $CootPos(i-1)$ is the second Coot.

Step 4: Adjusting the position according to the group leaders

Few coots are in the front position. They are responsible for leading their entire group, and the remaining coots must change their position and move toward the leaders of their group. Depending on the leader's control, each coot bird will change its position. The Coot's position is updated according to the average position of the leaders present in the group. According to the group leader's k, the Coot (i) can alter their position. The leader selects the next position of the Coot using the following expression (8).

$$CootPos(i) = LeaderPos(k) + 2 \times R1 \times \cos(2R\pi) \times (LeaderPos(k) - CootPos(i)) \quad (8)$$

Where, $LeaderPos(k)$ indicates the selected leader position, the present position of Coot is represented as $CootPos(i)$, R is a random number in the interval [− 1, 1], R1 is a random number ranges between 0 and 1, and π represents the same pi value as 3.14.

Step 5: Leading the group by the leaders towards the optimal area (leader movement)

Leaders can adjust their stance toward the target in order to steer the group toward a common goal (optimal area). To update the leader's position using the following expression (9).This expression searches for improved positions in the region of the current optimal point. Sometimes, in order to find better places, leaders must move away from their current best position. This equation is a nice way to get closer to and away from the optimal location. The following expression is used for providing an optimal location in a better way.

$$LeaderPos(i) = \begin{cases} B \times R3 \times \cos(2R\pi) \times (gBest - LeaderPos(i) + gBest) & R4 < 0.5 \\ B \times R3 \times \cos(2R\pi) \times (gBest - LeaderPos(i) - gBest) & R4 > 0.5 \end{cases} \quad (9)$$

Where, R is a random number in the interval [− 1, 1], R3 and R4 are random numbers in the interval [0, 1], $gBest$ is the best position ever found, π denotes the same pi value as 3.14.

(b) Coot Optimization for finding optimal trust value:

The proposed model utilized coot optimization to find the optimal trust value in order to reduce the packet travel time. The steps involved in the coot optimization for the selection of trust value are discussed as follows.

Step 1: Initialization

The initialization stage is considered as a population analysis in this swarm intelligence, and here the swarm is considered as a coot bird. The candidate solution or population is used to initialize the variable-related problems. The mobile node during the transmission of packet travel time from source to destination is considered the problem in this research. The initialization of packet travel time is given as in equation (10).

$$PT = \{pt_1, pt_2, pt_3 \ldots . pt_n\} \quad (10)$$

In equation (6), pt illustrates the population of packet travel time of each node present in the MANET takes time for packet transmission among source nodes to the destination nodes.

Step 2: Fitness Function

The packet travelling time from the source to destination is given as the fitness to find the best trust value among the nodes in the MANET. The optimal trust value is selected in order to reduce the packet travel time.

$$Fitness = min\{packet\ travel\ time\} \quad (11)$$

The packet travel time is estimated by the destination node. The packet travel time is also termed transmission time, and it is known as the time taken for the packet to transmit from source to destination or the time taken from the beginning of the transmission until receiving the acknowledgement from the same node. The amount of time taken from the start of transmission to receiving acknowledgement from the same node. The destination node's formula for calculating packet travel time is given in equation (12)

$$packet\ travel\ time = CT - OT \quad (12)$$

In equation (3), OT denotes the packet's beginning timestamp, CT the current timestamp, and $packet\ travel\ time$ refers to the packet's travel time from source to destination.

Step 3: Updating

To find out the optimal trust value within the nodes, the value of packet travel time related to nodes gets updated in each iteration using equation (9).

Step 4: Termination

The selection process finally gets terminated once the optimal trust value is found.

The attack nodes and normal nodes are categorized based on the attained trust value of each node in the MANET. If the nodes attain a trust value above 5, then the nodes are considered the normal nodes, and if the attained trust value is below 5, the nodes are considered as attack nodes. Then the attacked node, i.e. attacked packets, are further taken for analyzing the types of attacks in the nodes.

Second stage: Prediction of Different Types of Attacks in MANET

In this stage, the attacked nodes are considered to forecast kinds of attacks in MANET. The process involved in forecasting various types of attacks present in MANET is described below.

(a) Min-Max normalization

The raw packet features of attack nodes are preprocessed using min-max normalization [19]. Min-Max Normalization transforms x to x' by converting each value of features to a range between 0 and 1, and this is also known as (0–1) Normalization. If the data had negative values, the range would have been between -1 and 1. The formula for Min-Max Normalization is:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (13)$$

The normalized value is denoted as x', x represents the original value, $\max(x)$ denotes the maximum value of x, and $\min(x)$ indicates the minimum value of x. The packet features are normalized in this stage that is used to assist in the effective reduction of dimension.

(b) Latent Semantic Indexing (LSI)

In this method, the dimensionality of the attack packet features is reduced with the help of Latent semantic Indexing (LSI). The dimension of the Term Document Matrix (TDM) is diminished by the idea of singular value decomposition (SVD) utilized by Latent Semantic Indexing (LSI) [20]. The issue of synonymy and polysemy words are noticed by LSI is the main objective equation (14). The singular value decomposition (SVD) method develops the semantic space of LSI in linear algebra. Retrieving and demonstrating the uniqueness of semantic of phrases for vector spacing and deduction of dimension utilized the LSI.

$$TDM = USV^T \quad (14)$$

V is the right singular vector of TDM and represents the $n \times n$ unitary matrix. U represents a $m \times m$ unitary matrix with the left-singular vector of TDM as its column, and S represents a $m \times n$ diagonal matrix of singular values. The first k diagonal values of S are kept by LSI, while the remainder is zero. It is said that the diagonal value of S is ordered descending. In addition, the first k columns and rows of the U and T matrices are conserved. The compact form of TDM is calculated using the equation below (15).

$$TDM_k = U_K S_K V_K^T \quad (15)$$

From equation (15), k denotes the first k columns and rows. According to the above equation, the dimension of the attack packet features is reduced, which can remove the classification burdens and provide an accurate prediction.

(c) Classification of Different Types of Attacks

The dimension-reduced attack packet features are taken as the input of the classification process for predicting different types of attacks present in the MANET. A detailed description of the classification process is illustrated in this section

Overview of Long Short-Term Memory (LSTM)

Hochreiter and Schmidhuber introduced the LSTM model, which was a modified version of the Recurrent Neural Network, to address the concerns of vanishing and bursting gradients and to reduce the loads [21]. The LSTM tactic can maintain information over long periods of time due to its unique structure, cell state, and gates that govern how data goes in the various layers [31]. The LSTM model's structure is shown in figure 2. The following are the equations for each stage of the LSTM model.

$$i_t = \sigma(W_i x_i + U_i h_{t-1} + b_i) \quad (16)$$
$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \quad (17)$$
$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (18)$$
$$\tilde{C}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (19)$$
$$C_t = f_t \otimes C_{t-1} + i_{t-1} \otimes \tilde{C}_t \quad (20)$$
$$h_t = o_t \otimes \tanh(C_{t-1}) \quad (21)$$

Where, W_i, W_f *and* W_o represent the weights connecting input, forget and output gate with the input; U_i, U_f*and* U_o denote the weights from input, forget and output gates to the hidden states, respectively; b_i, b_f*and* b_o are input, forget, and output gate bias vectors, respectively; i_t, f_t *and* o_t represents the input gate, forget gate and the output, respectively. σ Represent the logistic sigmoid function, h_t refers to the current hidden state, C_t represent the current cell state, $\tilde{C}_t$ is the cell input; $tanh$ is the hyperbolic tangent function and $\otimes$ is element wise multiplication. The idea behind this framework is that the cell state acts as a memory unit, remembering necessary details through the various gate operations. The information added to the cell with the help of current input and historical information is filtered by the input gate filter, while the forget gate discards certain previous information, allowing the cell to clear the value it contains; finally, the output gate selectively outputs the information.

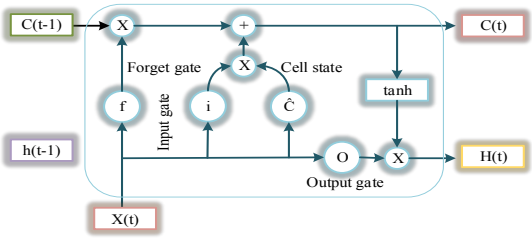

Figure 2. Structure of LSTM

Overview of K-Nearest Neighbor (KNN)

K-nearest neighbour (KNN) classification is one of the simplest and most common classification methods, and the performance of this classifier still competes with the most complex classifier. As a result, if a sample's classification is unknown, it can be predicted by looking at the classification of its closest neighbours. Given an unknown sample and a training set, all distances between it and all the samples in the training set can be computed. The sample closest to the unknown sample in the training set correlates to the distance with the lowest value. As a result, the unknown sample's classification can be based on that of its nearest neighbour [22]. The following are the steps involved in the KNN classifier:

Training phase: The training samples and their class labels are saved; no missing data or non-numeric data are permitted.

Classification phase: The following processes are used to classify each test sample using the majority vote of its neighbours [32].

- Step 1: A specified distance function or similarity measure is used to calculate distances from the test sample to all stored training samples.
- Step 2: The test sample's KNNs are chosen, where K is a specified small integer.
- Step 3: The test sample is assigned to the KNN class that is most frequently used. To put it another way, a test sample is assigned to class c if it is the most common class label among the K training samples closest to it. If K=1, the test sample is allocated the class of the nearest neighbour.

(d) Prediction of Different Types of Attacks using Hybrid LSTM-KNN

This proposed method introduces a hybrid LSTM-KNN classification model for the prediction of different types of attacks among the attack nodes. A hybrid LSTM-KNN classification for the detection of threats in MANET is proposed in this study paper. When comparing hybrid LSTM-KNN to other ML and DL approaches, it produced better results. Because the LSMT model extracts features efficiently and more quickly than the KNN model, multi-class instances can be handled organically in the detection process. This is the rationale cited in the part on the proposed methodology for selecting hybrid LSTM-KNN. Figure 3 illustrates the hybrid LSTM-KNN classification model for the prediction of different types of attacks. The LSTM model has four layers: input, hidden, FC, and softmax. This method included KNN classification instead of the FC layer and softmax layer in the LSTM model. FC layer is the classification layer in LSTM that is modified into KNN to create a hybrid LSTM-KNN model. Initially, the dimensionality-reduced packet is split into two sets such as a training set (containing 80% of data) and a testing set (containing 20% of data).

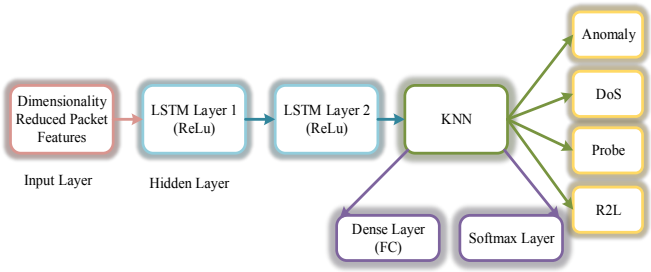

Figure 3. Architecture of the hybrid LSTM-KNN model

The training dataset is given as the input of the LSTM model, and the input layer is responsible for gathering the information from the previous stage, i.e. dimensionality reduction. The data is then passed from the input layer to the next layer, which is the hidden layer, also known as a gated cell or gated unit in the LSTM. It has four layers that interact with one another to produce the cell's output and state. After that, both these things are passed onto the next hidden layer. After that, the outcome of the hidden layer is fed into the classification layer, and this proposed model considers KNN as the FC layer and softmax layer. The KNN classification model classifies the attack packet features. After completing the training process, the remaining 20% of the data are given to the LSTM-KNN model for testing the model. Finally, the hybrid classification model produces four types of attacks in the MANET such as anomaly, DoS, Probe and R2L.

Third stage: Finding the Frequency of Attack

After predicting the different types of attacks, the occurrence of attack, whether it is frequent or rare, is found based on packet features such as Packet transmission rate, Hop count, Trust values, Number of miss-transmitted packets, and Number of packets delivered successfully within a time interval. These packet features are given as input into fuzzy interference systems in order to classify the occurrence of attack, whether it is frequent or rare.

(a) Fuzzy Inference System

Fuzzy logic can deal with uncertainty, which is why it's used in intrusion detection systems. Intrusion detection features can be viewed using fuzzy variables or linguistic phrases, allowing for the determination of normal and abnormal network activity. Fuzzy rules based on if-then-else rules are used to specify all network scenarios for the occurrence of attacks, whether they are frequent or infrequent. The fuzzy inference system (FIS) is a fuzzy rule-based system that is in charge of making decisions [23].

(b) Fuzzification

Fuzzification is the process of converting an input device into linguistic terms or a fuzzy set. The membership function can only be determined using the stated fuzzy set. The membership can be represented using a variety of curves. Triangular or trapezoidal-shaped functions are the most common among these. Every input region will be graphed using the membership function. Smooth mapping necessitates the overlapping of membership functions. Because of this conversion into linguistic value, developing rules will be easy, even for complex operations [33]. Packet transmission rate, trust values, and the number of packets transmitted successfully within a time interval are the inputs for this suggested fuzzy model. These three factors are transformed into linguistic values and combined to generate a function that can be represented as $PTT = (PT_{source} \sim PT_{desttination}), TV = (CT \sim OT)$ and $NPD = (NP_{source} \sim NP_{desttination})$. Based on this criterion, the packet travel time, trust value and number of packets delivered are assumed. Using this considered degree, the membership function for both packet travel time, trust value and number of packets delivered is plotted in a graphical form.

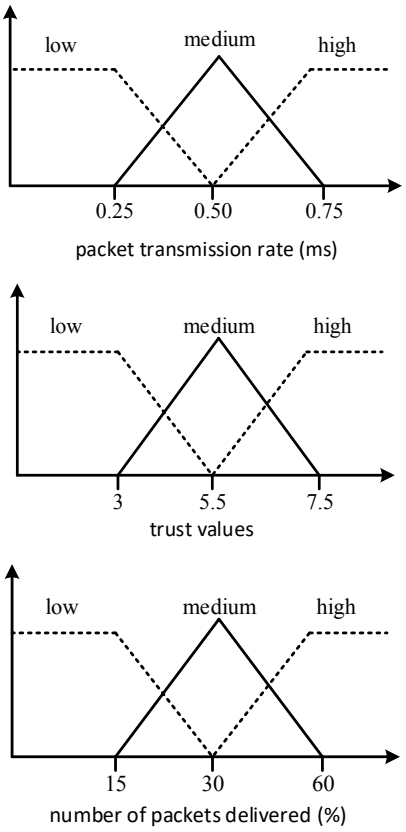

Figure 4. Membership Function Graphed for Packet Transmission Rate, Trust Value and Number of Packets Delivered

The membership function for packet transmission rate, trust value, and the number of packets sent is drawn using the anticipated linguistic value. The packet transmission rate is divided into three categories: low, medium, and high. A distance value of 0.25-0.75 (ms) is considered medium, and a value of 0.75 to 1 is considered high. A value of less than 3 is considered low, a value of 3 to 7.5 is considered medium, and a value greater than 7.5 is considered high. When it comes to the number of packets delivered, less than 15 is considered low, 15-60 is considered medium, and more than 60 is considered high. Based on this membership function, the rules for relay node selection are generated in the following knowledge base stage.

(c) Knowledge rule

At this general expertise level, the membership function that describes the input variables is converted into a fuzzy rule. The ability of an expert system to arrive at a solution from a set of facts is seen as a key attribute. In most cases, the collected data is transformed into a conditional statement, such as If else, and used as evidence. The reasoning rules, which might be either production or interference rules, are framed based on the gained information about evidence. Each created rule will link two sets of facts, such as premises and

consequents. With each reality and rule, optimism factors are presented to describe ambiguity in knowledge. In table 2, the membership function is used to generate rules for input variables such as packet transmission rate, trust values, and the number of packets sent.

Table 2. Rule Generated Based on Membership Function

Packet Transmission Rate	Trust Values	Number Of Packets Delivered	Output
Low	Low	Low	Frequent
Low	Low	Medium	Frequent
Low	Medium	High	Rare
Low	Medium	Low	Frequent
Medium	High	Medium	Rare
Medium	High	High	Rare
Medium	Low	Low	Rare
Medium	Low	Medium	Frequent
High	Medium	High	Frequent
High	Medium	Low	Rare
High	High	Medium	Frequent
High	High	High	Rare

The previous table explains how to frame rules to reach a conclusion about the alleged fact. According to this table 2, if the computed packet transfer rate between the sender and destination mobile nodes is low, then mobile network attacks are infrequent. Suppose the trust value is high among the source node and destination, then the node has attacked in rare cases. The number of packet transmission ranges is high among the source mobile node and destination mobile node, then the attacks present in the network are a rare condition. In the MANET, the source and destination node contain low packet transmission time, high trust value and a high number of packets delivered then. The attacks occur in rare cases only. The source and destination nodes contain high packet transmission time, low trust value and a low number of packets delivered. Then attacks occur frequently on a mobile network. If the mobile features have medium ranges means, the attacks occur either sometimes rare or frequent.

(d) De-fuzzification

Defuzzification is the ultimate stage of a fuzzy model. The output received as a fuzzy set is transformed into crisp during defuzzification. To put it another way, it's the process of transforming imprecise data into exact data. The output of a fuzzy rule is expressed in frequent and rare categories. In this proposed method, the attack occurrence, whether it is frequent or rare, is evaluated based on considering several packet parameters such as Packet transmission rate, Hop count, Trust values, Number of miss-transmitted packets and Number of packets delivered successfully within a time interval. The systems feature a single output that specifies the verity level of a node's behaviour based on input criteria, such as frequent or rare. Table 2 shows the suggested system rules for evaluating the existence of a node.

3.2. Intrusion Prevention Technique

Fourth stage: Intrusion prevention mechanism

The intrusion prevention mechanism is designed in order to restrict the access of malicious nodes, i.e. attack nodes present in the network. Initially, the mobile nodes will register with the trusted authority by means of using a two-factor authentication scheme. The two-factor authentication scheme is designed by merging biometrics with OTP. Only after verifying these biometrics and OTP by a trusted authority the mobile users can access it. The files transmitted by the mobile user is encrypted using DNA cryptography for security purpose. It is again decrypted at the receiver end using the generated key.

Register with the Trusted Authority

Initially, the mobile nodes will register with the trusted authority through two-factor authentication. Generally, the trusted centre will enable secure communication between data owners. And the trusted centre is also responsible for generating the private key and public key needed for encoding and decoding the data. The key will be created based on the sensitivity level of the information. If the information is highly sensitive, then a strong key is generated for encrypting the data. On the other hand, if the information is less sensitive, then a moderate key is generated for data encryption. This private and public key will be generated in the trust centre at the time of registration. The trust centre stores the user id, name, password and email-id. Further, the personal information other than user ids such as user name, mail id and password of the mobile users will be encrypted based on a simple XOR encryption algorithm in order to secure the user information of the organization from the attackers.

DNA Cryptography

Initially, in this algorithm, the data which is to be stored is taken as input. Then, the input data which appears in a string is converted to ASCII values using the ASC II code conversion technique. Then, these ASCII values are converted into binary data with 8 bits. Then, the private key, which is created by the trusted centre, is also converted into binary values [24]. Following that, in the next step, the input binary value is XOR with the binary values of the private key. The resultant output obtained as the outcome of XOR is again converted into its corresponding ASCII values. The acquired ASCII value is then converted to its related binary values. Finally, to reach 1024 bits, some of the bits are added to the related binary values. This 1024-bit binary value is split into four parts to reach 256 bits. Then, a complementary rule is applied to these four portioned binary values to reach the final DNA sequence. This DNA sequence is the encrypted data which is stored. The numerical example used for illustrating the process of encryption and decryption is given below.

(a) Encryption Phase

The process of encryption is generally converting the input data into an indefinite form to secure it from attackers.

Step 1: The input data which is to be encrypted is considered (20-02-1990 ABCD)

Step 2: Then, it is converted to ASCII values using the ASCII code conversion technique (50 48 32 48 50 32 49 57 57 48 65 66 67 68)

Step 3: In this step, the ASCII values are converted to binary values (00110101 00110000 00110100 00111000 00110011 00110010 00110100 00111000 00110101 00110000 00110011 00110010 00110100 00111001 ……00110110 00111000)

Step 4: The gained binary value is EXOR with the binary value of key ((00110101 00110000 and 00111000 00111000) …….. (00110110 00111000 and 00110111 00110000)) = ((000010101 00001000)……. (00000001 00001000))

Step 5: The output obtained as a result of the EXOR operation is once again converted into ASCII value (78…...

Step 6: Similarly to step 3, the ASCII value is again converted to a binary value, and the obtained binary value must of 1024 bits (00110111 00111000…..00110101 00110101)

Step 7: In this step, 1024 bits are divided into 256 bits, and the complementary rule is provided for this 256-bit binary as ATCG (Adenine, Thymine, Cytosine and Guanine).

Step 8: The obtained DNA sequence is the final encrypted data which is stored in cloud.

(b)Decryption phase

In this phase, the reverse process of encryption takes place. The users who have valid authentication will get the secret public key from the data owner and then access the data through decrypting.

Tow factor Authentication Scheme

Recently, mobile users stored their data in secure servers by data owners. To improve the security of the stored data in servers, some authentication approaches were developed to remove the issues for the users at the time of the data accessing and transmitting the data. Initially, the mobile nodes register into the trusted centre, and it stores the users' id, name and email-id. If mobile nodes, i.e. users, need to access these data, they request the trusted centre, and then the trusted centre analyzes the user id, name and password. If these are matched, the trust centre sends OTP to the user. After that, the OTP is matched and then goes for biometric authentication. If the biometric matches, the data owner provides access to use the data. A detailed description is provided below.

OTP

The second authentication that acts as a barrier for the user is OTP. The OTP is termed as One Time Password. Generally, the OTP is incorporated along with two or multi-factor authentication schemes such as PIN and fingerprint verification. The OTP generation system utilizes randomness or else pseudo-randomness to create the seed or shared key. Then the cryptographic hash function is included along with the shared key to obtain a value. In this present work, after verifying the password, the trusted centre will be sent the OTP to the registered mail ID of the user. If the user enters OTP correctly, then the mobile user will access to transmit the data.

Biometric

The last and final barrier used to check the authenticated user is fingerprint verification. This fingerprint identification is a widely used authentication technique, and it is also termed biometrics. In this present work, for verification, the fingerprints stored in the trusted centre are matched with the fingerprint provided by the user at the time of accessing. The process flow enclosed in the authentication scheme is given in figure 3.

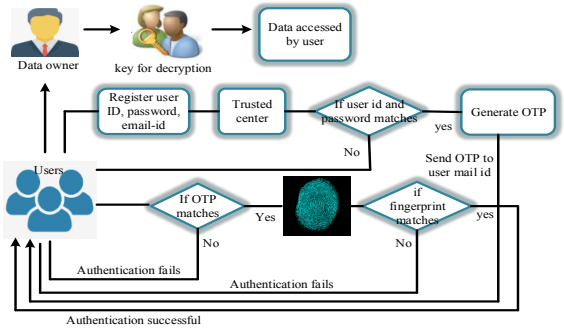

Figure 5. Process Flow of two factor Authentication approach

The two-factor authentication scheme is designed by merging biometrics with OTP. Only after verifying these biometrics and OTP by a trusted authority the mobile users can access it. The files transmitted by the mobile user is encrypted using DNA cryptography for security purpose. It is again decrypted at the receiver end using the generated key. Using this proposed intrusion detection and prevention model, the different types of attacks can be predicted effectively, and at the same time, the security during data transmission can also be improved.

4. Result and Discussion

The proposed Intrusion Detection and Prevention Technique in MANET are tested on an NS2 simulator with Intel Core Processor, CPU @ 2.70 GHz and memory at 8.00 GB RAM as system configuration. The experimental analysis is carried out with the help of mobile nodes that are randomly selected in the area of $1000 \times 1000m^2$ dimension and entirely 100 mobile nodes are selected in MANET for this experimental analysis. Several stimulation parameters are considered for this analysis which is given in below table 3.

Table 3. Simulation Parameters Considered for Analysis

Simulation parameter	Values
Simulator	NS-2.35
No. of Nodes	100
Packet size	100 bytes
Area of Simulation	1000mX1000m
Routing Protocol	ROACM
Channel Type	Wireless channel
Initial Energy	10Joules
Initial Transmit power	0.660
Initial Receiver power	0.395
destination location	950, 970

In the beginning, the mobile nodes are deployed in a specified area to find the route between the source and destination for effective transmission to avoid the occurrence of attacks. Figure 6 illustrates the mobile node deployment in the MANET. Entirely a hundred mobile nodes are selected for the detection and prevention of attacks, and these nodes have no proper pattern and are in a random manner. Figure 7 illustrates the data transmission through the gateway. In this stage, the packets are transmitted to the gateway with several packet features. Based on these features, set a threshold value for identifying the attack node from normal nodes. If below the threshold value means, the nodes are considered as attack nodes, and above the threshold value, contained nodes are considered the normal nodes. The threshold value, i.e. trust value of each node, is discovered by COOT optimization. Figure 8 illustrates the detected attack in the MANET. The yellow nodes are considered as the attacked nodes. Then the types of attacked nodes are identified by transmitting the data from source to destination, which is shown in figure 8.

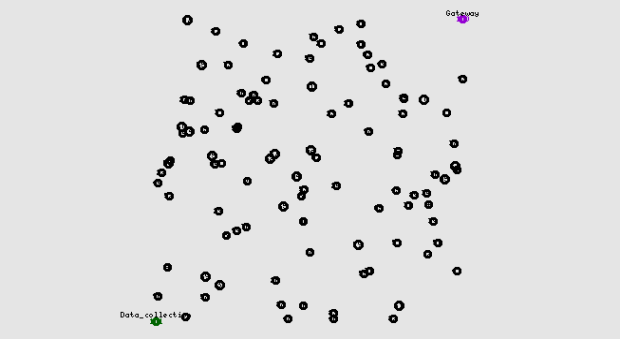

Figure 6. Node deployment in the MANET

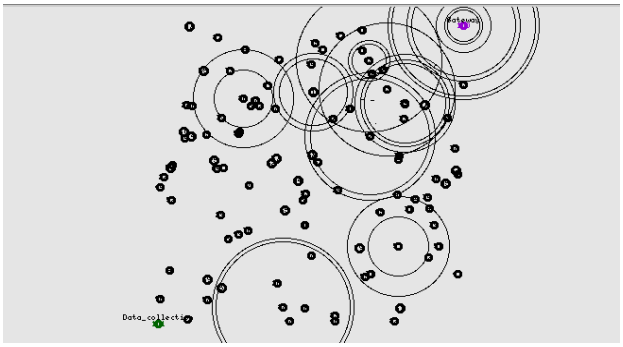

Figure 7. Data transmission through gateway

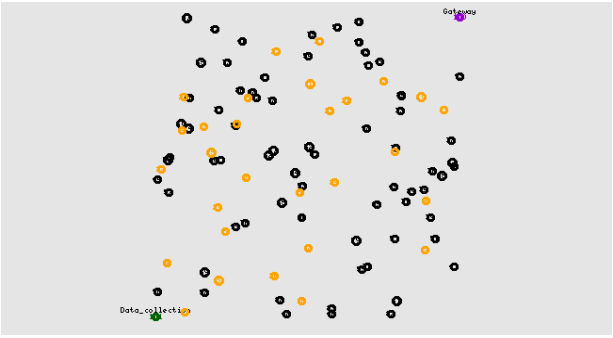

Figure 8. Detection of attack node in the MANET

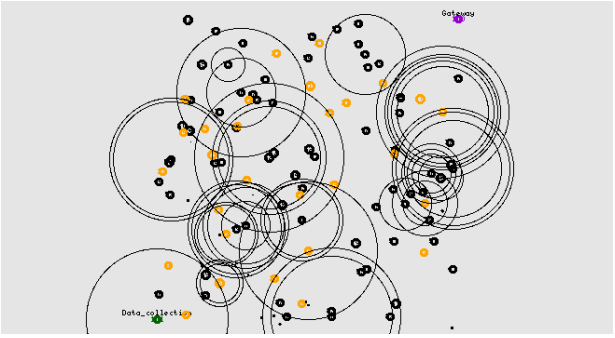

Figure 9. Data collection from attacked nodes

4.1. Experimental Analysis

After finding the attack nodes present in the MANET, the attack nodes are avoided and only suggested normal nodes to transmit data. In this way, effective data transmission is achieved in MANET. The performance of the proposed routing protocol is performed based on certain metrics such as average delay (ms), packet delivery ratio (%), packet loss ratio (%), throughput (Mbps) and attack detection rate (%).

Moreover, the performance of the proposed intrusion detection and prevention technique for effective routing is proved by comparing with some existing intrusion detection and prevention approaches, such as the Modified Zone Based Intrusion Detection System (MZBIDS), Alleviating the effects of the Black hole through Identification and Protection in MANET (ABIP-MANET) and Multi-level trust based intelligence intrusion detection system in MANET (MTIID-MANET) in MANET. The comparison analysis is shown in a graphical representation in this section.

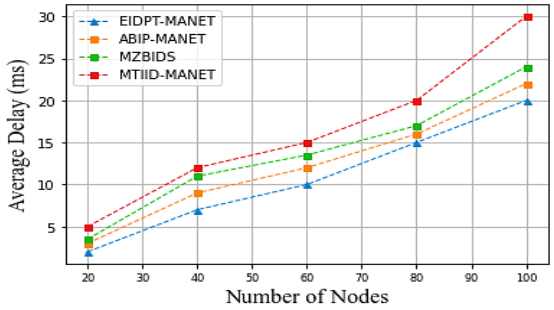

Figure 10. Comparison of average delay vs number of nodes

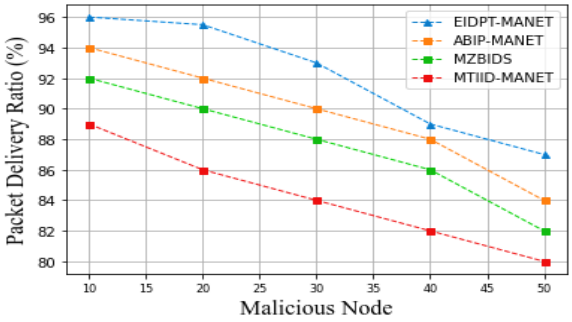

Figure 11. Comparison of packet delivery ratio vs malicious nodes

Figure 10 illustrates the comparison of average delay with a number of nodes. The proposed model attained 1ms at 20 nodes, 6ms at 40 nodes, 10ms at 60 nodes, 15ms at 80 nodes and 20ms at 100 nodes. The existing method of ABIP-MANET attained 2ms at 20 nodes, 8ms at 40 nodes, 12ms at 60 nodes, 16ms at 80 nodes and 23ms at 100 nodes. MZBIDS attained 4ms at 20 nodes, 12ms at 40 nodes, 14ms at 60 nodes, 16ms at 80 nodes and 24ms at 100 nodes. MTIID-MANET attained 5ms at 20 nodes, 13ms at 40 nodes, 15ms at 60 nodes, 20ms at 80 nodes and 30ms at 100 nodes. Compared to the proposed model of EIDPT-MANET attained less average delay (ms) than the existing methods. Figure 11 illustrates the comparison of packet delivery ratio vs malicious nodes. The graph is plotted among the number of malicious nodes on X-axes and the packet delivery ratio in percentage on Y-axes. The proposed model attained a packet delivery ratio of 96% at 10 nodes, 95% at 20 nodes, 93% at 30 nodes, 89% at 40 nodes and 87% at 50 nodes. The existing method of ABIP-MANET attained a packet delivery ratio of 94% at 10 nodes, 92% at 20 nodes, 90% at 30 nodes, 88% at 40 nodes and 84% at 50 nodes. MZBIDS attained packet delivery ratio of 92% at 10 nodes, 90% at 20 nodes, 88% at 30 nodes, 89% at 40 nodes and 87% at 50 nodes. MTIID-MANET attained packet delivery ratio of 89% at 10 nodes, 86% at 20 nodes, 84% at 30 nodes, 82% at 40 nodes and 80% at 50 nodes. Compared to the proposed model of EIDPT-MANET attained high packet delivery ratio (%) than the existing methods. Figure 12 illustrates the comparison of packet loss ratio vs malicious nodes. The graph is plotted among the number of malicious nodes on X-axes and the packet delivery ratio in percentage on Y-axes. The proposed model attained a packet loss ratio of 4% at 10 nodes, 4.5% at 20 nodes, 7% at 30 nodes, 11% at 40 nodes and 13% at 50 nodes. The existing method of ABIP-MANET attained a packet loss ratio of 6% at 10 nodes, 8% at 20 nodes, 10% at 30 nodes, 12% at 40 nodes and 16% at 50 nodes. MZBIDS attained packet loss ratio of 8% at 10 nodes, 10% at 20 nodes, 12% at 30 nodes, 14% at 40 nodes and 18% at 50 nodes. MTIID-MANET attained packet loss ratio of 11% at 10 nodes, 14% at 20 nodes, 16% at 30 nodes, 18% at 40 nodes and 20% at 50 nodes. Compared to the proposed model of EIDPT-MANET attained less packet loss ratio (%) than the existing methods.

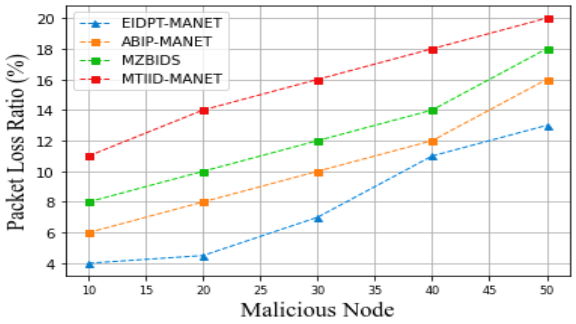

Figure 12. Comparison of packet loss ratio vs malicious node

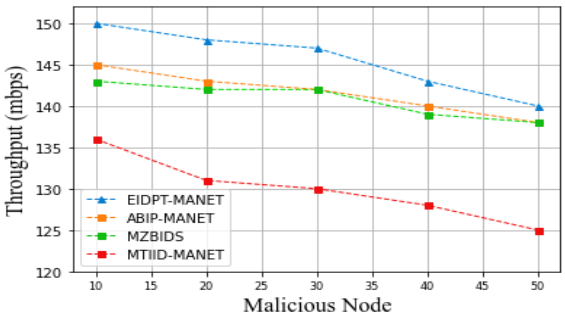

Figure 13. Comparison of throughput vs malicious node

Figure 13 illustrates the comparison of throughput vs malicious nodes. The graph is plotted among the number of malicious nodes on X-axes and throughput in Mbps on Y-axes. The proposed model attained a throughput of 150 (Mbps) at 10 nodes, 147 (Mbps) at 20 nodes, 146 (Mbps) at 30 nodes, 143 (Mbps) at 40 nodes and 140 (Mbps) at 50 nodes. The existing method of ABIP-MANET attained a throughput of 145 (Mbps) at 10 nodes, 143 (Mbps) at 20 nodes, 142 (Mbps) at 30 nodes, 140 (Mbps) at 40 nodes and 138 (Mbps) at 50 nodes. MZBIDS attained throughput 144 (Mbps) at 10 nodes, 143 (Mbps) at 20 nodes, 142 (Mbps) at 30 nodes, 138 (Mbps) at 40 nodes and 137 (Mbps) at 50 nodes. MTIID-MANET attained a throughput of 136 (Mbps) at 10 nodes, 131 (Mbps) at 20 nodes, 130 (Mbps) at 30 nodes, and 128 (Mbps) at 40 nodes and 125 (Mbps) at 50 nodes. Compared to the proposed model of EIDPT-MANET attained high throughput (Mbps) than the existing methods. Figure 14 illustrates the comparison of attack detection ratio vs types of attack nodes. The graph is plotted among the number of malicious nodes on X-axes and the attack detection ratio in percentage on Y-axes. The proposed model attained an attack detection ratio of 94% at DoS attack, 95% at Probe attack, 100% at R2L attack and 98% at the anomaly. The existing method of ABIP-MANET attained an attack detection ratio of 92% at DoS attack, 91% at Probe attack, 96% at R2L attack and 91% at the anomaly. MZBIDS attained an attack detection ratio of 90% at DoS attack, 90% at Probe attack, 91% at R2L attack and 88% at the anomaly. MTIID-MANET attained an attack detection ratio of 87% at DoS attack, 88% at Probe attack, 90% at R2L attack and 85% at the anomaly. Compared to the different types of attack detection, the proposed model of EIDPT-MANET attained high attack detection ratio (%) than the existing methods.

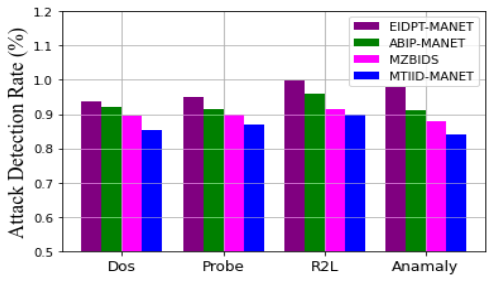

Figure 14. Comparison of attack detection rate vs types of attacks

The proposed classification scheme is implemented in python 3.8 software to validate its achievements. The proposed classification model is tested with several system configurations such as Intel core i5 processor, GPU of Nvidia GeForce GTX 1650 and 16GB Memory (RAM). The dataset is gathered from the deployed nodes in the MANET [25]. Initially, the packet information is gathered from the attacked mobile nodes to find the type of attacks present in the MANET. Then these gathered attacked packets are subjected to preprocessing to remove unwanted data using min-max normalization. After preprocessing the data, it is given to the feature extraction approach for reducing the dimensionality of the packet data through LSI. Finally, the dimension-reduced data are given into the hybrid LSTM-KNN model for classifying the types of attacks. After predicting the different types of attacks, the occurrence of attack, whether it is frequent or rare, is found based on packet features. These packet features are given as input into the fuzzy interference system in order to classify the occurrence of attack, whether it is frequent or rare. The intrusion prevention technique is designed to prevent hostile nodes from gaining access to the network. Thus a two factor authentication scheme is designed by merging biometrics with OTP. The files transmitted by the mobile user is encrypted using DNA cryptography for security purpose. Several evaluation criteria were used in the experimental analysis of the proposed intrusion detection and prevention methods. The metrics considered for this analysis are accuracy, precision, recall, error, specificity, f1_score, Negative Predictive Value (NPV), False Negative Rate (FNR), and False Positive Rate (FPR). Then the DNA cryptography algorithm is compared with several existing cryptography algorithms such as blowfish, DES, and AES with several parameters such as encryption time (sec), decryption Time (sec) and execution time (sec) in Table 4.

Table 4. Proposed vs existing ones parameters comparison.

Parameters	Proposed LSTM-KNN	KNN	LSTM	SVM	RF
Accuracy	96	75	89	80	72
Precision	93	75	80	75	70
Recall	82	40	70	45	65
Error	0.04	0.26	0.11	0.20	0.28
Specificity	98	80	92	81	85
F1_Score	85	72	77	71	70
Negative Predictive Value (NPV)	98	88	92	89	95
False Negative Rate (FNR)	0.18	0.23	0.20	0.24	0.28
False Positive Rate (FPR)	0.03	0.10	0.09	0.12	0.15

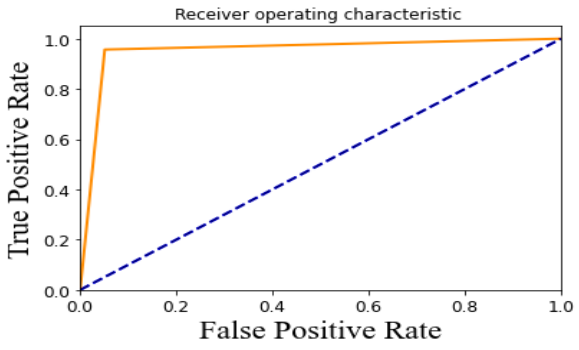

Figure 15. ROC attained for the proposed model

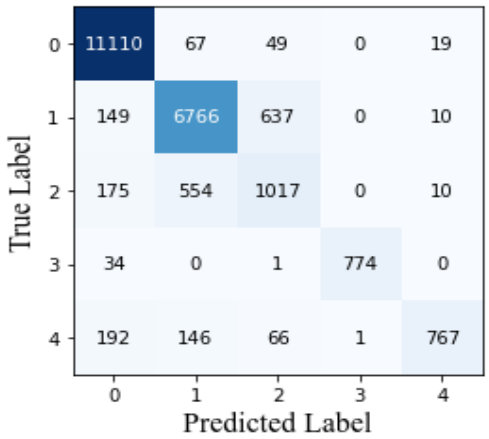

Figure 16. Confusion matrix attained for the proposed approach

Figure 15 illustrates the confusion matrix attained for the proposed model. The Area under the Curve (AUC) curves, or Receiver Operating Characteristics (ROC) curves, appear to be a performance metric for classification issues at various threshold levels. The AUC denotes the degree of separation, whereas the ROC denotes a probability curve. The system's performance measure is regarded well when the curve approaches "1." When the curve's value is "0," the system's performance is considered bad. The ROC value in the proposed method is "1," indicating that the system performs well. To provide an analytical assessment, a confusion matrix and performance measurements like as accuracy, detection rate, and so on are used. Confusion matrixes are useful because they provide a direct value evaluation of false positives, true positives, false negatives, and false positives, among other things. The values of false, true, negative, and positive are written as follows: False Positive represents the number of non-attacking incidents that were mistakenly labelled as attacks. The amount of attacks successfully predicted is referred to as True Positive. The number of negative attacks that are appropriately characterized as ordinary is known as True Negative. The number of assault instances that are wrongly labelled as normal is represented by False Negative. Figure 16 illustrates the confusion matrix attained for the proposed model. Class 0 predicts 11110 data, class 1 predicts 6766 classes, class 2 predicts 1017, class 3 predicts 774, and class 4 predicts 767 data correctly.

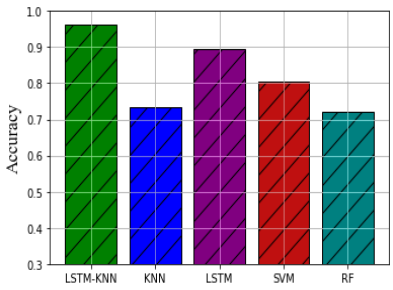

Figure 17. Comparison of accuracy

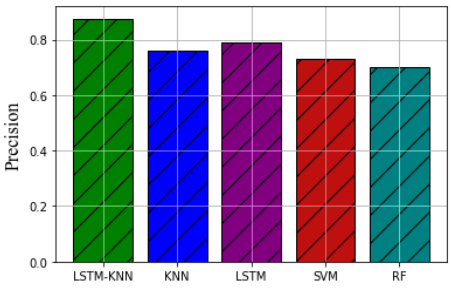

Figure 18. Comparison of precision

Figure 17 depicts a comparison analysis of accuracy (percentage) between the suggested hybrid LSTM-KNN and current strategies for detecting different types of malware attacks. The graphical representation is based on several machine learning and deep learning techniques, as well as the accuracy value in percentage on both the X and Y labels. The proposed hybrid LSTM-KNN method has a 96 percent accuracy, which is higher than existing methods such as KNN (75 percent), LSTM (89 percent), SVM (80 percent), and RF (72 percent). This can demonstrate that the suggested model performs better than other techniques. Figure 18 depicts a precision (percentage) comparison analysis of the proposed hybrid LSTM-KNN and existing strategies for detecting different types of malware attacks. The precision found for the proposed hybrid LSTM-KNN method is 93% which is greater compared to existing methods such as KNN is 75%, LSTM is 80%, SVM is 75%, and RF is 70%. This reveals that the proposed hybrid model attained a better precision value compared to other techniques.

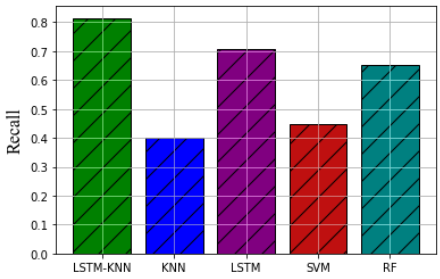

Figure 19. Comparison of recall

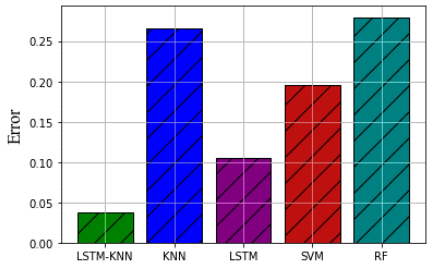

Figure 20. Comparison of error

The comparison analysis of recall (%) among the suggested hybrid LSTM-KNN and current strategies for the detection of types of malware attacks is shown in figure 19. The recall found for the proposed hybrid LSTM-KNN method is 82% which is greater compared to existing methods such as KNN is 40%, LSTM is 70%, SVM is 45%, and RF is 65%. This reveals that the proposed hybrid model attained a better recall value compared to other techniques. Figure 20 depicts a comparison analysis of error (percentage) between the proposed hybrid LSTM-KNN and existing strategies for detecting different types of malware attacks. The error found for the suggested hybrid LSTM-KNN method is 0.04, which is less compared to existing methods such as KNN is 0.27, LSTM is 0.11, SVM is 45%, and RF is 65%. This reveals that the proposed hybrid model works and attains a better error value compared to other techniques.

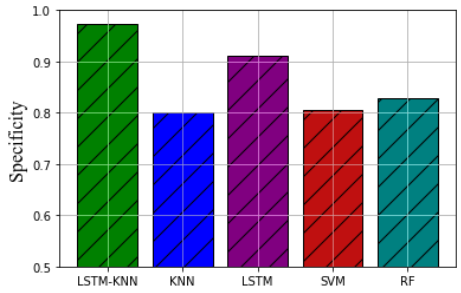

Figure 21. Comparison of specificity

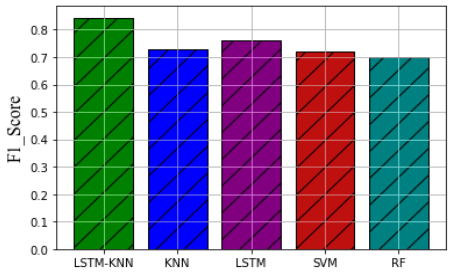

Figure 22. Comparison of f1_score

The comparison analysis of specificity (%) among the proposed hybrid LSTM-KNN and existing strategies for the detection of types of malware attacks is shown in figure 21. The specificity for the suggested hybrid LSTM-KNN method is 98% which is greater compared to current methods such as KNN is 80%, LSTM is 92%, SVM is 81%, and RF is 85%. This can demonstrate that the suggested hybrid model performs better than other techniques. The comparison analysis of f1_score (%) among the suggested hybrid LSTM-KNN and current strategies for the detection of types of malware attacks is shown in figure 22. The graphical illustration is based on several machine learning and deep learning techniques, and the value of f1_score found for the suggested hybrid LSTM-KNN method is 85% which is greater compared to current methods such as KNN is 72%, LSTM is 77%, SVM is 71%, and RF is 70%. This reveals that the proposed hybrid model works attained a better f1_score value compared to other techniques.

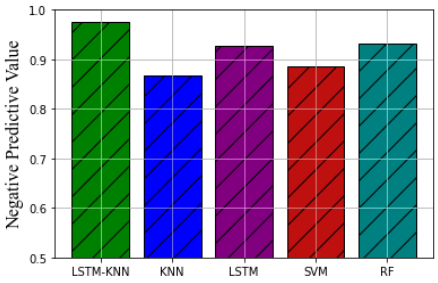

Figure 23. Comparison of Negative Predictive Value

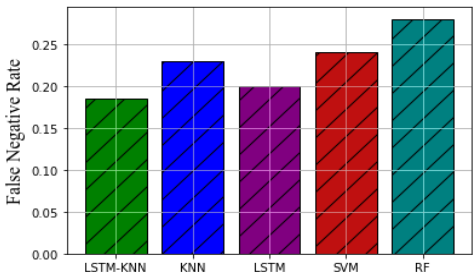

Figure 24. Comparison of False Negative Rate

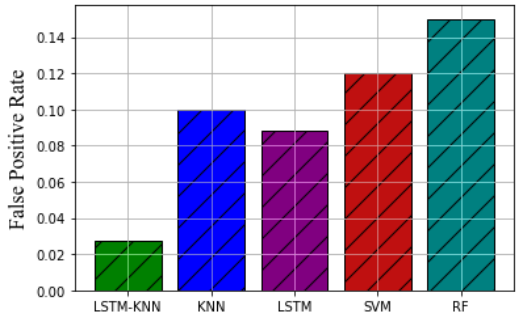

Figure 25. Comparison of false Positive Rate

The comparison analysis of NPV among the suggested hybrid LSTM-KNN and current strategies for the detection of types of malware attacks is shown in figure 23. NPV of the suggested hybrid LSTM-KNN method is 98% which is greater compared to current methods such as KNN is 88%, LSTM is 92%, SVM is 89%, and RF is 95%. This reveals that the proposed hybrid model works attained a better NPV value compared to other techniques. The comparison analysis of FNR among the suggested hybrid LSTM-KNN and current strategies for the detection of types of malware attacks is illustrated in figure 24. The FNR for the suggested hybrid LSTM-KNN method is 0.18, which is less compared to current methods such as KNN is 0.23, LSTM is 0.20, SVM is 0.24, and RF is 0.28. This reveals that the proposed hybrid model works attained less FNR value compared to other techniques. The comparison analysis of FPR among the suggested hybrid LSTM-KNN and current strategies for the detection of types of malware attacks is shown in figure 25. The graphical representation is based on several machine learning and deep learning techniques, and the value of FPR for the proposed hybrid LSTM-KNN method is0.03, which is less than existing methods such as KNN, LSTM, SVM, and RF. This reveals that the proposed hybrid model works attained less FPR value compared to other techniques.

Table 5. Comparison of proposed hybrid LSTM-KNN with other hybrid models

Methods	**Accuracy**	**False positive rate**	**False negative rate**	**error**
LSTM+attention+SVM	88.78	0.78%	11.71 %	11.22 %
LSTM+SVM	88.27%	18.67 %	12.57 %	11.73 %
GRU+SVM	84.15%	22.39 %	15.63 %	15.85 %
GRU+softmax	70.75%	4.08%	44.32 %	29.25 %
Proposed hybrid LSTM+KNN	96%	0.03%	0.18%	4%

The above table 5 shows the comparison of proposed hybrid LSTM-KNN with other hybrid models. The other hybrid models are, LSTM+attention+SVM, LSTM+SVM, GRU+SVM, GRU+softmax. When compared to the other hybrid models, the proposed attained better values than the other models which reveals the superiority of the proposed models performance.

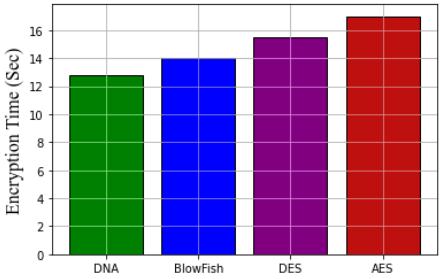

Figure 26. Comparison of Encryption Time (sec)

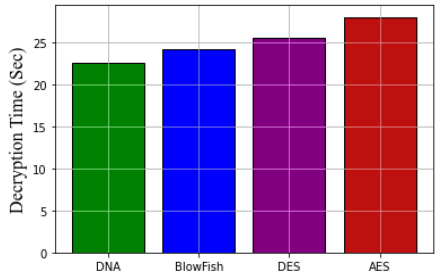

Figure 27. Comparison of Decryption Time (sec)

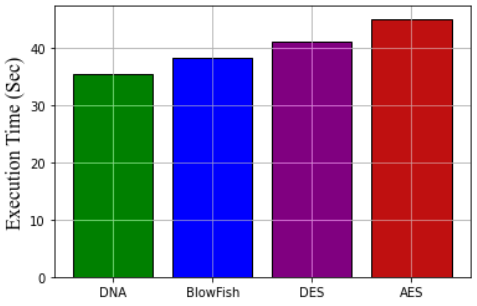

Figure 28. Comparison of execution time (sec)

Figure 26 illustrates the comparison of encryption time among proposed cryptography and existing cryptography technique. The encryption time of the proposed DNA cryptography takes 12.5 sec for encryption. Blowfish cryptography takes 14 sec for encryption time. DES cryptography takes 15.5 sec for encryption time. AES cryptography takes 17 sec for encryption time. This reveals the proposed cryptography takes less time compared to

existing techniques. Figure 27 illustrates the comparison of decryption time among proposed cryptography and existing cryptography technique. The decryption time of the proposed DNA cryptography takes 23 sec for decryption. Blowfish cryptography takes 24.5 sec for decryption time. DES cryptography takes 26.2 sec for decryption time. AES cryptography takes 27 sec for encryption time. This reveals the proposed cryptography takes less time compared to existing techniques. Figure 28 illustrates the comparison of execution time between proposed cryptography and the existing cryptography technique. The decryption time of the proposed DNA cryptography takes 35 sec for decryption. Blowfish cryptography takes 39 sec for decryption time. DES cryptography takes 41 sec for decryption time. AES cryptography takes 46 sec for encryption time. This reveals the proposed cryptography takes less time compared to existing techniques.

5. Conclusion

An efficient Intrusion Detection and Prevention Model Using COOT Optimization and Hybrid LSTM-KNN Classifier for MANET is presented for improving network security. The proposed intrusion detection and prevention approach entirely comprises four phases such as classifying normal nodes from attack nodes, predicting different types of attacks, finding the frequency of attack, and intrusion prevention mechanism. The attack nodes from the normal nodes are discovered through COOT optimization to find the optimal trust value in the initial stage. A hybrid LSTM-KNN model is introduced for the detection of different kinds of attacks in the network in the second phase. The second stage consists of preprocessing, features extraction and classification of different kinds of attacks. The third stage performs to classify the occurrence of attacks, whether the attack presents rare or frequent. The final stage is designed to restrict the attack nodes present in the network via a two-factor authentication scheme, and DNA cryptographic algorithm is used for security purposes. The proposed routing protocol is performed with an average delay of 2ms, a packet delivery ratio of 96%, a packet loss ratio of 4% and a throughput of 150 (Mbps). In addition, the proposed hybrid LSTM-KNN classification model is tested with several metrics which attained better performance, such as accuracy of 96%, the precision of 93%, recall of 82%, error value of 0.04, specificity of 98%, f1_score of 85%, Negative Predictive Value (NPV) of 98%, False Negative Rate (FNR) value is 0.18, False Positive Rate (FPR) value is 0.03. This reveals that the proposed security approach effectively mitigates the malicious attack in MANET.

Declarations

Funding: There is no funding provided to prepare the manuscript.
Conflict of Interest: The process of writing and the content of the article does not give grounds for raising the issue of a conflict of interest.

Data availability statement: If all data, models, and code generated or used during the study appear in the submitted article and no data needs to be specifically requested.
Code availability: No code is available for this manuscript.

References

[1] Abid I, Rasool, RM. Saleem M, Aleem M and Hanif R. Intrusion Detection Mechanism to Mitigate Intrusion In MANET.

[2] Sivanesh S and Dhulipala VRS. Analytical Termination of Malicious Nodes (ATOM): An Intrusion Detection System for Detecting Black Hole attack in Mobile Ad Hoc Networks. Wireless Personal Communications, 2021, 1-14.

[3] Sivanesh S and Dhulipala VR. Accurate and cognitive intrusion detection system (ACIDS): a novel black hole detection mechanism in mobile ad hoc networks. Mobile Networks and Applications, 2021, 26(4), 1696-1704.

[4] Valiveti, Ramakrishna S, Manglani A and Desai T. Anomaly-Based Intrusion Detection Systems for Mobile Ad Hoc Networks: A Practical Comprehension. International Journal of Systems and Software Security and Protection (IJSSSP), 2021, 12(2), 11-32.

[5] Makani, Ruchi, and Reddy BVR. Trust-based-tuning of Bayesian-watchdog intrusion detection for fast and improved detection of black hole attacks in mobile ad hoc networks. International Journal of Advanced Intelligence Paradigms, 2022, 21(1-2), 53-71.

[6] Popli, Renu, Sethi M, Kansal I, Garg A and Goyal N. Machine Learning Based Security Solutions in MANETs: State of the art approaches. In Journal of Physics: Conference Series, IOP Publishing, 2021, 1950(1), 012070

[7] Raj, Paul AA and Mozhi JKK. Real-Time Multi Level Behavioral Analysis Model for Efficient Intrusion Detection in Manet. Malaya Journal of Matematik, 2021, S1, 140-144.

[8] Zardari, Z. Ali, He J, Pathan MS, Qureshi S, Hussain MI, Razaque F, He P and Zhu N. Detection and prevention of Jellyfish attacks using kNN algorithm and trusted routing scheme in MANET. International Journal of Network Security, 2021, 23(1), 77-87.

[9] Farahani and Gholamreza. Black hole attack detection using K-nearest neighbor algorithm and reputation calculation in mobile ad hoc networks. Security and Communication Networks, 2021, 2021.

[10] Singh, Saurabh, Sharma S, Sharma S, Alfarraj O, Yoon B and Tolba A. Intrusion Detection System based Security Mechanism for Vehicular ad-hoc Networks for Industrial IoT. IEEE Consumer Electronics Magazine, 2021.

[11] Ahmed, Siraj NS and Acharjya DP. A framework for various attack identification in manet using multi-granular rough set. In Research Anthology on Securing Mobile Technologies and Applications, IGI Global, 2021, pp. 119-143.

[12] Srilakshmi, Uppalapati, Alghamdi S, Ankalu V V, Veeraiah N and Alotaibi Y. A secure optimization routing algorithm for mobile ad hoc networks. IEEE Access, 2022.

[13] Alghamdi and Saleh A. Novel trust-aware intrusion detection and prevention system for 5G MANET–Cloud. International Journal of Information Security, 2021, 1-20.

[14] Kowsigan M, Rajeshkumar J, Baranidharan B, Prasath N, Nalini S and Venkatachalam K. A novel intrusion detection system to alleviate the black hole attacks to improve the security and performance of the MANET. Wireless Personal Communications, 2021, 1-21.

[15] Kondaiah, Ramireddy and Sathyanarayana B. Trust factor and fuzzy-firefly integrated particle swarm optimization based intrusion detection and prevention system for secure routing of manet. International Journal of Computer Sciences and Engineering, 2018, 10(1), 13-33.
[16] Doss, Srinath, Nayyar A, Suseendran G, Tanwar S, Khanna A and Thong PH. APD-JFAD: Accurate prevention and detection of jelly fish attack in MANET. IEEE Access, 2018, 6, 56954-56965.
[17] Verma, Vanita and Jha VK Detection and Prevention of Vampire Attack for MANET. In Nanoelectronics, Circuits and Communication Systems, pp. 81-90. Springer, Singapore, 2021.
[18] Naruei, Iraj and Keynia F. A new optimization method based on COOT bird natural life model. Expert Systems with Applications, 2021, 183, 115352.
[19] Henderi, Henderi, Wahyuningsih T and Rahwanto E. Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor (kNN) Algorithm to Test the Accuracy of Types of Breast Cancer. International Journal of Informatics and Information Systems, 2021, 4(1), 13-20.
[20] Horasan and Fahrettin. Latent Semantic Indexing-Based Hybrid Collaborative Filtering for Recommender Systems. Arabian Journal for Science and Engineering, 2022, 1-15.
[21] Mohiyuddin, Aqsa, Javed AR, Chakraborty C, Rizwan M, Shabbir M and Nebhen J. Secure cloud storage for medical IoT data using adaptive neuro-fuzzy inference system. International Journal of Fuzzy Systems, 2021, 1-13.
[22] Pawar, Mohandas V and Anuradha. Detection and prevention of black-hole and wormhole attacks in wireless sensor network using optimized LSTM. International Journal of Pervasive Computing and Communications, 2021.
[23] Zardari, Ali Z, He J, Pathan MS, Qureshi S, Hussain MI, Razaque F, He P and Zhu N. Detection and prevention of Jellyfish attacks using kNN algorithm and trusted routing scheme in MANET. International Journal of Network Security, 2021, 23(1), 77-87.
[24] Kolate, Varsha, and Joshi RB. An Information Security Using DNA Cryptography along with AES Algorithm. Turkish Journal of Computer and Mathematics Education, 2021, 12(1S), 183-192.
[25] https://www.kaggle.com/kiranmahesh/nslkdd?select=kdd
[26] Zou, L., Wang, X., & Deng, L. (2021). Secure Data Fusion Analysis on Certificateless Short Signature Scheme Based on Integrated Neural Networks and Elliptic Curve Cryptography. EAI Endorsed Transactions on Scalable Information Systems, 9(34).
[27] Nouman, M., Ullah, K., & Azam, M. (2021). Secure Digital Transactions in The Education Sector Using Blockchain. EAI Endorsed Transactions on Scalable Information Systems, 9(35).
[28] Yin, J., Tang, M., Cao, J., You, M., Wang, H., & Alazab, M. (2022). Knowledge-driven cybersecurity intelligence: Software vulnerability co-exploitation behaviour discovery. IEEE Transactions on Industrial Informatics.
[29] You, M., Yin, J., Wang, H., Cao, J., Wang, K., Miao, Y., & Bertino, E. (2022). A knowledge graph empowered online learning framework for access control decision-making. World Wide Web, 1-22.
[30] Yin, J., Tang, M., Cao, J., Wang, H., You, M., & Lin, Y. (2022). Vulnerability exploitation time prediction: an integrated framework for dynamic imbalanced learning. World Wide Web, 25(1), 401-423.
[31] Jacobsson, A., & Gustavsson, C. (2003). Prediction of the number of residue contacts in proteins using LSTM neural networks. Halmstad University, 9.
[32] Liao, Y., & Vemuri, V. R. (2002). Use of k-nearest neighbor classifier for intrusion detection. Computers & security, 21(5), 439-448.
[33] Abbod, M. F., von Keyserlingk, D. G., Linkens, D. A., & Mahfouf, M. (2001). Survey of utilisation of fuzzy technology in medicine and healthcare. Fuzzy Sets and Systems, 120(2), 331-349.