Research Article

International Journal of Advanced Technology and Engineering Exploration, Vol 11(119)
ISSN (Print): 2394-5443 ISSN (Online): 2394-7454
http://dx.doi.org/10.19101/IJATEE.2023.10102615

Trust-based secure and optimal route selection in MANET utilizing multiple agent-based reinforcement learning

Syed Zeeshan Hussain[*] and Shalini Sharma
Department of Computer Science, Jamia Millia Islamia University New Delhi, India

Received: 18-December-2023; Revised: 20-October-2024; Accepted: 24-October-2024

Abstract

A mobile ad-hoc network (MANET) is a wireless network that has a set of moving nodes merged, with no constant infrastructure where nodes are self-configured. Secure routing is essential for preventing the mobile devices from other vulnerabilities, and thus, efficient characteristics in MANET are exploited to obtain an effective secure routing. The existing techniques have the drawbacks of security and network traffic. In this research, trust-based secure and optimal route selection by multiple agent-based reinforcement learning (MA-based RL) is proposed. Initially, optimum routes are selected through using MA-based RL algorithm via a secure communication. The proposed algorithm reduces reliability, packet delivery ratio (PDR), and breaks service. The performance of the developed algorithm is determined through performance measures of PDR, throughput, delay, and energy consumption with several nodes. The proposed algorithm attains high PDR 94.2%, 93.1%, 92.4% and 90.8% for 50, 100, 150, 200 nodes respectively, which is comparatively effective than the previous methods of two-tier security mechanism (TTSM), trust based adaptive genetic algorithm (TAGA), adaptive trust-based secure and optimal route selection utilizing hybrid fuzzy optimization (ATSORS – HFO), and trust-based topology hiding multipath routing protocol (T-TOHIP).

Keywords

Mobile ad-hoc network, Multiple agent, Optimal route selection, Reinforcement learning and Secure communication.

1.Introduction

Mobile ad-hoc network (MANET) gives an automatic structure of different systems connected along wireless links and devices connected within the MANET, known as nodes [1]. The MANET nodes have characteristics of mobility, multi-hop packet transmissions, and infrastructure-less nature [2]. With the help of a direct communication link, all nodes communicate with each other [3]. When a huge distance divides the nodes, then the communication is carried out along the intermediate nodes [4]. The MANET identifies the applications in different fields like mobile phones, military, space applications, and so on [5]. MANET establishes a simultaneous communication flow between the source and the destination nodes with good security and quality of service (QoS) [6–9]. The major challenge in MANET routing is finding a secure routing path between MANET nodes, which depend on each other for establishing wireless communication links [10–12].

The MANET protocol which has multiple path routing ignores the communication link breakage among the nodes [13]. Multiple path routing establishes many paths between the source and the destination node for communication, alongside benefits like routing reliability and load balancing [14–16]. The disadvantage of multiple path routing protocols is exposure to topology [17]. Multiple path routing simultaneously modifies the nodes routing table and keeps the network open to hackers for obtaining data from the network easily [18–20]. The MANET nodes are related to different attacks in the process of multiple path routes [21]. The different attacks in MANETs are sybil, wormhole, black hole, and rushing [22]. Every MANET node contains its factors of security and communication between nodes that depend on the trust among neighbours, which establishes reliability in MANET [23, 24]. MANET nodes communicate with their neighbour nodes based on the trust created for some time [25].

The trust-based node selection requires the consideration of delay, distance, efficiency, and

[*]Author for correspondence

neighbour nodes' energy. In multiple path routing, the nodes routing tables are simultaneously added because of modification in routing paths. This minimizes the nodes' energy in the network, and hence the nodes' energy in MANET requires simultaneous monitoring.

The routing protocols have disadvantages such as communication between nodes, QoS, energy constraints, huge rates of error, non-scalability, and insecure. One of the major issues in MANET are security and QoS which require a better routing. The content of the packet is modified by the attacker while the packet is forwarded to the destination. The above procedure reduces reliability, packet delivery ratio (PDR), and breaks service.

Optimum routes are successfully selected by using multiple agent-based reinforcement learning (MA based RL) with a model of energy and delay. The MA-based RL algorithm offers secure communication by removing malicious nodes for optimal routing. The efficacy of the developed algorithm is determined based on the performance measures of PDR, throughput, delay, and energy consumption.

The rest of this manuscript is described as follows: Section 2 describes the related works, section 3 provides an explanation of the developed methodology for optimal route selection and secure communication, while section 4 and 5 presents the performance analysis of the developed technique, and section 6 describes the conclusion of the research.

2.Literature review

Yin et al. [26] suggested the energy-aware trust algorithm based on ad hoc on-demand distance vector (AODV) protocol and multi-path routing method named as (EATMR) which enhanced the network security. The suggested method included two primary phases, initially nodes were clustered depended on open-source development model algorithm (ODMA) and next clustering-based routing was employed. The optimum and secure routing was determined depended on different parameters such as trust, energy, hop-count and distance.

Srinivas and Patnaik [27] introduced quantum worm swarm optimization-based clustering with secure routing protocol (QGSOC-SRP) for MANET. Initially, the introduced algorithm derived the fitness function utilizing node degree, trust factor, energy and distance for optimum selection of secure cluster head (CH). Next, secure route protocol (SRP) was used the oppositional gravitational search algorithm (OGSA) was employed for optimum selection of routes for base station (BS). For enhancing the effectiveness of gravitational search algorithm (GSA), the OGSA was derived depended on opposition-based learning concept to initialize population and generation jumping.

Ragesh and Kumar [28] presented the system to secure information in internet of things (IoT) communication was developed. The presented system included two methods, initially the method with combination of fuzzy and particle swarm optimization (F-PSO) was used for determining the secure route. Next, the cryptographic method utilizing learning with errors over rings (R-LWE) encryption scheme was employed to encrypt data. The trust score of each node in network was measured by utilizing trustworthiness. The secure path was chosen optimally by F-PSO and then identify secure way by utilizing R-LWE method.

Alamelumangai and Suresh [29] developed the firebug optimized modified bee colony (FOMBC) algorithm that attained highest network security with less energy consumption in WSN. The secure node selection and orthogonal routing (OR) were the two phases, the developed method carried out. Initially, secure node between initial nodes of network was chosen depended on characteristics like delay, trust and QoS. Next, FOMBC method was utilized for identifying the optimum path depended on factors of distance, latency, QoS and trust.

Reddy et al. [30] implemented the energy trust-based approach (ETA) integrated the energy, trust and reliable routing. The implemented method identified the trusted nodes depended on indirect, trust, past experience and estimated its energy values of routing was performed with trusted nodes whose residual energy extended determined threshold. The effectiveness of routing was improved through measuring trust adopted nodes in network. The implemented method ensured the effective routing among source and destination with trustworthy and energy efficiency of intermediate nodes.

Bangotra et al. [31] developed the trust-based secure intelligent opportunistic routing protocol (TBSIOP). The developed protocol utilized three attributed for executing the likelihood of nodes being malicious. These attributes were forwarding the data packets, acknowledgement and energy depletion and utilized for trust computation. Based on trust calculation, relay selection of developed protocol prevented malicious nodes from being selected as relay nodes.

Akwirry et al. [32] suggested the multi-tier management system to address the malicious vehicles in vehicle ad hoc network by utilizing three security problems. The initial tier of suugested method employed the vehicles in network, the trust value depended on behaviour like delay, packet loss and previous behaviour history of vehicle. The next tier is to prevent the watchdogs, that was performed through behaviour history of watchdogs. The final security tier is to prevent the integration of data utilized for calculating the trust value.

Kamarunisha and Vimalanand [33] presented the delay centric speed and directional routing (DMDR) method. The presented method considered factors such as delay among neighbours and nodes' mobility speed with their direction in selection of transmission path. The presented method identified the paths available in destination and executed the delay sensitive route support (DSRS) that represented the route suitability to deliver packet on time, mobility-based transmission support (MTS) that represented the route availability in accordance with node speed. The directional reaching support (DRS) that represented the count of nodes in further position tom reached the destination.

Singh et al. [34] introduced the optimal fuzzy clustering and trust-based routing (OFC-TR) minimize the energy consumption, latency and improve the network security and lifetime. The introduced method was attained in three phases. The first phase was to organize and choose CH by utilising the improved fuzzy c-means (IFCM) technique that resolved the problem of unequal distribution through employing every sensor of cluster membership. The next segment included the measuring of trust value by utilizing the fuzzy cognitive medium (FCM) that considered scores of indirect and direct trusts. The next segment included the routing by utilizing bacteria foraging algorithm (BFA) that was introduced to successful optimum control, estimation of harmonic and reduction of transmission loss.

Prasad [35] implemented enhanced energy efficient – secure routing (EEE-SR) protocol in MANET. The protocol decided multi paths in the network for selecting optimum routes, and also introduced a detection system for the security of MANET to increase performance and reliability of the network. The method provided effective network routing with less energy degradation. The model had the problems of scalability and security.

Lakshmi and Vaishnavi [36] introduced trust and anonymous model (TAM) for effective and secure routing in MANET. The introduced protocol was developed along the two-tier security mechanism (TTSM). At the initial stage, trustable nodes were chosen depending on their capacity to progress the controlled messages. In the second stage, the identity of the actual node was hidden and data was transferred with chosen trusted nodes with fake individualities produced through a factorial recursive function. The introduced model enabled secure transmission while malicious nodes was not able to find processing nodes in the routing process. However, the introduced model lacked integrity in messages.

Han et al. [37] presented an energy aware and trust-based routing protocol for wireless sensor networks (WSNs) utilizing an adaptive genetic algorithm (AGA), known as the trust-based adaptive genetic algorithm (TAGA) model. The presented model was integrated with a trust security mechanism and AGA that considered both energy saving and security. The presented model improved security by developing an adaptive trust model for evaluating the wide trust value of every node, so as to avoid both general and particular trust attacks. The presented model efficiently reduced the effect of malicious nodes and the number of loss packets by maximizing the usage of network energy. The computation cost of the presented model enhanced the energy consumption.

Reddy et al. [38] developed energy efficient master auditor node with trust based secure routing (EE-MAN-TbSR) protocol in secure route selection and information transmission in WSNs. This method aided in selecting an energy-efficient and trustable path, making it feasible for optimal trust setting. The method obtained a balancing of energy and minimization. But, the execution of energy was not considered with network layer security in this model.

Bai [39] suggested a routing manager based secure rate analysis (RMBSRA) model for secure routing protocol. The suggested model had a centralized algorithm developed with three main elements known as the routing manager, neighboring and routing table. These elements were responsible for efficient packet transferring in the shortest route, without information loss. The suggested model maximized multiple cast routing by balancing the allocation of bandwidth and traffic on the network. But this method failed to predict the resource availability and reliability.

Ravi et al. [40] presented adaptive trust-based secure and optimal path selection utilizing hybrid fuzzy optimization (ATSORS–HFO) method for secure route selection in MANET. Initially, optimum paths were chosen with the support of fuzzy butterfly optimization (FBO) and adaptive chaotic grey wolf optimization (ACGWO) technique, ensuring secure communication. The presented method obtained high throughput and was superior in secure routing. The disadvantage of the model was that it had huge execution costs.

Pari and Sudharson [41] implemented enhanced trust-based secure route protocol (ETBSRP) method in secure path selection for MANET. The initial and second characteristics were captured and obtained routing through measuring. The characteristics of trust were obtained through integrating all physical and logical trust from every node. The method used cryptographic process to increasing secure data transmission. The method obtained low packet drop and delay with huge throughput, but the method had low security and trust-aware protocols.

Ambekar and Kolekar [42] introduced trust-based topology hiding multipath routing (T-TOHIP) for secure routing in MANETs. The introduced routing protocol contained four models for determining security such as delay, energy mobility and trusted model. The introduced model defined secured routing between sender and receiver depending on the chosen neighbor nodes, and at last, data communication was processed along selected multiple paths. The introduced model avoided node attacks like black holes and rushing attacks. The routing was carried out by trust nodes however, when nodes were divided by huge distances, the trust factor among nodes was varied.

Patil et al. [43] implemented a monarch-cat swarm optimization (M-CSO) algorithm that was a combination of monarch butterfly and cat swarm optimization for secure route selection in MANET. The implemented method performed on two aspects of initially choosing secure nodes and selecting opportunistic nodes between chosen secure nodes. Further, these nodes were selected depended on fitness parameters of trust, distance, delay and connectivity. The implemented method enhanced the data transmission reliability. However, the method failed to estimate clustered WSN with many sink nodes.

Usturge and Pavan [44] suggested a DEroute algorithm for secure route selection in MANET. Initially, the nodes were taken and allocated in MANET, and the following process of routing was performed for data transmission to identify the route. Furthermore, direct, indirect and historical trust was taken for the identification of secure node. However, the method had huge communication overhead.

Muruganandam and Renjit [45] developed a particle swarm optimization (PSO) algorithm for secure route selection in MANET. The method was adopted for efficient cluster head selection and detection of malicious node. This method created a trust between merged nodes which enhanced the security and propagated the authentic and trust content within the network. The suggested method attained high network lifetime. Nonetheless, the method majorly depended on a particular route scheme and platform.

From the overall analysis, the existing methods had limitations when the nodes were divided by huge distances, varied trust factor among nodes, security and scalability issues, and less trust-aware protocol, while failing to estimate clustered WSN with many sink nodes. They also majorly depended on a particular route scheme and platform.

3.Methods

A trust-based secure route communication approach in MANET is proposed in this research, as illustrated in *Figure 1*. The *Figure 1* shows the route optimization and security in MANET, by using MA-based RL for ensuring the network performance effectively and securely. The RL agents continuously adapt based on energy levels, delay and security measures for choosing optimum routes for communication. The test score analyses trust and untrust nodes, while the optimum routes are selected by MA-based RL and give secure communication.

3.1Route selection

The optimum routes are chosen by utilizing MA-based RL that learns the network's present state and selects the optimal paths accordingly. The delay and energy-optimal routes are chosen by utilizing multi reinforcement learning (MRL). The proposed work depends on the supposition that the entire mobile nodes are unchangeable and are aware of their location with a single BS. The transmission of data is processed, leading to the transmission of minimization of delay when the throughput is increased. The reward functions are selected by trust nodes, ensuring the path transmission is more trustworthy and reliable. The proposed method assigns multiple agents to choose the best possible route.

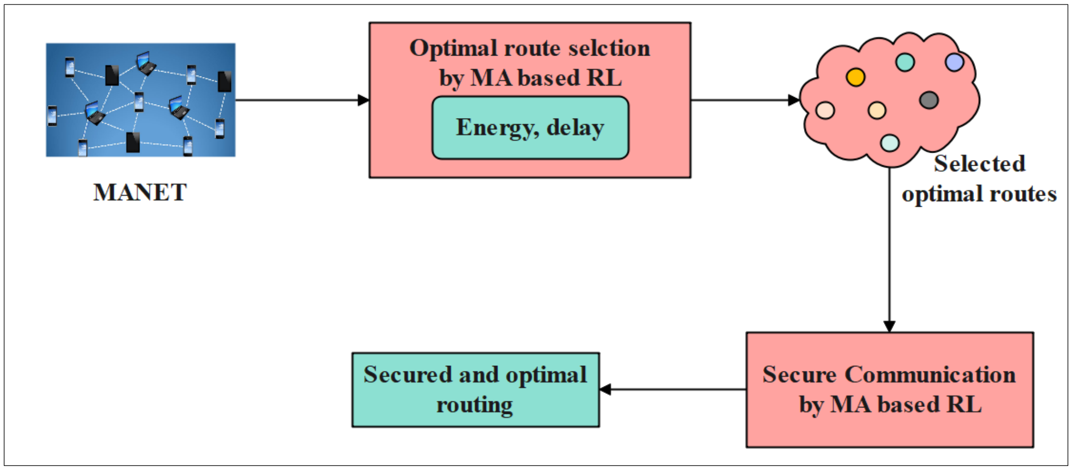
MANET
Optimal route selction by MA based RL
Energy, delay
Selected optimal routes
Secure Communication by MA based RL
Secured and optimal routing

Figure 1 Overall process of the proposed model

3.1.1Reinforcement learning (RL)

The RL is defined as Markov decision mode and is described as (S, A, P, R, γ). Here, the S and A represent the spaces of status and action, respectively [46, 47]. The state in RL environment describes different conditions of the network like the position of node and routing decision in a given time. The action in RL environments describes probable actions that an agent of RL takes to the network influence, like the routing decisions. The reward in RL environment describes feedback mechanism for RL agent depending on the actions taken and the resulting state. The values of parameters are learning rate 0.01 and discount factor 0.99. The P and R represent the transition of state and reward functions, respectively. The following rewards are detected through factor γ. For each phase, the RL agent chooses the action a_{tp} for the state environment s_{tp} along policy p that is represented as $a_{tp} = p(s_{tp})$. The environment reacts to agent with reward r_{tp} and transfers to a further stage s_{tp+1}. Therefore, the agent objective is utilized for detecting good policy p^*, as well as for enhancing the reward. The state s_t represents the present position, the a_t represents the action and the r_t represents the reward which provides positive for moving closer to destination or give negative reward. After every action, environment gives following position (new state). The process of RL is represented in *Figure 2.*

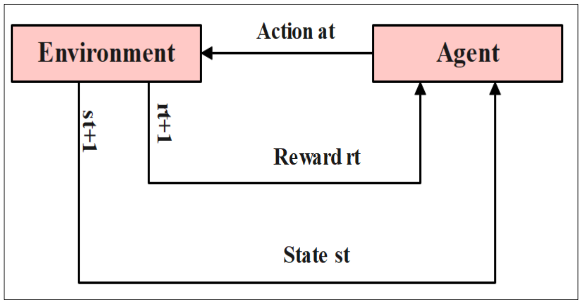

Figure 2 Process of RL

The S_A function value is defined as Q-function (Q_{fn}) and a good scheme is chosen by using the Equation 1.

$$p^* = \overset{max}{p}\ Q\ (s, a) \qquad (1)$$

Therefore, the efficiency of the RL method depended on Q_{fn}, hence executed through the selection of reward functions. This method assigns trustable reward functions, and so the executed routes are reliable and trustworthy.

3.1.2MA-based RL

The optimal routing algorithm is proposed based on RL by assigning multiple agents and the process is illustrated in *Figure 3*. The RL method is utilized where agents choose superior possible path by considering the present status of network. The proposed method assigns three trustable reward functions, are discussed.

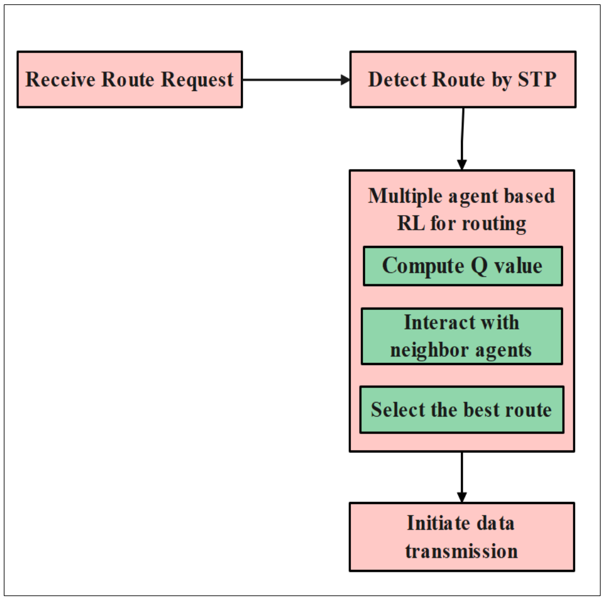
Receive Route Request
Detect Route by STP
Multiple agent based RL for routing
Compute Q value
Interact with neighbor agents
Select the best route
Initiate data transmission

Figure 3 Process of MA-based RL

When the mobile node is required to communicate with various nodes, the route request R_{Req} is sent to agents. Various agents are located in the network to enhance the routing procedure. The nodes share their Q-values with neighbours for aligning its process of decision making. The agents cooperate through sharing its Q-values that learned from the experience of each other to make optimal route decisions. The spanning tree protocol (STP) is assigned to identifying nodes that overcome the issue of network cycling, and a good route is selected by the RL method. By integrating the STP with MA-based RL, the network attains a balance between adaptability and stability, and leverages the strengths of both approaches for optimizing the routing decisions dynamically. The R_{Req} is a packet combined with the unique identifier of the node, desired address, message, and timestamp. While R_{Req} packet is received through the agent that identifies a good possible path between the source and the destination, depending on the Q-value. The network behaviour is learned by an agent for a certain time, while an optimal path is chosen depending on the knowledge gained. Because of the combination of trustable reward functions, the optimal and reliable paths are chosen to improve the QoS by minimizing the delay transmission and enhancing throughput.

3.1.3Reward function computation

The reward function is a major stage in the RL method and the agents monitor state changes when the environment is tracked. By obtaining a reward, an agent that learns a path behaves in a particular state. So, the reward function directly affects the network operation, making it a challenging task. There are three reward functions assigned through the proposed method, depending on the path's energy level, PDR, and packet consistency (PC). The boundary of energy nodes is essential for carrying out the routing superior, and so these metrics are selected. Likewise, the PC and PDR are considered for ensuring the nodes' reliability to measure the route reliability.

Energy level

The initial reward function executes the energy of all nodes through a particular path. The agent selects path from every possible route. Here, N represents the total number of nodes in a particular path. By utilizing the total available node energy, the approximate lifetime of the path is executed. The path distributes appropriately when the level of energy is enough for the transmission of data, and the mathematical formula is given as Equation 2.

$$RF_1 = \sum_{i=1}^{N} \frac{ANE_i}{N} \qquad (2)$$

Where, N represents the total count of nodes, r represents the total count of routes, and ANE_i represents nodes' available energy across r. By utilizing RF_1, the routes' approximate lifetime is determined.

PDR

PDR is one more significant metric which determines the mobile nodes' willingness to forward packets. The PDR is higher when node moves forwarded to receive packets effectively. On the other hand, while nodes do not forward received packets, PDR is less. The PDR is for communication, taken as a reward function, and the mathematical formula is given in Equation 3.

$$RF_2 = \sum_{i=1}^{N} \frac{STP_i}{TTP_i} \qquad (3)$$

Where, STP represents the successfully transmitted packets, and TTP represents the total number of transmitted packets.

Mobile nodes are considered reliable when the PDR is reasonable. While many of the mobile nodes employ a path, the route does not behave normally, and the selection of path decreases performance. So, it is superior to regulate PDR of nodes in a particular route.

PC

The PC is utilized as one more measure for the developed method. The PC function spotlights the actual behaviour of mobile nodes. For example, nodes process certain malicious functions upon packets like packet tampering, deletion, etc. To prove originality of the mobile node, it is required to examine node behaviour during packet transmission. The communication happens directly among the source and destination nodes. Therefore, intermediate nodes are involved in malicious activities. The proposed method analyzes the consistency of packet by evaluating the MD5 hash code. The mathematical formula of RF_3 execution is given by Equations 4 and 5.

$$RF_3 = \sum_{i=1}^{N} \frac{PC_i}{N} \qquad (4)$$

$$PC = \begin{cases} 1 & if\ hc_a(p) = hc_b(p) \\ 0 & if\ hc_a(p) \neq hc_b(p) \end{cases} \qquad (5)$$

When a packet is transferred to the alternative node, the hash code is executed in present and past nodes, whereas when the hash nodes are similar, it represents the packets as similar.

3.1.4Computation of Q-value

The value of Q is set to 0 at the initial stage and gets enhanced through the acquired rewards. A reward is evaluated through three reward functions described above. The value of Q is evaluated by using Equation 6.

$$Q_{val}(k) = (1 - \alpha) + \alpha \times \gamma \times R(k) \qquad (6)$$

Where, α represents the learning rate, γ represents the discount factor, and R represents the reward. The convergence criteria are when the Q-value stabilizes and does not change over 200 iterations. The learning rate of 0.01 and discount factor of 0.99 to update the Q-values, the agents of RL effectively learn the optimum routs in the network. Also, the technique of Q-learning is merged along function approximately to overcome the memory issue of RL.

The proposed method assigns multiple agents for selecting a good path and those agents communicate around further rewards at time t, and add Q value at time $t + 1$. In the route selection stage, agent a_i receives certain further rewards from their neighbor agents N_{agn} at time t. At time $t + 1$, agent a_i receives delayed rewards and learns better matchable actions for a particular state. The major advantage of using multi-agents is that the agent determines on their own when considering the neighbors' decision. This stage improves the performance of communication through reliability. The delayed rewards predict the future reward and environment of dynamic operation, which are both taken into consideration. The algorithm for the proposed RL for optimal route selection is given below:

Algorithm for proposed RL based route selection

```
Input: mobile nodes
Output: route Selection
//Node Side
Begin
Send $R_{Req}$ packet to agent;
Check energy backup;
If energy backup is enough
    Do
        Receive packets to forward;
        Forward Packets;
    End Do;
End
//Agent Side
Begin
Broadcast ID to whole agents and nodes;
Acquire detected routes;
Evaluate Q-value;
Interact with neighbor agents;
Choose route with optimal Q value;
Exclude dead node;
Repeat process;
End;
```

Therefore, the RL based routing is formulated. When the agent receives R_{Req}, the agent identifies whole probable routes with the support of STP. The superior probable path is chosen through the path with optimal $Q - value$ that depends on three reward functions. All three reward functions are trustable and therefore, the optimum paths are selected. The proposed algorithm considers energy backup, PDR, and PC to ensure that superior routes are chosen for data transmission, by enhancing the energy efficiency and network lifetime. The MA-based RL algorithm employs multiple agents across network, that allows effective management of various node densities. While node density is high, algorithm provides optimal routes through sharing Q-values between agents, enhance the routing performance. This decentralized algorithm enables the fast-routing optimization, improves throughput and minimize the delay. By sharing the Q-values between agents, RL agents fastly learns from every experience, that enhances the convergence to optimum routing decisions. This fastens the routing adjustments in dynamic environments, minimize the delay and increase PDR as well as minimize the packet loss.

3.2Secure communication

Optimal routes are chosen by utilizing the reward functions of MA-based RL. The communication is processed directly between the source and destination nodes with intermediate nodes, and therefore, it involves malicious activities. In secure communication, malicious nodes are detected and removed from the selected optimal routes and the routing is secured.

4.Results

The developed algorithm is simulated on python with the system requirements as follows, random access memory (RAM): 16GB, processor: Intel core i7. The performance of the developed technique is determined through performance measures of PDR, throughput, delay, and energy consumption. The numerical expression for performance measures is represented.

PDR

PDR is defined as the number of packets received at the destination, divided by the number of packets sent from the source. The numerical expression is given as Equation 7.

$$Packet\ delivery\ ratio = \frac{\sum Packet\ received}{\sum Packet\ transmitted} \quad (7)$$

Throughput

Throughput refers to the number of packets successfully delivered to the desired node from actual source. The numerical expression is given as Equation 8.

$$Throughput = \frac{Delivered\ Packets}{Simulation\ Time} \quad (8)$$

Delay

Delay is defined as the measuring time occupied for transferring data packets from the source to the destination node. The numerical expression is given as Equation 9.

$$Delay = \frac{\sum Transaction\ time - Receiving\ time}{F} \quad (9)$$

Energy consumption

Energy Consumption is defined as the quantity of energy that is consumed for data transfer and process by node. Energy consumption for every node is measured through each set of transmissions at a specific time. *Table 1* represents the simulation parameters.

Table 1 Simulation parameters

Parameters	Values
Network size	50 to 200
Node Density	100 m × 100 m
Mobility Model	Random waypoint
Communication range	250 m
Simulation duration	300 s
Area of deployment	1000 m × 1000 m

4.1Quantitative and qualitative analysis

The performance of the developed MA based RL technique is determined by using the performance measures. The existing algorithms used in estimation are TTSM [36], TAGA [37], adaptive trust-based secure and optimal route selection utilizing hybrid fuzzy optimization (ATSORS – HFO) [40], and T-TOHIP [42]. Every RL agent monitors traffic patterns and detects anomalies to indicate the security threats. By decentralizing this process across multiple agents, the MANET swiftly identifies and mitigates the potential attacks. The performance of the proposed MA based RL method varies effectively with node density. In high node density, it improves the routing efficiency, security, scalability and energy efficiency, further leveraging abundant routes and collaborative agent learning. In low node density, the performance of the network is challenged through limited paths and high node strain. However, the proposed MA based RL still manages to adapt and optimize its robustness and adaptability.

Table 2 describe the performance of the developed algorithm with PDR. The developed algorithm acquires high PDR of 94.2%, 93.1%, 92.4%, and 90.8% respectively for 50, 100, 150 and 200 nodes. The developed method exhibits effective performance than the existing algorithms like TTSM, TAGA, adaptive trust-based secure and optimal path selection using hybrid fuzzy optimization (ATSORS-HFO) and T-TOHIP that attain lesser PDR for different nodes.

Table 3 display the performance of the developed algorithm in throughput. The developed algorithm acquires high throughput of 550kbps, 535kbps, 515kbps and 505kbps correspondingly for 50, 100, 150 and 200 nodes. The developed algorithm demonstrates more effective performance than the existing algorithms like TTSM, TAGA, ATSORS-HFO and T-TOHIP which obtain reduced throughput for different nodes.

Table 4 demonstrate the performance of the developed algorithm with delay. The proposed algorithm minimizes delay of 11.9ms, 12.2ms, 12.9ms, and 13.6ms respectively for 50, 100, 150 and 200 nodes. The developed algorithms represent effective performance than existing algorithms like TTSM, TAGA, ATSORS-HFO and T-TOHIP that obtain high delay for different nodes. *Table 5* present the performance of the developed algorithm in terms of energy consumption. The technique demands lesser energy of 325.5J, 329.3J, 332.8J, and 338.1J simultaneously for 50, 100, 150 and 200 nodes. The developed algorithm represents a commendable outcome than the previous algorithms namely, TTSM, TAGA, ATSORS-HFO, and T-TOHIP, which exhibit high energy consumption for various nodes.

Table 2 Performance of developed algorithm in terms of PDR

			PDR (%)		
No. of nodes	**TTSM**	**TAGA**	**ATSORS-HFO**	**T-TOHIP**	**Proposed MA based RL**
50	88.0	87.6	90.9	92.1	94.2
100	87.3	86.5	89.1	90.9	93.1
150	86.8	85.7	88.3	89.3	92.4
200	86.0	84.8	87.7	88.6	90.8

Table 3 Performance of the developed algorithm in throughput

			Throughput (kbps)		
No. of nodes	**TTSM**	**TAGA**	**ATSORS-HFO**	**T-TOHIP**	**Proposed MA based RL**
50	475	483	490	512	550
100	468	476	482	505	535
150	460	469	473	494	515
200	451	457	465	487	505

Table 4 Performance of proposed algorithm in delay

			Delay (ms)		
No. of nodes	**TTSM**	**TAGA**	**ATSORS-HFO**	**T-TOHIP**	**Proposed MA based RL**
50	17.3	16.2	15.1	14.7	11.9
100	18.2	17.0	15.9	15.0	12.2
150	18.9	18.1	16.8	15.9	12.9
200	19.6	19.0	18.0	16.7	13.6

Table 5 Performance of the developed method in energy consumption

			Energy consumption (J)		
No. of nodes	**TTSM**	**TAGA**	**ATSORS-HFO**	**T-TOHIP**	**Proposed MA based RL**
50	340.5	335.4	337.1	329.3	325.5
100	345.1	342.7	337.8	331.6	329.3
150	351.2	348.3	344.7	337.8	332.8
200	358.4	359.2	350.3	346.4	338.1

4.2Comparative analysis

In this section, the performance of the developed algorithms is compared with that of the existing algorithms like EEE-SR [35], ETBSRP [41] and EE-MAN-TbSR [38], as displayed in *Table 6*. The developed MA based RL method obtains a high PDR of 93.1% with less delay of 12.2ms for 100 nodes. This algorithm represents an effective performance than the existing algorithm EEE-SR [16] that obtains a PDR of 82% with a delay 44ms, aside from ETBSRP [22] which obtains a PDR of 91% with a delay of 12.2ms.

Table 6 Comparative analysis

Author	Method	No. of nodes	PDR (%)	Delay (ms)
Prasad [35]	EEE – SR	100	82	44
Pari and Sudharson [41]	ETBSRP		91	14.5
Reddy et al. [38]	EE-MAN-TbSR		96.5	N/A
Proposed model	MA based RL		93.1	12.2

5.Discussion

In this research, trust-based secure and optimal route selection through MA based RL is proposed depending on the trust nodes. Initially, optimal paths are selected through using MA-based RL algorithm, resulting in a secure communication. The proposed method obtains high PDR and throughput with less delay and superior performance in routing. The performance of MA based RL is evaluated with TTSM, TAGA, ATSORS-HFO, and T-TOHIP by varying the nodes from 100 to 500. Moreover, compared the MA based RL method with existing methods like EEE-SR [35], ETBSRP [41] and EE-MAN-TbSR [38]. The benefits of developed method and limitations of the previous algorithms are noted in this section. The EEE-SR [33] method has limitations with security and scalability. The ETBSRP [39] method has limitations of less security and trust-aware protocols. To overcome these limitations, a MA based RL method is proposed for secure routing in this research. The proposed method avoids malicious nodes in the routes and ensures a secure route. This method also obtains high PDR and throughput with less delay and commendable performance in routing. The developed MA based RL method obtains a high PDR of 93.1% with less delay of 12.2ms for 100 nodes.

5.1Limitations

In this research, routing is performed depending on the trust among nodes but, while nodes are divided with a large distance, the trust factor among nodes gets varied. Moreover, when the count of nodes in the cluster is increased rapidly, the reliability of the model is also affected. Furthermore, the lack of authentication facilities that detect negative information in the network is to be administered, alongside countering to ensure the identification of some kind of attacks in the route path. These are the recorded drawbacks of the current research. A complete list of abbreviations is listed in *Appendix I*.

6.Conclusion and future work

Secure routing is an essential way to protect mobile devices from various vulnerabilities, and the efficient characteristics of MANET obtain an effective secure routing path. In this research, trust-based secure and optimal route selection by MA based RL is proposed depending on the trust nodes. Initially, optimum routes are selected by using the MA-based RL algorithm for a secure communication. The proposed method obtains high PDR and throughput with less delay and superior performance in routing. The proposed model attains high PDR of 94.2%, 93.1%, 92.4% and 90.8% for 50, 100, 150 200 nodes, respectively. This is comparatively effective than the previous methods namely, TTSM, TAGA, ATSORS–HFO and T-TOHIP. In the future, an efficacious optimization algorithm can be used for secure and optimal routing.

Acknowledgment

None.

Conflicts of interest

The authors have no conflicts of interest to declare.

Data availability

None.

Author’s contribution statement

Syed Zeeshan Hussain: Conceptualization, validation, formal analysis, data curation, supervision and project administration. **Shalini Sharma**: Conceptualization, methodology, software, validation, investigation, resources, writing—original draft preparation, review and editing.

References

[1] Su B, Zhu B. TBMOR: a lightweight trust-based model for secure routing of opportunistic networks. Egyptian Informatics Journal. 2023; 24(2):205-14.

[2] Duong TV. An improved method of AODV routing protocol using reinforcement learning for ensuring QoS in 5G-based mobile ad-hoc networks. ICT Express. 2024; 10(1):97-103.

[3] Yamini KA, Stephy J, Suthendran K, Ravi V. Improving routing disruption attack detection in MANETs using efficient trust establishment. Transactions on Emerging Telecommunications Technologies. 2022; 33(5):e4446.

[4] Rathod JA, Kotari M. Secure and efficient message transmission in MANET using hybrid cryptography and multipath routing technique. Multimedia Tools and Applications. 2024:1-24.

[5] Yu X, Li F, Li T, Wu N, Wang H, Zhou H. Trust-based secure directed diffusion routing protocol in WSN. Journal of Ambient Intelligence and Humanized Computing. 2022; 13:1405-17.

[6] Das MV, Premchand P, Raju LR. Security enhancing based on node authentication and trusted routing in mobile ad hoc network (MANET). Turkish Journal of Computer and Mathematics Education. 2021; 12(14):5199-211.

[7] Chiejina E, Xiao H, Christianson B, Mylonas A, Chiejina C. A robust Dirichlet reputation and trust evaluation of nodes in mobile ad hoc networks. Sensors. 2022; 22(2):1-32.

[8] Sharma B, Saxena D. Design and analysis of energy efficient service discovery routing protocol in MANETs. SN Computer Science. 2023; 4(5):495.

[9] Rao TV, Swamy VK, Karthigeyan KA, Gopalakrishnan S, Kalaichelvi T, Koteswari S. Energy efficient trust based data communication using AODV protocol in MANET. Journal of Advanced Research in Applied Sciences and Engineering Technology. 2023; 32(1):390-405.

[10] Srilakshmi U, Alghamdi SA, Vuyyuru VA, Veeraiah N, Alotaibi Y. A secure optimization routing algorithm for mobile ad hoc networks. IEEE Access. 2022; 10:14260-9.

[11] Hajiee M, Fartash M, Eraghi NO. Trust-based routing optimization using multi-ant colonies in wireless sensor network. China Communications. 2021; 18(11):155-67.

[12] Gripsy JV, Kanchana KR. Relaxed hybrid routing to prevent consecutive attacks in mobile ad-hoc networks. International Journal of Internet Protocol Technology. 2023; 16(2):92-8.

[13] Sivapriya N, Mohandas R. Optimal route selection for mobile ad-hoc networks based on cluster head

selection and energy efficient multicast routing protocol. Journal of Algebraic Statistics. 2022; 13(2):595-607.

[14] Thapar S, Purohit A, Kanwer B, Jaiman A, Mounika A, Madhumala VS. An approach to detect wormhole attack in mobile ad hoc networks using direct trust based detection approach. International Journal of Intelligent Systems and Applications in Engineering. 2023; 11(6s):276-83.

[15] Suresh KR, Manimegalai P, Vasanth RPT, Dhanagopal R, Johnson SA. Cluster head selection and energy efficient multicast routing protocol-based optimal route selection for mobile ad hoc networks. Wireless Communications and Mobile Computing. 2022; 2022(1):1-12.

[16] Pathak A, Al-anbagi I, Hamilton HJ. An adaptive QoS and trust-based lightweight secure routing algorithm for WSNs. IEEE Internet of Things Journal. 2022; 9(23):23826-40.

[17] Yang Z, Li L, Gu F, Ling X, Hajiee M. TADR-EAODV: a trust-aware dynamic routing algorithm based on extended AODV protocol for secure communications in wireless sensor networks. Internet of Things. 2022; 20:100627.

[18] Alappatt V, Joe PPM. Trust-based energy efficient secure multipath routing in MANET using LF-SSO and SH2E. International Journal of Computer Networks and Applications. 2021; 8(4):400-11.

[19] Bondada P, Samanta D, Kaur M, Lee HN. Data security-based routing in MANETs using key management mechanism. Applied Sciences. 2022; 12(3):1-16.

[20] Thirunavukkarasu V, Senthil KA, Prakasam P. Cluster and angular based energy proficient trusted routing protocol for mobile ad-hoc network. Peer-to-Peer Networking and Applications. 2022; 15(5):2240-52.

[21] Rajeswari AR, Lai WC, Kavitha C, Balasubramanian PK, Srividhya SR. A trust-based secure neuro fuzzy clustering technique for mobile ad hoc networks. Electronics. 2023; 12(2):1-16.

[22] Tu J, Tian D, Wang Y. An active-routing authentication scheme in MANET. IEEE Access. 2021; 9:34276-86.

[23] Tej DN, Ramana KV. MSA-SFO-based secure and optimal energy routing protocol for MANET. International Journal of Advanced Computer Science and Applications. 2022; 13(6):306-13.

[24] Zarzoor AR, Abbas TM. Securing data conveyance for dynamic source routing protocol by using SDSR-ANNETG technique. In machine intelligence for smart applications: opportunities and risks 2023 (pp. 213-25). Cham: Springer Nature Switzerland.

[25] Reddy MV, Srinivas PV, Mohan MC. Assessing node trustworthiness through adaptive trust threshold for secure routing in mobile ad hoc networks. International Journal of Advanced Computer Science and Applications. 2022; 13(4):224-31.

[26] Yin H, Yang H, Shahmoradi S. EATMR: an energy-aware trust algorithm based the AODV protocol and multi-path routing approach in wireless sensor networks. Telecommunication Systems. 2022; 81(1):1-9.

[27] Srinivas M, Patnaik MR. Clustering with a high-performance secure routing protocol for mobile ad hoc networks. The Journal of Supercomputing. 2022; 78(6):8830-51.

[28] Ragesh GK, Kumar A. Trust-based secure routing and message delivery protocol for signal processing attacks in IoT applications. The Journal of Supercomputing. 2023; 79(3):2882-909.

[29] Alamelumangai M, Suresh S. Firebug optimized modified bee colony algorithm for trusted WSN routing. IETE Journal of Research. 2024; 70(5):4903-16.

[30] Reddy MV, Srinivas PV, Mohan MC. Energy efficient routing with secure and adaptive trust threshold approach in mobile ad hoc networks. The Journal of Supercomputing. 2023; 79(12):13519-44.

[31] Bangotra DK, Singh Y, Selwal A, Kumar N, Singh PK. A trust based secure intelligent opportunistic routing protocol for wireless sensor networks. Wireless Personal Communications. 2022; 127(2):1045-66.

[32] Akwirry B, Bessis N, Malik H, Mchale S. A multi-tier trust-based security mechanism for vehicular ad-hoc network communications. Sensors. 2022; 22(21):1-20.

[33] Kamarunisha M, Vimalanand S. An efficient delay centric speed and directional routing (DMDR) for improved routing in MANET. SN Computer Science. 2024; 5(2):235.

[34] Singh CE, Priya SS, Kumar BM, Saravanan K, Neelima A, Gireesha B. Trust aware fuzzy clustering based reliable routing in Manet. Measurement: Sensors. 2024; 33:101142.

[35] Prasad R. Enhanced energy efficient secure routing protocol for mobile ad-hoc network. Global Transitions Proceedings. 2022; 3(2):412-23.

[36] Lakshmi GV, Vaishnavi P. A trusted security approach to detect and isolate routing attacks in mobile ad hoc networks. Journal of Engineering Research. 2024; 12(3):379-86.

[37] Han Y, Hu H, Guo Y. Energy-aware and trust-based secure routing protocol for wireless sensor networks using adaptive genetic algorithm. IEEE Access. 2022; 10:11538-50.

[38] Reddy DMK, Sathya R, Lakshmi VVAS. An energy efficient master auditor node with trust based secure routing in wireless sensor networks. International Journal of Intelligent Systems and Applications in Engineering. 2023; 11(3):519-29.

[39] Bai PK. RMBSRA: routing manager based secure route analysis mechanism for achieving secure routing protocol in IOT MANET. International Journal of Computer Networks and Applications. 2022; 9(2):150-9.

[40] Ravi S, Matheswaran S, Perumal U, Sivakumar S, Palvadi SK. Adaptive trust-based secure and optimal route selection algorithm for MANET using hybrid fuzzy optimization. Peer-to-Peer Networking and Applications. 2023; 16(1):22-34.

[41] Pari SN, Sudharson K. An enhanced trust-based secure route protocol for malicious node detection. Intelligent Automation & Soft Computing. 2023; 35(2):2541-54.

[42] Ambekar RK, Kolekar UD. T-TOHIP: trust-based topology-hiding multipath routing in mobile ad hoc network. Evolutionary Intelligence. 2022; 15(2):1067-81.

[43] Patil PA, Deshpande RS, Mane PB. Trust and opportunity based routing framework in wireless sensor network using hybrid optimization algorithm. Wireless Personal Communications. 2020; 115:415-37.

[44] Usturge S, Pavan KT. DEroute: trust-aware data routing protocol based on encryption and fuzzy concept for MANET secure communication in IOT. Information Security Journal: A Global Perspective. 2023; 32(5):331-46.

[45] Muruganandam S, Renjit JA. Real-time reliable clustering and secure transmission scheme for QoS development in MANET. Peer-to-Peer Networking and Applications. 2021; 14(6):3502-17.

[46] Valero JM, Sánchez PM, Pérez MG, Celdrán AH, Pérez GM. Toward pre-standardization of reputation-based trust models beyond 5G. Computer Standards & Interfaces. 2022; 81:103596.

[47] Venkateswaramma PV, Reddy IR. Node pattern state and trust-rate base route selection for reliable data transmission in mobile ad hoc networks. International Journal of Computers and Applications. 2021; 43(9):874-80.

Syed Zeeshan Hussain earned his Ph.D. from Jamia Millia Islamia, New Delhi, India, and his Master of Computer Applications (MCA) from IGNOU, New Delhi. He is currently a Professor in the Department of Computer Science at Jamia Millia Islamia, New Delhi, India. His research interests include Computer Networks, Network Security, Web Technology and Applications, Object-Oriented Computing, and Scripting Languages.
Email: szhussain@jmi.ac.in

Shalini Sharma is pursuing a Ph.D. in Computer Science from Jamia Millia Islamia University, New Delhi, India. She earned her Master of Technology in Information Technology from Guru Gobind Singh University, Delhi, in 2011. She worked as an Assistant Professor at Sharda University from 2011 to 2015. Since 2015, she has served in IT organizations, including Infogain India, IRIS Software, and Accenture.
Email: shalinisharma1980@gmail.com

Appendix I

S. No .	Abbreviation	Description
1	ACGWO	Adaptive Chaotic Grey Wolf Optimization
2	AGA	Adaptive Genetic Algorithm
3	ATSORS-HFO	Adaptive Trust-Based Secure and Optimal Route Selection Utilizing Hybrid Fuzzy Optimization
4	AODV	Ad Hoc On-Demand Distance Vector
5	BFA	Bacteria Foraging Algorithm
6	BS	Base Station
7	DMDR	Delay Centric Speed and Directional Routing
8	DRS	Directional Reaching Support
9	DSRS	Delay Sensitive Route Support
10	EEE-SR	Enhanced Energy Efficient – Secure Routing
11	ETA	Energy Trust-Based Approach
12	EE-MAN-TbSR	Energy Efficient Master Auditor Node with Trust based Secure Routing
13	ETBSRP	Enhanced Trust-Based Secure Route Protocol
14	FBO	Fuzzy Butterfly Optimization
15	FCM	Fuzzy Cognitive Medium
16	FOMBC	Firebug Optimized Modified Bee Colony
17	F-PSO	Fuzzy and Particle Swarm Optimization
18	GSA	Gravitational Search Algorithm
19	IFCM	Improved Fuzzy C-Means
20	IoT	Internet of Things
21	MANET	Mobile Ad-Hoc Network
22	M-CSO	Monarch-Cat Swarm Optimization
23	MA-based RL	Multiple Agent-based Reinforcement Learning
24	MRL	Multi Reinforcement Learning
25	MTS	Mobility-Based Transmission Support
26	ODMA	Open-Source Development Model Algorithm
27	OFC-TR	Optimal Fuzzy Clustering and Trust-Based Routing
28	OGSA	Oppositional Gravitational Search Algorithm
29	OR	Orthogonal Routing
30	PC	Packet Consistency
31	PDR	Packet Delivery Ratio
32	PSO	Particle Swarm Optimization
33	QGSOC-SRP	Quantum Worm Swarm Optimization-Based Clustering with Secure Routing Protocol
34	QoS	Quality of Service
35	RAM	Random Access Memory
	RL	Reinforcement Learning
36	R-LWE	Learning with Errors Over Rings
37	RMBSRA	Routing Manager Based Secure Rate Analysis
38	SRP	Secure Route Protocol
39	STP	Spanning Tree Protocol
40	TAM	Trust and Anonymous Model
41	TAGA	Trust based Adaptive Genetic Algorithm
42	TBSIOP	Trust-Based Secure Intelligent Opportunistic Routing Protocol
43	T-TOHIP	Trust-based Topology Hiding Multipath Routing Protocol
44	TTSM	Two-Tier Security Mechanism
45	WSNs	Wireless Sensor Networks