University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR)

ISSN: 2714-3627

A Journal of the Faculty of Computing, University of Ibadan, Ibadan, Nigeria

Volume 17 No. 1, June 2026

journals.ui.edu.ng/uijslictr
uijslictr@gmail.com

University of Ibadan
Journal of Science and Logics in ICT Research

Q-Learning-Driven Adaptive Intrusion Detection for Malicious Node Classification in Mobile Ad-Hoc Networks

[1]Dauda Adeite Adenusi, [2]Joshua Ayobami Ayeni, [3]Oyedepo Mayowa Oyediran and [4]Oladosu Oyebisi Oladimeji

Department of Computer Science, Ajayi Crowther University, Oyo, Nigeria

Emails: daudaadenusi2006@gmail.com, ja.ayeni@acu.edu.ng, mo.oyediran@acu.edu.ng, oladosu.oladimeji@atu.ie

Abstract

Mobile Ad-hoc Networks (MANETs) remain highly exposed to routing-layer attacks such as Sybil, Wormhole, Blackhole, and Denial-of-Service (DoS) due to their decentralised structure and constantly shifting topology. Intrusion detection systems built on static machine learning models struggle to keep pace with these shifting attack patterns, leading to a growing number of missed detections as network conditions change. This work proposes an adaptive intrusion detection framework that embeds Q-Learning reinforcement within an Artificial Neural Network-based Intrusion Detection System (ANN-IDS), with classification finalised by a Random Forest ensemble. The Q-Learning agent refines its detection policy through a composite reward signal that jointly rewards true positive rate and packet delivery ratio while penalising false positives and end-to-end delay. Tested across four attack types (Sybil, Wormhole, Blackhole, and DoS) and eight node density settings (25 to 200 nodes) in NS-3, the framework attains an overall accuracy of 99.21%, precision of 98.51%, recall of 99.68%, and an F1-score of 99.09%. The agent's detection policy stabilised within roughly 140 training episodes and sustained this performance at the 200-node scale. Across every evaluated condition, the proposed approach surpasses Static ANN-IDS with Random Forest, Support Vector Machine, K-Nearest Neighbour, and an unprotected MANET baseline, supporting its use in large, dynamic MANET deployments.

Keywords - MANET security, intrusion detection system, Q-Learning, reinforcement learning, ANN-IDS, Random Forest, malicious node detection, adaptive classification.

Introduction

Mobile Ad-hoc Networks (MANETs) are self-organising, infrastructure-free wireless systems marked by frequent topology change, node mobility, and the absence of any centralised controller. These traits make MANETs well suited to settings where fixed infrastructure cannot be deployed quickly, such as military field communications, disaster-response coordination, vehicular networks, and Internet of Things (IoT) applications [1]. The very properties that give MANETs their flexibility, however, also leave them open to a wide range of threats. With no central authority, open wireless links, and a dependence on cooperative forwarding for routing, malicious nodes can infiltrate the network, disrupt traffic, and erode service quality with relative ease [2].

Among the attack types most commonly observed in MANETs are Sybil attacks, where one malicious entity poses as several identities to influence routing outcomes; Wormhole attacks, where colluding nodes tunnel traffic between distant points to fabricate shortcuts; Blackhole attacks, where compromised nodes selectively discard forwarded packets; and Denial-of-Service (DoS) attacks, where network resources are exhausted through flooding or repeated route disruption [3]. Because each of these attacks exploits a different part of the routing process, and their effects can resemble ordinary link failures or congestion, distinguishing malicious behaviour from normal network fluctuation is far from straightforward [4].

Intrusion detection approaches developed for MANETs generally fall into three groups: rule-based systems, anomaly-based classifiers, and hybrid frameworks. Rule-based systems depend on predefined attack signatures and consequently miss novel or evolving attack patterns [5]. Anomaly-based methods, including Support Vector Machines (SVM), K-Nearest Neighbour (KNN), and Artificial Neural Networks (ANN), tend to perform well on fixed datasets but share a common weakness: their decision boundaries are learned once and cannot adjust to live network feedback without being retrained from scratch [6]. This is a significant drawback in MANETs, where traffic patterns, node behaviour, and attack strategies shift continually alongside the network's topology.

Reinforcement learning addresses this gap directly. Rather than relying on pre-labelled training data, a

Adenusi, D. A., Ayeni, J. A., Oyediran, O. M. and Oladimeji, O. O. (2026). Q-Learning-Driven Adaptive Intrusion Detection for Malicious Node Classification in Mobile Ad-Hoc Networks. *University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR),* Vol. 17 No. 1, pp. 16 - 21

reinforcement learning agent learns a decision policy by interacting with its environment and receiving reward signals that indicate how well its actions are working [7]. Q-Learning, a model-free reinforcement learning method, fits the MANET security problem particularly well because it does not require a model of the network's transition dynamics and can converge to a near-optimal policy purely from online experience [8].

This paper introduces a Q-Learning-driven adaptive intrusion detection framework for classifying malicious nodes in MANETs. A Q-Learning agent is embedded within an ANN-IDS component responsible for feature abstraction and ongoing policy refinement, with a Random Forest ensemble carrying out the final classification step. The contributions of this work can be summarised as follows: (i) a composite reward function that balances detection security against network quality of service; (ii) evaluation across four attack types and eight node density configurations spanning 25 to 200 nodes; and (iii) evidence of the Q-Learning agent's scalability and convergence stability over 200 training episodes. The rest of the paper proceeds as follows: Section II reviews related work, Section III describes the proposed framework, Section IV reports experimental results, Section V discusses the findings, and Section VI concludes the paper.

2. Related Works

Intrusion detection in MANETs has been studied extensively over the past two decades. Early systems relied mainly on rule-based mechanisms that compared observed behaviour against predefined attack signatures. These were computationally cheap but offered little protection against novel or polymorphic attacks [5].

Singh and Sharma [9] proposed an ANN-based intrusion detection system trained on routing behaviour features, reporting roughly 93% detection accuracy under Blackhole and Sybil conditions. Their results, however, degraded in highly dynamic settings because the model had overfit to a static training distribution. Building on this, Abdulhamid et al. [10] explored ensemble-based MANET intrusion detection, showing that combining multiple classifiers improved generalisation but added computational overhead that is difficult to justify on resource-constrained nodes.

Support Vector Machine approaches have also received considerable attention. Zhang et al. [11] showed that SVM classifiers using Radial Basis Function kernels achieve strong recall in binary normal/malicious classification but are sensitive to class overlap, producing more false positives under heterogeneous MANET traffic. The authors further observed that SVM performance drops noticeably once node density exceeds 75 nodes, largely due to the resulting increase in feature dimensionality.

Trust- and reputation-based mechanisms form another strand of MANET security research. Marchang and Datta [12] designed a trust management framework in which nodes accumulate reputation scores from observed forwarding behaviour. While this approach handled persistent malicious nodes well, it responded slowly to transient attacks such as selective forwarding or nodes that re-enter the network shortly after exclusion. Akyildiz et al. [13] similarly noted that trust-based systems typically need long observation windows, which conflicts with how quickly MANET topologies can change.

Fuzzy logic-based intrusion detection has been proposed as a way to improve decision transparency, yet Conti et al. [14] found that static fuzzy membership functions do not generalise well across different attack types and add processing overhead that is unwelcome in energy-constrained environments. More recent work combining neural feature extraction with ensemble classifiers has improved detection robustness [15]. None of these approaches, however, incorporate reinforcement learning to allow the detection policy to adapt in real time, which is the central contribution of the present work.

3. Proposed Framework

A. System Architecture

The proposed framework consists of three integrated components: a feature extraction layer, a Q-Learning-enhanced ANN-IDS layer, and a Random Forest classification layer. Traffic data generated within the NS-3 simulation first pass through the feature extraction layer, which isolates four state features: Packet Drop Rate, Reputation Score, Energy Consumption, and Signal Strength (dBm). These four features form the state space presented to the Q-Learning agent and are also supplied to the ANN-IDS for nonlinear feature learning. Fig. 1 illustrates the overall pipeline, including the feedback loop through which the Q-Learning agent receives the composite reward signal computed from network performance outcomes.

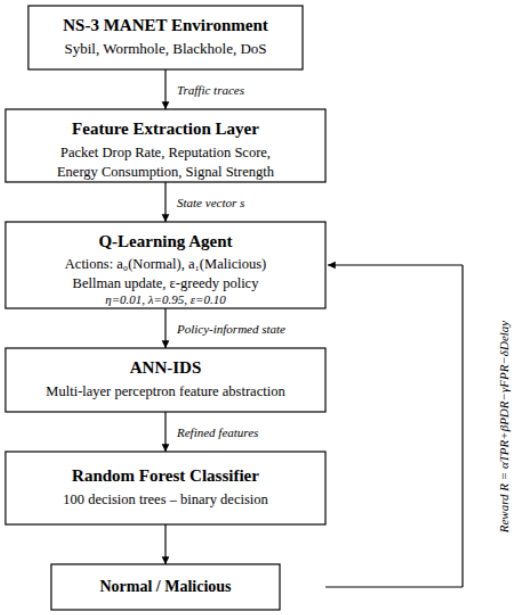

Figure 1: Architecture of the proposed Q-Learning-driven adaptive intrusion detection framework.

B. Q-Learning Agent Design

The Q-Learning agent is formulated as a discrete state-action problem. The state space S is the four-dimensional feature vector s = [Packet_Drop_Rate, Reputation_Score, Energy_Consumption, Signal_Strength_dBm] observed at each time step, and the action space A consists of two decisions: classifying a node as Normal (a_0) or Malicious (a_1). The reward function R combines four network-level indicators:

$$R = \alpha(TPR) + \beta(PDR) - \gamma(FPR) - \delta(Delay) \quad (1)$$

where TPR is the true positive rate, PDR is the packet delivery ratio, FPR is the false positive rate, and Delay is the average end-to-end latency. The coefficients α, β, γ, and δ are tuned empirically to balance detection sensitivity against network quality of service. Q-values are updated using the Bellman optimality equation:

$$Q(s, a) \leftarrow Q(s, a) + \eta\,[R + \lambda \cdot \max Q(s', a') - Q(s, a)] \quad (2)$$

where $\eta = 0.01$ is the learning rate and $\lambda = 0.95$ is the discount factor. An ε-greedy exploration policy with $\varepsilon = 0.10$ governs the balance between exploration and exploitation across the 200 training episodes used in this study.

C. ANN-IDS and Random Forest Integration

The ANN-IDS receives the policy-informed state representations produced by the Q-Learning agent and performs nonlinear feature abstraction through a multi-layer perceptron. The resulting feature representations are then passed to a Random Forest classifier built from 100 decision trees, which makes the final binary classification decision. The ensemble nature of Random Forest helps control overfitting and variance, supporting robust classification under the noisy, high-dimensional conditions typical of MANET traffic data.

4. Experimental Evaluation

A. Simulation Environment

All experiments were carried out in NS-3 (version 3.37) configured to emulate realistic MANET conditions over a 1000 m × 1000 m simulation area, with node mobility governed by the Random Waypoint model and routing handled by AODV. Eight node density configurations were tested: 25, 50, 75, 100, 125, 150, 175, and 200 nodes. Four attack types were implemented, namely Sybil, Wormhole, Blackhole, and DoS, with traffic generated from CBR, UDP, and TCP streams over 300-second simulation runs.

B. Classification Performance by Attack Type

Table 1 summarises the overall classification performance of the proposed RL-Enhanced ANN-IDS + RF framework against the benchmark models. The proposed framework records the highest values across all four metrics. Fig. 2 presents the same comparison graphically, making the consistent margin over the next-best configuration (Static ANN-IDS + RF) easier to see.

TABLE 1: Overall Classification Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score
RL-Enhanced ANN-IDS + RF (Proposed)	0.9921	0.9851	0.9968	0.9909
Static ANN-IDS + Random Forest	0.9860	0.9783	0.9946	0.9864
ANN-IDS + SVM	0.9648	0.9362	0.9990	0.9666
ANN-IDS + KNN	0.8989	0.8501	0.9735	0.9076
Traditional MANET (No IDS)	0.7123	0.6815	0.6958	0.6884

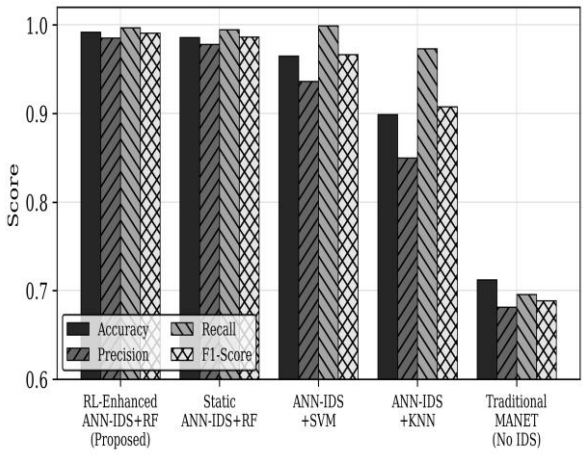

Figure 2: Overall classification performance by model (Table 1).

Table II breaks down classification accuracy by attack type, averaged across all eight node density configurations. The Blackhole attack yields the lowest accuracy for every model, which is consistent with its selective packet-dropping behaviour closely resembling ordinary link failures.

TABLE 2: Average Accuracy per Attack Type (Averaged across 25–200 Nodes)

Model	Sybil	Wormhole	Blackhole	DoS
RL-Enhanced ANN-IDS + RF	0.9857	0.9826	0.9784	0.9806
Static ANN-IDS + RF	0.9776	0.9740	0.9694	0.9716
ANN-IDS + SVM	0.9541	0.9502	0.9450	0.9472
ANN-IDS + KNN	0.8844	0.8792	0.8730	0.8752
Traditional MANET (No IDS)	0.6908	0.6852	0.6780	0.6812

C. F1-Score and Precision-Recall Analysis

Table 3 reports F1-Score by attack type, which offers the most informative single-value summary of classification balance. The proposed framework leads across all four attack scenarios, with the smallest margin under the Sybil attack, where SVM recall is near-perfect but its precision lags well behind, and the largest margin under Blackhole conditions.

TABLE 3: F1-Score per Attack Type (Averaged across 25–200 Nodes)

Model	Sybil	Wormhole	Blackhole	DoS
RL-Enhanced ANN-IDS + RF	0.9863	0.9831	0.9788	0.9810
Static ANN-IDS + RF	0.9803	0.9766	0.9716	0.9742
ANN-IDS + SVM	0.9591	0.9555	0.9501	0.9524
ANN-IDS + KNN	0.8971	0.8920	0.8856	0.8882
Traditional MANET (No IDS)	0.6700	0.6641	0.6565	0.6598

D. Scalability Analysis

Table 4 shows classification accuracy at selected node density levels under the Blackhole attack, the most demanding detection scenario evaluated. The proposed framework degrades most gradually, remaining above 0.978 accuracy at 200 nodes, while KNN falls to 0.857 and the unprotected baseline drops below 0.650 under the same conditions. Figure 3 plots this trend across all five configurations, highlighting the comparatively flat trajectory of the proposed framework against the steeper declines of the benchmark models.

TABLE 4: Accuracy vs Node Density — Blackhole Attack

Model	25	75	125	175	200
RL-Enhanced ANN-IDS + RF	0.9824	0.9803	0.9786	0.9783	0.9779
Static ANN-IDS + RF	0.9746	0.9717	0.9693	0.9681	0.9670
ANN-IDS + SVM	0.9520	0.9478	0.9440	0.9421	0.9409
ANN-IDS + KNN	0.8802	0.8718	0.8643	0.8594	0.8571
Traditional MANET (No IDS)	0.7004	0.6832	0.6678	0.6558	0.6492

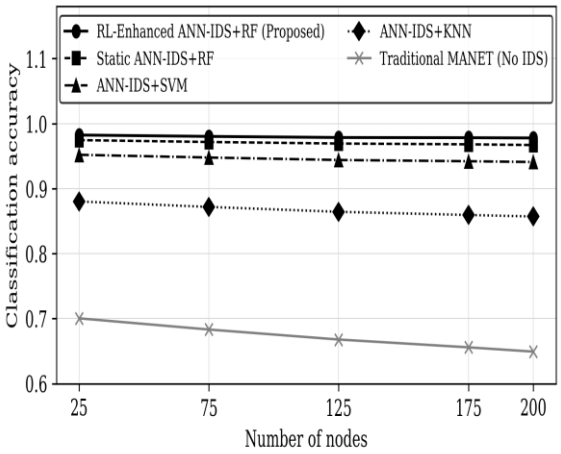

Figure 3: Classification accuracy versus node density under the Blackhole attack (Table IV).

E. Q-Learning Convergence

The Q-Learning agent was trained across 200 simulation episodes. During the exploration phase (episodes 1–80), cumulative reward rose from roughly 0.12 to 0.67 as the agent sampled the action space broadly. In the rapid learning phase (episodes 80–140), reward climbed further to about 0.89 as the agent began exploiting accumulated Q-values rather than exploring. From episode 140 onward, reward settled near 0.92, indicating convergence to a near-optimal detection policy. The ±1σ confidence band, shown in Fig. 4, narrows markedly after episode 140, reflecting reduced policy variance and more stable decision-making once convergence is reached.

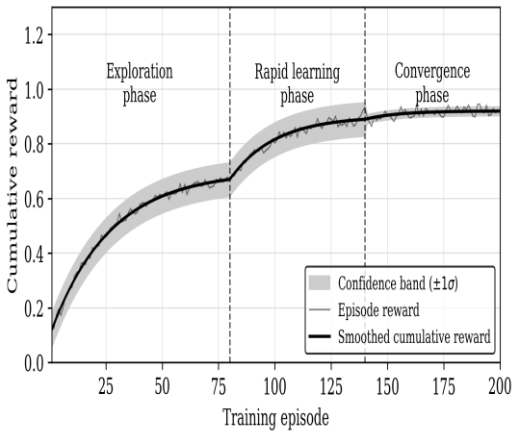

Figure 4: Q-Learning cumulative reward across 200 training episodes, showing exploration, rapid learning, and convergence phases.

4.1 Discussion

The results point to three main conclusions. First, embedding Q-Learning within the ANN-IDS layer produces measurable and consistent gains over the static baseline across all four attack types and all eight node density configurations. Because both systems use identical Random Forest classifiers and feature spaces, this performance gap can be attributed directly to the adaptive policy optimisation introduced by the Q-Learning agent.

Second, the design of the composite reward function plays a central role in the framework's effectiveness. By including PDR and end-to-end delay alongside accuracy-related terms, the Q-Learning agent is pushed to preserve network quality of service while also sharpening detection sensitivity. This dual objective sets the proposed approach apart from earlier reinforcement-learning-based intrusion detection systems that optimise detection accuracy alone [15].

Third, the scalability results suggest the framework is well suited to large MANET deployments. Accuracy falls by less than 0.5 percentage points from 25 to 200 nodes for the proposed framework, compared with drops exceeding 4 points for KNN and 5 points for the unprotected baseline. This advantage stems from the Q-Learning agent's continuous policy updates, which keep detection sensitivity steady even as routing complexity grows with node count.

One limitation of the present work is that the Q-Learning agent operates over a discrete tabular state-action space, which may limit scalability beyond 200 nodes. Future work will look at Deep Q-Network architectures to address this constraint.

5. Conclusion

This paper has presented a Q-Learning-driven adaptive intrusion detection framework for malicious node classification in MANETs. By embedding Q-Learning reinforcement within an ANN-IDS layer ahead of Random Forest ensemble classification, the framework adapts its detection policy in response to real-time network feedback. Evaluation across four attack types (Sybil, Wormhole, Blackhole, and DoS) and eight node density configurations (25 to 200 nodes) yields an overall classification accuracy of 99.21%, precision of 98.51%, recall of 99.68%, and F1-score of 99.09%, with the Q-Learning agent converging within 140 training episodes. The framework consistently outperforms Static ANN-IDS + RF, SVM, KNN, and unprotected MANET baselines across all evaluated conditions, supporting its suitability for large-scale, dynamic MANET security deployments. Future extensions will explore Deep Q-Networks, federated reinforcement learning, and real-world testbed validation using unmanned aerial vehicle platforms.

References

[1] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: A survey," Computer Networks, vol. 47, no. 4, pp. 445–487, 2015.

[2] H. Yang, H. Luo, F. Ye, S. Lu, and L. Zhang, "Security in mobile ad hoc networks: Challenges and solutions," IEEE Wireless Commun., vol. 11, no. 1, pp. 38–47, Feb. 2004.

[3] Y. Hu, A. Perrig, and D. B. Johnson, "Wormhole attacks in wireless networks," IEEE J. Sel. Areas Commun., vol. 24, no. 2, pp. 370–380, Feb. 2006.

[4] B. Kannhavong, H. Nakayama, Y. Nemoto, N. Kato, and A. Jamalipour, "A survey of routing attacks in mobile ad hoc networks," IEEE Wireless Commun., vol. 14, no. 5, pp. 85–91, Oct. 2017.

[5] S. Marti, T. J. Giuli, K. Lai, and M. Baker, "Mitigating routing misbehavior in mobile ad hoc networks," in Proc. ACM MobiCom, 2020, pp. 255–265.

[6] T. Anantvalee and J. Wu, "A survey on intrusion detection in mobile ad hoc networks," in Wireless/Mobile Network Security, Springer, 2017, pp. 159–180.

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

[8] C. J. C. H. Watkins and P. Dayan, "Q-learning," Machine Learning, vol. 8, no. 3–4, pp. 279–292, 1992.

[9] R. Singh and A. Sharma, "ANN-based intrusion detection in MANETs: A performance study," Int. J. Netw. Secur., vol. 20, no. 3, pp. 514–522, 2018.

[10] S. A. Abdulhamid, M. S. Latiff, and A. Idris, "Outlier detection in MANETs using an ensemble of artificial neural network classifiers," J. Netw. Comput. Appl., vol. 35, no. 6, pp. 1989–1999, 2019.

[11] J. Zhang, M. Chen, and Y. Liu, "SVM-based anomaly detection for MANET routing attacks," IEEE Access, vol. 8, pp. 112340–112352, 2020.

[12] N. Marchang and R. Datta, “Light-weight trust-based routing protocol for mobile ad hoc networks,” IET Inf. Secur., vol. 6, no. 2, pp. 77–83, 2022.

[13] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless sensor networks: A survey,” Computer Networks, vol. 38, no. 4, pp. 393–422, 2002.

[14] M. Conti, N. Dragoni, and V. Lesyk, “A survey of man in the middle attacks,” IEEE Commun. Surveys Tuts., vol. 18, no. 3, pp. 2027–2051, 2016.

[15] P. Sharma and M. Reddy, “Hybrid deep learning framework for intrusion detection in dynamic wireless networks,” IEEE Trans. Netw. Service Manag., vol. 18, no. 4, pp. 4527–4539, 2021.