WSN Clustering with Decision Transformer and Dynamic Cluster Head Selection

Homin Oh
Department of Artificial Intelligence
Ajou University
Suwon Korea
gom2553@ajou.ac.kr

Young-June Choi
Department of Artificial Intelligence
Ajou University
Suwon Korea
choiyj@ajou.ac.kr

ABSTRACT

This paper proposes a novel clustering optimization method for wireless sensor networks (WSNs) that integrates a reinforcement learning–based decision transformer (DT) with a dynamic cluster head selection strategy. The primary objective is to extend the network's operational lifetime by jointly optimizing the First Node Death (FND) and the average node lifetime. This approach enhances both the initial stability and long-term performance of the network. At each round, information such as the number of alive nodes, residual energy, node locations, and inter-node communication distances is utilized to minimize energy imbalance, thereby maintaining efficient clustering throughout network operation. Extensive simulations and experiments demonstrate that the proposed method consistently outperforms existing techniques across various network sizes, node densities, and energy configurations. These results highlight the method's adaptability, efficiency, and scalability. Moreover, it requires no additional parameter tuning to sustain optimal clustering performance, making it well suited for practical WSN applications. Overall, the proposed method offers a robust and scalable solution for efficient WSN management.

Keywords – Clustering, Decision Transformer (DT), lifetime, Reinforcement Learning (RL), wireless sensor network (WSN)

I. INTRODUCTION

Wireless sensor networks (WSNs) are composed of a number of sensor nodes, each collecting and processing environmental data with limited energy and transmitting it to a Base Station (BS). Energy consumption in the data transmission process directly affects the network life, and the clustering technique is used as a representative way to improve it. Clustering is a structure that divides the network into several clusters, and one Cluster Head (CH) in each cluster aggregates data and transmits it to the BS, which distributes traffic load and reduces the transmission distance to increase energy efficiency. However, the efficiency varies depending on various factors such as CH selection criteria, node distribution, distance, energy state, and network change. Therefore, existing fixed clustering has limitations in that excessive load is concentrated on some CHs because it is difficult to adapt to the dynamic environment, and network life is shortened due to energy imbalance and data loss.

Early representative clustering techniques sought to shorten the transmission distance and reduce energy consumption by selecting CHs according to the node's residual energy or probabilistic criterion. However, existing clustering-based routing protocols of such a simple approach have the following limitations. First, energy consumption of some nodes is concentrated because the energy distribution of the entire network is not comprehensively considered [1], [2]. Second, the distance or spatial imbalance between nodes is not sufficiently reflected, resulting in load imbalance and shortening network life among CHs [3], [4]. Third, it is difficult to adapt in real time to environmental changes such as network topology changes or residual energy changes of nodes [5], [6]. Accordingly, there is a need for an adaptive clustering technique that can operate efficiently even in a dynamic network environment.

To cope with such problems, recently, clustering based on Reinforcement Learning (RL) has been actively investigated. RL is a method of learning the optimal behavioral policy based on reward signals without labeled data and is suitable for WSN clustering in that it can operate efficiently even in uncertain and nonlinear environments. However, traditional RL methods frequently cause inefficient CH selection while performing random exploration in the early stages of learning. In the case of WSNs, the energy of the node is consumed every round, so the initial wrong exploratory choice immediately deepens on the energy imbalance and leads to poor performance of the entire network. In addition, as the number of nodes in the network increases, the state space expands exponentially, resulting in a scalability problem that rapidly slows down the learning speed [7].

This study proposes a RL-based WSN clustering framework that combines a Decision Transformer (DT) and Dynamic Cluster Head selection [8]. DT converts the sequential decision-making problem of RL into a transformer structure and utilizes Replay Buffer and Target Network to ensure learning stability and efficiency. Since learning is done based on experience data stored in Replay Buffer, it minimizes unnecessary energy waste by random exploration and improves initial energy efficiency, First Node Death (FND), and average node lifetime at the same time. In addition, it provides high adaptability to environmental changes, so it can be applied immediately without re-learning when the number of nodes, batch area, and transmission/reception energy model are changed.

The proposed model dynamically selects a CH using an equation considering the node's location, distance, and energy state in each round, while maintaining an appropriate CH ratio in the initial stage to prevent energy imbalance in the second half and maximize the average node lifetime. The DT structure can operate efficiently even in large-scale WSN environments due to its fast-learning speed and low computational complexity. The proposed model of this study aims to overcome the limitations of existing RL-based clustering and achieve scalability, adaptability, and energy efficiency in various network environments at the same time. In addition, this study shows an increase of up to 893.46% in

FND performance and up to 174.41% in the average node lifetime compared to the existing model.

II. RELATED WORK

The clustering technique of WSNs has recently developed into an intelligent model using RL, starting from a simple probabilistic method in the early days. In this section, we will look at LEACH, HEED, and GEECS, which are representative classical clustering techniques, LEACH-RLC and DQN, which are RL-based techniques, and finally, the Decision Transformer (DT) model proposed in this study.

A. Classical clustering techniques

In classical WSNs, clustering techniques play a key role in extending network life and improving energy efficiency. Representative protocols include Low Energy Adaptive Cluster Hierarchy (LEACH), Hybrid Energy-Efficient Distributed Cluster (HEED), and Greedy Energy-Efficient Cluster Scheme (GEECS). LEACH maintains overall energy consumption at a constant level by autonomously becoming a CH with nodes at a certain probability and distributing traffic loads through periodic CH re-elections. However, if all nodes have the same initial energy, performance is degraded in large or non-uniform environments because distance-based transmission costs or environmental imbalances are not considered [9].

To improve this, the hybrid energy efficient distributed clustering (HEED) repeatedly selects CH by considering the residual energy and communication cost of the node together. HEED improves energy efficiency and network life by allowing clusters to be configured with only local information, but there is a limitation in that control message overhead occurs in the repetitive CH selection process and the cluster reconfiguration period is fixed, resulting in poor adaptability to dynamic environmental changes [3]. The proposed GEECS then modeled the CH selection process as a multi-armed bandit (MAB) problem and combined the improved ε-greedy search strategy with K- means-based clustering to improve energy efficiency. However, this approach suffers from some nodes are likely to be over selected during the ε-greedy search, and that adaptability may be limited if the network environment is rapidly changing [10].

B. Reinforcement Learning-Based Cluster Techniques

To overcome these limitations, RL based clustering techniques have been proposed. RL is a method of learning optimal policies that maximize rewards through interaction with the environment, which has the advantage of enabling autonomous decision-making even in nonlinear and abnormal WSNs environments. Among them, Low-Energy Adaptive Clustering Hierarchy with RL-based Controller (LEACH - RLC) defines the CH selection problem as the Markov Decision Process (MDP) and is designed to learn optimal behaviors that maintain network connectivity while minimizing energy consumption. In addition, energy distribution can be managed in a more balanced way than the existing stochastic approach by combining it with a Mixed Integrated Linear Programming (MILP)- based optimization model, but due to the centralized learning structure and limited state representation, the scalability and learning efficiency are reduced in large networks [11].

Later, Deep Q-Network (DQN)-based clustering has been improved to enable learning even in a continuous state-action space by combining deep learning with RL and approximating Q-values with neural networks. DQN effectively reduced transmission distance and energy consumption by dynamically determining CHs by learning node location, energy state, and distance information as inputs [12]. However, there are limitations in which energy imbalance and FND performance fluctuations occur due to unstable exploration in the early stages of learning. Subsequently, LEACH-RLC secured initial stability but had limitations in improving overall lifespan, and DQN improved average lifespan but showed contradictory characteristics due to low learning stability.

In summary, classical models such as LEACH and HEED, GEECS laid the foundation for early WSN clustering with a simple and efficient structure, but there was a limitation in the lack of dynamic environment response. RL-based techniques that emerged to overcome this improved performance through autonomous learning but still did not completely achieve the balance of stability and scalability. Accordingly, a RL model combined with a Transformer-based structure was applied to WSN clustering.

III. SYSTEM MODEL

The WSN covered in this study is composed of sensor nodes with limited energy resources, and the network life is determined by data transfer and clustering strategies between nodes. For the proposed RL-based clustering research, the system model of WSNs is largely defined as a network model and an energy model.

A. Network Model

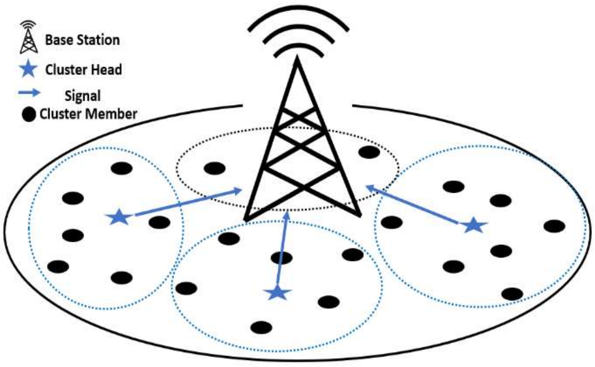


Fig. 1. WSNs Network Model.

In this study, a WSN composed of N sensor nodes as shown in Fig. 1 is considered. It is assumed that sensor nodes are $L \times L$ square sensing area, and a single BS exists in a fixed position outside the sensing area. All sensor nodes have limited initial energy and are composed of nodes with the same hardware and communication capabilities. Each node periodically senses its surroundings and then transmits data to the CH of its cluster. The CH removes duplicates through the data aggregation process collected within the cluster, merges them into one packet, and transmits it to the BS.

The network model considered in this study was tested under the following assumptions.

- The node has a fixed position and the same initial energy. In addition, the sensor node may perform data collection and intra-cluster transmission functions.

- The network is clustered on a round-by- round basis, and in each round, a RL-based model dynamically selects a CH, nodes send data to the CH, which aggregates and delivers it to the BS.

- All nodes can communicate directly with the BS, minimizing the transmission distance through clustering, and calculating the data transmission distance as Euclidean distance.

B. Energy Model

In WSNs, energy consumption of each node occurs in the process of data transmission and reception, and based on this, the network life is evaluated. In this study, an electromagnetic displacement model is used, and the energy consumption of each node is defined as follows. This model considers both a free-space model and a multi-path fading model and applies different path loss indices according to the transmission distance. The transmission energy $E_{tx}(k,d)$ when transmitting bit-sized data by a distance is defined as follows:

$$\boldsymbol{E_{tx}(k,d)} = \begin{cases} \boldsymbol{kE_{elec} + k\varepsilon_{fs}d^2, d < d_0,} \\ \boldsymbol{kE_{elec} + k\varepsilon_{mp}d^4, d \geq d_0.} \end{cases} \tag{1}$$

In this case, the received energy is as follows:

$$\boldsymbol{E_{rx}(k) = kE_{elec},} \tag{2}$$

where, E_{elec} is the energy consumed by the transmission and reception circuits, and E_{fs}, E_{mp} is the transmission amplification coefficient in the free space and multi-path model, respectively. Also, the critical distance d_0 is $d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}}$. Based on this energy model, this study performs the selection of the CH based on RL considering the residual energy and transmission efficiency of the node in each round and is designed to maximize the network life.

IV. METHOD

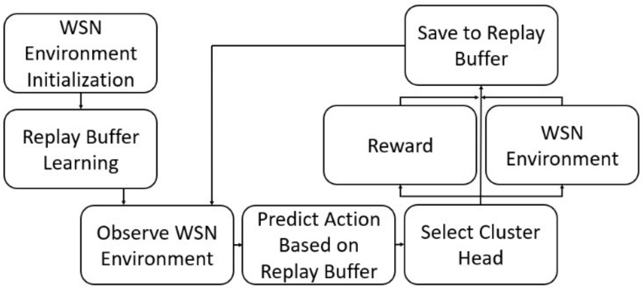


Fig. 2. Decision Transformer Framework.

Table Ⅰ. Reinforcement Learning Hyperparameters

Discount factor (γ)	0.99
Learning rate	10^{-3}
Batch size	32
Replay buffer size	10,000
Exploration rate (ϵ)	0.05

Fig. 2 shows how the Decision Transformer model works in the WSN environment. Unlike other RL methods, DT uses replay buffer to predict behavior based on a predetermined result from the initial exploration and receive rewards accordingly. In addition, for dynamic CH selection, the method which is given in the LEACH model is recalculated according to the current state.

Table I, lists the core reinforcement learning hyper-parameters that govern the learning stability, fusion behavior, and exploration efficiency of the proposed model.

A WSN is a design consideration where network life is important due to limited energy resources, and clustering techniques are widely used for efficient energy management, which evenly distributes energy consumption among nodes and optimizes data transmission paths to improve the overall stability and sustainability of the network. In addition, the effectiveness of clustering depends on certain conditions such as cluster density and distance from the node's BS. Therefore, CH selection considering distance and space between nodes is important in the clustering process [13], [14].

The following formula modifies and applies the formula that determines the number of CHs to fit the dynamic clustering environment, thereby selecting the number of CHs and improving energy efficiency in the data transmission process [15].

$$\boldsymbol{K_{opt} = \sqrt{\frac{n_{alive}}{2\pi}} \cdot \sqrt{\frac{E_{fs}}{E_{mp}}} \cdot \frac{M}{d_{BS}}} \tag{3}$$

In this case, the network size according to the current node follows the following formula M.

$$\boldsymbol{M = \sqrt{M_x \cdot M_y}} \tag{4}$$

$$\boldsymbol{M_y = 2\max_i |y_i - y_{BS}|} \tag{5}$$

$$\boldsymbol{M_y = 2\max_i |y_i - y_{BS}|} \tag{6}$$

K_{opt} refers to the number of CHs, which is changed in real time by the current number of surviving nodes of N_{alive}, the size of the network according to the current node of M, and the average distance between the node of d_{BS} and the BS, thereby maintaining the WSN clustering efficiency over time. Also, each node n_i periodically calculates a distance $d(n_i, x)$ to the BS and a CH set $CH = CH_1, CH_2, \ldots, CH_{k_{opt}}$. After that, the node selects an object that satisfies the following minimum distance condition and becomes a Cluster Member (CM) accordingly or communicates directly to the BS. Equation (7) is a key factor in minimizing the transmission of energy and extending the lifetime of the entire network.

$$\boldsymbol{x^* = \arg\min_{x \in (CH \cup BS)} d(n_i, x)} \tag{7}$$

The reward function of RL proposed later is designed based on energy balance and energy efficiency and aims to improve the stability and lifespan of the network by optimizing both elements at the same time. The formula of the reward function is as follows:

Energy balance: It is a key factor in FND delay and overall network life extension in WSNs. If energy consumption is unbalanced, some nodes are consumed early, accelerating the FND time, and reducing network connectivity and data transmission reliability. Therefore, maintaining energy balance is a prerequisite for distributing

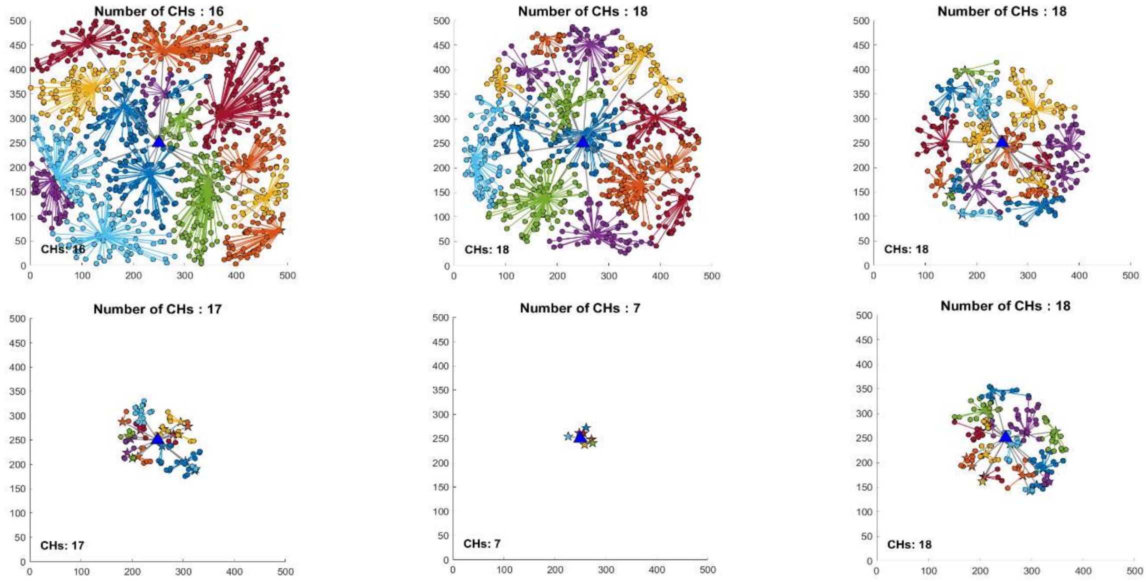


Fig. 3. Clustering process.

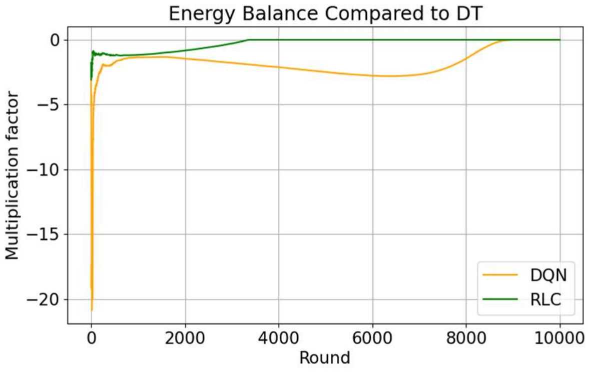


Fig. 4. Relative energy levels of existing models, DQN and LEACH-RLC, compared to the proposed model.

WSNs loads and securing communication stability.

$$\boldsymbol{E_{balance} = Var(E_i) + \max(E_i) - \min(E_i)} \quad (8)$$

Energy efficiency: It is important to extend network life by minimizing energy consumption in WSNs with limited battery resources. Efficient energy use reduces communication distance and retransmission between nodes and distributes the energy burden each round by optimizing the CH selection and data transmission process. This simultaneously improves the stability of data transmission and network operation efficiency; thus, enabling sustainable operation of WSNs with energy balance.

$$\Delta \boldsymbol{E_{avg} = \overline{E}_{prev} - \overline{E}_{current}, \frac{1}{N}\sum_{i=1}^{N} E_i} \quad (9)$$

In this study, the state includes the location, residual energy, and cluster information of each node; therefore, the action is defined as the selection of the CH determined by the number of dynamic clusters per round. The reward function is designed to reflect alive nodes, energy balance, and data transmission efficiency at the same time, and the performance of the proposed model is evaluated in terms of FND and average node lifetime.

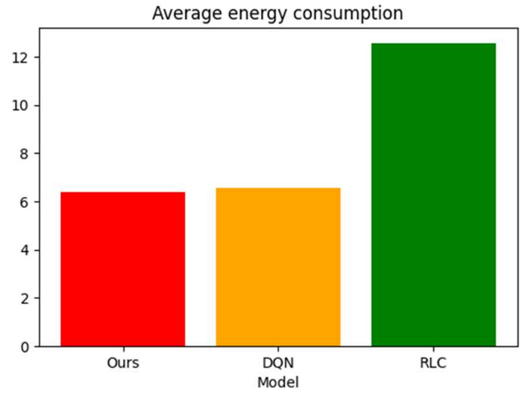
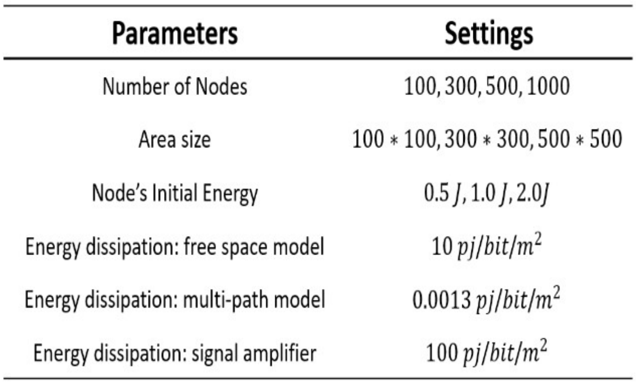


Fig. 5. Average consumed energy of our model, DQN, and LEACH-RLC.

V. EXPERIMENT

Table II. WSN experimental environment.



Parameters	Settings
Number of Nodes	100, 300, 500, 1000
Area size	$100 * 100, 300 * 300, 500 * 500$
Node's Initial Energy	$0.5\,J, 1.0\,J, 2.0J$
Energy dissipation: free space model	$10\,pj/bit/m^2$
Energy dissipation: multi-path model	$0.0013\,pj/bit/m^2$
Energy dissipation: signal amplifier	$100\,pj/bit/m^2$

The experimental environment was conducted in the same environment as Table Ⅱ, except for the number of nodes, space size, and initial energy of the nodes.

The clustering process proceeds as shown in Fig. 3 and according to the above-defined equation, the number of CHs is dynamically adjusted according to the number of surviving

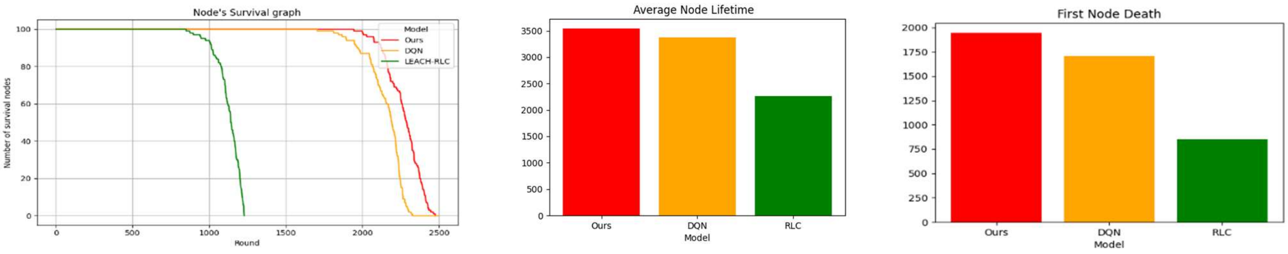


Fig. 6. The number of survived nodes, average node lifetime, and first node death for the proposed model, DQN, and LEACH-RLC (100 nodes, 100×100 m², initial energy of 0.5J)

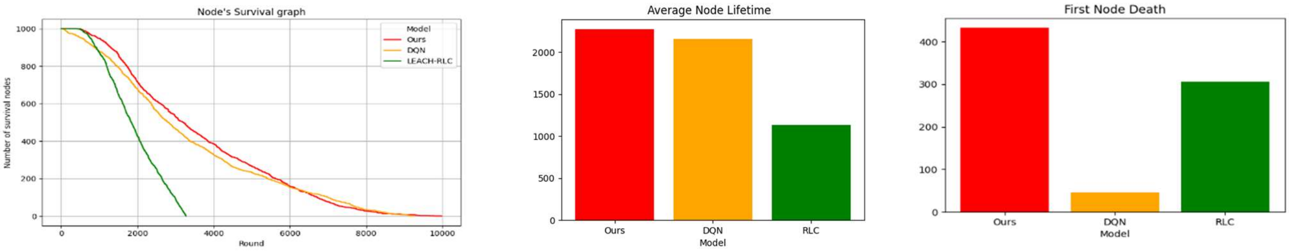


Fig. 7. The number of survived nodes, average node lifetime, and first node death for the proposed model, DQN, and LEACH-RLC (1,000 nodes, 500×500 m², initial energy of 2.0J)

Table III. FND rounds and average alive nodes of the proposed model, DQN, and LEACH-RLC for various environments.

WSN Environments		FND (Round)					Average Node Lifetime (Round)				
Number of Nodes	Area Size/ Initial Energy	Ours	DQN	LEACH-RLC	Performance Improvement over Ours		Ours	DQN	LEACH-RLC	Performance Improvement over Ours	
					DQN	RLC				DQN	RLC
100	100*100/0.5J	**1946**	1707	852	114.00%	228.40%	**2273.91**	2155.10	1130.45	105.51%	201.15%
	300*300/1.0J	**503**	165	276	304.85%	182.25%	**2160.33**	1872.81	964.62	115.35%	223.96%
	500*500/2.0J	**106**	62	92	170.97%	115.22%	**1720.33**	1168.94	626.91	147.17%	274.41%
300	100*100/0.5J	**2123**	1770	954	119.94%	222.54%	**2538.05**	2348.11	1188.20	100.42%	198.45%
	300*300/1.0J	**783**	88	770	889.77%	101.69%	**2977.97**	2581.34	1697.22	115.37%	175.46%
	500*500/2.0J	**207**	88	206	235.23%	100.49%	**2439.10**	2581.34	1515.07	94.49%	156.77%
500	100*100/0.5J	**2188**	1973	946	110.90%	231.29%	**2387.74**	2393.49	1206.07	99.76%	197.98%
	300*300/1.0J	**1082**	523	947	104.68%	177.13%	**3246.14**	3100.96	1832.60	104.68%	177.13%
	500*500/2.0J	**278**	65	290	427.69%	95.86%	**2980.03**	2979.25	1880.86	100.02%	158.44%
1000	100*100/0.5J	**2277**	1675	940	135.94%	242.23%	**2418.55**	2422.89	1216.58	99.82%	198.80%
	300*300/1.0J	**1520**	153	1012	993.46%	150.20%	**3578.60**	3438.78	1914.15	104.07%	186.96%
	500*500/2.0J	**433**	46	306	934.78%	141.50%	**3547.24**	3377.13	2261.93	105.04%	156.82%

nodes in the state of WSNs, the distance between the node and the BS, and the space in which the node exists. Each node is connected to the nearest CH or BS to transmit data, and the CH aggregates the collected data and delivers it to the BS.

In Fig. 4, The reward values of DQN fluctuated up to approximately 20 times more than those of DT during the initial learning stage due to unstable policy evaluation in early RL [16]. It refers to a phenomenon in which the reward estimate is greatly affected due to the insufficient learning of policies and value functions in the initial stage of RL. When performing RL-based clustering in a WSN environment, the initial FND performance is also unstable because some nodes are selected as CHs under unfavorable conditions in the initial exploration stage, or because the value function estimation is inaccurate, when clustering selection with low energy efficiency is repeated. In addition, the initial exploration process that does not sufficiently reflect various states such as node location, residual energy, and communication distance also affects instability of FND. Therefore, prior learning using the DT's replay buffer plays an important role in the initial FND stabilization of RL-based clustering in WSNs.

Fig. 5 shows the average energy consumption per round of each model. The main reason why the average node lifetime of the RLC model is lower than that of the DT model is that the energy consumption is about twice as high. High energy consumption acts as a major factor in reducing the overall survival rate of the network by accelerating the depletion of residual energy of the node, which is consistent with the results of previous studies that show that energy efficiency is directly related to node viability.

As can be seen in Fig. 6 and Fig. 7, the existing DQN model showed a similar level of FND and average node lifetime to the proposed model in a small space, but the FND performance sharply deteriorated in an environment where some nodes are likely to die even in the initial stage due to the large space. This is due to the initial learning instability of DQN and the problem of some nodes being excessively selected during the exploration process. The LEACH-RLC model has limited performance due to its low energy efficiency in a small space but maintained minimal FND performance even when the space increased. This is because early energy consumption of some nodes was mitigated through MILP-based optimization and dynamic CH selection.

On the other hand, the proposed model stably achieved high FND and average node lifetime regardless of changes in space size and network density. It is shown that node survival and overall network life can be optimized simultaneously even in various WSN environments by incorporating dynamic clustering and an energy-balance-aware reward function that distributes the communication load evenly among nodes and prevent excessive energy consumption even in the initial exploration stage, i.e., clustering by calculating equations according to the network state at every moment without additional parameter adjustment. Therefore, the results of this experiment support that the proposed model is a practical and scalable solution for WSN clustering.

Table III, presents FND rounds, average node lifetime, and performance improvements across experiments. The proposed model consistently outperforms existing methods, adapting effectively to changes in node count and network size. This demonstrates its high efficiency, adaptability, and scalability for WSN clustering.

VI. CONCLUSION

In this study, RL-based WSN clustering framework was proposed to effectively overcome the limitations of the existing static clustering and DQN-based models. The proposed model improved not only energy efficiency but also overall network life by integrating the dynamic cluster head selection and multi-target reward structure using decision transformer. As a result of the experiment, the performance improved stably compared to the existing model in the First Node Death (FND) and average node lifetime, and consistent results were confirmed for various numbers of nodes and space sizes. In addition, the proposed model effectively distributed energy depletion due to round progress by considering energy efficiency and balance at the same time and showed the performance of extending the FND and average node lifetime by up to twice or more compared to the previous model. High adaptability and robustness were secured even in various environmental conditions such as node density, placement area, and initial energy change. Taking together, the framework of this study proved that it operates stably even in dynamic environments while simultaneously improving energy efficiency and network life and provides a foundation that can be practically applied to large-scale IoT and smart sensor networks in the future.

VII. ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2023R1A2C1003783) and Institute of Information & communications Technology Planning & Evaluation (IITP) under the Artificial Intelligence Convergence Innovation Human Resources Development (IITP-2025-RS-2023- 00255968) grant funded by the Korea government(MSIT).

REFERENCE

[1] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy efficient communication protocol for wireless microsensor networks," in *Proc. 33rd Annu. Hawaii Int. Conf. System Sciences*, vol. 2, 2000, pp. 8020-8027.

[2] S. Lindsey and C. S. Raghavendra, "PEGASIS: Power efficient gathering in sensor information systems," in *IEEE Aerospace Conf. Proc.*, vol. 3, 2002, pp. 1125–1130.

[3] O. Younis and S. Fahmy, "HEED: A hybrid, energy efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Trans. Mobile Comput.*, vol. 3, no. 4, pp. 366–379, 2004.

[4] M. Ye, C. Li, G. Chen, and J. Wu, "EECS: An energy efficient clustering scheme in wireless sensor networks," in *Proc. 24th IEEE Int. Performance, Computing, and Communications Conf. (IPCCC)*, 2005, pp. 535–540.

[5] A. A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Computer Commun.*, vol. 30, no. 14–15, pp. 2826–2841, 2007.

[6] S. R. Kashaf, N. Javaid, Z. A. Khan, and I. A. Khan, "TSEP: Threshold sensitive stable election protocol for WSNs," in *10th Int. Conf. Frontiers of Information Technology*, 2012, pp. 164–168.

[7] T. Sharma, S. S. Singh, and R. K. Sharma, "ReLeC: A reinforcement learning based clustering enhanced protocol for efficient energy optimization in wireless sensor networks," *Computational Intelligence and Neuroscience*, vol. 2022, Art. 3337831, 2022.

[8] L. Chen, K. Lu, and Z. Wang, "Decision Transformer: Reinforcement learning via sequence modeling," in *Advances in Neural Information Processing Systems*, vol. 34, 2021, pp. 15084–15097.

[9] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An application specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 660–670, 2002.

[10] N. E. H. Bourebia and C. Li, "A greedy energy efficient clustering scheme based on reinforcement learning for WSNs," *Peer-to-Peer Networking and Applications.*, vol. 15, pp. 2572–2588, 2022.

[11] F. F. Jurado Lasso, J. F. Jurado, and X. Fafoutis, "A centralized reinforcement learning framework for adaptive clustering with low control overhead in IoT networks," *arXiv preprint*, arXiv:2401.15767, 2024. [Online]. Available: https://arxiv.org/abs/2401.15767

[12] C. Yan, Y. Deng, and Y. J. Choi, "A novel deep reinforcement learning based clustering scheme for WSN," in *IEEE Global Commun. Conf. (GLOBECOM), IoT and Sensor Networks Symposium*, 2023, pp. 1–6.

[13] K. Khedhiri, I. Ben Omrane, D. Djabour, and A. Cherif, "Clustering for lifetime enhancement in wireless sensor networks," *Telecom*, vol. 6, no. 2, Art. 30, 2025.

[14] V. K. Sunanda and J. V., “Survey on dynamic clustering for energy efficient data aggregation technique using secure data encoding scheme for WSN,” *SSRN Electron. J.*, Feb. 15 2014. Available: https://ssrn.com/abstract=4876252

[15] A. B. M. Al Islam, C. S. Hyder, H. Kabir, and M. Naznin, "Finding the optimal percentage of cluster heads from a new and complete mathematical model on LEACH," *Wireless Sensor Network.*, vol. 2, no. 2, pp. 129–140, 2010.

[16] J. Bjorck, C. P. Gomes, and K. Q. Weinberger, "Is high variance unavoidable in RL? A case study in continuous control," *arXiv preprint*, arXiv:2110.11222, 2021. [Online]. Available: https://arxiv.org/abs/2110.11222