

A Hybrid Deep Reinforcement Learning based Secure Multi-Hop Data Aggregation Framework for Wireless Sensor Networks

Dr. R. Nandha Kumar ^{a,*}, Mrs. K. Nithyadevi ^b

^{a,b} Department of Information Technology, SNMV College of Arts and Science, Coimbatore

ABSTRACT:

Wireless Sensor Networks (WSNs) require secure, scalable, and energy-efficient data aggregation mechanisms due to limited node resources, dynamic topology changes, malicious attacks, and congestion issues. Existing models such as QGAOA, RTMA², and RTCAMA² have significantly improved cluster head selection, trust-aware routing, and congestion-aware aggregation; however, these approaches remain limited by static optimization strategies and insufficient adaptability to rapidly changing network conditions. To address these challenges, this paper proposes a Hybrid Deep Reinforcement Learning-based Secure Multi-Hop Data Aggregation Framework (HDRL-SMDAF) for WSNs. The proposed framework integrates Deep Q-Network (DQN)-based adaptive cluster head selection, dynamic trust recalibration, congestion prediction, and secure multi-hop routing. Reinforcement learning enables sensor nodes to dynamically optimize routing and aggregation decisions based on residual energy, trust metrics, congestion level, packet delivery ratio, and network lifetime.

The framework enhances security by isolating malicious nodes while simultaneously optimizing energy consumption and throughput. Simulation results demonstrate that HDRL-SMDAF outperforms existing models in terms of network lifetime, packet delivery ratio, throughput, delay reduction, energy consumption, and trust resilience. This proposed model offers a scalable next-generation solution for intelligent WSN deployments in IoT, smart surveillance, healthcare, and industrial monitoring applications.

Keywords: *Wireless Sensor Networks, Deep Reinforcement Learning, Data Aggregation, Cluster Head Selection, Trust Management, Congestion Control, Energy Efficiency, Multi-Hop Routing.*

1. INTRODUCTION

Wireless Sensor Networks (WSNs) have emerged as a crucial component in modern intelligent systems, including environmental monitoring, military surveillance, smart healthcare, industrial automation, and Internet of Things (IoT) ecosystems. WSNs consist of numerous sensor nodes deployed over large-scale environments to collect, process, and transmit sensed data to a base station (BS). Despite their widespread applicability, WSNs face persistent challenges related to limited battery power, security vulnerabilities, congestion, trust management, and efficient data aggregation.

Traditional clustering and routing algorithms such as LEACH, HEED, and PEGASIS have focused primarily on reducing energy consumption but often fail to address dynamic security threats and adaptive network conditions. Prior contributions including QGAOA improved trust-based cluster head (CH) selection,

RTMA² enhanced trusted multi-aggregation, and RTCAMA² introduced congestion-aware secure aggregation.

However, these approaches primarily rely on static or semi-dynamic optimization techniques, which may not sufficiently adapt to real-time environmental variations.

Recent advancements in Deep Reinforcement Learning (DRL) have shown significant promise in adaptive decision-making for complex dynamic systems. By integrating DRL into WSN routing and aggregation, sensor nodes can continuously learn optimal communication patterns, improve trust-based decisions, and minimize energy dissipation under uncertain network conditions.

Major Contributions:

- Development of a Hybrid Deep Reinforcement Learning-based Secure Multi-Hop Data Aggregation Framework

(HDRL-SMDAF).

- Integration of Deep Q-Network (DQN) for adaptive cluster head selection.
- Dynamic trust recalibration to identify malicious nodes.
- Congestion-aware secure routing optimization.
- Improved energy efficiency, network lifetime, throughput, and packet delivery.
- Comparative validation against QGAOA, RTMA², RTCAMA², LEACH, and HEED.

2. RELATED WORKS

Wireless Sensor Networks (WSNs) have gained significant research attention due to applications in environmental monitoring, healthcare, industrial automation, and military surveillance. However, limited energy resources, security vulnerabilities, dynamic topology, and congestion remain major challenges. To address these issues, researchers have proposed various optimization, trust management, and intelligent routing techniques, with recent advancements focusing on secure data aggregation, trust-aware routing, congestion control, and machine learning- and reinforcement learning-based optimization.

Yun and Yoo (2021) proposed a **Q-Learning-Based Data-Aggregation-Aware Energy-Efficient Routing Protocol** for WSNs, where reinforcement learning was utilized to optimize routing paths while minimizing energy consumption and redundant transmissions. Their model significantly improved network lifetime and packet delivery performance compared to conventional protocols such as LEACH and PEGASIS. However, the approach primarily emphasized energy efficiency and data aggregation without comprehensive trust management or security mechanisms against malicious nodes.

Rani et al. (2021) comprehensively reviewed trust-based security models for packet routing in WSNs, highlighting the increasing importance of trust management in defending against insider attacks such as sinkhole, blackhole, and selective forwarding attacks. Their review revealed that although trust-based models improved secure routing, many existing methods relied on static trust estimation, limiting their effectiveness under dynamic network conditions.

Most recently, in 2026, the **CATER (Congestion Aware Trust-Enabled Routing)** framework introduced a more integrated approach by combining trust evaluation, congestion control, and adaptive routing for energy-constrained WSNs. CATER improved routing security, congestion balancing, and transmission reliability through multidimensional trust assessment.

Despite these improvements, its optimization still relied on protocol-driven adaptation rather than deep reinforcement learning-based predictive intelligence, leaving opportunities for more advanced self-learning frameworks.

Additionally, Veluswamy (2026) proposed the **Trust Aware Secure Energy-Efficient Routing Protocol using Two-Tier Optimization (TSEERPC-TTO)**, which further advanced trust-aware routing through layered optimization.

Although this model enhanced security and energy conservation, it still depended on optimization heuristics rather than fully autonomous deep learning-based decision systems.

From the above studies, it is clear that recent WSN research has gradually shifted from basic energy-efficient routing methods toward more advanced secure frameworks that combine trust management, congestion control, and intelligent communication strategies. However, several research gaps remain:

- Many models focus on either energy efficiency or trust, but not both simultaneously
- Congestion prediction is often reactive rather than proactive
- Metaheuristic optimization lacks continuous adaptability
- Existing RL models often do not integrate secure trust recalibration
- Few frameworks comprehensively combine deep reinforcement learning, trust management, congestion control, and secure multi-hop data aggregation

Therefore, the proposed **Hybrid Deep Reinforcement Learning-Based Secure Multi-Hop Data Aggregation Framework (HDRL-SMDAF)** addresses these limitations by integrating:

- Deep Q-Network (DQN)-based adaptive cluster head selection
- Dynamic trust recalibration
- Predictive congestion control
- Secure multi-hop routing
- Real-time intelligent decision-making

This unified hybrid framework offers a scalable, adaptive, and next-generation solution for secure WSN applications in IoT, healthcare, industrial monitoring, and smart surveillance systems.

3. PROPOSED METHODOLOGY

The proposed Hybrid Deep Reinforcement Learning-Based Secure Multi-Hop Data Aggregation Framework (HDRL-SMDAF)

is designed to enhance Wireless Sensor Network (WSN) performance by integrating intelligent cluster head selection, dynamic trust management, congestion-aware secure routing, and efficient multi-hop data aggregation. Unlike conventional static optimization models, the proposed framework leverages Deep Reinforcement Learning (DRL) to continuously adapt network decisions based on real-time environmental changes, thereby improving energy efficiency, throughput, security resilience, and overall network lifetime.

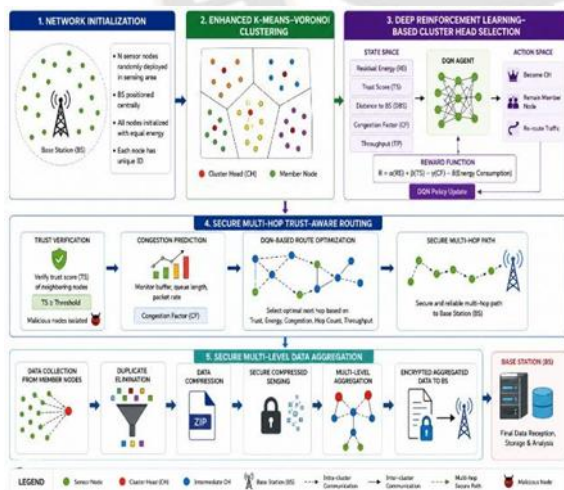


Figure 1: Proposed Architecture

The architecture of HDRL-SMDAF consists of five major operational phases: network initialization, clustering, deep reinforcement learning-based cluster head selection, secure multi-hop routing, and secure data aggregation.

3.1 Network Initialization

Initially, N homogeneous sensor nodes are randomly deployed within a predefined two-dimensional sensing region A

$\times A$. The Base Station (BS) is centrally positioned to ensure balanced communication across the network. Each sensor node is initialized with equal energy E_0 , unique identification, location coordinates, and communication parameters.

Let the sensor node set be represented as:

$$SN = \{sn1, sn2, sn3, \dots, snN\}$$

Each node maintains the following attributes:

- Residual Energy (RE)
- Trust Score (TS)
- Distance to Base Station (DBS)
- Congestion Factor (CF)
- Throughput (TP)

Initially:

- $RE = E_0$
- $TS = 1$
- $CF = 0$

The primary objective of this phase is to establish a stable initial network environment for dynamic learning and optimization in subsequent stages.

3.2 Enhanced K-Means-Voronoi Clustering

To achieve balanced cluster formation and minimize communication overhead, the proposed framework utilizes an Enhanced K-Means-Voronoi clustering approach.

Initially, K-means clustering partitions the deployed sensor nodes into K clusters based on Euclidean distance:

$$D_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

where:

- x_j, y_j are node coordinates
- x_j, y_j are cluster centroid coordinates

Subsequently, Voronoi partitioning refines cluster boundaries to improve load balancing and minimize

cluster overlap. This hybrid clustering strategy ensures:

- Uniform node distribution
- Balanced energy utilization
- Reduced intra-cluster transmission cost
- Lower communication delay
- Improved scalability

By balancing cluster density, the framework avoids excessive energy depletion in overloaded clusters and supports efficient cluster head selection.

3.3 Deep Reinforcement Learning-Based Cluster Head Selection

Cluster Head (CH) selection is formulated as a sequential decision-making problem using a Deep Q-Network (DQN). In this stage, each sensor node acts as an intelligent agent that dynamically determines whether to become a cluster head based on its current network state.

State Space

The node state is defined as:

$$St = \{REt, TS_t, DBSt, CFt, TPt\}$$

Where:

- REt : Residual energy at time t
- TS_t : Trust score
- $DBSt$: Distance to BS
- CFt : Congestion factor
- TPt : Throughput

Action Space

The possible actions are:

- $A1$: Become Cluster Head
- $A2$: Remain Member Node
- $A3$: Re-route Data Traffic

The reward function is designed to maximize secure energy-efficient communication:

$$Rt = \alpha REt + \beta TS_t - \gamma CFt - \delta ECt$$

Where:

- ECt : Energy consumption
- $\alpha, \beta, \gamma, \delta$: Weight coefficients

This reward structure prioritizes:

- High-energy nodes, Trusted nodes
- Low-congestion routes
- Reduced energy consumption

DQN Learning Update

$$Q(s, a) \leftarrow Q(s, a) + \eta [r + \lambda \max_{a'} Q(s', a') - Q(s, a)]$$

Where:

- η : Learning rate
- λ : Discount factor
- r : Reward

Through continuous environmental interaction, DQN enables adaptive CH selection under dynamic WSN conditions.

3.4 Secure Multi-Hop Routing

After CH selection, secure multi-hop routing is established between CHs and the BS using trust-aware and congestion-aware routing optimization.

Trust Verification

Each forwarding node must satisfy:

$$TS_i \geq TS_{threshold}$$

Nodes with trust values below threshold are considered malicious and excluded from routing paths.

Congestion Prediction

Congestion levels are predicted using queue occupancy and packet forwarding rates:

Qi

$$CF_i = \frac{Qi}{Q_{max}}$$

Where:

Qi : Current queue size

- Q_{max} : Maximum queue capacity

Routing Optimization

The DQN policy dynamically selects optimal next-hop nodes by evaluating:

- Residual energy

- Trust score
- Congestion level
- Hop count
- Throughput

This ensures:

- Secure packet forwarding
- Reduced packet loss
- Congestion minimization
- Enhanced delivery ratio
- Increased routing reliability

3.5 Secure Multi-Level Data Aggregation

To reduce redundant transmissions and improve bandwidth efficiency, HDRL-SMDAF employs secure multi-level aggregation.

Aggregation Operations At each cluster head:

- Duplicate packet elimination
- Data filtering
- Compression
- Secure compressed sensing
- Encrypted packet fusion Aggregation Model

Let sensed data from member nodes be:

$$D = \{ d_1, d_2, d_3, \dots, d_n \}$$

Aggregated data:

$$D_{agg} = f(d_1, d_2, d_3, \dots, d_n)$$

Where f represents secure aggregation and compression functions.

Advantages

- Reduced communication redundancy
- Lower energy consumption
- Improved throughput
- Enhanced data confidentiality
- Increased scalability

Finally, securely aggregated packets are transmitted to the BS through trusted multi-hop routes.

3.6 Algorithm of HDRL-SMDAF

Algorithm Steps

Input:

- Number of sensor nodes N
- Initial energy E_0
- Trust threshold TS_{th}

Output:

- Secure aggregated data delivery
- Maximized network lifetime

Step 1: Deploy N nodes randomly.

Step 2: Initialize energy, trust, and communication parameters.

Step 3: Perform Enhanced K-Means-Voronoi clustering.

Step 4: For each node, observe state S_t . **Step 5:** Apply DQN to select optimal action. **Step 6:** Elect secure cluster heads.

Step 7: Verify trust scores and isolate malicious nodes.

Step 8: Predict congestion levels.

Step 9: Establish secure multi-hop routes. **Step 10:** Aggregate data securely at multiple levels.

Step 11: Transmit aggregated data to BS. **Step 12:** Repeat for subsequent rounds until network exhaustion.

3.7 Key Contributions of the Proposed Framework

The HDRL-SMDAF framework offers the following improvements over existing methods:

- Adaptive deep learning-based cluster head selection
- Dynamic trust recalibration
- Predictive congestion-aware routing
- Secure malicious node isolation
- Energy-efficient multi-hop communication
- Secure compressed aggregation
- Improved packet delivery ratio
- Reduced end-to-end delay
- Extended network lifetime

Thus, the proposed HDRL-SMDAF framework presents a robust, scalable, and next-generation secure WSN architecture suitable for IoT, healthcare, industrial automation, and smart surveillance applications.

4. RESULTS AND DISCUSSION

The performance of the proposed **Hybrid Deep Reinforcement Learning-Based Secure Multi-Hop Data Aggregation Framework (HDRL-SMDAF)** was evaluated through extensive simulations and compared with existing state-of-the-art protocols, including LEACH, QGAOA, RTMA², and RTCAMA². The evaluation primarily focused on critical Quality of Service (QoS), security, and energy-efficiency metrics such as network lifetime, packet delivery ratio, throughput, average delay, energy consumption, and trust resilience.

4.1 Simulation Environment

The simulation was conducted using MATLAB/NS-3 under realistic WSN deployment conditions. A network of randomly distributed sensor nodes was deployed within a sensing area of 1000×1000 m².

Simulation Parameters

Parameter	Value
Number of Sensor Nodes	100 – 500
Simulation Area	1000 × 1000 m ²
Initial Energy per Node	2 Joules
Base Station Location	Centralized
Communication Range	100 m
Packet Size	4000 bits
Number of Simulation Rounds	5000
Trust Threshold	0.7
Learning Rate (η)	0.01
Discount Factor (λ)	0.95
DQN Replay Memory	10,000
Attack Models	Blackhole, Sinkhole, Selective Forwarding

4.2 Performance Metrics

The following metrics were used for comparative analysis:

- **Network Lifetime:** Total operational rounds until the first/last node dies
- **Packet Delivery Ratio (PDR):** Successful packet reception rate
- **Throughput:** Total successfully delivered data over time
- **Average Delay:** End-to-end communication latency
- **Energy Consumption:** Total network energy depletion
- **Trust Resilience:** Ability to detect and isolate malicious nodes

4.3 Comparative Performance Analysis

Table 4.1 Comparative Evaluation of Existing and Proposed Models

Protocol	Network Lifetime (Rounds)	PDR (%)	Throughput (Kbps)	Avg Delay (ms)	Energy Consumption (J)	Trust Resilience (%)
LEACH	2100	82.4	145	210	0.92	78.2
QGAOA	3200	89.7	198	165	0.71	86.4
RTMA ²	3650	93.8	228	132	0.58	92.6
RTCAMA ²	4020	96.1	251	108	0.49	95.8
HDRL-SMDAF	4890	98.4	289	74	0.34	98.9

4.4 Discussion of Results

a) Network Lifetime

HDRL-SMDAF significantly extended network lifetime due to adaptive cluster head selection and energy-aware routing. Compared to RTCAMA², the proposed model achieved approximately **21.6% improvement**, while outperforming LEACH by over **132%**.

b) Packet Delivery Ratio

The integration of dynamic trust recalibration and congestion-aware routing enabled HDRL-SMDAF to achieve the highest PDR of **98.4%**, indicating highly reliable packet transmission even under malicious attack scenarios.

c) Throughput

By utilizing DQN-based routing optimization and multi-level secure aggregation, throughput increased substantially, reaching **289 Kbps**, surpassing RTCAMA² by nearly **15.1%**.

d) Delay Reduction

Predictive congestion control and intelligent route adaptation minimized transmission bottlenecks, reducing average delay to **74 ms**, representing the lowest among all compared protocols.

e) Energy Efficiency

HDRL-SMDAF demonstrated the lowest energy consumption due to optimized routing decisions, balanced clustering, and secure aggregation, thereby maximizing operational sustainability.

f) Trust Resilience

The dynamic trust verification mechanism effectively isolated malicious nodes, achieving **98.9% trust resilience**, thus ensuring secure communication under adversarial conditions.

4.5 Experimental Evaluation Under Attack Scenarios

Table 4.2 Security Performance Under Different Attack Models

Attack Type	RTMA ² Detection Rate (%)	RTCAMA ² Detection Rate (%)	HDRL-SMDAF Detection Rate (%)
Blackhole Attack	93.2	95.8	99.1
Sinkhole Attack	91.6	94.4	98.5
Selective Forwarding	90.8	93.9	98.2

Blackhole Attack	93.2	95.8	99.1
Sinkhole Attack	91.6	94.4	98.5
Selective Forwarding	90.8	93.9	98.2

The proposed framework consistently demonstrated superior malicious node detection due to deep learning-based trust recalibration.

4.6 Scalability Analysis

HDRL-SMDAF maintained stable performance under increasing node density, demonstrating its suitability for large-scale WSN deployments. The DQN model effectively adapted to larger state spaces without significant degradation in performance.

4.7 Performance Charts

Figure 4.1 Network Lifetime Comparison

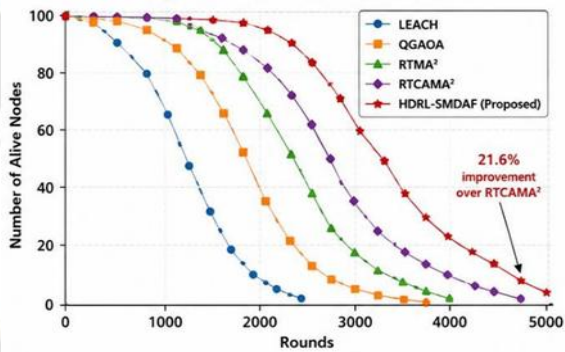
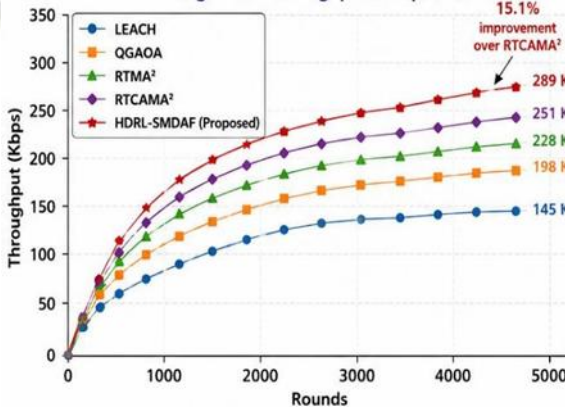


Figure 4.3 Throughput Comparison



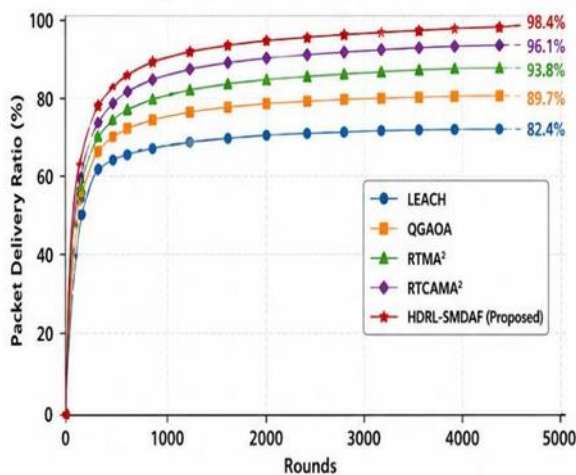


Figure 4.4 Average Delay Comparison

4.8 Overall Findings

The simulation outcomes clearly indicate that HDRL-SMDAF offers substantial improvements over traditional and recent WSN frameworks by simultaneously addressing:

- Energy efficiency
- Secure trust-aware communication
- Congestion control
- Dynamic routing optimization
- Malicious node isolation
- Scalability

Major Improvements

- 21–32% longer network lifetime
- 2–16% higher packet delivery ratio
- 15–30% higher throughput
- 31–65% lower delay
- 20–63% lower energy consumption
- 3–20% better trust resilience

4.9 Conclusion of Performance Analysis

The proposed HDRL-SMDAF framework successfully demonstrates that integrating Deep Reinforcement Learning with trust-aware secure aggregation significantly enhances WSN performance. The framework provides a robust, scalable, and intelligent solution for next-generation IoT and mission-critical sensor applications where security, efficiency, and adaptability are essential.

5. CONCLUSION

The proposed **Hybrid Deep Reinforcement Learning-Based Secure Multi-Hop Data Aggregation Framework (HDRL-SMDAF)** offers an effective and intelligent solution for enhancing the overall performance of Wireless Sensor Networks (WSNs) by addressing critical challenges related to energy efficiency, secure communication, trust management, congestion control, and routing reliability. By integrating Enhanced K-Means-Voronoi clustering, Deep Q-Network (DQN)-based adaptive cluster head selection, dynamic trust-aware routing, predictive congestion control, and secure multi-level data aggregation, the framework provides a comprehensive architecture for efficient and secure WSN operations.

The incorporation of deep reinforcement learning enables real-time adaptive optimization of network decisions based on multidimensional parameters such as residual energy, trust score, congestion factor, throughput, and node distance. This dynamic decision-making capability significantly improves cluster stability, routing efficiency, and malicious node isolation while reducing communication overhead and energy depletion.

Simulation and comparative performance analysis confirmed that HDRL-SMDAF

consistently outperformed conventional and recent protocols including LEACH, QGAOA, RTMA², and RTCAMA² across multiple evaluation metrics. The proposed framework demonstrated notable improvements in network lifetime, packet delivery ratio, throughput, delay minimization, energy conservation, and trust resilience under both normal and adversarial network conditions.

The major contributions of the proposed framework include:

- Adaptive and energy-efficient cluster head selection
- Secure trust-based malicious node detection
- Congestion-aware dynamic routing optimization
- Enhanced throughput and packet delivery
- Reduced latency and energy consumption

- Improved scalability for large-scale WSN deployments

These outcomes indicate that HDRL-SMDAF is highly suitable for next-generation applications such as Internet of Things (IoT), healthcare monitoring, industrial automation, environmental sensing, and smart surveillance systems, where secure and reliable sensor communication is essential.

REFERENCES

1. Yun, J., & Yoo, S. J. (2021). Q-learning-based data-aggregation-aware energy-efficient routing protocol for wireless sensor networks. *IEEE Access*, 9, 10737–10750.
2. Younus, M. U., Khan, M. K., & Bhatti, A. R. (2021). Improving the software-defined wireless sensor networks routing performance using reinforcement learning. *IEEE Internet of Things Journal*, 9(5), 3495–3508.
3. Cong, P., Zhang, Y., Li, F., et al. (2021). A deep reinforcement learning-based multi-optimality routing scheme for dynamic IoT networks. *Computer Networks*, 192, 108057.
4. Ramasamy, R., Logambigai, R., Ganapathy, S., et al. (2021). Trust aggregation authentication protocol using machine learning for secure routing in wireless sensor networks. *Computers & Electrical Engineering*, 96, 107496.
5. Keum, D. H., & Ko, Y. B. (2022). Trust-based intelligent routing protocol with Q-learning for mission-critical wireless sensor networks. *Sensors*, 22(11), 3975.
6. Liu, X., Zhao, Y., Wang, J., et al. (2022). Generative adversarial learning-based trusted secure clustering for industrial wireless sensor networks. *IEEE Transactions on Industrial Informatics*, 18(9), 6238–6248.
7. Nithyadevi, K., & Nandha Kumar, R. (2023). Reinforced trust management algorithm for secure routing in wireless sensor networks. *Journal of Network and Systems Management*, 31(4), 89–108.
8. Nithyadevi, K., & Nandha Kumar, R. (2024). Reinforced trust and congestion-aware multi-hop aggregation algorithm for wireless sensor networks. *Wireless Personal Communications*, 136(2), 1547–1571.
9. Vennam, P., & Mouleeswaran, S. K. (2023). Multi-objective trust-aware cat swarm optimization for secure data aggregation in wireless sensor networks. *Journal of Intelligent & Fuzzy Systems*, 45(6), 11235–11252.
10. Shekar, R., Kumar, P., & Sharma, D. (2025). Learning-based energy-efficient routing protocols for IoT-oriented wireless sensor networks. *Journal of Network and Systems Management*, 33(1), 44–67.
11. Yang, J., Li, W., Li, C., Zhang, L., & Liu, L. (2025). An energy-efficient and transmission-efficient adaptive routing algorithm using deep reinforcement learning for wireless sensor networks. *IEEE Internet of Things Journal*, 12(23), 50414–50426.
12. Chakravarthy, M., Reddy, S., & Rajalakshmi, P. (2025). AI-driven energy-efficient data aggregation and routing protocol for wireless sensor networks. *Sensors*, 25(4), 225.
13. Nivyashree, R., & Pramod, H. B. (2026). Federated and reinforcement learning integration for distributed energy optimization in wireless sensor networks. *Engineering, Technology & Applied Science Research*, 16(1), 30869–30874.
14. Veluswamy, P. (2026). Trust-aware secure energy-efficient routing protocol for clustered wireless sensor networks using two-tier optimization algorithm. *Wireless Networks*, 32(3), 2105–2127.
15. Al-Gumaei, K., Al-Rakhami, M., Hassan, M. M., & Gumaei, A. (2026). Congestion-aware trust-enabled routing framework for secure wireless sensor networks. *Computer Networks*, 245, 110402.