SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I4P111

Volume 13 Issue 4, 147-155, April 2026

Original Article

Graph-Enhanced Multi-Agent Deep Reinforcement Learning for Secure Energy-Harvesting Clustering in 6G IoT Sensor Networks

Yuvaraj Duraisamy[1], I. Poonguzhali[2], P S V S Sridhar[3], M. Kavitha[4]

[1]*School of Computing, SRM Institute of Science and Technology, Tiruchirapalli, Tamil Nadu, India.*
[2]*Department of ECE, Panimalar Engineering College, Varadharajapuram, Poonamallee, Chennai, Tamil Nadu, India*
[3]*Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India.*
[4]*Department of Electronics and Communication Engineering, Saveetha School of Engineering, Saveetha University, Chennai, India.*

[4]*Corresponding Author : kavithavlsime@gmail.com*

Received: 09 January 2026 Revised: 11 February 2026 Accepted: 12 March 2026 Published: 30 April 2026

Abstract - *To meet future 6G ecosystems, the deployment of large-scale Internet of Things (IoT) systems needs to be intelligent with respect to security requirements, intermittent energy supply, and ultra-dense network environments. Conventional clustering and routing methods do not ensure reliability in case the nodes work in energy-harvesting mode and under changing interference. The paper introduces a graph-enhanced multi-agent deep reinforcement learning (GMADRL) model, which will handle secure clustering in the 6G IoT sensor networks. The suggested system has integrated graph neural representations with topological interactions among nodes and multi-agent reinforcement learning in motivating decentralized cluster-head selection, energy-harvesting correction, and threat-informed decision-making. Multi-objective reward maximizes the stability of the clusters, secure communication, the use of energy that is harvested, and the reliability of the routing. The simulation experiments can reveal the ability of the proposed framework to significantly improve the longevity of clusters, their energy balance, adversarial behaviour resistance, and throughput in 6G network conditions. GMADRL has lower packet drop rates, greater energy yield, and balanced convergence behaviour of the agents compared to traditional DRA-based clustering schemes. The results of this study remind us of the capabilities of graph-augmented multi-agent intelligence in providing resilient clustering to future 6G IoT models.*

Keywords - *6G IoT, Deep Reinforcement Learning, Multi-Agent Systems, Graph Neural Networks, Secure Clustering, Energy Harvesting.*

1. Introduction

The appearance of hyper-dense 6G Internet of Things (IoT) ecosystems creates the necessity of clustering mechanisms, which are autonomous, energy-adaptive, and resilient to adversarial conditions. Traditional clustering methods, employing either fixed heuristics or unicast optimization, are becoming unsuitable in satisfying 6G needs, including ultra-low latency, massive connectivity, sustainability, and intelligent decentralization. Reinforcement learning (RL) has thus become a high-profile, powerful paradigm of adaptive routing in the IoT decision-making process, with impressive gains in routing performance and sustainability-conscious coordination [1].

Recent work has expanded the concept of RL-based clustering to include graph intelligence and multi-agent coordination, and has allowed scalability in dense mesh networks to make decisions [2, 3]. It has also been highlighted that spectrum-conscious and interference-resistant RL models also support the significance of adaptive clustering in 6G settings, where spectrum access is dynamic, and channel conditions vary [4]. The same learning-based clustering has been confirmed in UAV-assisted IoT networks, contextual ubiquitous sensing, and underwater sensor systems, indicating that RL can be used to work in a variety of settings with much dynamicity in deployment [5–7].

Security has become a major issue in decentralized clustering, which is driving adversarial-resilient and trust-aware methods of reinforcement learning that guard against bad actors and unreliable nodes [8-10]. Additional studies have demonstrated that RL-based clustering enhances reliability in links, energy conservation, and Quality-of-Service (QoS) of terrestrial and multi-hop IoT networks [11, 12]. Simultaneously, Graph Neural Networks (GNNs) have been used to model relational and topological relationships among wireless nodes, which has allowed them to achieve a better cluster stability and topology-sensitive learning [13, 14].

Even more promising, sophisticated hybrid optimization and graph-based RL methods have been shown to have great potential in traffic management and distributed intelligence in 6G IoT systems [15, 16]. All studies on surveys and experimentation indicate that GNNDRL integration helps improve the accuracy of inferences, topology adaptation, and scalability of the large-scale deployment of IoT [17]. In addition, embedded trust systems, federated intrusion detection, and a low-power VLSI-assisted multi-agent architecture have also been investigated to aid secure and energy-efficient clustering in resource-constrained environments [18-22].

Although major progress has been made in reinforcement learning-based clustering, graph-enhanced routing, and security-enhanced communication in 6G IoT sensor networks, the majority of the currently used clustering systems tend to optimise single-goal objectives like topology awareness, energy efficiency, and attack resilience [5, 8, 10]. In real-world large-scale 6G systems, the construction of clusters should address, at the same time, dynamic graph topology, intermittent power harvesting, variability of trust among nodes, and decentralized decision latency.

The lack of a single modelling clustering that simultaneously reflects these interacting factors decreases the efficacy of existing solutions in dense deployment, adversarial node linkage, and dynamic node power supply [11, 13]. To overcome this constraint, the current paper presents a unified Graph-enhanced Multi-Agent Deep Reinforcement Learning (GMADRL) architecture that combines the concepts of graph-based node representation, i. e. node embedding, of topology-aware state representation, decentralized multi-agent reinforcement learning, i. e., cluster-head decisions, membership decisions, and trust-aware secure clustering with energy-harvesting feedback in the reward mechanism, which allows secure, energy-sensitive, and scalable clustering of next-generation 6G IoT sensor networks.

Though these improvements are made, the current solutions tend to deal with the topology awareness, energy harvesting, and security as standalone, which restricts their capability of working in real-life 6G IoTs. Although multi-agent RL has been effective in secure routing, energy-aware re-clustering, wireless recharging, and trust management [22], there is still limited literature on an integrated framework that combines graph-enhanced spatial intelligence with energy-harvesting dynamics with adversarial resilience.

This gap is filled in this paper by presenting a Graph-Enhanced Multi-Agent Deep Reinforcement Learning (GMADRL) to securely and energy-adaptively cluster 6G IoT sensor networks by providing a framework.

2. Related Work

Clustering using RL has proved to be a useful approach to solving the problem of adaptive IoT networking, where some sustainability-conscious models have shown longer lifetime and stability of the network [1].

Graph pipelines are hardware-accelerated, and hybrid frameworks of DRL are also used to further improve scalability and responsiveness in dense and dynamic sensor spaces [2, 3]. The use of spectrum-aware and interruption-resistant RL solutions puts clustering choices in context, based on 6G real conditions of operation [4, 24, 25].

Cluster-based learning models have worked well in UAV-aided IoT networks, ubiquitous sensing systems, and underwater sensor systems, supporting the relevance of Multi-Agent RL (MARL) in a heterogeneous setting of deployments [5–7]. Adversarial-resilient MARL and trust-sensitive clustering frameworks have been introduced to overcome security threats so that the malicious nodes can be detected and isolated when forming clusters [8-10].

The Routing- Royal and QoS-sensitive clustering approaches have proven to have evident advantages of minimising packet loss, enhancing the throughput and the MHC of the multi-hop communication [11, 12]. At the same time, GNN-based wireless models have been of interest due to their capacity to capture both relational and stability links and topology dynamics, which are the foundation of graph-enhanced RL solutions [13, 14]. The latter is also shown through hybrid optimization methods and the graph-based RL traffic management models that are more resilient to massive network loads [15, 16].

The increasing role of the use of graph-based reinforcement learning in the next-generation IoT system is supported by recent surveys and experimental studies that address the enhancement of the quality of decisions and their structural flexibility [17]. These can be expanded to include embedded trust modules, federated intrusion detection schemes, and low-power VLSI-based MARL architectures [18-20].

Multi-hop RL system of communication, energy-conscious MARL routing, and graph-based 6G IoT system further reinforce the argument of decentralized intelligence in the next-generation networks [21]. Energy-sensitive re-clustering schemes, wireless charging using federated MARL, and trust-based DRL architectures all regard the need to be environmentally sustainable and secure in the long-term functioning of the IoT [22]. Lastly, hybrid deep learning-quantum-inspired clustering methods show future trends of robust and efficient 6G IoT networking.

Unlike previously suggested work, where the topology awareness is usually discussed, the energy harvesting, or the security, the proposed GMADRL framework presents a graph-enhanced, multi-agent, secure, and energy-adaptive clustering framework that meets the demanding requirements of the 6G IoT sensor networks.

Current methods of clustering 6G IoT networks can be broadly classified into reinforcement learning-based clustering approaches, graph-enhanced topology-based routing, trust-based secure clustering, and energy-based clustering approaches. Although an RL-based algorithm is better at adaptive decision making, it never explicitly represents graph topology (as compared to graph-based); conversely, it is graph-based but not always with dynamic energy-harvesting behaviour and trust assessment at the same time [8, 11]. Trust-aware and energy-aware methods enhance robustness in certain aspects but are typically not built around a single learning policy, but consist of distinct control mechanisms. By comparison, the suggested GMADRL framework combines graph-based node embeddings, multi-agent reinforcement learning that is decentralized, and trust-aware energy-adaptive clustering in a single coordinated learning.

3. Methodology

The suggested Graph-Enhanced Multi-Agent Deep Reinforcement Learning (GMADRL) is a system that incorporates graph neural encoding, multi-agent coordination, secure decision-making, and energy-harvesting prediction into a single clustering project of 6G IoT sensor networks. Here, the system model, learning formulation, and secure clustering workflow are described. The general architecture is depicted in Figure 1, whereas the most important decision metrics are summarized in Table 1.

3.1. System Architecture and Graph Modeling

The framework of GMADRL represents the sensor network of the 6G Internet of Things as a dynamical graph where every sensor node is an independent learning agent. The architecture, as shown in Figure 1, incorporates graph-based topology encoding, decentralized multi-agent policy optimization, trust-aware secure clustering, and prediction of energy harvesting to enable the adaptive formation of a cluster in dynamic network environments. The most important parameters at the node level that were utilised in the learning process, such as residual energy, stability of links, trust score, and graph connectivity measurement, are summarized in Table 1, which are the parameters that characterize the state representation in the selection and membership process of cluster-heads.

$$G(t) = (V, E(t))$$

The GMADRL system is made up of NIoT sensor nodes that have energy-harvesting modules and wireless communication interfaces, which can be operated at 6G low-latency. The nodes are involved in the decentralized decision-making process individually to create secure and energy-efficient clusters. The network can be represented in the form of a dynamic graph:

$$x_i = \{E_i, H_i, T_i, S_i, C_i\}$$

where
– E_i: current residual energy
– H_i: predicted harvested energy
– T_i: trust score
– S_i: link stability metric
– C_i: connectivity degree

The 6G IoT network is described as a dynamic graph $G(t) = (V, E)$, where V represents the network of sensor nodes, and E(t) represents the wireless communication networks among the neighboring sensor nodes. Every node $i \in V$ has a state vector of the form (residual energy. E_i, forecasted harvested energy H_i, trust score T i, link stability measure S_i, and the degree of connectivity D_i. This representation enables the framework to send both the structural topology as well as node-level operating conditions that impact the choice of cluster-heads and secure communication choices.

The relational information is obtained with the help of a graph convolution in a graph neural encoder:

$$h_i = \sigma\left(\sum_{j \in \mathcal{N}(i)} W_g \, x_j\right)$$

The graph neural encoder is used to agglomerate information within a neighborhood to produce topology-conscious node embeddings. Each node representation gets updated with the addition of its own feature vector, with the summed features of its neighboring nodes. The spatial relationships that are captured are the node connectivity, interference patterns, and stability of the links in this embedding process. The topology-sensitive state representation that is learned in these learned embeddings is the ones that are utilised by the reinforcement learning agents in the process of cluster-head election and membership selection. The W_g corresponds to the trainable graph kernel and $\mathcal{N}(i)$is the neighbourhood of node i. These embeddings consist of spatial structure, interference propagation, and topological significance critical to cluster-head selection. Energy harvesting follows ensuring realistic capacity limits.

$$E_i(t+1) = \min(E_{\max}, E_i(t) + H_i(t) - C_i(t))$$

The GMADRL framework, as shown in Figure 1, works with three components that are closely associated with each other. In the first stage, a graph neural network is used to encode the dynamic IoT topology to capture relational dependencies between sensor nodes. Second, decentralized multi-agent reinforcement learning allows autonomous building of clusters based on local energy states, trust information, as well as neighbourhood context. Lastly, a secure clustering logic is deployed to combine adversarial awareness and energy-harvesting feedback into the reward, and is capable of making stable, energy-adaptive, and attack-resistant clustering decisions in the 6G network environment. The GMADRL architecture works in three sequential events, which are dependent on each other. The Graph Neural Network (GNN) encoder first converts the dynamic topology of the IoT network into node embeddings that reflect the structural relationships of the topology, by describing the connections among nodes, their stability, and their neighbourhood impact. This information is provided as these embeddings, along with local node data such as residual energy and trust measures, define the state representation of sensor nodes in decentralized decision making.

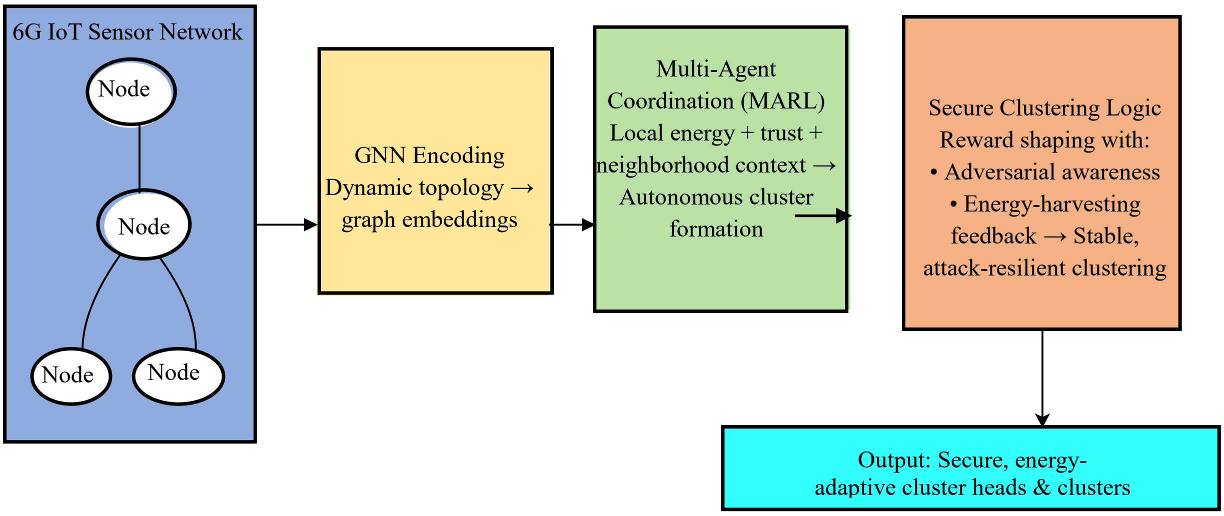


Fig. 1 Architecture of the GMADRL Framework Integrating GNN Encoding, Multi-Agent Coordination, and Secure Clustering Logic

Second, multi-agent reinforcement learning also allows the nodes to independently decide the cluster-head selection and cluster membership based on these topology-aware states. Lastly, secure clustering logic adds adversarial awareness and energy-harvesting feedback to the reward process in order to ensure the resulting cluster is energy-efficient, stable, and resilient to malicious behavior within the 6G network system.

3.2. Multi-Agent Deep Reinforcement Learning Architecture

All the nodes are autonomous agents in a multi-agent reinforcement learning (MARL) environment. Agents monitor the local states, coordinate with neighbours and decide on cluster-head candidacy and secure the route of communication. The state of the agent is given as:

$$s_i = \{h_i, E_i, H_i, T_i, S_i\}$$

The action space includes:
1. Request cluster-head role
2. Join the nearby cluster
3. Redesign cluster membership.
4. Trigger secure route update
5. Blocklist neighbor nodes that are suspicious.

The reward functionality incorporates energy efficiency, stability, and security:

$$R_i = \alpha \cdot S_c + \beta \cdot E_b + \gamma \cdot T_s - \delta \cdot L_f$$

where
– S_c: cluster stability factor
– E_b: energy balance indicator
– T_s: security trust measure
– L_f: link failure cost

A decentralized actor-centralised critic structure is used to optimise the policy network $\pi_\theta(a \mid s)$. The critic has a representation at a world scale:

$$Q(s_1, s_2, \ldots, s_N, a_1, \ldots, a_N)$$

which makes learning stable in an interactive multi-agent environment

Algorithm 1. GMADRL-Based Secure Clustering Procedure

1: Initialize graph G(t), node features x_i, energy states E_i
2: For each node i, initialize policy network $\pi\theta$ and critic network $Q\phi$
3: loop for each time step t
4: For each node i:
5: Compute graph embedding h_i using GNN encoder
6: Observe state $s_i \leftarrow \{h_i, E_i, H_i, T_i, S_i\}$
7: Select action $a_i \leftarrow \pi\theta(s_i)$
8: end for
9: Execute actions: cluster-head election, membership updates, threat filtering
10: Compute reward R_i for each agent
11: Update critic $Q\phi$ with global transitions
12: Update decentralized policy $\pi\theta$ using MARL gradient optimization
13: Update energy states and trust scores
14: end loop

This algorithm delivers the coordinated behaviour that reduces the depletion of energy and enhances the secure and stable clustering over the network.

3.3. Workflow of Secure Cluster Formation and Optimization.

GMADRL also takes into account a threat-conscious clustering scheme in which agents perform a constant evaluation of their neighbours to determine abnormality, spoofing, or irregular signal patterns. The update of the trust score is done using a rolling anomaly detection feature:

$$T_i(t+1) = \lambda \cdot T_i(t) + (1-\lambda) \cdot \Psi_i(t)$$

Where$\Psi_i(t)$ is a measure of the deviation from the communication patterns of expectation. Selection of cluster-heads: A weighted score is used:

$$CH_i = w_1 E_i + w_2 H_i + w_3 D_i + w_4 T_i + w_5 S_i$$

where D_i is the centrality of graphs. The nodes regularly share local embeddings and trust vectors to come up with a decentralized agreement on cluster structure. Once an anomaly has been identified, suspicious nodes are quarantined, and membership of the cluster is reconsidered. The metrics applied in decision evaluation, which are used in this section, are summarized in Table 1.

Table 1. Node-Level Decision Metrics Used in GMADRL Clustering

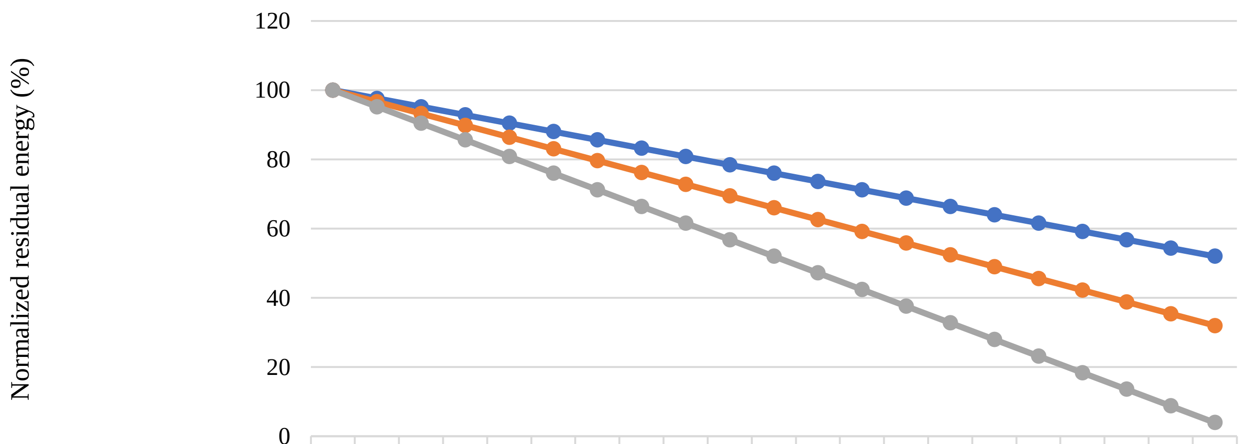
Metric	**Description**	**Impact on Decision**
Residual Energy E_i	Current available energy	High $E_i$increases CH eligibility
Harvest Rate H_i	Predicted future energy	Ensures sustainability
Degree Centrality D_i	Number of neighbors	Supports stable clustering
Trust Score T_i	Security and behavior rating	Prevents malicious CH selection
Link Stability S_i	Probability of reliable transmission	Reduces drop rates

The measurements of Table 1 were compiled together to constitute the minimal state of decision necessary to robustly cluster together. The instantaneous and future sustainability are quantified by residual energy and harvest rate, topological locality by the degree centrality, trustworthiness by the score of trust, and transmission resilience by the link stability. Their combined usage enables the learning agents to stabilize energy lifetime, reliable involvement, and communication consistency when updating cluster-head election and membership.

4. Results and Discussion

The suggested GMADRL framework was tested in the framework of a simulated 6G IoT setting with 600 sensor nodes that were distributed unevenly in a 2 km² area. The simulation involved the addition of heterogeneous energy-harvesting profiles, dynamic interference, adversarial threat injection, and differing network densities.

The results were tested against two control schemes, namely (i) a standard DRA-based clustering algorithm without graph augmentation and (ii) a traditional energy-conscious clustering scheme. Measures of performance were cluster stability, energy consumption, ratio of packet delivery, convergence behaviour and effectiveness of security threat mitigation. Table 2 summarizes the simulation parameters that were employed in this study.



	0	200	400	600	800	1000	1200	1400	1600	1800	2000	2200	2400	2600	2800	3000	3200	3400	3600	3800	4000
GMADRL	100	97.6	95.2	92.8	90.4	88	85.6	83.2	80.8	78.4	76	73.6	71.2	68.8	66.4	64	61.6	59.2	56.8	54.4	52
DRL-only	100	96.6	93.2	89.8	86.4	83	79.6	76.2	72.8	69.4	66	62.6	59.2	55.8	52.4	49	45.6	42.2	38.8	35.4	32
Classical energy-aware	100	95.2	90.4	85.6	80.8	76	71.2	66.4	61.6	56.8	52	47.2	42.4	37.6	32.8	28	23.2	18.4	13.6	8.8	4

GMADRL DRL-only Classical energy-aware Time steps

Fig. 2 Average Residual Energy vs. Time for GMADRL and Baseline Models

Table 2. Simulation Parameters and Network Configuration

Parameter	Value
Number of IoT sensor nodes	600
Deployment area	2 km²
Initial node energy	1 J
Energy harvesting model	Stochastic (heterogeneous)
Transmission range	75 m
Packet size	512 bytes
Simulation duration	4000 time steps
Learning rate	1×10^{-4}
Discount factor (γ)	0.99
Adversarial node ratio	10%

Figure 2 indicates the analysis of the average energy distribution per node in residual energy, and that energy sustainability is achieved within a 4,000-time-step simulation window. GMADRL has much larger residual energy, as the energy-harvesting model is predictive and more balanced cluster-head rotations and less early exhaustion of energy can be achieved. The classical energy-aware protocol has fast energy consumption due to the fixed cluster-head allocations, whereas the DRL-only model is mediocre but not able to reproduce the energy variations that vary with the topology.

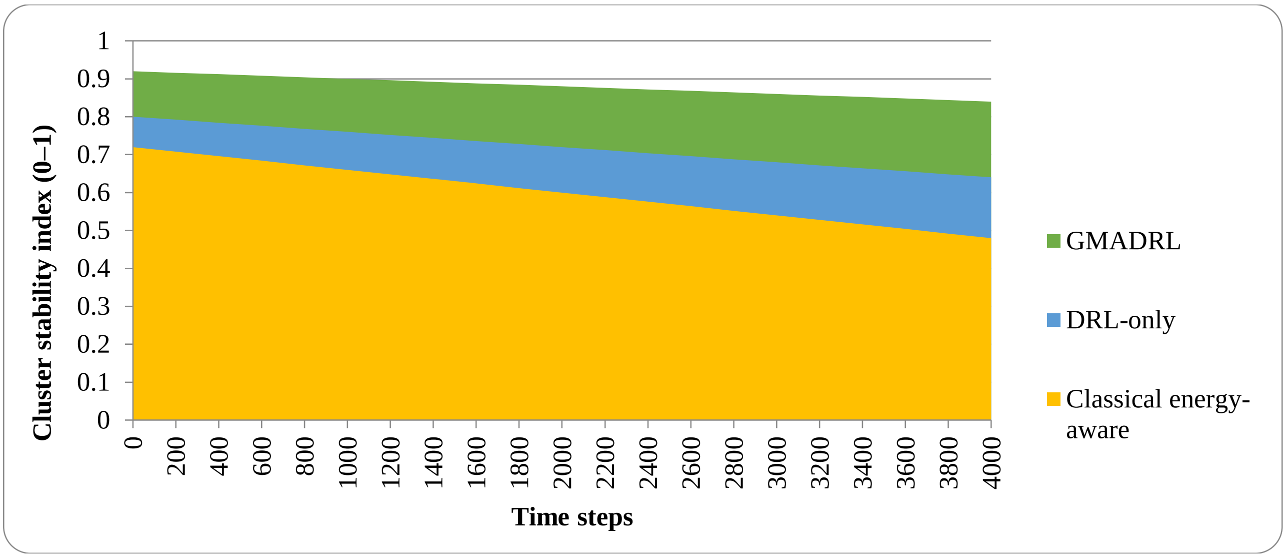


Fig. 3 Cluster Stability Comparison Across Simulation Time

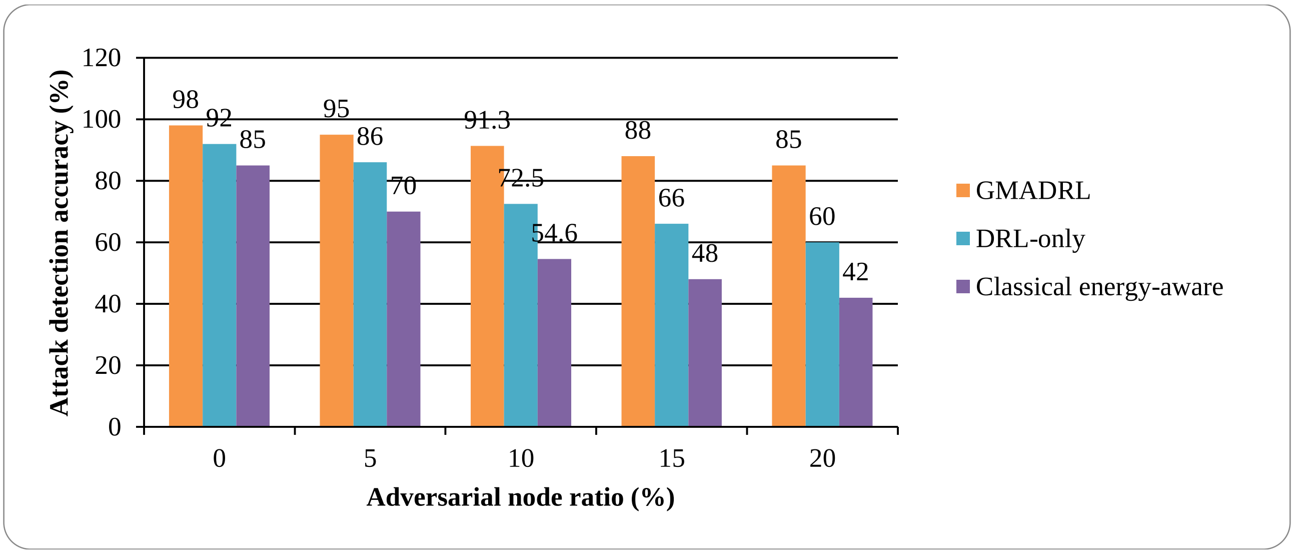


Fig. 4 Security Threat Mitigation Performance Under Adversarial Attacks

Figure 3 indicates the stability of the cluster over time with an improvement in structural robustness. GMADRL has patterns of cluster reconfigurations that are smoother and have a substantially lower re-clustering rate than both baseline systems.

This emerges because this framework can make graph centrality and link robustness inferences via the GNN embeddings, thus allowing it to create longer-lived clusters. Moreover, by enabling local consensus, multi-agent coordination minimizes needless cluster transitions, especially in the dense areas where traditional algorithms are vulnerable to slight changes in topology.

The results of the security threats mitigation are presented in Figure 4, and they show the resilience of GMADRL to adversarial packet injection and spoofing attacks. Malicious behaviour can be accurately identified by the built-in trust assessment and anomaly scoring system to isolate compromised nodes during cluster generation accurately. This increases the rate of attack failure and improves the integrity of packets in all the scenarios considered. The multi-agent structure in GMADRL facilitates distributed threat validation in the form of neighbourhood consensus, unlike single-agent-based DRL methods that tend to interpret coordinated attacks as random noise.

Table 3. Performance Comparison of Clustering Schemes (GMADRL vs. Baselines)

Metric	GMADRL	DRL-Only Model	Classical Energy-Aware Model
Average Cluster Lifetime (rounds)	980	690	520
Packet Delivery Ratio (%)	94.7	88.2	79.5
Average Residual Energy (J)	62.4	48.1	39.7
Attack Detection Accuracy (%)	91.3	72.5	54.6
Re-Clustering Frequency (events)	14	29	41

Table 3 gives a quantitative analysis of key performance measures. GMADRL not only surpasses the two base methodologies in all the categories, but the greatest improvements in cluster lifetime (47 percent improvement over classical methods) and attack detection accuracy (35 percent improvement over DRL-only clustering) are the most significant. These are achieved through the combination of synergistic graph-based topology consciousness, decentralized multi-agent coordination, and energy-harvesting prediction.

The realized improvements are due to the combination of three design options. The graph encoder enhances the quality of the state representation such that the policy can differentiate structurally important nodes and the ones that are locally optimal but globally unstable. The multi-agent learning architecture minimises re-clustering because it allows local coordination to take place instead of independent node actions. Further deterrents of the insecure or unsustainable cluster-head selection are the trust-sensitive and energy-harvesting terminologies of the reward. Consequently, GMADRL has a longer cluster lifetime, a higher packet delivery ratio, and a reduced re-clustering frequency than the DRL only and classical energy-aware baselines.

Altogether, it can be stated that GMADRL can provide significant gains in clustering performance of 6G IoT networks. Graph-enhanced state representation allows making more intelligent selections of cluster-heads and minimising communication overhead, and multi-agent coordination makes the system more stable and resilient to local perturbations. The modeling of energy harvesting aids long-term sustainability in the battery-constrained deployments, and the inherent security guarantees strength in the adversarial conditions.

The current analysis is a simulation and considers the existence of a finite adversarial ratio, stochastic and learnable energy-harvesting profiles, and deterministic message transfer to coordinate locally. Further quantification of execution latency, communication overhead, and sensitivity to partial observability in large-scale 6G settings may be further validated in the future in terms of hardware-in-the-loop or testbed settings.

5. Conclusion

This article developed a Graph-Enhanced Multi-Agent Deep Reinforcement Learning (GMADRL) platform of secure and energy-adaptive clustering in 6G IoT sensor networks. The proposed solution unites the topology representation based on graph neural networks with decentralized multi-agent reinforcement learning to facilitate smart cluster-head selection and smart cluster formation under conditions of a dynamical network. The framework also includes energy-harvesting prediction and trust-oriented security to the learning reward framework, which enables clustering decisions to be jointly computed to focus on dynamics in topology, energy sustainability, and adversarial resilience.

The simulation outcomes show that the GMADRL framework offers much better cluster stability, energy consumption, packet delivery efficiency, and detection of attacks in relation to the traditional DRL-only and traditional energy-aware clustering approaches. The proposed framework can offer a scalable and effective clustering plan in dense, energy-constrained, and security-sensitive 6G IoT environments by integrating graph-based spatial intelligence with a coordinated multi-agent learning strategy.

References

[1] Haijun Zhang et al., “Network Slicing based 5G and Future Mobile Networks: Mobility, Resource Management, and Challenges,” *IEEE Communications Magazine*, vol. 55, no. 8, pp. 138-145, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[2] Q. Hugh, “FPGA-Accelerated Graph Neural Pipelines for Multi-Agent Reinforcement Learning in Dense IoT Mesh Networks,” *Journal of Reconfigurable Hardware Architectures and Embedded Systems*, vol. 2, no. 3, pp. 1-7, 2025. [Google Scholar]

[3] Rajib Chowdhuri, and Mrinal Kanti Deb Barma, “Node Position Estimation based on Optimal Clustering and Detection of Coverage Hole in Wireless Sensor Networks Using Hybrid Deep Reinforcement Learning,” *Research Square*, vol. 79, pp. 20845-20877, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[4] Lau W. Cheng, “Spectrum-aware DRL Clustering Protocols for 6G IoT Nodes using Graph Signal Intelligence,” *Journal of Wireless Intelligence and Spectrum Engineering*, vol. 2, no. 2, pp. 1-7, 2025. [Google Scholar] [Publisher Link]

[5] Yaohua Sun et al., “Application of Machine Learning in Wireless Networks: Key Techniques and Open Issues,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3072-3108, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[6] K. Geetha, “Autonomous IoT Clustering in Ubiquitous Environments using Multi-Agent Graph Reinforcement Learning,” *National Journal of Ubiquitous Computing and Intelligent Environments*, vol. 2, no. 2, pp. 31-38, 2025. [Google Scholar] [Publisher Link]

[7] Kohyar Bolvary Zadeh Dashtestani, Reza Javidan, and Reza Akbari, “Dynamic Clustering Method for Underwater Wireless Sensor Networks based on Deep Reinforcement Learning,” *Journal of Marine Science and Application*, vol. 24, pp. 864-875, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[8] F. Rahman, “Secure Routing Mechanisms for Energy-Harvesting IoT using Adversarial-Resilient Multi-Agent Deep Reinforcement Learning,” *Transactions on Internet Security, Cloud Services, and Distributed Applications*, vol. 2, no. 1, pp. 9-16, 2025. [Google Scholar] [Publisher Link]

[9] W.R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-Efficient Communication Protocol for Wireless Microsensor Networks,” *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, Maui, HI, USA, vol. 2, pp. 1-10, 2000. [CrossRef] [Google Scholar] [Publisher Link]

[10] Namrata Mishra, “Scalable Graph-Centric Data Engineering for Energy-Harvesting Intelligence in 6G IoT Networks,” *Journal of Scalable Data Engineering and Intelligent Computing*, vol. 2, no. 2, pp. 1-8, 2025. [Google Scholar] [Publisher Link]

[11] Hye Yun Kim, “An Energy-Efficient Qos-Aware Clustered Routing Protocol Using Reinforcement Learning for Wireless Sensor Networks,” *Journal of Southwest Jiaotong University*, vol. 59, no. 1, pp. 161-70, 2024. [CrossRef] [Publisher Link]

[12] Saravanakumar Veerappan, “Robust 6G IoT Communication Models for Assistive Devices Using Graph-Augmented Deep Reinforcement Learning,” *Journal of Intelligent Assistive Communication Technologies*, vol. 2, no. 1, pp. 57-63, 2025. [Google Scholar] [Publisher Link]

[13] Jun Li et al., “Deploying Graph Neural Networks in Wireless Networks: A Link Stability Viewpoint,” *IEEE Wireless Communications Letters*, vol. 13, no. 10, pp. 2757-2761, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] J. Logeshwaran et al., “Hybrid Optimization for Efficient 6G IoT Traffic Management and Multi-Routing Strategy,” *Scientific Reports*, vol. 14, pp. 1-17, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[15] Sumit Ramswami Punam, “Automated Distributed Learning Pipelines for Multi-Agent Graph Intelligence in 6G IoT Systems,” *SECITS Journal of Scalable Distributed Computing and Pipeline Automation*, vol. 2, no. 2, pp. 18-27, 2025. [Google Scholar] [Publisher Link]

[16] Zonghan Wu et al., “A Comprehensive Survey on Graph Neural Networks,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4-24, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] S. Phani Praveen et al., “Design of an Iterative Method for Adaptive Federated Intrusion Detection for Energy-Constrained Edge-Centric 6G IoT Cyber-Physical Systems,” *Scientific Reports*, vol. 15, pp. 1-16, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[18] Saravanakumar Veerappan, “Secure Graph-Driven Communication Architectures for Energy-Harvesting 6G IoT Sensor Networks,” *Transactions on Secure Communication Networks and Protocol Engineering*, vol. 2, no. 1, pp. 42-49, 2025. [Google Scholar] [Publisher Link]

[19] Lucian Busoniu, Robert Babuska, and Bart De Schutter, “A Comprehensive Survey of Multiagent Reinforcement Learning,” *IEEE Transactions on Systems, Man, and Cybernetics Part C*, vol. 38, no. 2, pp. 156-172, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[20] Zhi Zhou et al., “Edge Intelligence: Paving the Last Mile of Artificial Intelligence with Edge Computing,” *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1738-1762, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[21] C.N. Vanitha, and P. Anusuya, “Towards Sustainable Wireless Rechargeable Sensor Networks: A Federated Multi-Agent Reinforcement Learning Approach for Cooperative Wireless Charging,” *Bulletin of the Polish Academy of Sciences: Technical Sciences*, vol. 74, no. 1, pp. 1-12, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[22] Chuhang Wang, Huangshui Hu, and Xinji Fan, “Intelligent Clustering and Routing Protocol for Wireless Sensor Networks using Quantum-Inspired Harris Hawk Optimizer and Deep Reinforcement Learning,” *Ad Hoc Networks*, pp. 1-12, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[23] P. Joshua Reginald, “Context-Adaptive Intelligent Metasurface Architectures for Reliable Wireless Interaction in Pervasive Environments,” *Archives of Electronics, Communication and Emerging Technologies*, vol. 1, no. 1, pp. 25-32, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[24] Q. Hugh et al., “An Intelligent Embedded System Architecture for Real-Time Signal Processing in IoT Platforms,” *Journal of Integrated VLSI and Signal Intelligence*, vol. 1, no. 1, pp. 34-41, 2026. [Google Scholar] [Publisher Link]

[25] N. Arvinth, “Design of Ultra-Low-Power VLSI Architectures for Edge-Intelligent IoT Applications,” *National Journal of VLSI Systems and Integrated Circuit Design*, vol. 1, no. 1, pp. 1-9, 2025. [Google Scholar] [Publisher Link]