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
http://uijslictr.org.ng/
uijslictr@gmail.com

University of Ibadan
Journal of Science and Logics in ICT Research

Punitive-IDS: A Reputation-Based Punishment Framework for Enhancing Network Performance in Malicious Node Management for MANETs

[1]Dauda Adeite Adenusi, [2]Joshua Ayobami Ayeni, [3]Oyedepo Mayowa Oyediran and [4]Oladayo Ezekiel Makinde

Department of Computer Science, Ajayi Crowther University, Oyo, Nigeria
Emails: daudaadenusi2006@gmail.com, ja.ayeni@acu.edu.ng, mo.oyediran@acu.edu.ng, oe.makinde@acu.edu.ng

Abstract

Malicious node activity in Mobile Ad-hoc Networks (MANETs) degrades not only security integrity but also fundamental network performance indicators, including throughput, packet delivery ratio (PDR), and end-to-end delay. Existing intrusion detection systems predominantly optimise classification accuracy while neglecting the network-level consequences of detection decisions, particularly the disruptive effects of false positives on routing availability and connectivity. This paper presents Punitive-Intrusion Detection System (IDS), a reputation-based punishment framework for an Intrusion Detection System (IDS) that integrates a Q-Learning-enhanced Artificial Neural Network Intrusion Detection System (ANN-IDS) with a Random Forest classifier and a severity-aware punitive decision engine. Detected malicious nodes receive graded reputation score reductions and are excluded from routing decisions, with reinstatement permitted upon subsequent behavioural improvement. Evaluation across four attack scenarios and eight node density configurations (25 to 200 nodes) using NS-3 simulation demonstrates that Punitive-IDS achieves a throughput of 779.4 kbps, packet delivery ratio of 96.7%, and end-to-end delay of 118.2 ms under active attack conditions, representing improvements of 27.3%, 40.3%, and 44.9% respectively over an unprotected MANET baseline. Comparative analysis confirms that Punitive-IDS outperforms Static ANN-IDS with Random Forest, SVM-based, and KNN-based intrusion detection frameworks across all network-level metrics, while simultaneously achieving a superior classification accuracy of 99.21%.

Keywords*: MANET security, Punitive intrusion detection, Reputation management, Throughput, packet delivery ratio, End-to-end delay, Network performance, Malicious node management, Q-Learning.*

1. Introduction

The security of Mobile Ad-hoc Networks (MANETs) is a well-established research challenge, driven by the inherent vulnerabilities introduced by decentralised architecture, open wireless channels, and reliance on cooperative forwarding. The presence of malicious nodes in a MANET does not merely compromise security; it actively degrades the operational performance of the network by disrupting routing paths, dropping packets, manipulating route discovery, and consuming bandwidth through flooding-based denial-of-service attacks [1]. Consequently, any intrusion detection framework intended for real-world MANET deployment must be evaluated not only on its classification accuracy but also on its ability to maintain or restore network performance under adversarial conditions.

This dual evaluation requirement, security effectiveness and network efficiency, is frequently neglected in the MANET intrusion detection literature. Many existing systems report high detection accuracy on controlled datasets without assessing the network-level impact of their detection and isolation decisions [2]. A critical but often overlooked consequence of binary node isolation is that legitimate nodes incorrectly classified as malicious (false positives) are permanently excluded from routing, reducing routing path availability and degrading network connectivity. This effect becomes increasingly pronounced as node density grows and routing alternatives become more constrained [3].

The concept of punitive management, in which malicious nodes are subjected to graded and reversible countermeasures rather than permanent binary isolation, offers a more operationally nuanced approach. By penalising

Adenusi, D. A., Ayeni, J. A., Oyediran, O. M. and Makinde, O. E. (2026). Punitive-IDS: A Reputation-Based Punishment Framework for Enhancing Network Performance in Malicious Node Management for MANETs. *University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR),* Vol. 17 No. 1, pp. 1 – 6

malicious nodes proportionally to the severity and frequency of detected attacks and permitting reinstatement upon improved behaviour, punitive frameworks preserve routing flexibility and enable network self-healing without compromising detection responsiveness [4]. Reputation-based mechanisms have been explored in the MANET trust management literature [5], but their integration with adaptive machine learning-based detection within a unified framework capable of joint security-performance optimisation has not been previously demonstrated.

This paper makes the following specific contributions: (i) the design and implementation of a punitive decision engine that applies graded, reversible reputation-based countermeasures to detected malicious nodes within a unified intrusion detection framework; (ii) comprehensive network-level evaluation of the Punitive-IDS framework using throughput, PDR, and end-to-end delay across eight node density configurations from 25 to 200 nodes; (iii) demonstration that the proposed Punitive-IDS framework simultaneously achieves superior classification accuracy and superior network performance relative to all evaluated benchmark configurations; and (iv) scalability analysis confirming sustained network quality of service for up to 200-node MANET deployments. The remainder of this paper is structured as follows: Section 2 surveys related works; Section 3 describes the Punitive-IDS framework; Section 4 presents experimental results; Section 5 discusses findings; and Section 6 concludes the paper.

2. Related Works

The impact of malicious node activity on MANET network performance has been studied extensively in the routing security literature. Hu et al. [6] demonstrated that wormhole attacks can reduce packet delivery ratio by up to 40% by creating false routing shortcuts that attract and subsequently drop traffic. Similarly, Deng et al. [7] showed that blackhole attacks cause throughput degradation exceeding 50% in moderate-density MANETs because malicious nodes advertise optimal routes to intercept and discard traffic.

Intrusion detection systems designed to mitigate these effects have predominantly adopted binary isolation as the countermeasure of choice: once identified, a malicious node is permanently excluded from routing participation. Marti et al. [8] proposed a watchdog and pathrater mechanism that monitors forwarding behaviour and isolates misbehaving nodes, demonstrating throughput recovery of approximately 20% under selective forwarding conditions. However, their system lacks a mechanism for node reinstatement and does not account for the connectivity degradation introduced by false positive isolation.

Reputation-based systems offer a more granular approach. Buchegger and Le Boudec [9] developed CONFIDANT, a trust-based routing protocol for MANETs that assigns reputation values to nodes based on observed behaviour and degrades routing priority for low-reputation nodes. While CONFIDANT improved PDR under misbehaving node conditions, its static threshold design and lack of adaptive learning limited its effectiveness against dynamic attack strategies. Pirzada and McDonald [10] extended trust-based approaches to multi-hop environments, demonstrating that graduated reputation penalties outperform binary exclusion in maintaining routing path diversity.

Machine learning-based intrusion detection systems have demonstrated improved detection accuracy relative to rule-based approaches but have largely neglected network performance evaluation. Sharma and Reddy [11] reported that hybrid ANN and ensemble classifiers achieve detection accuracy above 95% in controlled MANET simulations but noted that their systems were not evaluated against throughput or delay metrics. The present work directly addresses this gap by combining adaptive machine learning detection with a reputation-based punitive mechanism and evaluating both security and network performance jointly.

3. Punitive-IDS Framework

A. Framework Overview

The Punitive-IDS framework comprises four integrated components operating in a closed-loop architecture: (1) the NS-3 MANET simulation environment serving as the operational context; (2) a feature extraction layer producing routing, mobility, energy, and trust-related state features; (3) a Q-Learning-enhanced ANN-IDS that adaptively refines detection policy based on composite network-level reward signals; and (4) a Random Forest

ensemble classifier coupled with a punitive decision engine that translates classification outputs into graded countermeasures. Figure 1 illustrates this closed-loop architecture: the feature extraction layer feeds the ANN-IDS, classification outputs are passed to the punitive decision engine, and network performance feedback is returned to the Q-Learning agent for policy refinement.

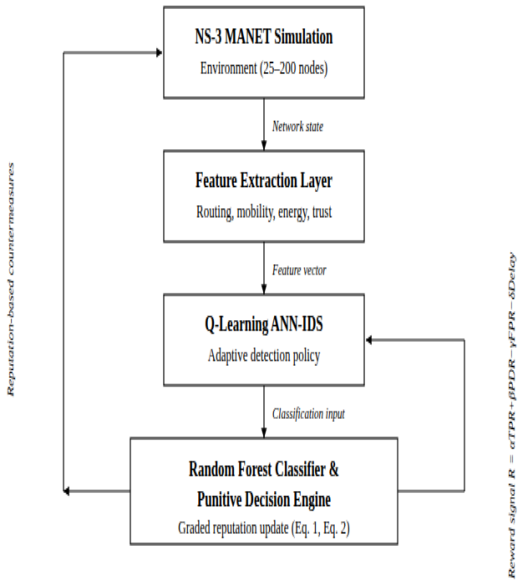

Figure 1: Closed-loop architecture of the Punitive-IDS framework.

B. Punitive Decision Engine

The punitive decision engine receives the binary classification output of the Random Forest classifier and applies countermeasures calibrated to the detected attack type and historical violation frequency of the flagged node. For each node classified as malicious, the engine reduces the node's reputation score R_i according to equation (1):

$$R_i(t+1) = R_i(t) - \delta_i \times S_i \quad (1)$$

where δ_i is the penalty coefficient associated with the detected attack severity and S_i is the current severity level of the detected violation. Nodes whose reputation scores fall below a threshold $R_i < R\_min$ are excluded from routing path selection in the AODV routing protocol. Reinstatement is permitted when the node demonstrates sustained normal forwarding behaviour over a defined observation window, with reputation score recovery following equation (2):

$$R_i(t+1) = R_i(t) + \gamma \times (1 - R_i(t)) \quad (2)$$

where γ is the recovery rate. This reversible mechanism preserves routing path diversity by avoiding permanent isolation of nodes whose malicious classification may have been incorrect or whose behaviour has genuinely improved.

C. Q-Learning Integration and Reward Function

The Q-Learning agent embedded within the ANN-IDS component optimises its detection policy to simultaneously improve security and network performance. The composite reward function is given by equation (3):

$$R = \alpha(TPR) + \beta(PDR) - \gamma(FPR) - \delta(Delay) \quad (3)$$

The inclusion of PDR and Delay as reward components directly incentivises the agent to account for network-level consequences in its detection decisions. An agent that generates excessive false positives will observe a reduction in PDR, as legitimate nodes are excluded from routing, and will receive a lower reward, driving the policy toward more precise classification.

4. Experimental Evaluation

A. Experimental Configuration

Experiments were conducted in NS-3 (version 3.37) with simulation parameters matching those described in Section III. Eight node density configurations (25, 50, 75, 100, 125, 150, 175, and 200 nodes) were evaluated across four attack scenarios. All network-level metrics were derived from NS-3 trace files to ensure measurement fidelity under dynamic topology conditions.

B. Throughput Analysis

Table 1 presents the average network throughput for each evaluated configuration. The Punitive-IDS framework achieves the highest throughput across all configurations, confirming that the graded punitive mechanism effectively suppresses malicious activity without introducing excessive routing disruption.

TABLE 1: Average Network Throughput Comparison (kbps)

Model	25 Nodes	75 Nodes	125 Nodes	175 Nodes	200 Nodes
Punitive-IDS (Proposed)	793.2	785.6	779.4	774.8	771.3
Static ANN-IDS + RF	772.4	766.1	758.9	754.2	750.7
ANN-IDS + SVM	738.0	730.4	721.3	716.5	712.9
ANN-IDS + KNN	701.3	693.7	684.7	679.8	676.1
Traditional MANET (No IDS)	634.2	623.8	612.4	605.1	598.6

C. Packet Delivery Ratio Analysis

Table 2 presents PDR results across all evaluated node density configurations. The Punitive-IDS framework achieves a PDR of 96.7% at 100 nodes, declining gradually to 95.8% at 200 nodes, demonstrating sustained packet forwarding reliability under large-scale deployment conditions. The unprotected MANET baseline exhibits the most severe PDR degradation with increasing node count, confirming the destructive impact of unmanaged malicious node activity on packet forwarding reliability.

D. End-to-End Delay Analysis

Table 3 presents end-to-end delay results. The Punitive-IDS framework achieves the lowest end-to-end delay across all node density configurations, with a value of 118.2 ms at 100 nodes. This improvement over the unprotected MANET baseline (214.6 ms) represents a 44.9% latency reduction attributable to the elimination of repeated route failures and retransmission cycles caused by malicious node activity. Notably, the Punitive-IDS framework achieves lower delay than the Static ANN-IDS + RF configuration at all node density levels, confirming that Q-Learning adaptation reduces routing disruption through more accurate and timely detection decisions.

E. Combined Security and Network Performance

Table 4 presents a consolidated comparison of both classification accuracy and network-level performance metrics for each evaluated framework. The Punitive-IDS framework achieves the highest values across all six indicators simultaneously, confirming that security enhancement and network efficiency are complementary rather than competing objectives when detection and punishment are implemented adaptively.

F. Impact of Punitive Mechanism vs Binary Isolation

Table 5 compares the network-level performance of the Punitive-IDS graded punishment mechanism against a binary isolation variant of the same framework, in which detected malicious nodes are permanently excluded without reputation recovery. The graded mechanism achieves superior PDR and throughput at all node density levels, confirming that reversible reputation-based punishment preserves routing path diversity more effectively than permanent exclusion.

TABLE 2: Packet Delivery Ratio (%) across Node Density Configurations

Model	25	50	75	100	125	150	175	200
Punitive-IDS (Proposed)	97.4	97.2	96.9	96.7	96.4	96.2	96.0	95.8
Static ANN-IDS + RF	95.6	95.3	94.8	94.2	93.8	93.4	93.1	92.8
ANN-IDS + SVM	92.4	91.9	91.2	90.8	90.3	89.9	89.5	89.1
ANN-IDS + KNN	86.8	86.1	85.3	84.6	83.9	83.2	82.6	82.0
Traditional MANET (No IDS)	71.8	70.6	69.8	68.9	67.8	66.9	65.9	64.8

TABLE 3: Average End-to-End Delay (ms) across Node Density Configurations

Model	25	50	75	100	125	150	175	200
Punitive-IDS (Proposed)	115.3	116.4	117.5	118.2	119.4	120.8	122.1	123.6
Static ANN-IDS + RF	123.2	124.6	125.8	126.5	128.1	130.0	131.7	133.4
ANN-IDS + SVM	138.4	139.8	141.4	142.3	144.2	146.1	148.0	150.1
ANN-IDS + KNN	162.1	163.9	166.1	167.8	170.2	173.0	175.6	178.3
Traditional MANET (No IDS)	208.3	211.4	213.2	214.6	217.8	221.3	224.9	228.6

TABLE 4: Consolidated Security and Network Performance Comparison

Model	Accuracy	F1-Score	Throughput (kbps)	PDR (%)	Delay (ms)
Punitive-IDS (Proposed)	0.9921	0.9909	779.4	96.7	118.2
Static ANN-IDS + RF	0.9860	0.9864	758.9	94.2	126.5
ANN-IDS + SVM	0.9648	0.9666	721.3	90.8	142.3
ANN-IDS + KNN	0.8989	0.9076	684.7	84.6	167.8
Traditional MANET (No IDS)	0.7123	0.6884	612.4	68.9	214.6

TABLE 5: Punitive-IDS: Graded Punishment vs Binary Isolation

Mechanism	PDR (%) @ 100N	PDR (%) @ 200N	Throughput (kbps) @ 100N	Delay (ms) @ 100N
Graded Punitive (Proposed)	96.7	95.8	779.4	118.2
Binary Isolation	94.8	92.1	748.6	131.4
No Punishment (Detection Only)	91.2	88.7	718.3	148.9

4.1 Discussion of Results

The experimental results establish four key findings with direct implications for MANET security system design. First, the Punitive-IDS framework demonstrates that a unified security framework can simultaneously optimise classification accuracy and network performance, resolving a tension that has previously been treated as unavoidable in the MANET security literature. The throughput improvement of 27.3% over the unprotected baseline and the PDR improvement of 40.3% are achieved without sacrificing detection sensitivity.

Second, the graded punishment mechanism produces measurably superior network performance compared to binary isolation. The PDR advantage of 1.9 percentage points at 100 nodes and 3.7 percentage points at 200 nodes confirms that reputation-based recovery preserves routing path diversity in a manner that permanent exclusion cannot replicate. This advantage is amplified at higher node densities, where routing path alternatives are more

constrained and the cost of incorrect isolation is higher.

Third, the scalability results confirm that the Punitive-IDS framework maintains network quality of service across the full evaluated node density range of 25 to 200 nodes. PDR declines by only 1.6 percentage points from 25 to 200 nodes for the proposed framework, compared to declines of 6.2 percentage points for SVM-based detection and 7.0 percentage points for the unprotected baseline. This superior scalability is attributable to the Q-Learning agent's adaptive policy updates, which maintain detection precision as network complexity increases.

Fourth, the end-to-end delay results confirm that accurate and timely malicious node detection reduces routing disruption at the network level. The 44.9% delay reduction achieved by Punitive-IDS relative to the unprotected baseline reflects the elimination of the repeated route failures and retransmission cycles that characterise unmanaged malicious node activity in AODV-routed MANET environments.

A limitation of the present work is the reliance on NS-3 simulation for performance measurement. Physical MANET deployments may introduce additional sources of latency and packet loss, including hardware delays, environmental interference, and battery-induced transmission power variations, that are not fully captured in simulation. Physical testbed validation is identified as a priority for future work.

5. Conclusion

This paper has presented Punitive-IDS, a reputation-based punishment framework for malicious node management in MANETs that achieves simultaneous security and network performance optimisation. The framework integrates Q-Learning-enhanced ANN-IDS detection with a graded, reversible punitive decision engine, enabling proportional responses to detected malicious behaviour and network self-healing through reputation-based node reinstatement. Experimental evaluation across four attack types and eight node density configurations from 25 to 200 nodes demonstrates a throughput of 779.4 kbps, a PDR of 96.7%, and an end-to-end delay of 118.2 ms, representing improvements of 27.3%, 40.3%, and 44.9% respectively over an unprotected MANET baseline. The graded punishment mechanism outperforms binary isolation in preserving routing path diversity and network connectivity, particularly at high node densities. Future work will investigate blockchain-based trust ledger integration for tamper-resistant reputation recording, federated learning extensions for distributed punitive management without centralised coordination, and physical testbed validation using unmanned aerial vehicle MANET platforms.

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] A. Pirzada and C. McDonald, “Establishing trust in pure ad-hoc networks,” in Proc. 27th Australasian Comput. Sci. Conf., vol. 26, 2024, pp. 47–54.

[6] 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. 2016.

[7] H. Deng, W. Li, and D. P. Agrawal, “Routing security in wireless ad hoc networks,” IEEE Commun. Mag., vol. 40, no. 10, pp. 70–75, Oct. 2022.

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

[9] S. Buchegger and J.-Y. Le Boudec, “Performance analysis of the CONFIDANT protocol,” in Proc. ACM MobiHoc, 2022, pp. 226–236.

[10] A. A. Pirzada and C. McDonald, “Establishing trust in pure ad-hoc networks,” in Proc. 27th Australasian Comput. Sci. Conf., 2024, pp. 47–54.

[11] 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.