Minimizing node failure in Wireless Rechargeable Sensor Network by utilizing GraphSAGE and Deep Q-Learning

Vuong Dinh An
Supervisor: Assoc. Prof. Huynh Thi Thanh Binh

School of Information and Communication Tecnology

Wireless Sensor Networks (WSNs)

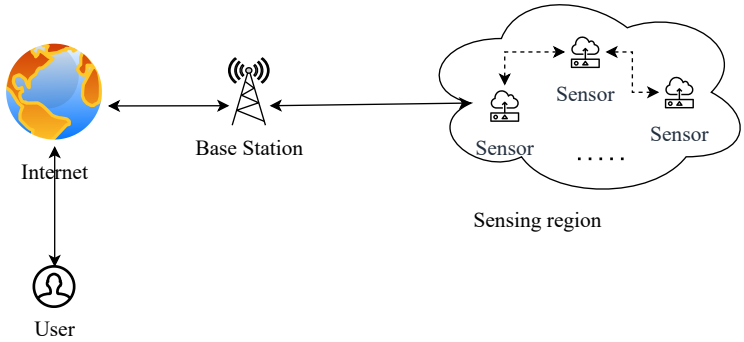


Figure 1: An example architecture of WSNs

Applications of WSNs

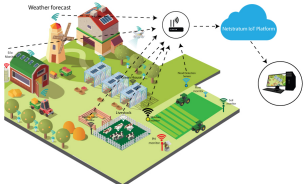


Figure 2: Agriculture

Figure 3: Healthcare system

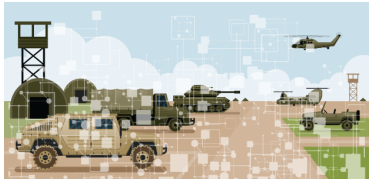


Figure 4: Military

Figure 5: Smart city

Challenges in WSNs

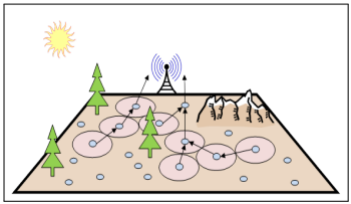


Figure 6: Coverage

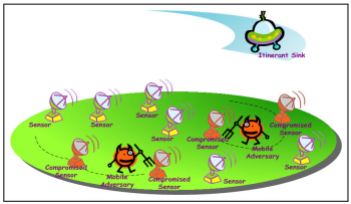


Figure 7: Security

Figure 8: Energy

Wireless Rechargeable Sensor Networks (WRSNs)

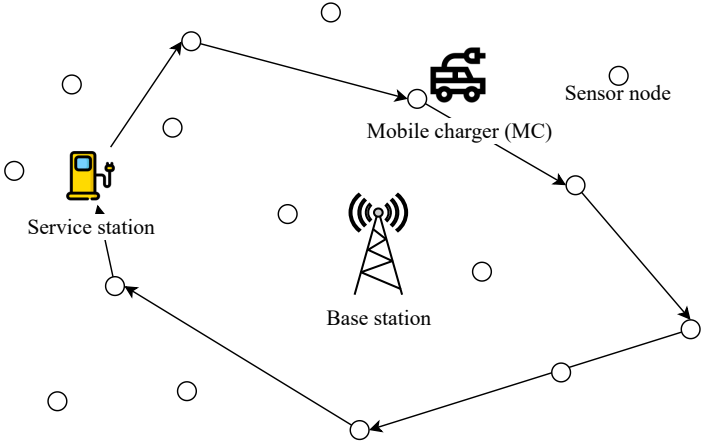


Figure 9: An example of WRSNs

The charging scheme optimization problem

- The sensors' lifetime is decided by the MC's charging scheme[1].
- A charging scheme consists of two essential factors:

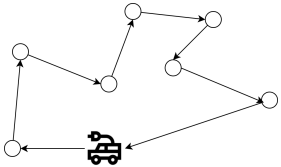


Figure 10: Charging path

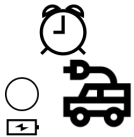


Figure 11: Charging time

[1] Z. Lyu, Z. Wei, J. Pan, H. Chen, C. Xia, J. Han, and L. Shi, "Periodic charging planning for a mobile wce in wireless rechargeable sensor networks based on hybrid pso and ga algorithm," Applied Soft Computing, vol. 75, pp. 388–403, 2019

Related Works

- Cao et al.[2] designed a reinforcement learning algorithm to determine optimal charging route.
- Lyu et al.[3] proposed a periodic scheme to maximize the ratio of the MC's vacation time at depot.
- Kaswan et al.[4] proposed a charging scheme which objective is to minimize the charging delay based on gravitational search .

[2] X. Cao, W. Xu, X. Liu, J. Peng, and T. Liu, "A deep reinforcement learning- based on-demand charging algorithm for wireless rechargeable sensor net- works," Ad Hoc Networks, vol. 110, p. 102 278, 2021.

[3] Z. Lyu, Z. Wei, J. Pan, H. Chen, C. Xia, J. Han, and L. Shi, "Periodic charging planning for a mobile wce in wireless rechargeable sensor networks based on hybrid pso and ga algorithm," Applied Soft Computing, vol. 75, pp. 388–403, 2019

[4] A.Kaswan, A.Tomar and P.K.Jana, "An efficient scheduling scheme for mobile charger in on-demand wireless rechargeable sensor networks," Jounal of Network and Computer Applications, vol.114, pp. 123-134, 2018

Related Works

Common limitations:

- Most works had unpractical assumptions: the MC's unlimited battery capacity and unified average energy consumption of sensors.
- Most works completely ignored the dynamics of the network, which is unreasonable since the network is constantly changing.
- Most works concentrated on small-sized network with insignificant number of sensors.

The main goals of the thesis:

- Investigate large-scale dense networks using multi-point charging scheme.
- Minimize the node failures of WRSNs.

Network settings

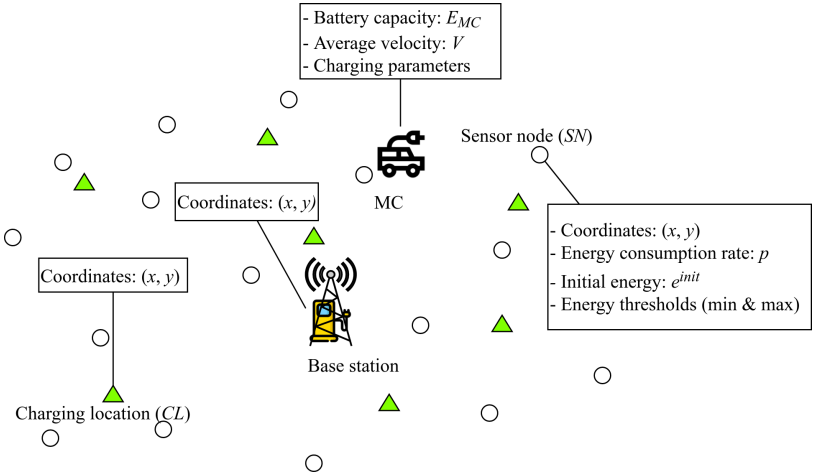


Figure 12: The input information of the network

Problem statement

Input: A Wireless Rechargeable Sensor Network.

Output:

- The MC's charging scheme: a sequence of charging locations.

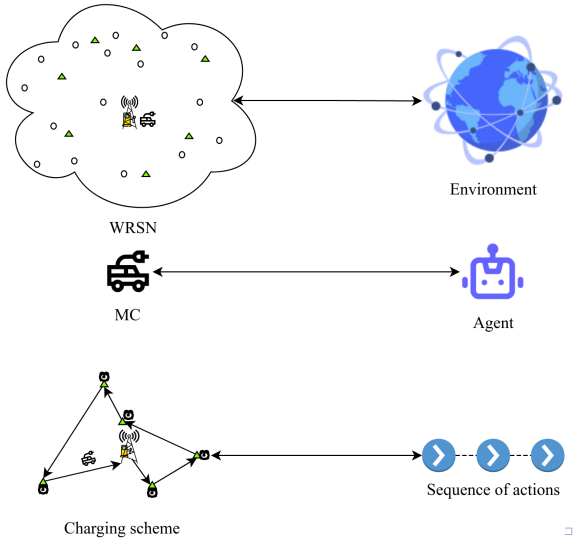
Objective:

- Minimize the number of energy-depleted sensors (node failures).
- Maximize the received energy of sensors.

Constraints:

- Total time spent for charging and travelling must not exceed the time period.
- Total energy spent for charging and travelling must not exceed the battery capacity of the MC.
- During charging, the energy of each sensor must stay in the predetermined range.

View this problem as a decision-making problem



Markov Decision Process (MDP) formulation (1/2)

State. Each timestep, the agent observes the state as the form of:

$$\hat{S}_t = \begin{bmatrix} e_1^t & e_2^t & \dots & e_n^t \\ d_{CL_j,SN_1} & d_{CL_j,SN_2} & \dots & d_{CL_j,SN_n} \end{bmatrix} \tag{1}$$

Action:

$$A_t = \text{index of next location} \tag{2}$$

Reward: The reward is defined as the difference between two states obtained by a transition:

$$R_{t+1} = F_{t+1} - F_t \tag{3}$$

Objective: The objective has the form of:

$$F_t = \frac{\min_{i \in \{1,2,\dots,n\}} e_i}{\max_{i \in \{1,2,\dots,n\}} e_i} \tag{4}$$

Markov Decision Process (MDP) formulation (2/2)

Episode: $(\hat{S}_0, A_0, \hat{S}_1, A_1, \ldots, \hat{S}_T)$, where $\hat{S}_0, \hat{S}_T$ are the state when the MC stays at the base station, $(A_0, A_1, \ldots)$ is the sequence of charging locations.

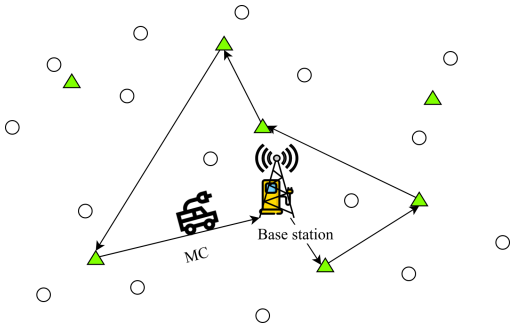


Figure 13: An example of an elapsing episode

Overview

The learning process of the agent is segregated into:

- Learning an embedding model of the state space by utilizing GraphSAGE[5].
- Training the agent to output the optimal policy by utilizing Deep Q-Learning[6].

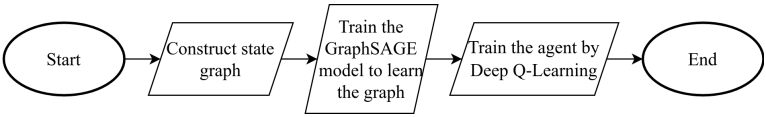


Figure 14: Abstract flowchart of the proposed algorithm

[5]W. Hamilton, Z. Ying, et al., "Inductive representation learning on large graphs," NIPS, vol. 30, 2017.

[6]V. Mnih, K. Kavukcuoglu, D. Silver, et al., "Human-level control through deep reinforcement learning," nature, vol. 518, no. 7540, pp. 529–533, 2015.

Construct a graph representing the state space

- For the i-th charging location, construct a cluster of nodes $\mathcal{C}_i = \{C_i^0, C_i^1, C_i^2, \ldots, C_i^{\frac{T}{\Delta}}\}$, where Δ is the discrete unit time interval.
- There will be an edge from C_i^t to $C_j^{t'}$ if:

$$t\Delta + \frac{d_{CL_i, CL_j}}{V} \in [t'\Delta, (t'+1)\Delta] \quad (5)$$

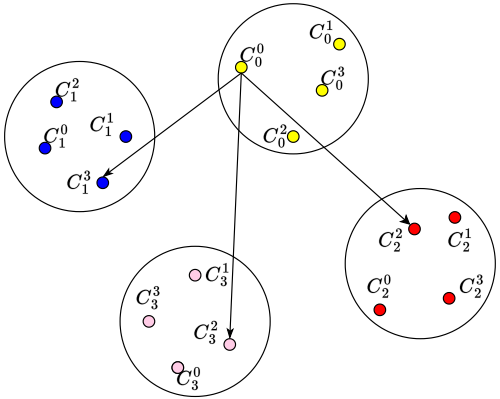


Figure 15: An example of constructing edges

Learning the graph

- GraphSAGE is utilized to learn the constructed graph.
- After learning, the model is able to transfer the initial vector of the state:

$$\hat{S}_t = \begin{bmatrix} e_1^t & e_2^t & \dots & e_n^t \\ d_{CL_j,SN_1} & d_{CL_j,SN_2} & \dots & d_{CL_j,SN_n} \end{bmatrix} \tag{6}$$

to a more compact vector S_t, which will be fed into the observation of the agent.

Training the agent (1/2)

- A Deep Q-Learning-based algorithm is adopted to control the agent with the goal of maximizing total cumulative rewards.
- For each timestep, the neural network takes an input as: $x_t = S_t$ and output the estimate of the action-value function: $\hat{Q}(S_t, a)$.

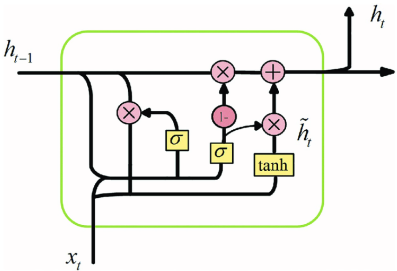


Figure 16: The neural network architecture (GRU) of the agent

Training the agent (2/2)

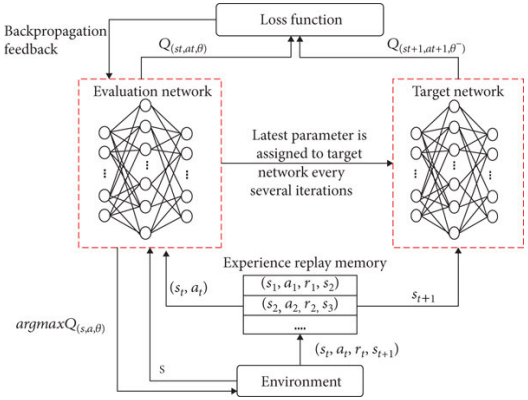


Figure 17: The process of training the agent using Deep Q-Learning [7]

[7] Li, Lincan et al. (2020). A Smart Cache Content Update Policy Based on Deep Reinfocement Learning. WCMC. 2020. 10.1155/2020/8836592.

Simulation settings (1/2)

- The algorithm is called GSADQL (Hybrid **G**raph**SA**GE-based and **D**eep **Q**-**L**earning Algorithm). We compare our proposal with most relevant works: INMA [8], and GR-GACS[9].
- The used metric is the node failure ratio, i.e. the ratio of number dead nodes to the total number of sensors:

$$\text{The node failure ratio (\%)} = \frac{\text{Number of energy-depleted sensors in a period}}{\text{Number of deployed sensors}} \times 100\% \tag{7}$$

[8] J. Zhu, Y. Feng, M. Liu, G. Chen, and Y. Huang, "Adaptive online mobile charging for node failure avoidance in wireless rechargeable sensor networks," Computer Communications, vol. 126, pp. 28–37, 2018.

[9] T. T. Huong, H. T. T. Binh, P. Le Nguyen, D. C. T. Long, V. D. An, et al., "Optimizing charging locations and charging time for energy depletion avoidance in wireless rechargeable sensor networks," in 2020 IEEE Congress on Evolutionary Computation (CEC), IEEE, 2020, pp. 1–8.

Simulation settings (2/2)

Server	Google Colab Pro
GPU	P100
Programing Language	Python
RAM	24GB

Table 1: Environment details

Parameter	**Value**
E_{max}	10800 J
E_{min}	540 $J = 0.05 \times E_{max}$
U	5 J/s
E_{MC}	108000 J
P_M	1 J/s
V	5 m/s

Table 2: Parameters of energy model

Parameter	Value
γ	0.95
Target Update	10
Number of episodes	50
$\mathcal{T}$	200000
Δ	40

Table 3: Base parameters of GSADQL

Learning process of the GSADQL agent

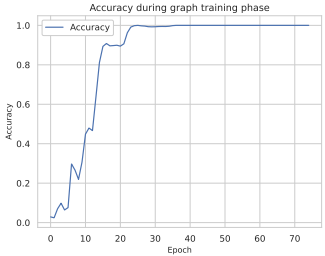


Figure 18: Accuracy of GraphSAGE model when learning the state space

Figure 19: Loss of GraphSAGE model when learning the state space

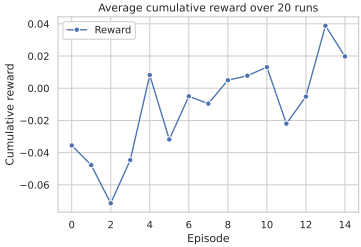


Figure 20: Accumulative reward of the agent

Impact of Δ on the performance of the proposed algorithm

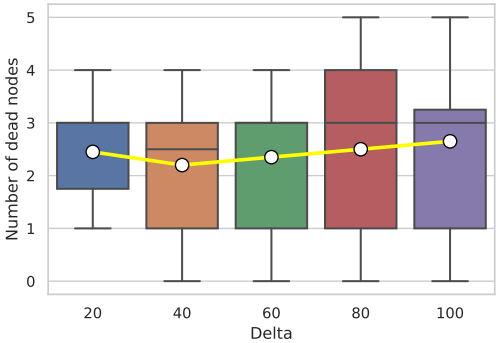


Figure 21: Interquartile range of number of dead nodes in 20 runs per changing Δ

Experimental evaluation

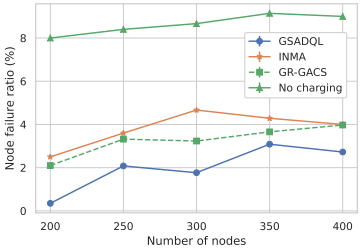


Figure 22: Impact of the number of sensors with respect to the three mentioned distributions.

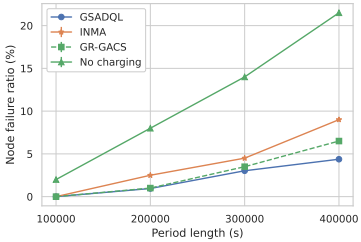


Figure 23: Impact of the average energy consumption rate of sensors

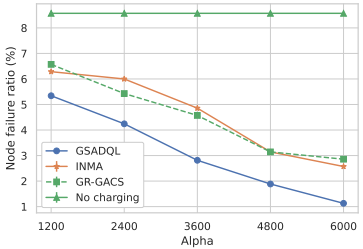
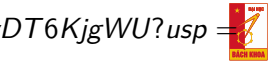


Figure 24: Impact of the average energy consumption rate of sensors

Conclusion

- This thesis tackles the problem of minimizing the energy depletion in WRSN by leveraging an online multi-node charging scheme that optimizes the charging route.
- This thesis proposes a novel MDP formulation for the nodes failure avoidance problem.
- This thesis proposes a GraphSAGE and Deep Q-Learning hybrid algorithm, which finds the optimal policy for the modelled MDP.
- The results of thorough experiments suggest that the proposal outperforms other state-of-the-art methods which have the goal of minimizing the number of energy-depleted sensors.
- Parts of this thesis were accepted as papers of 2020 IEEE Congress on Evolutionary Computation (CEC): **Huong, T. T., Binh, H. T. T., Le Nguyen, P., Long, D. C. T., & An, V. D. (2020, July). Optimizing charging locations and charging time for energy depletion avoidance in wireless rechargeable sensor networks. In 2020 IEEE Congress on Evolutionary Computation (CEC) (pp. 1-8). IEEE.**

- All the settings and programmings can be found at: https://colab.research.google.com/drive/1pSus$_E$d − zLqA07EWjktEQmSvDT6KjgWU?usp = sharing



THANK YOU
FOR YOUR ATTENTION!