WJCMS

Wasit Journal of Computer and Mathematics Science

Journal Homepage: https://wjcm.uowasit.edu.iq/index.php/WJCM

e-ISSN: 2788-5879 p-ISSN: 2788-5879

# Design an Improved Trust-based Quality of Service Aware Routing in Cognitive Mobile Ad-Hoc Network

**Dr. Alok Singh Chauhan[1],* and H Mary Henrietta[2]**

[1]School of Computer Applications and Technology, Galgotias University, Greater Noida, India
[2]Saveetha Engineering College, Chennai, Tamil Nadu 602105. India.

*Corresponding Author: Dr. Alok Singh Chauhan

DOI:
Received: September 2023; Accepted: November 2023; Available online: December 2023

**ABSTRACT:** Mobile ad hoc networks (MANETs) are wireless networks that can be configured at will. It has no infrastructure and centralized control, so it is only suitable for provisional communications. In a dynamically topological and resource-constrained network, ensuring QoS and security is challenging. MANETs are dynamic networks, so navigating them can be challenging and more susceptible to attacks. MANET requires significant memory, speed, and transmission bandwidth for conventional security measures like cryptographic techniques.
Consequently, these methods are unsuitable for identifying malicious behaviour or self-centered nodes. Nodes that are malicious, selfish, or malfunctioning can be identified based on the trust method, which calculates how much trust exists between them. A trust-based QOS-aware routing protocol is proposed in this paper to calculate trust in MANET (I-TQAR). The tree important performance metrics are considered for result validation such as delay, throughput and packet delivery ratio (PDR). I-TQAR offers significantly improved performance in all areas compared to the existing TQR and TQOR protocols.

**Keywords:** AdHoc Network, MANET, Quality of Service, Routing

## 1. INTRODUCTION

Wireless mobile nodes spontaneously form temporary networks in MANETs. People and vehicles can access the internet wirelessly in such areas without an existing communications infrastructure [1]. Mobile ad hoc networks have radio ranges that allow direct communication between nodes; nodes outside these ranges communicate through intermediates [2], [3]. This type of wireless network is called a MANET since all communication nodes automatically form a wireless network [4]. There are various requirements for ad hoc network routing protocols, including scalability, safety, facility excellence, energy efficiency, multicasting, combining, and collaborating between nodes [5]. Security and service quality are considered here as qualitative properties. Mobile nodes form an autonomous distributed system when interconnected by wireless channels, as shown in Figure 1.

Recently, Quality of Service (QoS) has become a major concern for MANETs [6], [6], [7]. Throughput, bandwidth, jitter, and delay and delay are all QoS requirements that must be satisfied when using traditional QoS routing. A MANET environment does not guarantee QoS based on security requirements. The security of MANETs is an important consideration for QoS routing. Despite the importance of trust issues in QoS routing, little research has been conducted on resisting malicious behaviour and addressing trust issues [8]. A classic routing method is enhanced with trust and quality of service parameter estimation to enhance network security. In MANETs, an indirect trust degree is calculated by incorporating neighbours' recommendations and direct trust degrees are calculated by analyzing direct observations. We only consider link delay as a QoS constraint in this case since multi-QoS constraints are NP-complete. Until now, cognitive wireless

*Corresponding author: alokchauhan.1983@gmail.com

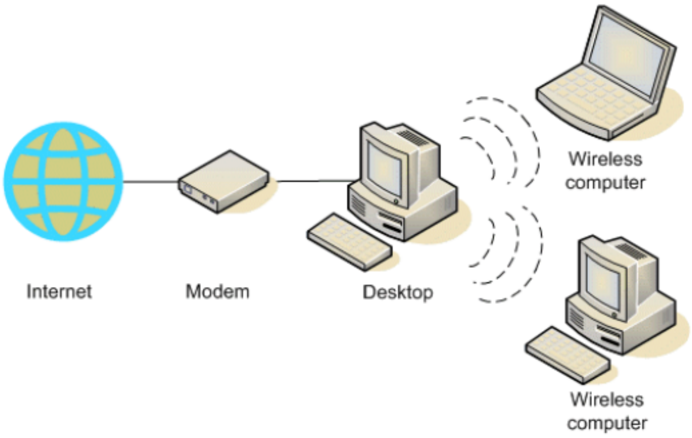

FIGURE 1. **Network topology of the ad-hoc network.**

networks have not considered trust management as a cognitive process. The trust-based routing protocol we use for secure routing in cognitive networks takes advantage of the cognitive properties of the cognitive layer. Trust learning makes it possible to detect blackhole and grayhole attackers more easily. Three optimization objectives are considered when solving the routing problem: delay, throughput, and PDR. MANETs have several important performance indicators. These include delay, throughput, and packet deliverability [9].

## 2. LITERATURE

MANETs are wireless networks that communicate over multiple hops between wireless nodes. MANET nodes must simultaneously act as routers and hosts. On-demand routing (reactive) and table-driven routing (proactive) are two types of MANET routing protocols [10]. Many situations lend themselves to on-demand routing protocols, which are more efficient than table-driven ones [11].

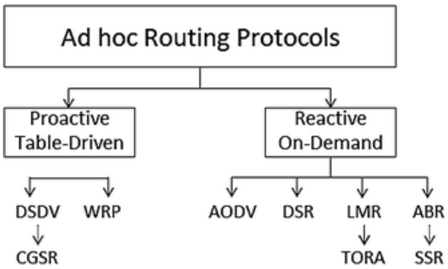

FIGURE 2. **Categorization of ad hoc routing protocols.**

Most MANETs assume that all routing protocols will be fully cooperative among participating nodes. Because MANETs are open, mobile, and have dynamic topologies and protocol weaknesses, attackers can target in various ways [2], [12]. MANETs have been proposed with several secure routing protocols [13–16]. It is common for them to assume centralized units or third parties, as a result of which MANETs cannot function independently.

Authentication, confidentiality, data integrity, and other security properties remain at risk from routing challenges, such as attacks [17]. Therefore, there is a distinction between internal and external nodal attacks. Attacks from external nodes cause the former, but attacks from grayholes and blackholes cause the latter [17–20]. Consequently, optimization plays a major role in achieving trust-enabled secure routing, providing a solution or two based on different objectives. A MANET has two types of trust: nodal and route trust [19]. Certain objectives must be met for simultaneous optimization to work in the real world. Bio-inspired algorithms also determine MANET routing paths. Several algorithms are used today, including PSO, ACO, bee colony optimization, and genetic algorithms. The ant-colony based route algorithm is one algorithm based on the ACO [21], HOPNET, AntHocNet [22], Ant-based Dynamic Zone Routing Protocol [23] and Hybrid ACO Routing [24] Foraging behaviour of ants is the basis for this theory [25].

MANET research mainly focuses on developing routing protocols that incur the smallest hop-count metric [26], incurring the least cost. Two categories of protocols can be distinguished: reactive protocols and proactive protocols.

Routes are discovered reactively by AODV nodes [27], [28] only when needed. In proactive routing mechanisms, OLSR continuously exchanges network topology information between its nodes to find routes [29], [30]. In contrast, hop-count metrics don't reflect a network's mobility or other contextual features when routing. Thus, a high mobility scenario may not be optimal for the route choice.

DSR was extended with a context-aware inference scheme to punish malicious accusers and accused. According to [31], selfish nodes could be detected by a context-aware mechanism. MANET environments with limited resources may not be able to use digital signatures. CORE (COllaborative REputation) is proposed by the author [32]. Inbuilt reputation functionality distinguishes direct, indirect, and functional reputations and a monitoring mechanism. A protocol is proposed to decide whether to cooperate or gradually isolate a node. Only positive reputation information is exchanged through this mechanism. However, if it does not have the option to submit negative feedback, it may be forced to rely on positive reports. The author proposes a new incentive-based approach to trust management called SORI (Secure and Objective Reputation-based Incentive) [33]. In addition to facilitating packet forwarding, one-way hash chain authentication schemes discourage selfish behaviour and encourage quantified objective measures.

This scheme may perform less well in a hostile environment because malicious nodes exist. The author [34] suggested that AODV (Adhoc On-Demand Distance Vector) can be considered an extension of AODV, in which the trust factor is used by one node, and another uses the security level. The approach taken by each node depends on its security level and trust factor. It proposes varying levels of encryption according to node trust factors rather than encrypting all routing information for every request. As a result, the approach conserves resources by adjusting security levels according to hostility; however, it does not assess trust levels.

## 3. METHODOLOGY

The algorithm uses the AODV protocol to implement Trust-Based QoS-aware Routing (I-TQAR) at the network layer [27]. Two cognitive phases make up the learning process in the cognitive layer of CN: the learning of paths (routings) and the learning of trust. RL methods of expected-SARSA [35], [36] are used to learn path segments. Considering the lower update variance and faster convergence of the expected SARSA, it is adopted [35]. As a result of improving an existing trust model, we use a model to learn about trust [37]. Nodes use RL methods to determine the best path for packet delivery based on their interactions with their environment. Further, nodes interact to learn each other's trustworthiness (trust learning). Figure 3 explains the path learning and trust learning phases, followed by the I-TQAR protocol.

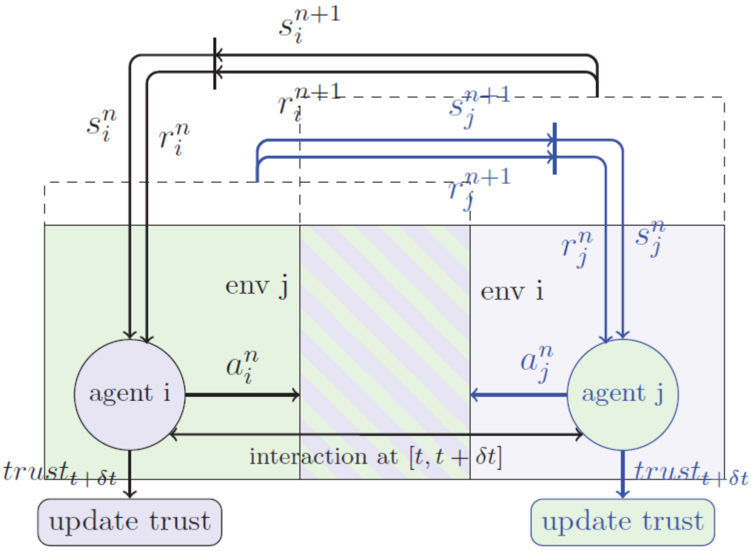

FIGURE 3. **Learning phases**

### A. Path learning phase

A RL agent is generally defined as a three-tuple $\{S, A, R\}$. A set of observed environment states is referred to as S. There are many actions that an agent can take. After time step $n$, agent $i$ observes the state $s_i^n \in S$ and decides to take action $a_i^n \in A$ and receive a reward $r_i^{n+1} \in R$, based on the knowledge it has about the environment. A subset of real numbers known as R represents the problem at hand. An agent's Q-value is a measure of the long-term rewards that can show expected as a result of their state-action pair when action $a_i^n \in A$ in state $s_i^n \in S$. State-action pairs $(Q(s_i^n, a_i^n))$ are tracked in Q-tables by each agent.

Our routing protocol is composed of the following RL components, assuming network nodes as agents:

- The environment of a node includes all other nodes except that particular one.
- Number of nodes currently holding RREQ packets in the network at time index $n$ in each state $s_i^n \in S\,(i \in (1, \ldots, k\}$
- RREQ is selected, at time index $n$, among the nodes directly connected to a forwarding network to $s_i$, among the nodes directly connected to a forwarding network to $s_i$, according to the action $a_i^n \in A$.
- $Q(s_i^n\ ,\ a_i^n)$ chooses an action based on its Q-value $a_i$ at state $s_i$ ending up at state n at time index n $s_i^{n+1}$. Agents update this value when the node select their next hop..
- There is a direct correlation between reward. i$r_i^n\ \in R$path quality and rate of travel. Nodes that select actions i$a_i^n$(select the next forwarding node.) can receive new rewards i$r_i^{n+1} \in R$based on the new state i$s_i^{n+1}$of the environment.

For each node (e.g. agent j) in the network to reach its end-to-end goal, it considers other nodes as part of its environment (e.g. agent i).

A node without network knowledge is assumed to have zero Q-value at the beginning of communication. As described in [35], This rule is based on weighted sums of old Q-values and learned Q-values, where learned Q-values represent immediate rewards, while old Q-values represent future rewards:

$$Q\left(s_i^n,\ a_i^n\right) \leftarrow (1-\propto).Q\left(s_i^n,\ a_i^n\right) + \propto \left(r_i^{n+1} + \gamma \textstyle\sum_{a_i^{n+1}} \pi(\left(s_i^{n+1},\ a_i^{n+1}\right).Q\left(s_i^{n+1},\ a_i^{n+1}\right)\right) \tag{1}$$

The learning rate, $\propto$, and discount factor, $\gamma$, are fixed in (1). Taking into account the $\varepsilon - greedy$ selection policy, the next intermediate forwarding node will be $\pi\left(s_i^{n+1},\ a_i^{n+1}\right)$. The immediate reward will be $r_i^{n+1}$.

A RL method based on Expected-SARSA is adopted in this study. The Future Q-value of expected SARSA is determined by weighing all possible actions following the formula below:

$$\textstyle\sum_{a_i^{n+1}} \pi(\left(s_i^{n+1},\ a_i^{n+1}\right).Q\left(s_i^{n+1},\ a_i^{n+1}\right) = (1-\varepsilon)\,.\underset{a_i^{n+1}}{max}\left(Q\left(s_i^{n+1},\ a_i^{n+1}\right)\right) + \varepsilon.mean_{a_i^{n+1}} Q\left(s_i^{n+1},\ a_i^{n+1}\right) \tag{2}$$

$r_i^{n+1}$ in Eqn. (1) a packet's delivery delay The reverse process occurs through the next intermediate forwarding node of its delivery delay through the next forwarding node. Using this strategy, you pick the best action based on its Q-value $a_i^{n+1}$ first through a chance of $1-\varepsilon$, and By selecting the average of all other actions with Q-values, you will arrive at the results you want.

A forwarding node's path quality affects the total link delay. Expected Transmission counts represent path quality [38]. Using packet loss ratios between neighbouring nodes estimates how many transmissions are required to deliver a packet to its destination.

Using hello packets, we obtain the value. Over a while $[t - \triangle t, t]$) A record of how many hello messages the machine broadcasts is kept (e.g. $h_j[t - \triangle t, t]$ a neighbour node's hello message number is recorded $j's$ broadcast (e.g. $h_j[t - \triangle t, t]$. In hello messages to neighbours, these values are conveyed. $ETX_{ij}^t$ according to the formula below, during time t, nodes i and j have the following relationship:

$$ETX_{ij}^t = \frac{1}{\frac{h_j}{h_i}} \tag{3}$$

As part of our computations, we ignore the queueing delay, which comes from transmission, propagation, and queuing delays. Thus, $d_{ij}^t$ is calculated based on $ETX$ and $t_{pkt}$ (ignoring propagation and queuing delays), as follows:

$$d_{ij}^t = ETX_{ij}^t \times t_{pkt} \tag{4}$$

Using this parameter, we can estimate how long a packet will take to reach a neighbouring node on average. The tradeoff between exploration and exploitation is a common challenge in RL algorithms. As a result of our analysis of all possible actions, we choose an action using a greedy strategy. An agent in $\varepsilon - greedy$ determines which action is optimal (one that maximizes Q-value); it takes action directly chance of $1 - \varepsilon$, and chooses consistently at random between the two actions with probability $\varepsilon$ where $0 < \varepsilon < 1$. Experimentally, we determine the appropriate value.

### B. Trust learning phase

- **Trust model:** Each node learns the trustworthiness of its neighbours during the trust learning phase of the algorithm. *agents i &j* interact through one another in time intervals $(t,\ t + \delta t]$. Consequently, each agent updates its neighbour's trust value in time intervals $\delta t$ seconds. A trust threshold $\lambda$ is used to determine a node's trustworthiness. Nodes that are considered trustworthy exceed or equal the threshold of trust. If node A is untrustworthy, it will remain isolated for the network's lifetime. Node B will also remain isolated since there will be no reconsideration. False decisions may lead to node isolation, which harms the final result. As part of our plans, we will address this issue.

Each node detects and isolates a black and grey hole using direct and indirect trust. Malicious nodes drop all packets intended for forwarding in black hole attacks. By participating in the routing process, it maintains its trustworthiness. While participating in the routing process, a malicious node selectively drops data packets randomly with an average probability of 0.5, unlike a black hole attack. There is an assumption that trust is asymmetric between neighbour nodes. A historical and current trust evaluation is also used to calculate the total/current trust. It is important to include previous trust evaluations in calculating current trust to prevent abrupt trust level changes caused by grayhole attacks.

**Trust computation:** Based on the interaction between two neighbours at a certain time t, direct trust is computed as follows:

$$DT^{t}_{ij} = \frac{f_j}{f_{i,j}} \tag{5}$$

There are two trust nodes $i$ is the trustee, node $j$ is the trustee, and node $i$ to node j is $f_{i,j}$ packets that are forwarded, in contrast, node $j$ forwards $f_j$ packets the time interval $(t,\ t + \delta t]$.

Based on neighbour recommendations, each node evaluates its trustee/target node. The trust value of each recommender is assigned by node $i$ based on its general reputation (total) since a recommender can be malicious. Based on the average weighted sum of the trust values of the recommenders, the indirect trust level of the trustee node is calculated as follows:

$$IT^{t}_{i,j} = \frac{\sum_{k\in N_i} T^{t-1}_{i,k} . T^{t-1}_{k,j}}{N_i} \tag{6}$$

A $T_{i,k}$ value representing the total trust between two nodes, whereas a $T_{k,j}$ an indication of how much a node trusts its neighbours. Alternatively, current trust is computed by weighing direct and indirect trusts at the node in question, as follows:

$$CT^{t}_{i,j} = w_1 DT^{t}_{i,j} + w_2 IT^{t}_{i,j}; \qquad w_1 + w_2 = 1,\ w_1 > w_2 \tag{7}$$

It is considered more reliable to obtain direct information directly from the node. In the formula below, the total/current trust is calculated by weighting the current and historical trusts:

$$T^{t}_{i,j} = w^{'}_1 CT^{t}_{i,j} + w^{'}_2 T^{t-1}_{i,j}; \quad w^{'}_1 + w^{'}_2 = 1, \quad w^{'}_1 > w^{'}_2 \tag{8}$$

The current trust is more important than the historical one because it is based on recent information.

## 4. RESULT ANALYSIS & DISCUSSION

Simulating I-TQAR, TQOR, and TQR protocols, which are recently proposed trust-based QoS-aware routing protocols, we examined their performance. Direct and indirect trust must be computed for packet forwarding through trusted routes to meet the QoS constraint of packet forwarding through trusted routes. A random mobility pattern was generated by our simulation using NS-2's 'setdest utility'. As nodes move randomly between 0 and the maximum specified speed (10 meters per second), nodes start randomly in a 1000m by 1000m area. A node's trust value can drop packets more often when malicious nodes are randomly distributed. The data transmitted was 512 bytes.

### A. Experimental Results Discussion

According to this definition, an end-to-end delay is the average time packets traverse between a source and a destination. This end-end delay can also include a retransmission, propagation or transfer delay. TQR and TQOR cause more delays than the proposed protocol I-TQAR. I-TQAR's network topology, which contains 6 malicious nodes, does not look like that. Our protocol prevents malicious nodes from interacting with each other. Considering the overall trust value between nodes in the proposed model, all interactions are based on that value**.**

It can be seen from Figure 5 that the PDR for the I-TQAR protocol decreases as the number of malicious agents increases. Communication from sources to destinations is slow as a result of packet loss. Despite maintaining a high PDR, our protocol maintains a normal network operation. We detect and mitigate attackers during the route discovery process,

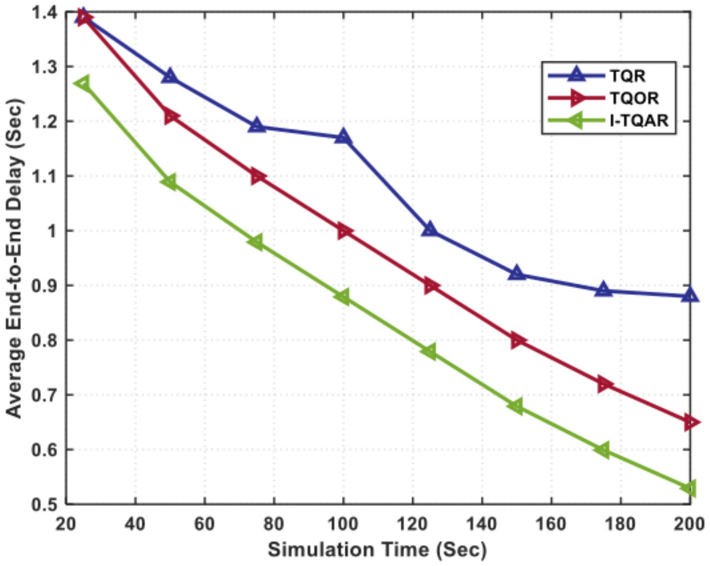

FIGURE 4. **Average end-end delay (Sec) versus simulation time (sec)**

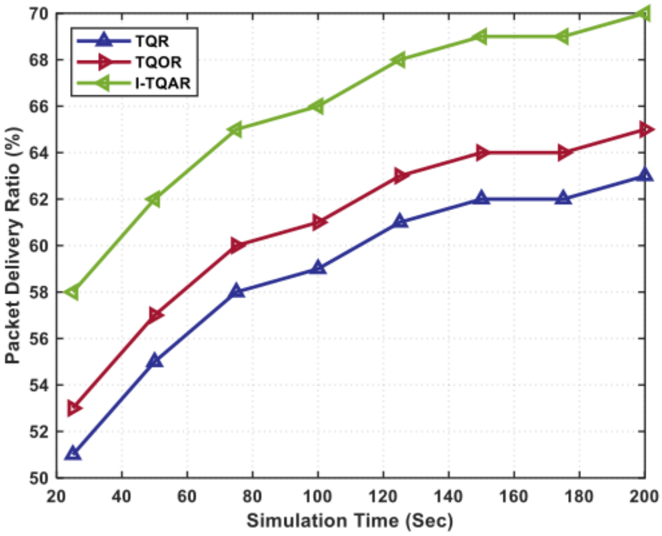

FIGURE 5. **Packet Delivery Ratio (%) versus simulation time (sec)**

so a combination of malicious attacks will not affect the PDR, as we usually communicate if any malicious attacks are present.

The information is sent from any source, the routing from all moving nodes determines the results. Generally, the simulation result is the accumulated data transmitted during the simulation period (irrespective of whether or not data have been sent and received). The figure below shows the simulation time comparison between the proposed I-TQAR model and current TQR and TQOR routing protocols.

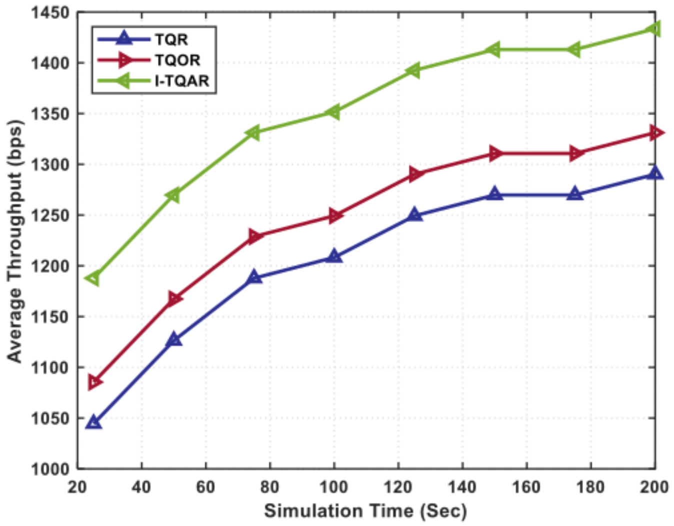

FIGURE 6. Time between end-of-simulation and simulation time (sec)

The propagation period and multiple delays in any communication system are reflected in latency. Because all nodes behave like clients and servers over the network channel, topology has a large end-to-end delay. Sending and receiving packets take time from end to end. When data is transmitted across a network, latency is how long it takes. With different values of network load, here is a typical time output. According to the proposed I-TQAR network topology, the maximum end-to-end delay is 1.26 seconds.

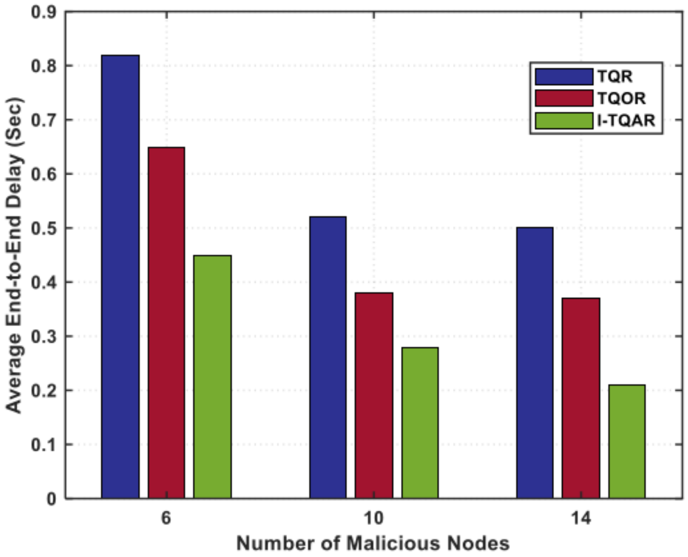

FIGURE 7. Number of malicious mobile nodes versus average end-end delay (Sec).

For a network to achieve its maximum throughput, its loss rate, or packet delivery ratio, must be determined. The number of malicious nodes and origin packets delivered by each protocol are calculated based on Figure 8. There is no relation between offered mobile nodes and the TQR, TQOR, and I-TQAR packet delivery ratio. The malicious node at 6 was served particularly well by routing protocols TQOR and I-TQAR, which delivered most of the original packets. TQR, however, delivered half of the original packets, as shown in Figure 8.

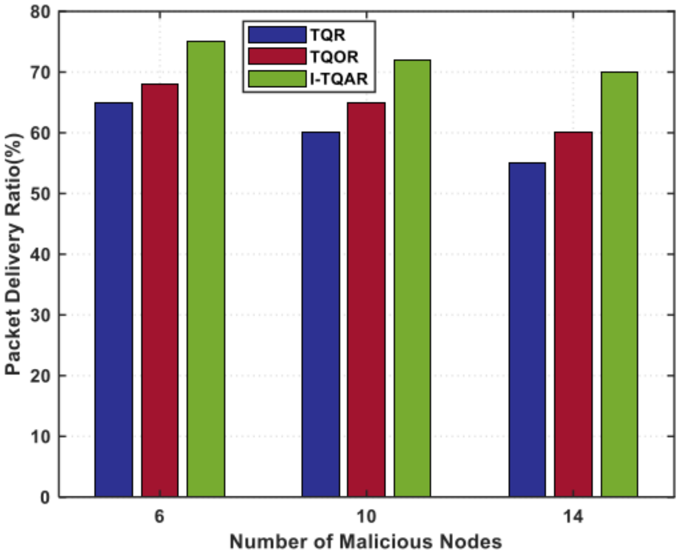

FIGURE 8. **Average end-end delay (Sec) versusnumber of malicious mobile nodes.**

## 5. CONCLUSION

Since wireless networks and mobile computing hardware can now support ad-hoc networking, researchers have been paying greater attention to the topic. Recently, several new routing protocols have been proposed for ad-hoc networking environments. Still, node-level performance comparisons and detailed performance information for each protocol were unavailable. Based on average end-end delays, average throughputs and packet delivery ratios (PDR), this paper compares TQR, TQOR, and I-TQAR. Therefore, this method outperforms those currently in use. An NS2 simulator monitors throughput, packet delivery ratios, delay between endpoints, and packet received ratios.

## FUNDING

None

## ACKNOWLEDGEMENT

None

## CONFLICTS OF INTEREST

The author declares no conflict of interest.

## REFERENCES

[1] M. S. Corson, J. P. Macker, and G. H. Cirincione, “Internet-based mobile ad hoc networking,” *IEEE Internet Comput*, vol. 3, no. 4, pp. 63–70, 1999.
[2] A. Attar, H. Tang, A. V. Vasilakos, F. R. Yu, and V. C. Leung, “A survey of security challenges in cognitive radio networks: Solutions and future research directions,” *Proc. IEEE*, vol. 100, pp. 3172–3186, 2012.
[3] N. Kumar, P. Rani, V. Kumar, S. V. Athawale, and D. Koundal, “THWSN: Enhanced energy-efficient clustering approach for three-tier heterogeneous wireless sensor networks,” *IEEE Sens. J*, vol. 22, no. 20, 2022.
[4] W. Li and J. Anupam, *Security Issues in Mobile Ad Hoc Networks-A Survey*. 2008.
[5] B. Bhola, “Quality-enabled decentralized dynamic IoT platform with scalable resources integration,” *IET Commun*, 2022.
[6] M. K. Gulati and K. Kumar, “QoS routing protocols for mobile ad hoc networks: a survey,” *Int. J. Wirel. Mob. Comput*, vol. 5, no. 2, pp. 107–118, 2012.

[7] N. Hussain, P. Rani, N. Kumar, and M. G. Chaudhary, "A deep comprehensive research architecture, characteristics, challenges, issues, and benefits of routing protocol for vehicular ad-hoc networks," *Int. J. Distrib. Syst. Technol. IJDST*, vol. 13, no. 8, pp. 1–23, 2022.

[8] N. Hussain and P. Rani, "Comparative studied based on attack resilient and efficient protocol with intrusion detection system based on deep neural network for vehicular system security," *Distributed Artificial Intelligence*, pp. 217–236, 2020.

[9] P. Rani and R. Sharma, "An Experimental Study of IEEE 802.11n Devices for Vehicular Networks with Various Propagation Loss Models," in *Advanced IoT Sensors, Networks and Systems* (A. K. Dubey, V. Sugumaran, , and P. H. J. Chong, eds.), vol. 1027, pp. 125–135, Springer Nature, 2023.

[10] J. Raju and J. J. Garcia-Luna-Aceves, "A comparison of on-demand and table driven routing for ad-hoc wireless networks," *2000 IEEE International Conference on Communications. ICC 2000. Global Convergence Through Communications*, pp. 1702–1706, 2000.

[11] J. W. De, G. Stênico, and L. L. Ling, "Topology Control for Ad-Hoc Networks: A comprehensive review for table driven and on-demand routing protocols," *Commun. Netw*, vol. 5, no. 03, pp. 239–246, 2013.

[12] Y. C. Hu and A. Perrig, "A Survey of Secure Wireless Ad Hoc Routing, Security and Privacy Magazine," *IEEE*, vol. 2, pp. 28–39, 2004.

[13] Y. C. Hu, D. B. Johnson, and A. Perrig, "SEAD: Secure efficient distance vector routing for mobile wireless ad hoc networks," *Ad Hoc Netw*, vol. 1, no. 1, pp. 175–192, 2003.

[14] Y. C. Hu, A. Perrig, and D. B. Johnson, "Ariadne: A secure on-demand routing protocol for ad hoc networks," *Proceedings of the 8th annual international conference on Mobile computing and networking*, pp. 12–23, 2002.

[15] A. Perrig, R. Canetti, J. D. Tygar, and D. Song, "The TESLA broadcast authentication protocol," *CryptoBytes*, vol. 5, no. 2, pp. 2–13, 2002.

[16] P. G. Argyroudis and D. O'mahony, "Secure routing for mobile ad hoc networks," *IEEE Commun Surv Tutor*, vol. 7, no. 1-4, pp. 2–21, 2005.

[17] R. M. Chintalapalli and V. R. Ananthula, "M-LionWhale: multi-objective optimisation model for secure routing in mobile ad-hoc network," *IET Commun*, vol. 12, no. 12, pp. 1406–1415, 2018.

[18] M. K. Garg, N. Singh, and P. Verma, "Fuzzy rule-based approach for design and analysis of a Trust-based Secure Routing Protocol for MANETs," *Procedia Comput. Sci*, vol. 132, pp. 653–658, 2018.

[19] J. Sathiamoorthy and B. Ramakrishnan, "Design of a proficient hybrid protocol for efficient route discovery and secure data transmission in CEAACK MANETs," *J. Inf. Secur. Appl*, vol. 36, pp. 43–58, 2017.

[20] N. Hussain, P. Rani, H. Chouhan, and U. S. Gaur, "Cyber Security and Privacy of Connected and Automated Vehicles (CAVs)-Based Federated Learning: Challenges, Opportunities, and Open Issues," in *EAI/Springer Innovations in Communication and Computing* (F. L. for IoT Applications, S. P. Yadav, B. S. Bhati, D. P. Mahato, , and S. Kumar, eds.), pp. 169–183, Springer International Publishing, 2022.

[21] M. Gunes, U. Sorges, and I. Bouazizi, "ARA-the ant-colony based routing algorithm for MANETs," *Proceedings. International Conference on Parallel Processing Workshop*, pp. 79–85, 2002.

[22] G. D. Caro, F. Ducatelle, and L. M. Gambardella, "AntHocNet: an adaptive nature-inspired algorithm for routing in mobile ad hoc networks," *Eur. Trans. Telecommun*, vol. 16, no. 5, pp. 443–455, 2005.

[23] A. M. Okazaki and A. A. Fröhlich, "Ant-based dynamic hop optimization protocol: A routing algorithm for mobile wireless sensor networks," *2011 IEEE GLOBECOM Workshops (GC Wkshps)*, pp. 1139–1143, 2011.

[24] D. Cañas, A. L. Orozco, L. J. García, P. S. Villalba, and Hong, "Hybrid ACO routing protocol for mobile ad hoc networks," *Int. J. Distrib. Sens. Netw*, vol. 9, no. 5, pp. 265485–265485, 2013.

[25] G. S. Pavani and R. I. Tinini, "Distributed meta-scheduling in lambda grids by means of Ant Colony Optimization," *Future Gener. Comput. Syst*, vol. 63, pp. 15–24, 2016.

[26] S. Corson and J. Macker *Mobile ad hoc networking (MANET): Routing protocol performance issues and evaluation considerations*, 1999.

[27] C. Perkins, E. Belding-Royer, and S. Das *Ad hoc on-demand distance vector (AODV) routing*, 2003.

[28] I. D. Chakeres and E. M. Belding-Royer, "AODV routing protocol implementation design," *24th International Conference on Distributed Computing Systems Workshops*, pp. 698–703, 2004.

[29] T. Clausen and P. Jacquet *Optimized link state routing protocol (OLSR)*, 2003.

[30] P. Jacquet, P. Muhlethaler, T. Clausen, A. Laouiti, A. Qayyum, and L. Viennot, "Optimized link state routing protocol for ad hoc networks," *Proceedings. IEEE International Multi Topic Conference*, pp. 62–68, 2001.

[31] K. Paul and D. Westhoff, "Context aware detection of selfish nodes in DSR based ad-hoc networks," *Proceedings IEEE 56th Vehicular Technology Conference*, pp. 2424–2429, 2002.

[32] P. Michiardi and R. Molva, "Core: a collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks," in *Advanced Communications and Multimedia Security: IFIP TC6/TC11 Sixth Joint Working Conference on Communications and Multimedia Security*, pp. 107–121, Springer, 2002.

[33] Q. He, D. Wu, and P. Khosla, "SORI: A secure and objective reputation-based incentive scheme for ad-hoc networks," *2004 IEEE Wireless Communications and Networking Conference*, pp. 825–830, 2004.

[34] R. K. Nekkanti and C. Lee, "Trust based adaptive on demand ad hoc routing protocol," *Proceedings of the 42nd annual Southeast regional conference*, pp. 88–93, 2004.

[35] H. V. Seijen, H. V. Hasselt, S. Whiteson, and M. Wiering, "A theoretical and empirical analysis of Expected Sarsa," *2009 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning*, pp. 177–184, 2009.

[36] R. S. Sutton and A. G. Barto *Reinforcement learning: An introduction*, 2018.

[37] B. Wang, X. Chen, and W. Chang, "A light-weight trust-based QoS routing algorithm for ad hoc networks," *Pervasive Mob. Comput*, vol. 13, pp. 164–180, 2014.

[38] D. S. D. Couto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multi-hop wireless routing," *Proceedings of the 9th annual international conference on Mobile computing and networking*, pp. 134–146, 2003.