International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.6992
***Article Received**: 28 March 2023 **Revised**: 16 May 2023 **Accepted**: 26 May 2023*

Secure Energy Aware Optimal Routing using Reinforcement Learning-based Decision-Making with a Hybrid Optimization Algorithm in MANET

K. Vinay Kumar[1,*], S.Venkatramulu[2],V.Chandra Shekar Rao[3],C.Srinivas[4], Sreenivas Pratapagiri[5] and B. Raghuram[6]
[1,2,3,4,5&6]Department of Computer Science and Engineering, Kakatiya Institute of Technology and Science, Warangal.
*Corresponding Author: kvk.cse@kitsw.ac.in

Abstract

Mobile ad hoc networks (MANETs) are wireless networks that are perfect for applications such as special outdoor events, communications in areas without wireless infrastructure, crises and natural disasters, and military activities because they do not require any preexisting network infrastructure and can be deployed quickly. Mobile ad hoc networks can be made to last longer through the use of clustering, which is one of the most effective uses of energy. Security is a key issue in the development of ad hoc networks. Many studies have been conducted on how to reduce the energy expenditure of the nodes in this network. The majority of these approaches might conserve energy and extend the life of the nodes. The major goal of this research is to develop an energy-aware, secure mechanism for MANETs. Secure Energy Aware Reinforcement Learning based Decision Making with Hybrid Optimization Algorithm (RL-DMHOA) is proposed for detecting the malicious node in the network. With the assistance of the optimization algorithm, data can be transferred more efficiently by choosing aggregation points that allow individual nodes to conserve power The optimum path is chosen by combining the Particle Swarm Optimization (PSO) and the Bat Algorithm (BA) to create a fitness function that maximizes across-cluster distance, delay, and node energy. Three state-of-the-art methods are compared to the suggested method on a variety of metrics. Throughput of 94.8 percent, average latency of 28.1 percent, malicious detection rate of 91.4 percent, packet delivery ratio of 92.4 percent, and network lifetime of 85.2 percent are all attained with the suggested RL-DMHOA approach.

Index Terms:- MANET, malicious nodes, Lifetime, Optimal routing, energy, decision making.

I INTRODUCTION

A Mobile Ad Hoc Network (MANET) is a group of two or more wireless devices that may dynamically create a network to exchange data without depending on any established infrastructure [1]. All nodes in such a network operate as routers or hosts, and the connection between them may change over time as nodes are added and removed [2]. MANETs offer a number of benefits over conventional networks, including the ease with which they may be put up and destroyed [3]. The mobile nodes in MANETs have a limited amount of battery power, so the energy supply for such communication devices is mainly reliant on rechargeable batteries. As a result, having an energy reserve becomes a requirement for MANET activities to run smoothly [4].

The dynamic character of the nodes' motion uses up a lot of the available energy. The node's lifetime, and by extension the network's lifetime, is prolonged when energy loss is decreased [5]. When a link between networks fails, unnecessary energy is lost. Because of the need to improve throughput and increase the nodes' residual energy lifetime, this consumption has an effect on network efficiency. As a result, it drives us to suggest a method for extending the network's lifespan by lowering energy consumption for routing [6]. The mobile nodes may be clustered into clusters to enhance the network lifespan, which is a feasible strategy to improve network stability and scalability.

The Cluster Head (CH) is an essential requirement to examine since the CH will serve as the architecture's co-coordinator [7]. By updating the routing tables following topological changes, it reduces routing overhead. The CH's job is to control the nodes in its own cluster and interact with other clusters while using the least amount of transmission power possible by preventing packet flooding [8]. Due to their dynamic structure and each node's role as a router, MANETs are not always reliable. It is also not assured that the path between two nodes is clean of malicious nodes. Link attacks can severely disrupt communications over wireless links between locations. Security in MANETs may suffer from limited resources [9]. Certain encipherment and decryption procedures require intensive computation to complete. A security solution addressing authentication, confidentiality, integrity, non-repudiation, and availability can be built on the foundation of the identified weaknesses and characteristics.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.6992
***Article Received**: 28 March 2023 **Revised**: 16 May 2023 **Accepted**: 26 May 2023*

It is not necessary to define a specific set of nodes in MANETs to carry out related activities such as network management, packet forwarding, and routing. Because of this, one of the network's nodes may be captured by an adversary, resulting in node misbehavior or non-cooperative behavior, as well as a malicious node that seeks to damage other nodes. MANET nodes may now be susceptible to a large variety of attacks, such as node capture and eavesdropping, as well as wormhole and Sybil attacks, sinkholes, and denial of service [10]. Important private data may be compromised or exploited as a result of these assaults, and the network may even be completely destroyed. Since a result, a fundamental difficulty with MANETs is the requirement for trust, as stable connection is regarded the network's backbone.

In MANETs, it is not required to identify a precise set of nodes in order to perform network administration, packet forwarding, and routing. Because of this, one of the network's nodes may be captured by an adversary, resulting in node misbehavior or non-cooperative behavior, as well as a malicious node that seeks to damage other nodes. Nodes may now be subject to a variety of assaults, including node capture and eavesdropping, wormhole and Sybil attacks, sinkholes, and denial of service attacks. Important private data may be compromised or exploited as a result of these assaults, and the network may even be completely destroyed. As a result, a fundamental difficulty with MANETs is the requirement for trust, as a stable connection is regarded as the network's backbone. The Ad-hoc On-Demand Distance Vector (AODV) protocol is a reactive MANET protocol that aims to improve scalability and performance by reducing control traffic distribution and eliminating data traffic overhead. Because AODV uses predetermined keys, it ensures that no malicious nodes exist in the network. Two forms of routing protocols are distance vector and dynamic source routing (DSR). Network hosts that use proactive routing approaches don't need to do route discovery, but they must convey more data to keep the routing tables up to date according to these standards. To put it another way, proactive routing systems can communicate more rapidly since packets don't need to be discovered first. The properties of reactive and proactive methods are combined in hybrid techniques in an attempt to improve network performance by reducing both the time and traffic overhead needed to distinguish the route for a packet. With the Secure Link State Protocol and the Zone Routing Protocol, networks are divided into clusters, with hosts only sharing information with hosts in the same cluster and clusters exchanging information only with other clusters. Recently, SDNs have become increasingly popular as a means of increasing network flexibility since a controller is assigned for routing purposes alone in these networks.

One method that has shown great promise for use in MANETs is deep reinforcement learning (DRL) for routing. A neural network is "trained" using the reward received from the environment after each activity to predict the result of the next activity, allowing the behaviors that maximize expected rewards to be carried out. Methods that fall under the umbrella of DRL include Deep Q-Learning (DQN), Dueling Deep Q-Learning (DDQN), and Policy Gradients (PG). PG predicts the probability of taking an action to maximize the reward, as opposed to the predicted profit like DQN and DDQN. For this reason, the DRL agent improves its interaction skills by learning to approximation the rules of the world. Given that a packet must pass through several hosts before arriving at its ultimate destination, and that PG has shown considerably better performance in such settings before the reward value was recognized, this makes it more ideal for the routing process. DRL is commonly used for routing in MANETs. However, the complicated calculations needed by neural networks are predicted to use a large amount of the restricted resources available on the MANET hosts. In general, MANET provides multi-hop routing and it is autonomous in nature. There is no centralized firewall present in MANETs so it is very essential to provide security. Because the network topology is dynamic as well as the node which is joining or leaving the network are dynamic that provides a ways to malicious activities in the network.

This paper presents a reinforcement learning technique for enhancing MANET trust by predicting the trustworthiness, reputation, and malevolent behavior of each node utilizing the AODV routing protocol. Due to their inherent weaknesses, wireless networks have been the focus of extensive study into ways to make them more resistant to DoS assaults. The most persistent denial of service assault is to completely deplete nodes' batteries, so while these strategies may protect a network from attacks that damage its availability in the short term, they can't defend it against attacks that last longer. This is an illustration of a resource depletion assault, in this case targeting battery life. We analyze how these so-called "vampire attacks," so named because they drain the life force of network nodes, can affect even the most secure routing algorithms. Vampire attacks, on the other hand, can affect any type of routing system, including distance vector, link-state, source routing, geographic, and beacon routing, because they exploit generic features shared by these classes. Instead of sending a deluge of data through the network, these attacks send as little as possible to achieve the largest energy drain and circumvent a rate-limiting solution. Due to the protocol

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.6992
***Article Received**: 28 March 2023 **Revised**: 16 May 2023 **Accepted**: 26 May 2023*

compliance of vampire attacks, they are hard to discover and prevent. Due to the characteristics of MANETs it is easy for the attackers to enter into the network. So it is very essential to secure the MANETs. So our main contribution is to protect the network from the malicious activities. The abbreviations used throughout the article are listed in Table 1.

Table 1. List of abbreviations

Acronym	Abbreviation
MANET	Mobile ad hoc networks
PSO	Particle Swarm Optimization
BA	Bat Algorithm
CH	Cluster Head
DRL	Deep Reinforcement Learning
DDQN	Deep Q-Learning
PG	Policy Gradients
AODV	Ad-hoc On-Demand Distance Vector
DSR	dynamic source routing
ACO	Ant Colony Optimization
SMA	Security Mobile Agent
SASR	Secure Atom Search Routing
KID	Knowledge and Intrusion Detection
TC-BAC	Trust and Centrality Degree Based Access Control

This paper has made the following contributions:

- Reinforcement Learning with Decision Making Algorithm (RL-DMA) is used to identify malicious activity in the network by adjusting to data transformations and obtaining a trust value that is relatively stable in spite of data exchange.
- In order to increase the energy efficiency of the network Tentative Cluster Head (TCH) is done which is based on the energy timer and trust value.
- Selection of optimal route using hybridization of Particle Swarm Optimization and Bat Algorithm (BA) with the maximum fitness function construction using cluster-wide distance, delay, and node energy.
- Parameters like routing overhead, end-to-end delay, energy economy, throughput, packet delivery ratio, malicious detection rate, and lifespan calculation are measured to assess the network's performance.

The paper is divided into following sections: Section II summarizes the literature on secure and optimum routing in Manets. Section III explains the proposed Secure Energy Aware Reinforcement Learning based Decision Making with Hybrid Optimization Algorithm (RL-DMHOA) for finding the malicious node in a network. The experimental analysis, presented with diagrams, is compared to three gold-standard procedures in Section IV. Section V provides a summary and suggestions for further research to wrap up the paper.

II RELATED WORKS

The MANET has become so extensively used in recent years that it has become necessary to provide nodes in the network with secure and energy-efficient communication even when they are not physically connected. Business users are generating massive volumes of data, and as a result, energy efficiency and network optimization are becoming more critical. There is a corresponding rise in the use of network bandwidth.

Nagendranath and Ramesh Babu [11] provided a MANET clustering strategy that is energy-efficient, stable, and secure. Cluster heads are selected using a fuzzy logic technique based on five variables: remaining energy level, node degree, distance, trust and mobility of the cluster nodes. There is also an extra Cluster Head to aid in the case that the main Cluster Head dies, transfers out of the cluster, or becomes insecure. Another standby CH (SBCH) is called upon to function as CH in this circumstance, and another SBCH is selected. This procedure aids in maintaining network availability while also providing an extra degree of protection. This solution requires a high level of security.

Rajeswari *et al.* [12] introduced trust-based secure energy aware clustering, a game-changing paradigm for removing malicious nodes from networks and building more stable and reliable cluster heads. It also includes two new algorithms: an energy-efficient trust-aware safe clustering method and a screening mechanism for untrustworthy recommendations. Both direct and indirect trust estimate methodologies are utilized to assess a node's degree of trust. The behavior of the node is used to assess the node's trust value. The filtering untrustworthy recommendation (FUR) algorithm's purpose is to improve clustering by preventing the trust distortion attack. As a consequence of the simulation findings, it is clear that the suggested work, trust-based secure energy-aware clustering (TSEAC), beats current work in terms of network lifespan improvement. This technique should concentrate on routing.

Robbi Rahim *et al.*[13] presented the Taylor-based Grey Wolf Optimization method, which combines the Taylor series with the Grey Wolf Optimization technique to discover the best hops for multi-hop routing. Multiple objective-based ways have been developed to accomplish safe, energy-conscious multi-hop routing, as shown in the suggested method. This technique should concentrate on routing.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.6992
***Article Received**: 28 March 2023 **Revised**: 16 May 2023 **Accepted**: 26 May 2023*

Alhasan Ameer *et al*.[14] proposed a Quality of Service (QoS) method for Internet of Things (IoT) applications In line with a four-parameter energy-aware trust model: communication trust, dependability trust, delay trust, and energy trust, the trust level of each node will be calculated, and each node will be classified depending on its degree of trust. In order to come up with the values for these parameters, we employed the k-mean clustering method. Results reveal that compared to other models, the proposed service model uses less energy because of the reduced communication costs. Security problems are not addressed in the proposed strategy.

Mallikarjuna Nandi and K Anusha [15] provided Adaptive Neuro-Fuzzy Inference System (ANFIS) based feature extraction and a classification model. The retrieved feature is trained and then categorized using the ANFIS classifier. In order to identify flooding risks in MANETs while maintaining energy economy, this paper presents the SMA2AODV protocol, which combines the Security Mobile Agent (SMA) with the Ad hoc On-demand Distance Vector (AODV) protocol. Particle swarm optimization with a fitness distance ratio and the Ant Colony Optimization (ACO) composite model were used to maximize energy efficiency. (FDR PSO). ACO-FDR PSO determines the most power-efficient path through the network and reduces energy usage to prolong the life of individual nodes. With this strategy, waiting is the name of the game.

Riasudheen *et al.* [16] presented a cloud-assisted routing strategy for MANETs that is energy-efficient. Energy consumption is reduced through fast local route recovery between mobile nodes and peer nodes. When a connection goes down, an overlay network is built that offers coverage while using less energy. Results from this suggested routing algorithm outperform current network models and routingprotocols when it comes to energy usage, residual energy, and network life span. Security should be the primary focus of this plan. Jubair *et al.* [17] developed a new bat optimized link state routing protocol. The optimum link state routing (OLSR) of a MANET and the Bat Algorithm both rely on transmitted and received signals to determine the optimal route. Because of these parallels, BOLSR was developed to predict the most efficient path between two points by analyzing their energy dynamics. The method should prioritize safe transportation. Isaac Sajan *et* al. [18] proposed Knowledge and Intrusion Detection (KID) protocol is based on the behavior of molecular dynamics and uses Secure Atom Search Routing (SASR) as its foundation. This technique utilizes the constraint and contact force of atoms to give an efficient solution for global optimization issues. In addition, SASR's efficiency is boosted by maintaining a healthy ratio of exploitation to exploration. Because the knowledge base analyses all data, the network's computational complexity is decreased and its lifespan is enhanced.

Singh *et al*. [19] proposed a EEMR (energy-efficient multipath routing) protocol for MANETs. The EEMR protocol uses a multi-objective lion optimization method to determine the best route, which consumes less energy. It is the primary goal of the EEMR protocol to find the most efficient method among numerous paths in order to meet quality standards such as latency and energy consumption.

Feroz Khan Company [20] This paper proposes an improved multi attribute based attack resistance (EMBTR) method to safely route nodes using known values. QoS factors like stability rate (SR), reliability rate (RR), and delayed time can improve network efficacy and defend against trust-related threats. (ET). Based on trusted metrics for network nodes, the method in the study excludes problematic nodes from the communication path to provide a trustworthy route. The proposed method beats state-of-the-art systems in discovery rate, energy economy, network security, and rogue node removal.

Salman Ali and coauthors [21] Mobile source and destination node locations are monitored across the entire system by dividing it up into zones. If the source and destination nodes are both mobile and in the same region, data packets can be sent directly between them using proactive routing. In a MANET, the nodes organize themselves. These locations are able to communicate with one another independent of any local networks. In this research, we compare the different MANET security attack types that have been documented so far. The unique qualities of MANETs were discussed, as were methods for measuring the effectiveness of a MANET's implementation.

Mobile ad hoc networks (MANETs) are infrastructure-free networks made of cellular mobile devices with limited battery life. MANETs' short battery life necessitates energy-awareness. MANET traffic methods aim to extend network lifetime by efficiently using limited energy. As a result, the network's layout is constantly changing, and links frequently fail due to movement limitations, wireless link sensitivity to external effects, and individual nodes' finite propagation range and leftover energy. Deepa Jeyaraj (2022) proposes SM-CSBO (sensory-modality-based cuckoo search butterfly optimisation), a new multipath routing protocol that uses a hybrid optimisation algorithm to find the best route between the source and the destination [22]. The aim is to choose the best way with the most secure links. A useful MANET routing method considers distance, normalised energy,

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.6992
***Article Received**: 28 March 2023 **Revised**: 16 May 2023 **Accepted**: 26 May 2023*

packet transit ratio, and control costs as part of the multi-objective function. With 150 nodes, the SM-CSBO algorithm model beats the PSO, SFO, CSO, and SSO algorithms by 5.8 percentage points, 30.4 percentage points, 36.7 percentage points, and 39.3 percentage points, respectively. It enhances network efficacy over the status quo, according to models.

Due to network layout dynamics, MANET routing is the biggest problem. MANET routing rests on route finding and route choosing. This study seeks the best MANET data transit way. Ravi S. (2023) suggests using test data to determine which peers can be trusted. Our mixed fuzzy optimization model, Adaptive Trust-based Secure and Optimal Route Selection, requires trustworthy nodes [23]. After Fuzzy Butterfly Optimization (FBOA) determines the most direct paths, an AES based on the Adaptive Chaotic Grey Wolf Optimization (ACGWO) method keeps chats private. Trust-based ACGWO selects the best K-paths from the results. To secure data transfers, the AES-ACGWO method checks the nodes' key and common code. Throughput, delay, packet loss ratio, packet delivery ratio, and discovery rate are used to evaluate and compare the proposed model to the state-of-the-art.

Considering the above flaws, we created a way to spot intruded nodes and provide the optimum propagation path. RL-DMA adapts to data changes and obtains a fairly accurate trust value independent of data sharing to detect network malice. Particle swarm optimization and the BAT algorithm determine the best path using cluster-wide distance, delay, and node energy as the optimum fitness function. Table.2 showing literature summary.

Table 2. literature summary

Ref No	Published Year	Approach name	Advantages	Disadvantages
[11]	2020	An energy efficient stable as well as secure clustering (EESSC)	Energy efficient and secure clustering	Need secure routing
[12]	2021	A new trust-based secure energy-aware clustering (TSEAC)	Improved the lifetime of the network	Need energy efficiency
[13]	2020	Taylor based Grey Wolf Optimization algorithm for multi-hop routing	Provides energy efficiency	Need to focus on security
[14]	2021	Quality of Service (QoS) approach to IoT application based on energy-aware trust model	Smaller energy consumption amount is achieved	Need to focus on security
[15]	2021	ANFIS (Adaptive Neuro-Fuzzy Inference System) classifier &Security Mobile Agent (SMA) with the Ad hoc On-demand Distance Vector (AODV)(SMA2AODV)	Provides energy efficient secure routing	More delay
[16]	2020	Energy-Efficient Cloud-Assisted Routing Mechanism (EECRM)	Reduce the energy consumption	Need to focus on security
[17]	2019	A new Bat Optimized Link State Routing (BOLSR) protocol	Better energy usage and bandwidth.	Security problems
[18]	2020	Secure Atom Search Routing (SASR) algorithm	Security during data transmission is provided	Need to focus on energy usage
[19]	2021	An energy-efficient multipath routing (EEMR)	Optimal path with less energy consumption	Security problems
[20]	2019	Multidimensional scaling-map (MDS-MAP) optimal routing	Energy efficient and secure routing	Need to focus on attacks

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.
Article Received: 18 March 2023 Revised: 25 April 2023 Accepted: 20 May 2023

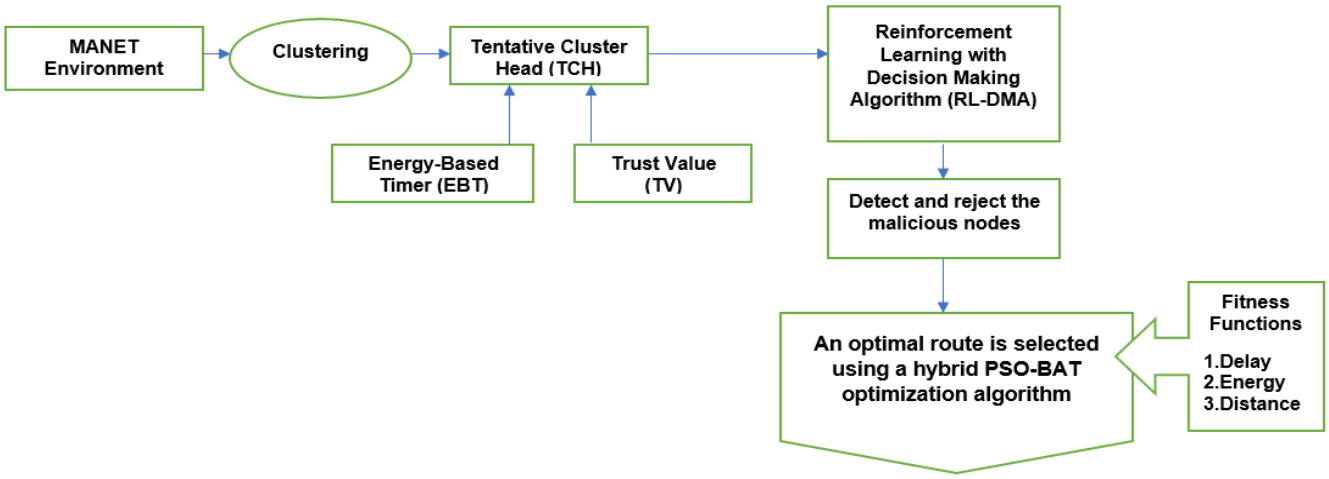

FIGURE 1a. System model for proposed method

III THE PROPOSED ROUTING ALGORITHM

3.1 SYSTEM MODEL

The Mobile Ad hoc Network is built in a unified fashion, with the cluster created and the cluster head chosen at the outset. Figure 1a depicts the system model for safe and efficient data transmission in a MANET, where malicious activity is discovered in the network using Reinforcement Learning with a Decision Making Algorithm (RL-DMA), and where the bad nodes are filtered out. Then, a hybrid PSO-BAT optimization method is used to choose the best possible path. This ensures the data is sent to the base station in an encrypted format. In the figure 1b the flow of the work is given.

3.2 FORMATION OF CLUSTERS AND SELECTION OF CLUSTER HEADS

Using an Energy Based Timer (EBT) and a Trust Value, a preliminary cluster head selection process is carried out. (TV) The value of a node in the running for Tentative Cluster Head(TCH) is based on its Trust score as a whole. TCH is chosen as the most trusted and powerful server [24].

In addition, depending on competition range, node degree, and head count, ultimate cluster head selection is proposed.

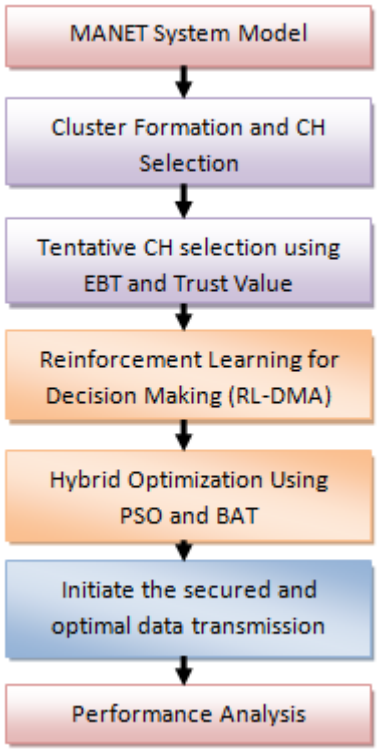

FIGURE 1b. Flow Diagram for proposed method

3.3 SELECTION OF TCH BASED ON ENERGY-BASED TIMER (EBT)

Each node's power is divvied up and used to power a clock. The amount of time the nodes have to wait is based on how much power they use. The waiting time is determined by two criteria: first-case nodes with more energy will have a shorter wait time, and second-case nodes with lower energy will have a longer wait time. The Tentative Cluster Head (TCH) would be the node whose timer value expires first. As a result of this procedure, high-energy nodes[5] are

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.
Article Received: 18 March 2023 Revised: 25 April 2023 Accepted: 20 May 2023

promoted to the position of provisional cluster leader. If not, the highest possible transfer energy would be at the same node as the cluster's epicenter. This energy-based timer's model description is as follows: Assume node i has k neighbors, and each node knows the mean energy level of its neighbors:$S_i = \{i_1, i_2, i_3, \ldots i_n \ldots . i_k, \}$ and i_n denotes the nth neighbor node. The average energy of node I can be calculated using the equation below (1)

$$Average\ Energy(i) = \begin{cases} \frac{1}{k}\sum_{n=1}^{k} Energy(i_n) & k > 0 \\ 0 & k = 0 \end{cases} \quad (1)$$

The Energy-based timer is used to pick TCH from the nodes. For any node ID S_i , The value of waiting time dependent on energy can be calculated as indicated in equation (2)

$$WaitTime\ (s_i) = \frac{AvgEnergyofs_iNeighbornode}{EnergyofS_i} \quad (2)$$

According to the equation above, as the node's energy grows, the waiting time reduces. The node with the most energy will have a shorter wait time. The Cluster Head is this selected component. As the waiting timer nears its conclusion, nodes that have received the tentative CH message from the chosen CH will exit the cluster head selection. Cluster distance, total energy (E_{total}), and Trust Value (TV)were used to pick the preliminary cluster. The procedure determines the greatest and minimum distance of nodes in each round based on the preliminary cluster selection criteria.

3.4 SELECTING A TCH BASED ON THE TRUST VALUE

The Trust Value (TV) measures the reliability of a component and is used to evaluate its overall usefulness. Data collection, node reconfiguration, and forwarding are also possible with its help. It provides a quantitative approach to determining which networks can be relied upon. The reliability of a component is measured and tracked for statistical and observational purposes. The confidence value is combined with the Energy Based Timer (EBT) to pinpoint the likely cluster's epicenter. Two methods are employed in tentative CH selection to attain the highest possible efficiency in cluster head selection. (EBT and TV). The following formula is used to determine a node's trustworthiness: (3).

$$Trust\ Value(TV)_{nodes} = \frac{N_{FD}}{N_{REC}} \quad (3)$$

Where NFD is the total number of packets sent and NREC is the total number of packets received. The trustworthiness of each node is calculated, and the node with the highest number is appointed as the acting cluster leader. After that, the final CH procedure is carried out. Finally, the EBT and TV deliver the TCH selection result.

3.5 REINFORCEMENT LEARNING WITH DECISION MAKING ALGORITHM (RL-DMA)

The Sink is the base station of the ad hoc network model, indicating the degree of direct interaction (or contact) between the current node and the Sink. Sink's communication coverage spans the whole ad hoc network. The purpose of normal nodes is to pass messages to sink; the closer current nodes and sink interact, the greater the contact intimacy value, as illustrated in equation (4).

$$Vc = \frac{Cs}{C} \quad (4)$$

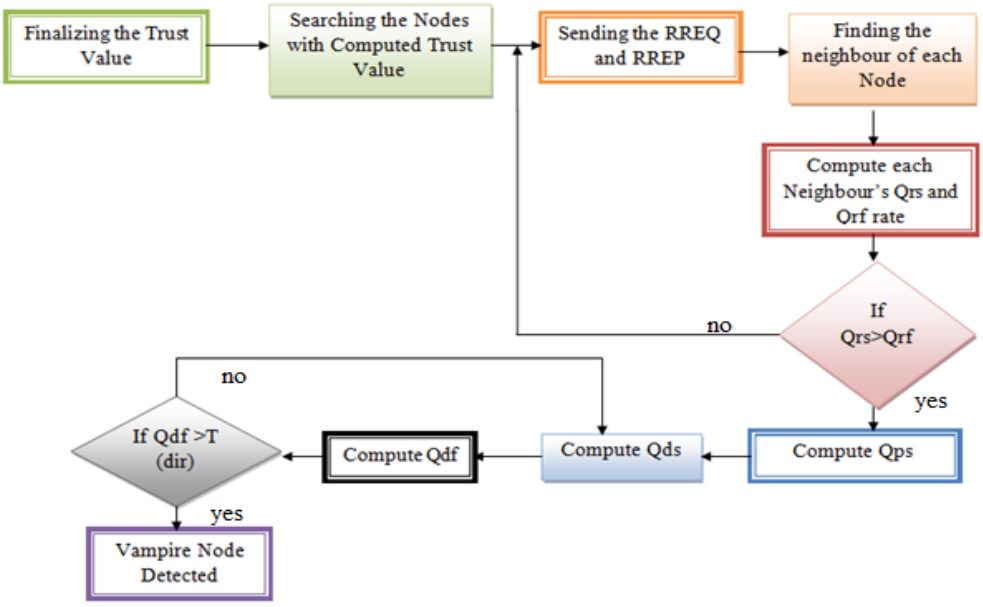

FIGURE 2. Workflow chart for malicious (vampire) nodes rejection

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.
Article Received: 18 March 2023 Revised: 25 April 2023 Accepted: 20 May 2023

Where Vc is the value of the number of contacts, CS is the rate at which the current node exchanges data with the drain, and C is the rate at which the current node exchanges data with all other nodes over the course of the network's existence. The location intimacy value quantifies how closely the present node is situated to the sink. The mobility path has the features of people's behavior, with nodes made up of portable gadgets. The closer a node is to the sink, the more partial its movement is to the sink, and the more likely it is to collide with the sink. As a result, the chance of transmission rises, as indicated in equation (5).

$$V_L = \frac{Cp}{Rp} \quad (5)$$

Where VL stands for location intimacy, Cp for node transfer packages, and Rp for the total amount of packets nodes receive from other nodes.

The nodes communicate with one another via route request (RREQ) and route respond (RREP) messages. The percentage of successful query requests is equal to the number of neighbor nodes that have gotten RREQ messages from the originating node. (Qrs). An RREQ message from the source node is used to calculate the query request failure rate by counting the number of neighbor nodes that have not gotten the message.(Qrf). To calculate Qps, we count the number of RREPs that the node that broadcast the RREQ got that were successful. Counting up the number of neighboring nodes that haven't replied to the query request allows us to calculate the query reply failure rate, abbreviated Qpf. The data success rate (Qds) is computed based on successfully transmitted data, whereas the data failure rate (Qdf) is derived based on data that did not reach its destination, as shown in equations (6),(7), and (8).

$$Qr = (Qrs-Qrf)/(Qrs+Qrf) \quad (6)$$

$$Qp = (Qps-Qpf)/(Qps+Qpf) \quad (7)$$

$$Qd = (Qds-Qdf)/(Qds+Qdf) \quad (8)$$

The intermediate variables Qr, Qp, and Qd are used to determine the nodes Request rate, Reply rate, and Data transfer rate. The values of Qr, Qp, and Qd have been normalized to lie between -1 and +1. If the numbers go outside of the normalized range, it indicates that the node's failure rate is high, and that the accompanying node may not be acceptable for routing is shown in Figure 2.

Setup phase- The MANET's timer for periodic transmission is kept in sync by each node. The neighbor node's timer checks the neighbor's expiration time and removes it from the table if it's legitimate. Data transmission from the source node begins once the neighbor table procedure is complete. Dropper nodes receive the data packet and discard it without sending it. In order to calculate the node's packet loss count, the data forwarder device engages in channel overhearing. This data is analyzed to determine the delay in information packet transmission as a function of the number of request and reply messages transmitted.

Detection phase - The parameters for validating the characteristics of nodes and the variation between them are found with the detection phase. Boltzmann input processing with the joint probability distribution is used to apply the structured learning process.

Learning and Adaptation phase- To accurately identify the nodes' behavior, reinforcement learning with adaptation is applied to the data collected in the detection phase. To tell the difference between network congestion and malicious behavior, the learning forecasts the cause of packet drop and delay. Validation of state-action policies informs the application of the forecast. In order to keep the data and the policy for state action up to current, a learning timer is used. The decision tree model is used to incorporate the adaptation into the detection process after the inference has been found in the learning system.

State Representation- MANET agents maintain two sets of data: one set is available to all nodes in the network, while the other is exclusive to the task of performing next-hop selection. To ensure that the computations required to extract the resulting features are not repeated, they are spread over the network. It then adds any additional data required to determine the packet's next hop and selects the appropriate node. There's no way to know exactly how much energy a node has left unless you run frequent reports. Each node that is capable of receiving and transmitting a packet is assigned one in the initial array. The residual normalized energy at each node over the course of an S-second time frame is shown in three additional arrays that are positioned in accordance with the measurement's location. As a consequence, the agent is able to monitor both energy use and the movement of nodes. This is followed by an array that describes the node's normalized range and an array of ones around its centre.

Training Procedure- Every node in a neural network must be trained simultaneously. Many MANETs must be employed in the training process to provide the agent flexibility. However, changing the MANET might cause the neural network to get confused during training. After training the neural network, a collection of 1000 randomly generated MANETs is constructed and stored on disc. The same networks used to analyse the agent's conduct allow for many MANETs to detect diverse situations while preserving reward levels. For each packet-routing action, MANET's

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.
Article Received: 18 March 2023 Revised: 25 April 2023 Accepted: 20 May 2023

lifespan is discounted, indicating recent activities that diminished a node's power. The reward value calculated for a set of activities taken to route a packet at the beginning of a MANET's life might be drastically altered by these actions. A moderate learning rate for backpropagation is also required to update the neural network weights and biases, limiting the impact of errors that may occur during packet routing. As a result, consecutive routing errors have less influence on a packet's efficiency.

3.6 SELECTION OF ENERGY AWARE OPTIMAL ROUTE USING HYBRID PSO AND BAT

This section provides a detailed explanation of how the hybrid PSO, and BAT Optimization Algorithm was used to find an energy-conscious optimum path. It is superior in finding the optimal solution in terms of accuracy and iteration. Consisting of a straightforward idea, straightforward programming, rapid convergence, and generally superior answer. A nomenclature of symbols with the definitions used is presented in table 3.

Table 3. Symbols and definitions

Symbols	definitions
X	Grey wolf
n	Size of the search space
I	Iteration
α, β, and δ	Fitness values
t	Reputation state
A	Coordination vector
$Fitness_{max}$	Multi-objective Fitness function
D(t)	Delay from sensor nodes
D_{norm}	Normalized delay
T (t)	Traffic rate
Z(t)	Cluster density

1) PSO ALGORITHM

PSO stands for particle swarm optimization [25]. When Kennedy and Eberhart first proposed PSO, they used it on a simple social sample and started with a variety of random answers (particles). In order to maximize the value of the cost function that is evaluated at each particle's position, each particle has a velocity that falls within a user-defined range. To find a better solution, the following technique is followed by each particle as it moves across the search space.

$$vk(j+1) = wvk(j) + s1n1[\text{ P bestk } - xk(j)] + s2n2[\text{ Gbest } - xkj] \quad (9)$$

$$xk(j+1) = xk(j) + vk(j+1) \quad (10)$$

Particle position (x) is represented by (v), velocity (k), and stable accelerations (s1, s2) are all equal to 2, and n1 and n2 are independent random integers (Pbest, Gbest) in the range (1, 1). Pbest is the best, least expensive position it has discovered so far. In this research, these equations were used to optimize the particle's position until the lowest possible expense was achieved.

2) BAT

Bats utilize microbats' echolocation to maintain their position and velocity, and the Bat method is used to handle optimization problems in a similar fashion [26],[27]. The technique iteratively adjusts the bats' locations in response to changes in the sound wave's speed and frequency in hertz. The locations of the bats are used as a representation of the problem's factors in network optimization. Each cycle will use the preceding cycle's velocity, frequency, and global statistics to calculate the current position and speed of every bat in the swarm. The following equations are used in the BAT technique to calculate new values for velocity and position:

$$F[j] = F[min] + (F[max] - F[min])\beta \quad (11)$$
$$Vt[j] = Vt - 1[j] + (Xt[j] - X[g])F[j] \quad (12)$$
$$Xt + 1[j] = Xt[j] + Vt + 1[j] \quad (13)$$

The microbat's sound wave frequency, in hertz, is represented by F (j), F (min), and F (max) at time t, and β is a random vector with values between 0 and 1. The current worldwide best solution, denoted by X (g), is shown below for each bat. The swarm's global best answer, X (g), is calculated after all of the iterations are complete.

3.7 THE HYBRID PSO-BAT ALGORITHM

A hybrid PSO-BAT routing method for MANETs is suggested in the following algorithmic steps in table 4:

Table 4. Algorithmic steps

Step 1	Specify the number of nodes, the number of connections, and the rate of speed sequentially
Step 2	Give the algorithm the outcomes of the scenarios as a dataset
Step 3	Compare the network parameters generated at random with the dataset.
Step 4	Collect evaluation data (delay, energy, distance).
Step 5	In PSO, use equations (9) and (10), and in BAT, use equations (12) and (13), to determine the position and speed of the present particle or bat agent.
Step 6	Once all of the updates have been completed, the procedure should be ended
Step 7	If no optimum solution is found after step 2, go on to the next agent and repeat steps 3–6 to optimize

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.
Article Received: 18 March 2023 Revised: 25 April 2023 Accepted: 20 May 2023

3.8 FITNESS FUNCTION WITH SEVERAL OBJECTIVES

The cluster hub (CH) is selected by applying the multi-objective fitness function to all of the cluster nodes. Objectives of the optimization fitness function for optimal path selection include minimizing the total travel time, average energy per node, and cluster density. The cluster head is defined as a node that meets this function with the highest value. Equation (14) expresses the multi-objective fitness function as follows

$$Fitness_{max} = \left\{\left[1 + \frac{D(t)}{D_{norm}}\right] + \left[1 - \frac{M(t)}{X*A*N}\right] + [c(t)] + [1 - T(t)] + [1 - z(t)]\right\} \quad (14)$$

where D(t) and D_{norm} represents the node delay and the adhoc network's normalized delay, respectively. The node-to-CH distance is denoted by M(t), the number of nodes in the cluster is denoted by X, and the number of CHs is denoted by A. N is the total amount of network nodes. C is the symbol for the energy contained in CHs.(t). The cluster node transmission rate, T(t), and cluster density, Z(t), are both functions of time.

The time it takes for the node to transmit the info is the delay. The best choice of CH is one that minimizes delay. Node latency is based on the node's Expected transfer Count (ETC) and the delay in propagation and transfer through the network. In equation (15), the node delay is stated as follows

$$Delay\ D(t) = \sum_{i=1}^{N} L_i(t)(\alpha + \beta_i) \quad (15)$$

where $L_i(t)$ is ETC of i^{th} node at time t, α signifies the time it takes for data to be transmitted through the network and β_i is the time it takes for the message to spread i^{th} node. ETC is calculated using the node's forwarded and received packet delivery ratio at time t. The ETC of the ith node is stated as follows in equation (16)

$$L_i(t) = \frac{1}{F_i(t) * R_i(t)} \quad (16)$$

At instant t, the ith node has forwarding and receiving packet delivery rates of Fi(t) and Ri(t), respectively.

Distance: The distance of nodes is the distance between one node and the next depending on the space of CH. This distance must be kept at a minimum in order to converse successfully. The distance is calculated using the nodes of the jth cluster, as stated in equation (17)

$$Distance, M(t) = \sum_{j=1}^{A} \sum_{i=1}^{N} \|S_i - H_j\|; \quad (17)$$

Energy: The CH is chosen as the node with the most energy. As indicated in equation, energy is provided (18)

$$Energy, C(t) = \frac{1}{A} \sum_{j=1}^{A} C_{t+1}(H_j) \quad (18)$$

where Ct+1(Hj) is the updated energy of CH as calculated by equation (19)

$$C_{t+1}(H_j) = C_t(H_j) - C_{dissipation}(H_j) \quad (19)$$

C_t (Hj) is the energy of CH at time t and $C_{dissipation}(H_j)$is dissipated energy of CH.

I) UPDATE THE ROUTE FOR SECURE AND OPTIMAL TRANSMISSION

A node's power consumption will vary depending on the surrounding environment. Routing information is used to find the quickest and most efficient path for the data transfer. It is understood by the network how much power is used in transmitting data. All nodes in the network receive DREQ data requests from the base station. The node sends a data reply DREP packet in response to the data request packet. When receiving information from the hub, a node must go through a series of steps before it can use that information. Using a private key, the Node verifies the validity of any transmitted data. If the packet's common key checks out, it continues on its way. The node always stays connected to the main station via a single shared key. Since both the sending and receiving nodes are in the same preparation state, data cannot be sent to a node that is already in use. In cases when the present node is not the intended recipient, the message is resent to the neighbor list. The base station determines the optimal path after collecting data from all nodes in line with the preceding phase. Once the base station has determined the most efficient way, it sends out a request for a new route. Expect a route to respond to this message by recognizing the packet sent from Node. An Error Packet (ERRP) is sent instead of a data response when the security key does not match.

Pseudo Code for the proposed work:

Step 1 – Construction of the system model.

Step 2 – Choosing a cluster leader based on a combination of energy efficiency and trustworthiness.

Step 3 – Reinforcement learning based decision making algorithm (RL-DMA)

Step 4 – RL-DMA process, setup, detection, learning & adaption, state representation and training

Step 5 – Hybrid optimization using PSO and BAT for optimal path finding

Step 6 – Route update for secured optimal data transfer

Step 7 – Performance analysis

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.
Article Received: 18 March 2023 Revised: 25 April 2023 Accepted: 20 May 2023

IV EXPERIMENTAL ANALYSIS

Parameters such as routing overhead, end-to-end delay, energy efficiency, throughput, average latency, malicious detection rate, packet delivery ratio, and network lifespan are used to generate the experimental outcome. Three state-of-the-art methods, the Taylor-based Grey Wolf Optimization method (TGWO)[28], the energy-aware trust model (EATM)[29], and the Adaptive Neuro-Fuzzy Inference System (ANFIS)[30], are compared to the proposed Reinforcement Learning based Decision Making with Hybrid Optimization Algorithm regarding these parameters. (RL-DMHOA). In table 5 we can see how the simulation is built up.

Table 5. Simulation Setup

Parameters	Ranges
Nodes that are normal	500
Nodes that are vampire nodes	100
Number of sink	1
Simulation area	1000*1000 m
Packet size	300KB
Transmission rate	200KB
Communication range	30m

As shown in equation (20), routing overhead (RO) can be thought of as the total number of maintenance and route finding routing packets sent.

$$RO = \frac{H}{P} \quad (20)$$

Where P is the total amount of packets being routed and H is the hop count. Routing overhead comparison between the current and suggested methods is shown in Table 6.

Table 6. Comparison for routing overhead of proposed method

Number of vampire nodes	TGWO	EATM	ANFIS	RL-DMHOA [proposed]
20	38	35	31	25
40	42	38	34	28
60	45	41	36	32
80	49	42	37	35
100	52	45	39	37

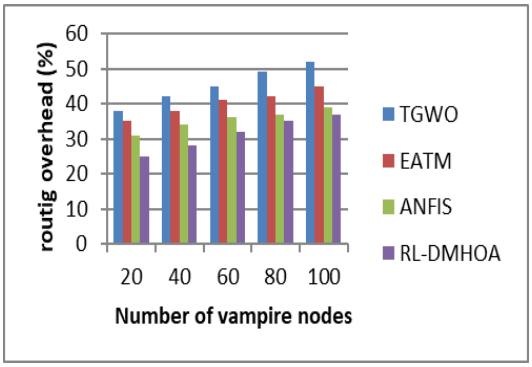

Figure 3. Routing overhead of proposed method

Figure 3 displays a comparison between the routing overhead of the suggested RL-DMHOA method and that of the currently used approaches. The percentage of routing overhead is shown along the y axis, and the overall number of vampire nodes is shown along the x axis. When compared to the current TGWO, EATM, and ANFIS methods, which yield 45.2%, 40.2%, and 35.4%, respectively, the proposed RL-DMHOA method gets 31.4%, which is 13.8% better. **End-to-end delay** measures how long it takes to travel through a network from its point of origin to its point of destination. Table 7 compares the end-to-end delay between existing methods and the proposed method.

.Table 7. Comparison for end-to-end delay of the proposed method

Number of vampire nodes	TGWO	EATM	ANFIS	RL-DMHOA [proposed]
20	39	32	25	16
40	42	35	29	19
60	45	39	32	22
80	49	42	35	26
100	51	46	39	32

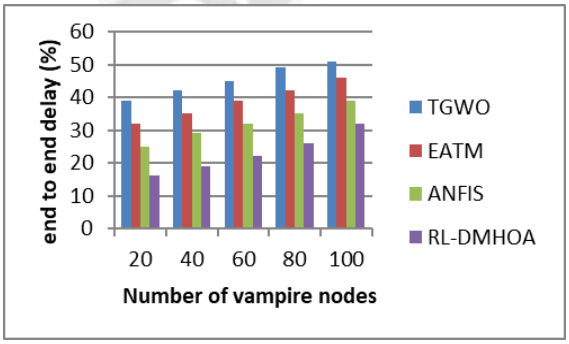

Figure 4. Proposed RL-DMHOA End-to-end delay

End-to-end delays are shown in Figure 4 for both the current methods and the suggested RL-DMHOA method. End-to-end delay percentages are plotted against the amount of vampire nodes used in the analysis along the Y axis. The proposed RL-DMHOA method achieves 23%, which is

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.
Article Received: 18 March 2023 Revised: 25 April 2023 Accepted: 20 May 2023

22.2% better than TGWO, 15.8% better than EATM, and 9% better than ANFIS, when compared to the existing TGWO, EATM, and ANFIS methods, which obtain 45.2%, 38.8%, and 32%, respectively.

Energy efficiency is the ratio of energy output to energy input, which is given as shown in the equation (21).

$$E=\frac{Wout}{Win}\times 100 \quad (21)$$

Table 8 compares the energy efficiency between existing methods and proposed RL-DMHOA method.

Table 8. Comparison for energy efficiency of the proposed method

Number of vampire nodes	TGWO	EATM	ANFIS	RL-DMHOA [proposed]
20	65	68	70	72
40	68	71	72	75
60	70	73	74	79
80	72	75	76	82
100	75	77	79	85

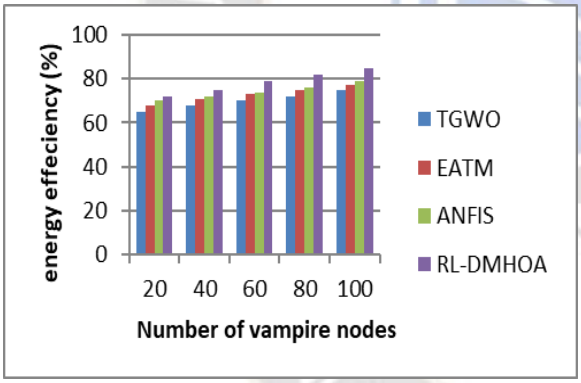

Figure 5. Energy efficiency for the proposed method

In Figure 5, the X-axis depicts the total number of vampire nodes used in the analysis, while the Y-axis displays the percentage of energy savings achieved using each of the three current methods and the proposed RL-DMHOA method. The suggested RL-DMHOA method achieves 78.6%, which is 8.6% better than TGWO, 6.2% better than EATM, and 4.2% better than ANFIS, respectively, when compared to the existing TGWO, EATM, and ANFIS methods, which achieve 70%, 72.8%, and 74.2%, respectively.

The throughput of a communication channel is the pace at which data can be transferred through it. The network channel has been used effectively to transmit bits or packets. Throughput, which is shown by the equation (22), is crucial in MANET apps.

$$\text{Throughput (bits/sec)} = \sum \frac{(n)*(avg)}{T} \quad (22)$$

Where, n= number of successful packets, Avg= average packet size, T= Total time spent in delivering the data. Table 9 compares the throughput between existing methods and proposed RL-DMHOA method.

Table 9. Comparison for throughput of the proposed method

Number of vampire nodes	TGWO	EATM	ANFIS	RL-DMHOA [proposed]
20	83	85	90	92
40	84	86	91	93
60	86	88	93	95
80	88	90	94	96
100	90	92	95	98

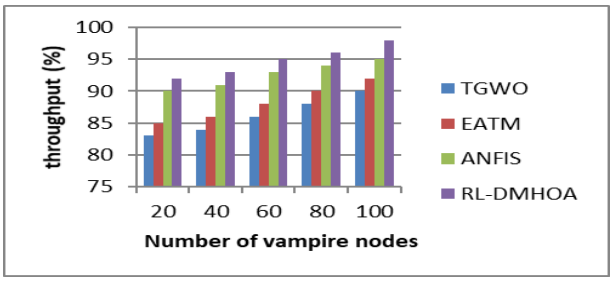

Figure 6. Throughput of the proposed method

Throughput is compared in Figure 6 between the RL-DMHOA technique and the existing methods, with the number of vampire nodes used in the analysis on the X axis and the percentage of throughput on the Y axis. The suggested RL-DMHOA method achieves 94.8%, which is 8.6% better than TGWO, 6.6% better than EATM, and 2.2% better than ANFIS, when compared to the existing TGWO, EATM, and ANFIS methods, respectively, which achieve 86.2%, 88.2%, and 92.6%.

The average latency between the pckets can be determined based on where you are in relation to the sink. Equation (23) describes the time it takes for a packet to travel along its assigned route Pmn during transmission.

$$P_{mn} = p_{mn} + \sum_{jN(n)} (Pnj \times Dnj) \quad (23)$$

Where, P_{mn} is data delivery delay, Pnj is probability that packet is forwarded through allocated path, N(n) is number of neighbor channel, Dnj is distance between two paths. Table 10 compares the average latency between existing methods and proposed RL-DMHOA method.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.
Article Received: 18 March 2023 Revised: 25 April 2023 Accepted: 20 May 2023

Table 10. Comparison for average latency of the proposed method

Number of vampire nodes	TGWO	EATM	ANFIS	RL-DMHOA [proposed]
20	38	34	31	24
40	40	35	34	26
60	45	39	36	29
80	49	42	39	30
100	50	45	42	32

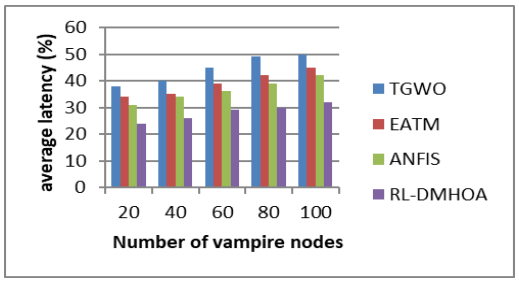

Figure 7. Proposed RL-DMHOA Average latency

Figure 9 compares the average latency of the current methods to that of the proposed RL-DMHOA method, where the X axis represents the number of vampire nodes used for analysis and the Y axis represents the average latency values as percentages. When compared to the existing TGWO, EATM, and ANFIS methods, which obtain 44.4%, 39%, and 36.4%, respectively, the proposed RL-DMHOA method gets 28.2%, which is 16.2% better.

Packet Delivery Ratio (PDR) is the average ratio of total packets successfully received (R) to total packets initially issued (S) shown in equation (24):

$$\text{PDR}= \sum_{0}^{N} \frac{R}{S} \qquad (24)$$

Table 11 compares the packet delivery ratio between existing methods and proposed RL-DMHOA method.

Table 11. Comparison for PDR of the proposed method

Number of vampire nodes	TGWO	EATM	ANFIS	RL-DMHOA [proposed]
20	77	80	85	87
40	79	82	87	90
60	81	86	90	92
80	82	88	92	95
100	86	90	93	98

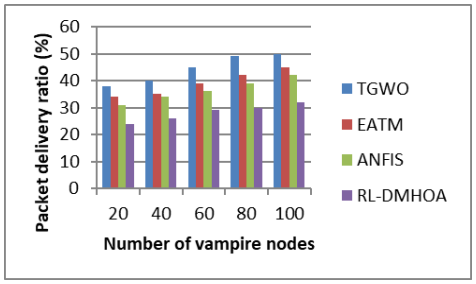

Figure 8. Proposed RL-DMHOA packet delivery ratio

In Figure 8, the X-axis depicts the number of vampire nodes used for analysis, and the Y-axis displays the values obtained as a percentage, comparing the packet delivery ratio of the suggested RL-DMHOA technique to that of the existing methods. When compared to the current TGWO, EATM, and ANFIS methods, which obtain 81%, 85.2%, and 89.4%, respectively, the proposed RL-DMHOA method achieves 92.4%, which is 11.4% better.

Network Lifetime (NL) is the amount of time that all nodes in the network can stay alive before one or more of them run out of energy. The formula is given as shown in equation (25).

$$\text{NL}=\sum_{T=1}^{T} Etx(k,d) + \sum_{r=1}^{r} Erx(k) \qquad (25)$$

Where,Etx(k, d) is the transmitted energy between node k and d,Erx(k) is the remaining energy at the destination side x(k).Table 12 compares the network lifetime between existing methods and proposed RL-DMHOA method.

Table 12. Comparison for network lifetime of the proposed method

Number of vampire nodes	TGWO	EATM	ANFIS	RL-DMHOA [proposed]
20	71	75	77	80
40	73	76	79	83
60	75	79	81	85
80	77	80	85	88
100	79	82	88	90

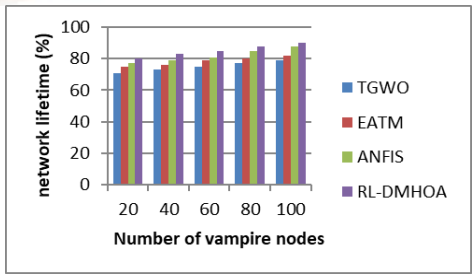

Figure 9. Proposed RL-DMHOA network lifetime

The number of vampire nodes used in the analysis is plotted along the X-axis, and the percentage values derived are

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.
Article Received: 18 March 2023 Revised: 25 April 2023 Accepted: 20 May 2023

shown along the Y-axis in Figure 9, which compares the network lifetime using the proposed RL-DMHOA technique to those using existing methods. When compared to the existing TGWO, EATM, and ANFIS methods, which produce 75%, 78.4%, and 82%, respectively, the proposed RL-DMHOA method achieves 85.2%, which is 10.2% better.

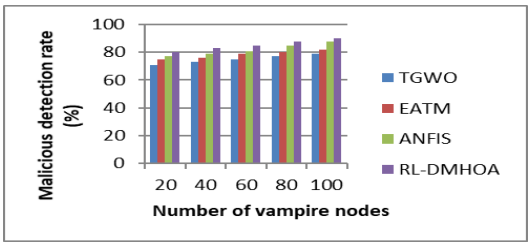

Figure 10. Detection rate of proposed RL-DMHOA

Figure 10 shows a comparison of the malicious detection rate between the current methods and the proposed RL-DMHOA method, with the number of vampire nodes used for analysis plotted along the X axis and the percentage values obtained along the Y axis. The suggested RL-DMHOA method outperforms the existing TGWO, EATM, and ANFIS methods by 11.6%, 7.4%, and 3.4%, respectively (91.4% vs. 80.8%, 84.6%, and 88%, respectively). The overall comparison analysis between the current methods and the proposed RL-DMHOA approach is shown in Table 13.

Table 13. Overall comparative analysis

Parameters	TGWO	EATM	ANFIS	RL-DMHOA [proposed]
Network Lifetime (%)	75	78.4	82	85.2
Energy Efficiency (%)	70	72.8	74.2	78.6
End to End Delay (%)	45.2	38.8	32	23
Throughput (%)	86.2	88.2	92.6	94.8
Routing Overhead (%)	45.2	40.2	35.4	31.4
Average Latency (%)	44.4	39	36.4	28.2
Malicious Detection Rate (%)	80.8	84.6	88	91.4
Packet Delivery Ratio (%)	81	85.2	89.4	92.4

V CONCLUSION

Wireless communications are increasingly used to establish connections and access services, so their existing design has grown limited. For this reason, ad-hoc networks have emerged to allow network nodes to join by transmitting packets. Because of this, network packet routing has been problematic due to a lack of infrastructure and node mobility. Making the most of the limited resources on the nodes is one of the main challenges with these networks. Our main motivation for writing this paper was to suggest a secure energy-aware clustering algorithm for MANET. To deal with the security issues posed by malicious nodes and to choose a trustworthy node as Cluster Head, we propose a Secure Energy Aware Reinforcement Learning based Decision Making with Hybrid Optimization Algorithm (RL-DMHOA) for protection against the vampire attack and the provision of the shortest path for packet transmission. (CH). To overcome the difficulty of finding the best answer, MANET makes use of both BAT and PSO. The simulation results show that the RL-DMHOA approach reliably guarantees 31.4% routing overhead, 23% end-to-end delay, 78% energy efficiency, 94.8 % throughput, 28.2% average latency, 91.4 % malicious detection rate, 92.4 % packet delivery ratio, 85.2 % network lifetime, 15.2% communication cost, and 28.6% trust computation error. Thus, the proposed method outperforms existing methods in speed and reliability. Future research may use sleep deprivation attacks and directional antenna attacks that cause service loss to build protection methods and studies.

REFERENCES

[1] N.Subramani et al., "An Efficient Metaheuristic-Based Clustering with Routing Protocol for Underwater Wireless Sensor Networks," Sensors,vol.22,no.2,pp. 415,2022.

[2] N. Veeraiah and B. Tirumala Krishna, ''Trust-aware fuzzyclus-fuzzy NB: Intrusion detection scheme based on fuzzy clustering and Bayesian rule,'' Wireless Networks, vol. 25, pp. 4021–4035, Jan. 2019.

[3] S.Rajendran et al.,"MapReduce-based big data classification model using feature subset selection and hyperparameter tuned deep belief network," Scientific Reports,vol.11,no.1,pp.1-10,2021.

[4] N. Veeraiah and B. T. Krishna, ''An approach for optimal-secure multi-path routing and intrusion detection in MANET,'' Evolutionary Intelligence, pp. 1–15, Mar. 2020.

[5] S. V. Kumar and V. AnurathaEnergy, ''Efficient routing for MANET using optimized hierarchical routing algorithm (Ee-Ohra),'' Int. J. Sci. Technol. Res., vol. 9, no. 2, pp. 2157–2162, Feb. 2020.

[6] G. M. Borkar and A. R. Mahajan, ''A secure and trust based on-demand multipath routing scheme for self-organized mobile ad-hoc networks,'' Wireless Networks, vol. 23, no. 8, pp. 2455–2472, Nov. 2017.

[7] A. Taha et al., " Energy efficient multipath routingprotocol for mobile ad-hoc network using the fitness function,'' IEEE Access, vol. 5, pp. 10369–10381, 2017.

[8] SriLakshmi Uppalapati, and B. S. Rao. "An overhead aware multipath routing protocol for improving relay node

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6s
DOI: https://doi.org/10.17762/ijritcc.v11i6s.
Article Received: 18 March 2023 Revised: 25 April 2023 Accepted: 20 May 2023

selection in MANET."International Journal,vol.8,no.1 2019.

[9] Mr. Bhushan Bandre, Ms. Rashmi Khalatkar. (2015). Impact of Data Mining Technique in Education Institutions. International Journal of New Practices in Management and Engineering, 4(02), 01 - 07. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/35

[10] U. Srilakshmi, S. Alghamdi, V. V. Ankalu, N. Veeraiah and Y. Alotaibi, "A secure optimization routing algorithm for mobile ad hoc networks," in IEEE Access, doi: 10.1109/ACCESS.2022.3144679.

[11] U. Srilakshmi et al., "An Improved Hybrid Secure Multipath Routing Protocol for MANET," in IEEE Access, vol. 9, pp. 163043-163053, 2021.

[12] M V S S Nagendranath and A. Ramesh Babu, A.R. "An efficient mobility aware stable and secure clustering protocol for mobile ADHOC networks," Peer-to-Peer Networking and Applications,vol.13, pp. 1185–1192,2020.

[13] A.R.Rajeswari et al.,"An efficient trust based secure energy-aware clustering to mitigate trust distortion attack in mobile ad-hoc network,"Concurrency and Computation: Practice and Experience,vol.33,no.12,2021.

[14] Raj, R., & Sahoo, D. S. S. . (2021). Detection of Botnet Using Deep Learning Architecture Using Chrome 23 Pattern with IOT. Research Journal of Computer Systems and Engineering, 2(2), 38:44. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/31

[15] Robbi Rahim et al.,"Taylor Based Grey Wolf Optimization Algorithm (TGWOA) for Energy Aware Secure Routing Protocol," International Journal of Computer Networks and Applications,vol.7,no.4,pp.93,2020.

[16] Ameer Alhasan et al., "An Energy Aware Qos Trust Model For Energy Consumption Enhancement Based On Clusters For IOT Networks," Journal of Engineering Science and Technology, vol.16, no.2,pp.957-976,2021.

[17] M.Nandi and K.Anusha, "An Optimized and Hybrid Energy Aware Routing Model for Effective Detection of Flooding Attacks in a Manet Environment," Wireless Pers Communications,2021.

[18] Russo, L., Kamińska, K., Christensen, M., Martínez, L., & Costa, A. Machine Learning for Real-Time Decision Support in Engineering Operations. Kuwait Journal of Machine Learning, 1(2). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/117

[19] H. Riasudheen et al., "An efficient energy-aware routing scheme for cloud-assisted MANETs in 5G," Ad Hoc Networks, vol.7,2020.

[20] M.A Jubair et al., "Bat Optimized Link State Routing Protocol for Energy-Aware Mobile Ad-Hoc Networks," Symmetry, vol.11,no.11,pp.1409,2019.

[21] R.Isaac Sajan and J.Jasper, " Trust based secure routing and the prevention of vampire attack in wireless ad hocsensor network," International Journal of Communication Systems,vol.33,no.8,2020

[22] O.Singh et al., "Multi-objective lion optimization for energy-efficient multi-path routing protocol for wireless sensor networks," International Journal of Communication Systems,vol.34,no.17.2021.

[23] A.B.Feroz et al., "A multi-attribute based trusted routing for embedded devices," Microprocessors and Microsystems, vol.89,2022.

[24] Salman Ali Syed, "A systematic comparison of mobile ad hoc network security attacks," Materials Today Proceedings, 2021.

[25] Deepa J, Yesudhasan J, AAS Aliar, "Developing multi-path routing protocol in MANET using hybrid SM-CSBO based on novel multi-objective function," International Journal of communication systems, vol.36, no.4, 2022**.**

[26] Ravi, S., Matheswaran, S., Perumal, U. et al. Adaptive trust-based secure and optimal route selection algorithm for MANET using hybrid fuzzy optimization. Peer-to-Peer Netw. Appl. **16**, 22–34 (2023). https://doi.org/10.1007/s12083-022-01351-2

[27] N. Veeraiah et al., "Trust Aware Secure Energy Efficient Hybrid Protocol for MANET," in IEEE Access, vol. 9, pp. 120996-121005, 2021.

[28] Yulia Sokolova, Deep Learning for Emotion Recognition in Human-Computer Interaction , Machine Learning Applications Conference Proceedings, Vol 3 2023.

[29] I. Koohi and V. Z. Groza, "Optimizing Particle Swarm Optimization algorithm," in Proc. IEEE 27th Canadian Conference on Electrical and Computer Engineering (CCECE), May 2014, pp. 1-5.

[30] I. Fister et al., "Bat algorithm: Recent advances," in Proc. IEEE 15th International Symposium on Computational Intelligence and Informatics (CINTI), Nov 2014, pp. 163-167.

[31] S.Bharany et al., "Energy-Efficient Clustering Scheme for Flying Ad-Hoc Networks Using an Optimized LEACH Protocol," Energies,vol.14,no.19,pp.6016,2021.

[32] Robbi Rahimet et al.,"Taylor Based Grey Wolf Optimization Algorithm (TGWOA) for Energy Aware Secure Routing Protocol," International Journal of Computer Networks andApplications,vol.7,no.4,pp.93,2020.

[33] Ameer Alhasanet et al., "An Energy Aware Qos Trust Model For Energy Consumption Enhancement Based On Clusters For IOT Networks," Journal of Engineering Science and Technology, vol.16, no.2, pp.957-976,2021.

[34] M.Nandi and K.Anusha, "An Optimized and Hybrid Energy Aware Routing Model for Effective Detection of Flooding Attacks in a Manet Environment," Wireless Pers Communications,2021.