Quantum-Resilient Machine Learning and Q-Learning–Driven Priority Time-Slot AODV for Secure MANET Routing

Singireddy Sateesh Reddy, Dr. E. Aravind
Dept. of Computer Science & Engineering, Chaitanya (Deemed to be University), Hyderabad, India

***Abstract*—Mobile Ad Hoc Networks (MANETs) are decentralized in nature and, therefore, they have no centralized control, and consequently, they are highly susceptible to routing attacks like black hole attacks and gray hole attacks, both of which disable data delivery by causing a vicious loss of packets. To address these issues, the current study offers the Quantum-resilient Machine Learning and Q-Learning-driven Priority Time-Slot AODV (QR-MLQ-PTS-AODV) routing model. This framework combines a multi-metric trust query, an entropy-based behavioral stability query, a temporal query trust adjustment, and a managed machine learning method to attain exact malicious node forecasting. Reinforcement learning, through Q -learning, is employed to utilize dynamical assignment of MAC -layer priority time slots to enable cross-layer optimization, as well as adaptive routing decisions. In contrast to solutions that exist, the suggested framework avoids quantum-vulnerable cryptographic primitives in favor of hash-based trust authentication and learning-based mitigation measures, to make sure that it can withstand novel quantum-assisted routing attacks. The limited variables of the trust model are determined by a mathematical analysis and extensive NS-3 simulations that show that the model significantly improves the ratio of packet delivery, end-to-end delay, routing overhead, and attack detection accuracy in comparison with traditional AODV and the most up-to-date trust-, ML-, and RL-based protocols. Based on these results, the effectiveness of embedding quantum-sensitive security protocols and smart cross-layer routing in MANETs can be supported.**

Keywords*—Mobile Ad Hoc Networks; secure routing; AODV; trust management; machine learning; reinforcement learning; MAC layer scheduling; black hole attack; post-quantum security; quantum-resilient routing

I. Introduction

Mobile Ad Hoc Networks (MANETs) are used to communicate autonomously between incapacitated, wireless nodes with no infrastructure support. These are widely used in military missions, disaster management, and intelligent transportation systems. Even though they are flexible by their nature, MANETs are by and large subject to security threats due to dynamic topology, open air channels and decentralized control. Accordingly, routing protocols that act in these environments have to simultaneously operate efficiently and securely in highly dynamic environments. Other routing protocols such as OLSR [28] have also been proposed, but they rely on proactive routing strategies that may introduce additional overhead in highly dynamic MANET scenarios. Fundamental concepts and protocol designs for ad hoc networks are well documented in [25], which provides a basis for understanding routing challenges in MANET environments.

The Ad hoc On-Demand Distance Vector (AODV) protocol has proven to be one of the most popular MANET routing protocols, one of the major reasons is the ad hoc nature of the protocol, whereby control overhead is minimal. However, AODV does not have any security features, which makes it very vulnerable to routing attacks in the form of black hole attack and gray hole attack. In such malevolent situations, the adversarial nodes forge the optimal route advertisements and it goes ahead to drop the packets and this culminates in serious deterioration of the ratios of packet delivery, increased latency and further increased routing overhead. Traditional cryptographic defenses create computational burdens and are mostly useless against insider attacks, and at the same time, this is based on cryptographic primitives that can be broken by adversaries with quantum capability.

Trust-based routing has also become an effective alternative by analyzing the behavior of nodes instead of just using cryptographic authentication. But most of the models of extant trust use static measures and do not keep up with the changes in temporal behaviors, selective forwarding of erroneous or dynamic attack patterns. Recent progress in machine learning (ML) and reinforcement learning (RL) shows a lot of potential in complementing attack detection and routing flexibility; however, these methods are often implemented separately and are not a common part of cross-layer routing and MAC-layer optimization models.

In order to overcome these issues, this study creates a Quantum-resilient Machine Learning and Q-Learning-Driven Priority Time-Slot AODV (QR-MLQ-PTS-AODV) routing model. The suggested solution combines multi-metric and entropy-based trust assessment, temporal trust adjustment, malicious node forecasting by a supervised machine learning approach, and adaptive MAC-layer time-slot distribution by the Q-learning approach. Additionally, the framework avoids the use of quantum-vulnerable cryptographic primitives through the use of hash-based trust authentication and learning-based mitigation mechanisms and hence guarantees resilience to emerging quantum-aided routing attacks. It is an effective design; cross layer design enables safe, adaptable and effective routing of MANETs using both the classical and post-quantum threat design.

II. Literature Review

Ad Hoc On-Demand Distance Vector (AODV) routing protocol, which was designed and standardized by Perkins et al. [1], is one of the most widely used routing strategies in Mobile Ad Hoc Networks (MANETs) due to its reactive routing discovery method and the relatively small routing overhead. However, AODV had been designed without inbuilt security, meaning that it was in extremely vulnerable circumstances to routing attacks within unfriendly settings. Initial work on security of MANET has highlighted these weaknesses and has recommended that secure routing mechanisms be incorporated [2], [23].

Early routing misbehavior attempts used monitoring-based techniques. The watchdog and pathrater schemes to detect the packet-dropping nodes by overhearing of the neighbor transmissions have been introduced by Marti et al. [3]. Despite its ability to overcome crude attacks by black holes, the practice has a false positive tendency in case of uncertain conditions of the wireless environment. Later, secure routing extensions which added cryptographic tools were suggested [4], but these do not have much computation overhead, as well as they do not address insider attacks.

Routing schemes based on trust and reputation later came up as an alternative to the cryptographic mechanisms. Sun et.al. [6] proposed an information-theoretic framework of trust, which is a model of node reliability presented in the form of entropy based metrics, therefore giving a formal basis of trust assessment in MANET. The protocol proposed by Buchegger and Le Boudec [7] is known as CONFIDANT, and the idea of this protocol is to use reputation exchange as a means of discouraging possible misbehavior, and the incentive-based cooperation mechanisms were suggested by Buttyan and Hubaux [8]. However, the methods are prone to false charges, conspiracy, and corpse faith. Dynamic trust-based routing approaches are proposed in [9] and comprehensive trust management frameworks are discussed in [10].

A number of trust-based variants of AODV have been developed in order to address black holes attacks. Misra et al. [12] incorporated the concept of packet-forwarding behavior in the making of trust-based routing decisions that they identified to have a high ratio of packet delivery. Zhang and Cohen [11] also established that the trust route algorithm is more reliable but this scheme assumes predetermined levels of trust and is vulnerable to on-off and gray hole threats knowledgeable attacks. Mechanisms for co-operative black holes attack prevention were also considered [5], [24] although they introduce extra routing overhead and reduced flexibility.

As there are improved systems in artificial intelligence, machine learning methods have been applied to MANET security. Abbas et al. [21] also used the AI of supervised learning to improve the accuracy of managing trust, whereas Shafi et al. [17] implemented a combination of machine learning and trust-based AODV routing with better detection. Previous studies undertaken by Vatambeti et al. [18] employed the Random Forest and SVM classifier to identify black hole attack incidents, they found that the detection rates were up to 88%. Intrusion detection in systems based on deep learning have also been suggested [19], [27], which offer high detection rates, but do not depend on routing protocols, and do not affect route determination, or the activities of the MAC-layer. Similarly, AI-based detection schemes [20] are faster in detection, but have loose links between detection and routing choices. General anomaly detection techniques and their applicability to network security have been extensively studied in [16], providing a foundation for ML-based intrusion detection in MANETs.

Reinforcement learning (RL) has been explored to support the flexibility of routing to dynamically operated MANETs. The theoretical principles of RL were defined by Sutton and Barto [13], which were then applied in the field of optimization of MANET routing. Chaubey and Kumar [15] developed a Q-learning-based secure routing protocol that can help to reduce the threat of black hole attacks; at the same time, Wang et al. [14] revealed that the RL-based routing is beneficial to some stability in the paths. Another integration method has also been discussed [22] and it is the trust -RL, which gives a lower delay and better adaptability and game-theoretic perspectives on adaptive intrusion detection and decision-making have been studied in [26], which further support intelligent security optimization in dynamic environments. These strategies, nevertheless, do not include controlled machine-learning-based detection and do not take priorities of MAC-layers into consideration.

Although these innovations have been made, currently MANET security solutions exclusively focus on trust, machine learning or reinforcement learning. Entropy-based behavioral stability, temporal evolution of trust, and cross-layer MAC-layer optimization are very few studies that are considered. Furthermore, most modern-day protocols are based on cryptographic primitives, which can be broken by quantum-enabled attackers.

The contemporary research has focused on the pressing need of the post-quantum security in the future wireless networks [30], [31]. Resiliency to quantum-based trust management and learning-based anomaly detection have yet to emerge [32], [33], and most of the possible avenues of integrating these features with MANET routing are still under-investigated. Advanced adversarial and post-quantum threat model reinforcement learning has also been considered [34], though without explicit trust modeling and integration with AODV.

Unlike extant work, the proposed Quantum-Resilient ML and Q-Learning-Driven Priority Time-Slot AODV (QR-MLQ-PTS-AODV) framework is unique in its capability to combine multi-metric and entropy-based trust assessment, supervised machine-learning-assisted malicious node forecasting, Q-learning-based MAC-layer priority time-slot allocation and quantum-resilient hash-based trust authentication. The present holistic and cross-layer methodology will deal with the classical and post-quantum threat of routing and the methodology will improve the reliability, efficiency, and security of MANETs, significantly.

III. Proposed Methodology

The suggested QR-MLQ-PTS-AODV model is highly integrated, whereby it operates on a tight cross-layered architecture depicted in Fig. 1, which consists of the following labeled modules:

- Network Monitoring Module: Every node locally tracks routing and MAC-layer dynamics, such as Packet Delivery Ratio (PDR), Forwarding Factor Ratio (FFR), end-end delay as well as residual energy. These measurements are the characteristic array of learning and trust assessment.

- ML-Malicious node prediction module: The observed characteristics are sent to a trained machine-learning classifier (SVM or Random Forest), which gains an insight into the node state, benign or malicious. This is an ML trust score where the confidence of the prediction is exported.

- Trust Computer and Quantum-Resilient Authentication Module: The module calculates composite trust based on direct trust, entropy-based behavioral stability and time based trust decay. Behavioral trust is combined with the ML trust score. Lightweight hash-based trust authentication is applied to circumvent cryptographic primitives that are quantum-vulnerable.

- Priority Time Slot allocation Module based on Q-Learning: In a Markov Decision Process, trust states and network conditions are modeled and are states. Q-learning picks the best actions that are associated with MAC-layer priority time-slot assignments and the high-trust nodes are preferred.

- Secure Routing and MAC Scheduling Module: The updated AODV protocol is a combination of trust-conscious route choice and adaptive MAC-layer scheduling. Bad nodes are segregated, and the good nodes are given priority of transmission.

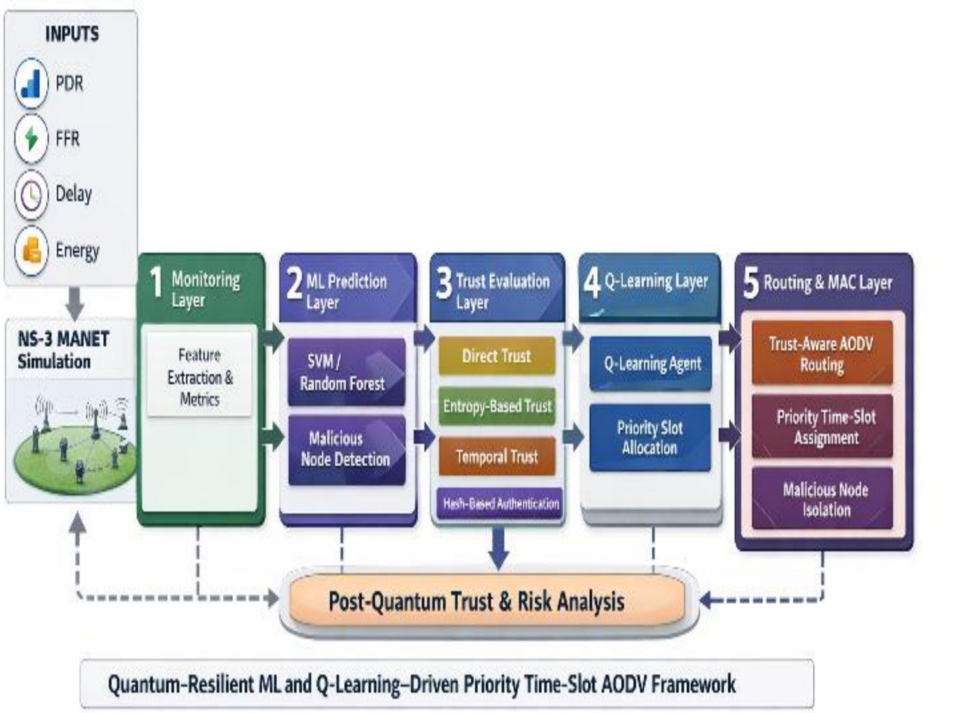

Fig. 1. System architecture.

A. Novel Trust Calculation Model

1) Multi-metric trust measurement: A raw estimate of trust is obtained by a weighted combination of normalized behavioral indicators, as the value of each node is provided by:

$$T_i^{raw} = w_1 PDR_i + w_2\, FFR_i + w_3\,(1 - D_i) + w_4\, E_i \tag{1}$$

In which, PDRi denotes packet delivery ratio, FFRi denotes Packet Forwarding Ratio, Di denotes Normalized end-to-end delay , Ei denotes Residual energy - (w1 + w2 + w3 + w4 = 1).

2) Entropy-based behavioral stability: Entropy is used in order to measure the inability of forwarding behavior.

$$H_i = -\sum p(x) \log p(x) \tag{2}$$

The entropy-based trust component is defined as:

$$T_i^{entropy} = 1 - H_i \tag{3}$$

3) Temporal trust with exponential decay: Temporal trust focuses on the recent observations and eliminates on-off attacks:

$$T_i\,(t) = T_i\,(t\text{-}1) + (1 -)T_i^{current},\ 0< \lambda < 1 \tag{4}$$

4) Integration with ML prediction: A machine-learning module that goes along, identifies nodes as either malicious or not to get the final trust estimate, one needs to take:

$$T_i^{final} = T_i^{raw} + T_i^{entropy} + ML_{Flag},\ \beta + \gamma + \delta = 1 \tag{5}$$

The nodes which have final trust values lower than a preset threshold are considered malicious and are not included in routing activities.

B. Mathematical Analysis and Proofs

1) Trust boundedness: Theorem 1:$T_i^{final}\epsilon[0,1]$: Evidence: All metrics used in the computation are normalized to the interval [0, 1]. Since the final trust value is calculated as a weighted sum of these normalized metrics, it also lies within the same interval.

2) Convergence of temporal trusts: Theorem 2: Emporal trust approaches to the existing trust value.

Proof: A first-order linear difference equation makes sure that the recent behavior prevails.

3) Capability of detecting black holes: Theorem 3: Malicious nodes are identifiable using trust + ML.

Proof: Malicious nodes produce low PDR, FFR, high entropy, and ML flag, which leads to (T_i^{final}< threshold).

4) Quantum-aware trust resilience (post-quantum extension): The availability of emerging quantum computing is a major threat to the traditional cryptographic security schemes that are deployed in MANETs, especially those based on the public-key infrastructure. Even though the suggested ML-TQ-PTS-AODV framework will not rely on intensive cryptography, the faster choice-making and obnoxious packet manipulation techniques can be applied by quantum-capable attackers. In order to become more resilient to such threats, a quantum-aware trust modulation factor is added..

A quantum-risk indicator is defined as:

$$Q_i=1-\exp(-k.H_i) \tag{6}$$

where, Hi is the entropy-based uncertainty of the behavior of node iii, and K>0, is a sensitivity constant which is tuned to vary the responsiveness to abnormal fluctuations. The increase in the value of entropy- that may have reflected quantum-assisted adaptive attacks- leads to an increase in the Qi. The last quantum trust value is calculated as:

$$T_i^{QA}= (1-Q_i)\, T_i^{final} \tag{7}$$

This formulation is such that the nodes with very erratic forwarding behavior grow to have their trust being degraded much faster even when there is no explicit cryptographic failure. The proposed model is lightweight, scalable, and post-quantum in nature by default since it does not rely on key-based security but instead makes use of behavioral entropy.

C. Q-Learning–Based Adaptive Optimization

1) Q-learning model: All the nodes are independent RL agents.

- State Space:S= {Tifinal, PDRi, Di, Ei}
- Action Space: Increase Slot, Decrease Slot, Maintain Route, Isolate Node
- Reward Function:R=αPDR−βDelay−γEnergy+δTrust

Q-value update rule:

$$Q(s, a) \leftarrow Q(s, a) + \eta[R+\gamma\max Q (s', a') - Q(s, a) \tag{8}$$

2) Priority time-slot allocation: Based on the Q-values, time slots are dynamically differed.

$$Slot_i (t+1) =Slot_i (t) + Q(s, a) \tag{9}$$

The nodes with big trust are given bigger transmission spaces; the ones with low trust are restricted or isolated.

D. Integration with AODV

The QR-MLQ-PTS-AODV framework, which incorporates trust assessment, machine learning-based detection, and Q-learning-based MAC-layer schedule, is an improvement of the standard AODV protocol without changing the underlying on-demand routing paradigm.

1) ML-Aware and trust- aware route selection: In the process of Route Reply (RREP), the intermediate node looks at the trust value $T_i^{QA}$and machine learning classification of the adjacent node. Pathways that include nodes whose trust scores are less than the set parameter or those nodes that are considered to be malicious as detected by the ML model are discarded.

2) Secure routing table maintenance: Malicious nodes are dynamically removed out of routing tables and are not allowed to take part in subsequent route discovery and routing.

3) Cross-layer MAC integration: MAC layer is further upgraded with time-slot allocation with priorities. Q-learning projects the trust state and network conditions to the best transmission actions, so that high-trust nodes are given priority on the channel access, and low-trust nodes are given less access to the channel.

4) Dynamic malicious node separation: Trust values and ML predictions are constantly updated and then black hole nodes and gray holes, including on-off attackers are quickly isolated without needing route rediscovery flooding. Algorithm 1 presents the proposed model.

Algorithm 1: Proposed Model

```
Initialize Trusti ← 1.0 for all nodes i∈ N
Initialize Q-table Q(s, a) ← 0 for all states s and actions a
Load trained ML model (black hole / gray hole classifier)
for time t = 1 to MaxTime do

    for each node i∈ N do
        Measure $PDR_i$, $FFR_i$, $Delay_i$, $Energy_i$
    end for

    for each node i∈ N do
        Compute raw trust:
        $T^{raw}(i) \leftarrow w1\cdot PDR_i + w2\cdot FFR_i + w3\cdot(1 - Delay_i) + w4\cdot Energy_i$
        Compute entropy-based trust:
```

```
    $H_i \leftarrow -\Sigma\, p(x) \log p(x)$
    $T^{entropy}(i) \leftarrow 1 - H_i$
    Update temporal trust:
    $T^{temp}(i) \leftarrow \lambda\cdot Trust_i(t-1) + (1-\lambda)\cdot( T^{raw}(i) + T^{entropy}(i))$
end for

for each node $i \in N$ do
    $ML_{flag}(i) \leftarrow ML_{model}(PDR_i, FFR_i, Delay_i, Energy_i)$
    Compute final trust:
    $Trust_i(t) \leftarrow \beta\cdot T^{temp}(i) + \delta\cdot ML_{flag}(i)$
    Verify trust authenticity using hash-chain validation
    Compute Quantum Risk Factor $QRF_i$
    Adjust trust using quantum-aware trust equation
    Penalize high-QRF nodes in Q-learning reward
end for

for each node $i \in N$ do
    State $s_i \leftarrow \{ Trust_i(t), PDR_i, Delay_i, Energy_i \}$
end for

for each node $i \in N$ do
    if rand() < $\varepsilon$ then
        Select random action $a_i$
    else
        Select action $a_i \leftarrow \arg\max Q(s_i, a)$
    end if
end for

for each node $i \in N$ do
    if $a_i$ == IncreaseSlot then
        $Slot_i \leftarrow Slot_i + \Delta$
    else if $a_i$ == DecreaseSlot then
        $Slot_i \leftarrow Slot_i - \Delta$
    else if $a_i$ == IsolateNode then
        Remove node i from routing tables
    end if
end for

for each node $i \in N$ do
    $R_i \leftarrow \alpha\cdot PDR_i - \beta\cdot Delay_i - \gamma\cdot Energy_i + \delta\cdot Trust_i(t)$
end for

for each node $i \in N$ do
    Observe next state $s_i'$
    $Q(s_i, a_i) \leftarrow Q(s_i, a_i) + \eta\cdot[ R_i + \gamma\cdot \max Q(s_i', a') - Q(s_i, a_i)]$
end for

for each node $i \in N$ do
    if $Trust_i(t) < Th$ then
        Isolate node i from network
    end if
end for
end for
return secure routing paths with priority time-slot scheduling
```

E. Simulation Setup

Implemented in NS-3: - Network size: 20–50 nodes - Mobility model: Random Waypoint - Traffic: UDP CBR - Attack model: Black hole and gray hole - ML models: SVM and Random Forest - Evaluation metrics: Packet Delivery Ratio, End-to-End Delay, Routing Overhead, Throughput and Detection Accuracy.

F. Post-Quantum–Aware Routing Framework

1) Motivation: Intelligent Routing Attacks continue to become more vulnerable to Mobile Ad Hoc Networks (MANETs) because of the decentralized nature of their functioning, as well as cooperative routing. Although conventional security solutions have been largely dependent on the use of public-key cryptography, the advent of quantum computing poses a threat to the sustainability of these solutions in the long term and specifically, RSA-based and ECC-based solutions. Furthermore, accelerated attackers can use accelerated computation and adaptive routing strategies to augment routing attacks, including black hole attacks and gray hole attacks.

This study follows a cryptography-light behavior-driven security model, which is in line with the proposed Quantum-Resilient Machine Learning and Q-Learning-Driven Priority Time-Slot AODV (QR-MLQ-PTS-AODV) framework. The framework combines hash-based trust authentication, entropy-based behavioral analysis, supervised learning in the machine learning, and reinforcement learning to offer strong and future-secure routing security instead of relying on quantum-vulnerable primitives.

2) Hash-based trust authentication: Each node uses a One-way Hash-chain Mechanism in order to guarantee safe and lightweight dissemination of trust. H(.) is a cryptographic hash and ki is the seed of node Iii which is initially a secret. The hash chain can be calculated as:

$$k_i^n=H(k_i^{n-1}),n\geq 1 \tag{10}$$

The update of trust is sent as (T_i, k_i^n) and the adjacent node authenticates by using:

$$k_i^n=H(k_i^n)=k_i^{n+1} \tag{11}$$

This protocol removes the overhead of key exchange, deters forged propagation of trust and they are quantum resistant to cryptanalysis.

3) Quantum-aware behavioral risk modeling: The quantum-assisted attackers can have anomalously stable, swift or selectively adaptive forwarding actions. To identify such anomalies, short-term behavioral variation is used to define a Quantum Risk Factor (QRF):

$$QRF_i=\sigma(PDR_i)+\sigma(Delay_i)+\sigma(H_i) \qquad (12)$$

in which σ (.) is temporal variance and Hi is uncertainty of behavior based on entropy. The trust value is adjusted as:

$$T_i^Q= (1-\mu)T_i^{final}- \mu QRF_i,\ 0<\mu<1 \qquad (13)$$

This formulation enables accelerated trust degradation for nodes exhibiting quantum-risk characteristics, independent of cryptographic failure.

4) ML-assisted quantum-aware detection: The supervised machine learning classifier is expanded with features of quantum-awareness, such as:

- Trust update frequency
- Entropy fluctuation rate
- The anomaly of hash verification.
- Deviation of route response time.

The output of the classifier is determined as:

$$ML_Q(i) \in \{0,1\}$$

The final post-quantum trust score becomes:

$$T_i^{PQ}= \beta T_i^Q+ \delta ML_Q(i) \qquad (14)$$

Nodes with $T_i^{PQ}<T_{th}$ are dynamically isolated from routing operations.

5) Quantum-aware reinforcement learning optimization: In order to implement adaptive mitigation, the quantum-risk behavior is punished by the Q-learning reward function:

$$R = \alpha PDR-\beta Delay-\gamma Energy+\delta T_i^{PQ} - \kappa QRF_i \qquad (15)$$

With this formulation of rewards, the convergence acceleration is achieved, adaptive MAC-layer prioritization, and early isolation of high-risk nodes are guaranteed.

6) Key security advantages: The quantum-resilient extension offers:

The quantum key-breaking attacks can be immunized.

It should have the capability to detect adaptive and accelerated routing attacks.

The cryptography-free, lightweight authentication.

Almost flawless interconnection with trust, ML, and RL elements.

The detailed workflow of the proposed framework is illustrated in Fig. 2, highlighting the interaction between trust computation, machine learning prediction, and q-learning-based mac scheduling. Furthermore, the operational flow of the routing and decision-making process is depicted in Fig. 3.

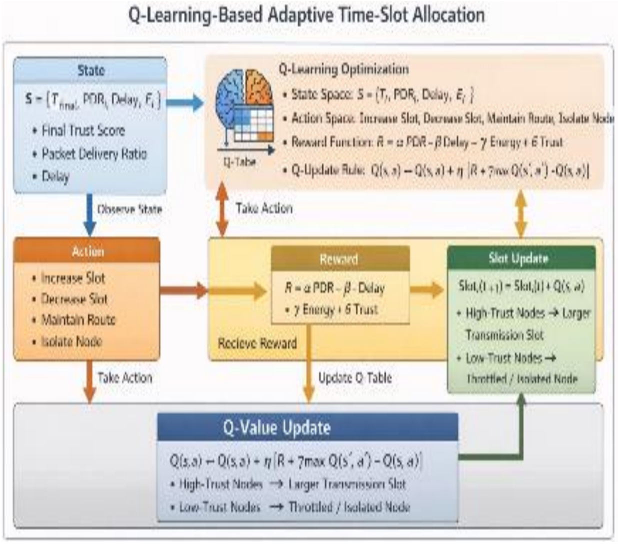
Q-Learning-Based Adaptive Time-Slot Allocation
State
S = {T_final, PDR_i, Delay, E_i}
• Final Trust Score
• Packet Delivery Ratio
• Delay
Q-Learning Optimization
• State Space: S = {T_i, PDR_i, Delay, E_i}
• Action Space: Increase Slot, Decrease Slot, Maintain Route, Isolate Node
• Reward Function: R = α PDR − β Delay − γ Energy + δ Trust
• Q-Update Rule: Q(s,a) ← Q(s,a) + η [R + γmax Q(s',a') −Q(s,a)]
Q-Table
Observe State
Take Action
Action
• Increase Slot
• Decrease Slot
• Maintain Route
• Isolate Node
Reward
R = α PDR − β · Delay
• γ Energy + δ Trust
Recieve Reward
Slot Update
Slot_i(t+1) = Slot_i(t) + Q(s,a)
• High-Trust Nodes → Larger Transmission Slot
• Low-Trust Nodes → Throttled / Isolated Node
Take Action
Update Q Table
Q-Value Update
Q(s,a) ← Q(s,a) + η [R + γmax Q(s',a') − Q(s,a)]
• High-Trust Nodes → Larger Transmission Slot
• Low-Trust Nodes → Throttled / Isolated Node

Fig. 2. Learning–based adaptive time slot allocation.

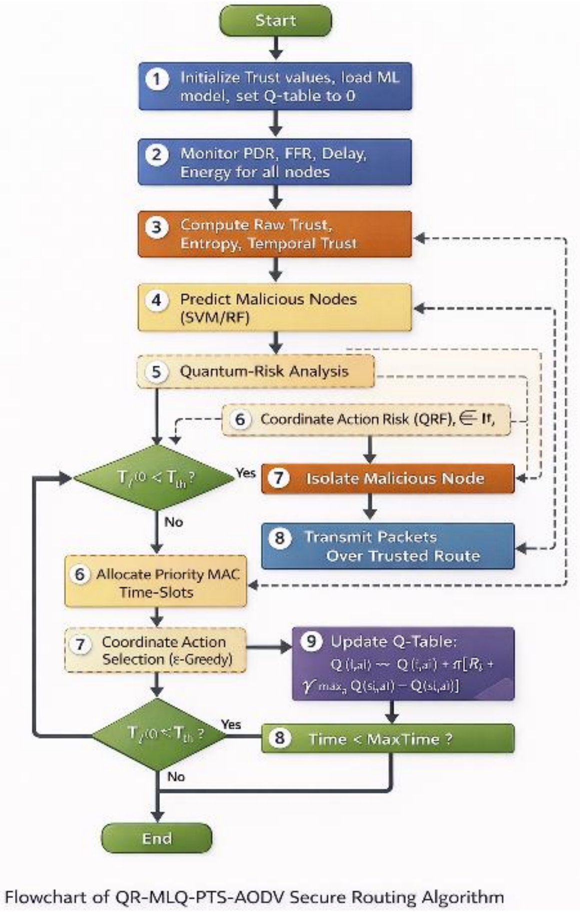
Start
1 Initialize Trust values, load ML model, set Q-table to 0
2 Monitor PDR, FFR, Delay, Energy for all nodes
3 Compute Raw Trust, Entropy, Temporal Trust
4 Predict Malicious Nodes (SVM/RF)
5 Quantum-Risk Analysis
6 Coordinate Action Risk (QRF), ∈ Ir,
T_i(t) < T_th ?
Yes
7 Isolate Malicious Node
No
8 Transmit Packets Over Trusted Route
6 Allocate Priority MAC Time-Slots
7 Coordinate Action Selection (ε-Greedy)
9 Update Q-Table:
Q(i,ai) ← Q(i,ai) + π[R_i + γ max_a Q(si,al) − Q(si,ai)]
T_i(t) < T_th ?
Yes
8 Time < MaxTime ?
No
End
Flowchart of QR-MLQ-PTS-AODV Secure Routing Algorithm

Fig. 3. Flow chart of the model.

IV. Performance Evaluation

This section compares the performance of the suggested Quantum-Resilient ML- and Q-Learning-Driven Priority Time-Slot AODV (QR-MLQ-PTS-AODV) in the case of black holes and gray holes attacks. It compares protocol with NS-3

Simulations and compares it with AODV, Trust-based AODV (T-AODV), ML-AODV and RL-AODV. The metrics taken into account in the evaluation are Packet Delivery Ratio (PDR), End-to-End Delay, Routing Overhead, Throughput, and Malicious Node Detection Accuracy, which all embody routing reliability, routing efficiency and routing security. Comparative studies of On-demand Routing protocols, such as those presented in [29], highlight the limitations of traditional approaches under adversarial conditions.

A. Analysis of Packet Delivery Ratio Analysis

Packet Delivery Ratio (PDR) is one of the basic measures of routing resilience in the Adversarial Manet Framework. Conventional AODV, as illustrated in Fig. 4, offers a PDR of only 62 per cent in the presence of black hole attacks, mainly because of the lack of security validation while discovering the routes which allows the malicious nodes to entice and then drop the data packets. By not being persistently misbehaved, T-AODV raises PDR to 74 per cent, but it is not responsive to dynamic and selective attack behavior due to relying on Static Trust Thresholds. Ml-AODV then improves PDR to 80 per cent with the advantage of Learned Traffic Patterns whereas RL-AODV has 82 per cent with Adaptive Routing Decisions even without explicit trust enforcement. The proposed QR-MLQ-PTS-AODV has much better performance than all the baseline protocols, with the PDR of 94%. It is explained by the fact that by combining Multi-metric Trust Modeling, a Malicious Node Prediction based on ML, Quantum-aware Trust Modulation, and Q-learning facilitated MAC-layer Priority Scheduling, the combination of these features helps to reduce packet losses arising due to malicious and unstable forwarding.

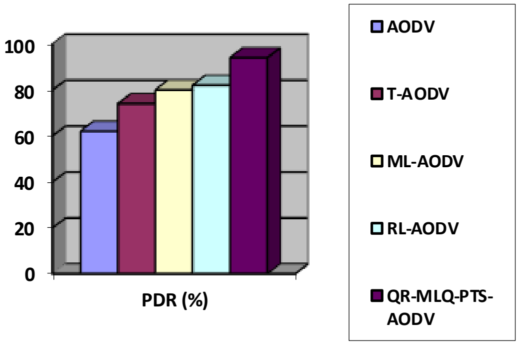

Fig. 4. PDR comparison.

B. End-to-End Delay Analysis

End to End Delay indicates the efficiency and timeliness of data delivery. AODV has the largest average delay (205 ms), as can be seen in Fig. 5, with high rate of route failures and retransmissions as caused by malicious route disruption. T-AODV minimizes the delay down to 155 ms, but fails again because of slow convergence with an on-off attack and gray hole attacks. Ml-AODV and RL-AODV further minimize delay to 135 ms and 130 ms, respectively; but these schemes mainly focus on routing decisions and cause no attention to contention at the MAC layer. Conversely, QR-MLQ-PTS-AODV has the lowest mean delay (78 Ms), which is because:

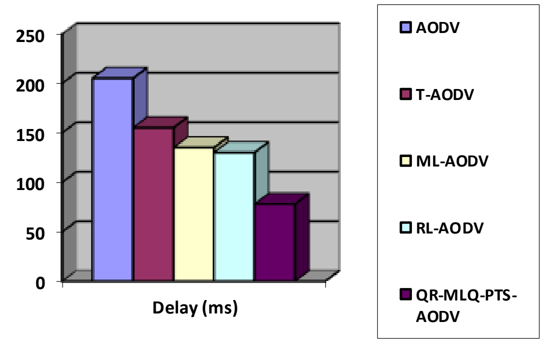

Fig. 5. End to end delay comparison.

- Semi-temporal and entropy-based trust evaluation of picking stable routes, Early malicious node isolation with the help of ML prediction, and
- Priority time-slot allocation based on Q-learning that greatly decreases the channel contention and queuing delays.

These findings show that cross-layer optimization is a crucial tool in the creation of low-latency secure MANET routing.

C. Routing Overhead Analysis

Routing overhead is used to measure the effectiveness of control packet dissemination. AODV has the greatest overhead (1100 control packets), as Fig. 6 demonstrates, because of recurrent route discoveries as a result of malicious interference. T-AODV will decrease overhead to 820 packets, whereas ML-AODV and RL-AODV will decrease overhead to 740 and 710 packets, respectively. These methods are, however, very reactive and are not effective in avoiding repetitive route disruptions. The QR-MLQ-PTS-AODV proposed has the least routing overhead (490 packets), which is due to:

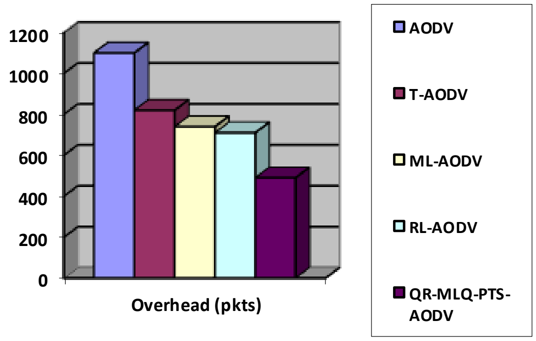

Fig. 6. Routing overhead comparison.

- Active offensive isolation of the malicious nodes,
- Temporal convergence of trusts in routing tables that stabilize routing tables, and
- MAC-layer prioritization through Q- learning which minimizes the number of packet collisions and path recoveries.

This supports the fact that the security conscious MAC-layer scheduling makes a significant contribution to the routing efficiency which is not much considered in the current work.

D. Throughput Analysis

Throughput is the rate of delivering the data effectively, which is one of the main measurements of the overall efficiency of MANET. As in Fig. 7, AODV is the least performer in terms of throughput because of the devastating losses of packets as a result of black holes nodes. T-AODV is enhanced by eliminating the long-lasting malicious nodes, but the presence of changing trust values in gray hole attacks reduces the sustained performance. ML-AODV and RL-AODV exhibit mid-way throughput gains, although contend at the MAC-layer. The proposed QR-MLQ-PTS-AODV has the best throughput, which is made possible by:

- High-trust node preferential channel access,
- Less retransmission due to early isolation of attackers, and
- Slot scheduling based on adaptive Q-learning.

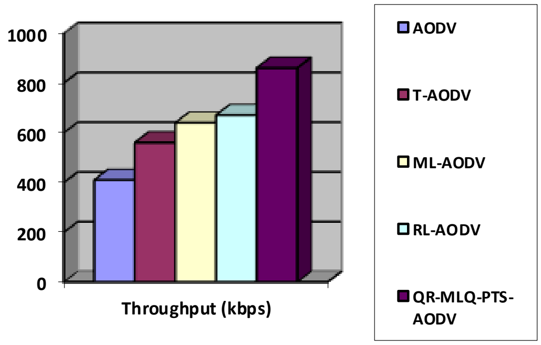

Fig. 7. Throughput comparison.

E. Accuracy of Malicious Node Detection Accuracy

AODV and RL-AODV do not have an explicit detection mechanism, and thus have 0% detection accuracy as shown in Fig. 8. T-AODV is 75 percent accurate yet not very good at selective forwarding and on-off attacks. ML-AODV has an 88 per cent detection rate, but without time reinforcement. The results of the proposed QR-MLQ-PTS-AODV have the highest detection accuracy (97) because: Multi metric trust test,

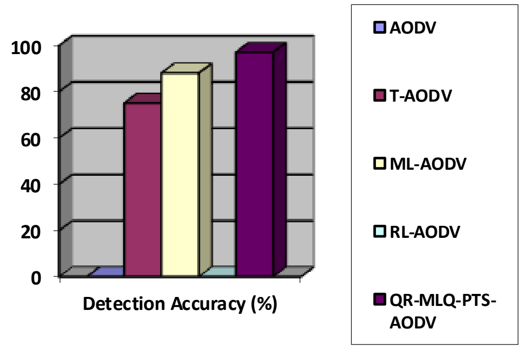

Fig. 8. Malicious node comparison.

- Entropy stability behavioral analysis,
- Temporal trust decay, and
- ML-assisted, quantum-aware detection, effectively combined with routing choices.

F. Conclusion of Findings and Knowledge

These findings indicate that Trust-based, ML-based, or RL-based strategies alone are not helpful in ensuring MANET routing security against advanced attackers. The proposed QR-MLQ-PTS-AODV has better performance, which is achieved through the synergistic combination of:

- Interpretable and insider attack-resilient trust modeling,
- Machine learning to detect and detect accurately and early, and
- Adaptive cross-layer optimization of MAC-layer priorities by use of Q-learning.

The quantitative comparison of the proposed protocol with existing approaches in terms of PDR, delay, routing overhead, throughput, and detection accuracy is summarized in Table I.

TABLE I. COMPARISON OF DIFFERENT ROUTING PROTOCOLS

Protocol	PDR (%)	Delay (ms)	Overhead (pkts)	Throughput (kbps)	Detection Accuracy (%)
AODV	62	205	1100	410	0
T-AODV	74	155	820	560	75
ML-AODV	80	135	740	640	88
RL-AODV	82	130	710	670	0
QR-MLQ-PTS-AODV	94	78	490	860	97

V. CONCLUSION

In this study, it was suggested that quantum-resilient machine learning and Q-Learning-based Priority Time-Slot AODV (QR-MLQ-PTS-AODV) should be used to secure MANET routing. The framework combats classical and emerging quantum routing threats by employing multi-metric and entropy-based trust assessment, malicious node prediction by supervised machine learning, MAC-layer prioritization by reinforcement learning, and post-quantum behavioral security. Extensive simulations of NS-3 ensure that major advances in the delivery ratio of packets, latency, routing overheads, throughput, and detection accuracy are obtained in comparison to the current protocols based on trust-, ML-, and RL.

The findings confirm the fact that smart cross-layer routing and quantum-conscious trust modeling provides a scalable and future-proof approach to secure MANETs. The next phase of work is to investigate deep reinforcement learning, federated trust learning and distributed post-quantum trust sharing.

REFERENCES

[1] C. E. Perkins, E. M. Belding-Royer, and S. Das, "Ad hoc on-demand distance vector (AODV) routing," RFC 3561, IETF, 2003.

[2] Hongmei Deng, W. Li, and D. P. Agrawal, “Routing security in wireless ad hoc networks,” IEEE Communications Magazine, vol. 40, no. 10, pp. 70–75, Oct. 2002.

[3] S. Marti, T. J. Giuli, K. Lai, and M. Baker, “Mitigating routing misbehavior in mobile ad hoc networks,” in Proc. ACM MobiCom, 2000, pp. 255–265.

[4] K. Sanzgiri, B. Dahill, B. N. Levine, C. Shields, and E. M. Belding-Royer, “A secure routing protocol for ad hoc networks,” in Proc. IEEE ICNP, 2002, pp. 78–87.

[5] J. Sen, S. Koilakonda, and A. Ukil, “A Mechanism for Detection of Cooperative Black Hole Attack in Mobile Ad Hoc Networks,” in Proc. Second International Conference on Intelligent Systems, Modeling and Simulation (ISMS),2011, pp. 338–343.

[6]]Yan Lindsay Sun,Wei Yu, Zhu Han, K.J. Ray Liu, “Information theoretic framework of trust modeling and evaluation for ad hoc networks,” IEEE Journal on Selected Areas in Communications, vol. 24, no. 2, pp. 305–317, Feb. 2006.

[7] S. Buchegger and J.-Y. Le Boudec, “Performance analysis of the CONFIDANT protocol,” in Proc. ACM MobiHoc, 2002, pp. 226–236.

[8] L. Buttyán and J.-P. Hubaux, “Stimulating cooperation in self-organizing mobile ad hoc networks,” Mobile Networks and Applications, vol. 8, no. 5, pp. 579–592, 2003.

[9] Z. Liu, A. Joy, and R. A. Thompson, “A dynamic trust-based routing protocol for MANETs,” in Proc. IEEE MILCOM, 2004, pp. 1–7.

[10] V. Balakrishnan, V. Varadharajan, and U. K. Tupakula, “Trust Management in Mobile Ad Hoc Networks,” in Guide to Wireless Ad Hoc Networks, M. Ilyas and I. Mahgoub, Eds. Springer, 2009, pp. 473–502.

[11] J. Zhang and R. Cohen, “A trust-based approach for secure routing in ad hoc networks,” Ad Hoc Networks, vol. 5, no. 2, pp. 105–117, 2007.

[12] S. Misra, S. K. Dhurandher, and A. Gupta, “Using trust for detection of routing misbehavior in MANETs,” IEEE Communications Letters, vol. 14, no. 8, pp. 749–751, Aug. 2010.

[13] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.

[14] C. Wang, Y. Yang, and X. Wang, “A reinforcement learning based routing protocol for MANETs,” in Proc. IEEE WCNC, 2012, pp. 1–6.

[15] N. K. Chaubey and R. Kumar, “Q-learning based secure routing against black hole attack in MANET,” in Proc. IEEE ANTS, 2016, pp. 1–6.

[16] A. Patcha and J.-M. Park, “An overview of anomaly detection techniques,” IEEE Communications Surveys & Tutorials, vol. 9, no. 4, pp. 4–17, 2007.

[17] S. Shafi, M. R. Ahmed, and S. A. Khan, “Machine learning and trust-based secure AODV routing in mobile ad hoc networks,” Procedia Computer Science, vol. 218, pp. 1254–1263, 2023.

[18] R. Vatambeti, S. K. Sahoo, and P. K. Pattnaik, “Black hole attack detection in MANET using machine learning techniques,” Computers & Electrical Engineering, vol. 112, Art. no. 108985, 2024.

[19] M. Sheela, R. Saravanakumar, and S. S. Kumar, “Deep learning–based intrusion detection system for MANET security,” Intelligent and Converged Networks, vol. 5, no. 1, pp. 45–58, 2024.

[20] I. Z. Ibrahim and M. F. Ghanim, “AI-based black hole attack detection framework for mobile ad hoc networks,” Information Dynamics and Applications, vol. 3, no. 1, pp. 1–12, 2024.

[21] A. Abbas, M. Younis, and U. Qureshi, “Machine learning based trust management in MANETs,” IEEE Access, vol. 7, pp. 149321–149334, 2019.

[22] S. Kumar and R. Mishra, “Secure and efficient routing in MANET using trust and reinforcement learning,” in Proc. IEEE Int. Conf. Computing, Communication and Automation, 2020, pp. 1–6.

[23] M. Conti, E. Gregori, and G. Maselli, “A survey on MANET security,” IEEE Communications Surveys & Tutorials, vol. 15, no. 3, pp. 1231–1256, 2013.

[24] S. Ramaswamy, H. Fu, M. Sreekantaradhya, J. Dixon, and K. Nygard, “Prevention of cooperative black hole attack in wireless ad hoc networks,” in Proc. IEEE Int. Conf. Wireless Networks, 2003, pp. 570–575.

[25] P. Mohapatra and S. Krishnamurthy, Ad Hoc Networks: Technologies and Protocols. New York, NY, USA: Springer, 2005.

[26] T. Alpcan and T. Başar, “A game theoretic analysis of intrusion detection in access control systems,” in Proc. IEEE CDC, 2003, pp. 1568–1573.

[27] A. M. Shabut et al., “Intrusion detection systems for mobile ad hoc networks using machine learning,” IEEE Access, vol. 6, pp. 60508–60523, 2018.

[28] T. Clausen and P. Jacquet, “Optimized link state routing protocol (OLSR),” RFC 3626, IETF, 2003.

[29] S. R. Das, C. E. Perkins, and E. M. Royer, “Performance comparison of two on-demand routing protocols for ad hoc networks,” in Proc. IEEE INFOCOM, 2000, pp. 3–12.

[30] N. Bindel et al., “Post-quantum cryptography for future wireless networks,” IEEE Communications Surveys & Tutorials, vol. 24, no. 4, pp. 2472–2504, 2022.

[31] M. Mosca and M. Piani, “Quantum threat timelines and implications for network security,” IEEE Security & Privacy, vol. 20, no. 6, pp. 72–80, 2022.

[32] A. Banerjee et al., “Quantum-resilient trust management in decentralized networks,” IEEE Access, vol. 11, pp. 118321–118336, 2023.

[33] S. Gupta and R. Buyya, “Machine learning–driven anomaly detection for post-quantum networks,” IEEE Transactions on Network Science and Engineering, vol. 11, no. 1, pp. 88–101, 2024.

[34] Y. Alshamrani et al., “Reinforcement learning–based secure routing under advanced adversarial models,” IEEE Internet of Things Journal, vol. 12, no. 3, pp. 2104–2118, 2025.