JOINT MOBILE SINK SCHEDULING AND DATA AGGREGATION IN ASYNCHRONOUS WIRELESS SENSOR NETWORKS USING Q-LEARNING

Surender Redhu, Pratyush Garg and Rajesh Hegde

Indian Institute of Technology Kanpur, India
Email: { redhu, pratyush, rhegde }@iitk.ac.in

ABSTRACT

Energy-efficient data aggregation is a challenging problem in asynchronous wireless sensor networks. Asynchronous behaviour of sensor nodes is generally due to adaptive duty cycling and it leads to information loss, buffer overflow and poor quality of services. To overcome these issues, a joint mobile sink scheduling and data aggregation scheme is proposed in this work. A reinforcement learning framework is developed herein for budgeting the energy of mobile sink while minimizing the information loss in each cluster of a clustered WSN. More specifically, a Q-learning approach is used to learn the network behaviour over time and compute adaptive halt-times for the mobile sink based on active number of nodes in each cluster. Experiments on joint mobile sink scheduling and data aggregation are conducted on a medium scale WSN. Experimental results indicate that proposed method minimizes the information loss in an asynchronous wireless sensor network. It is also observed that mobile sink performs the data gathering operation with limited energy consumption while maximizing network lifetime.

Index Terms— Data Aggregation, Mobile Sink Scheduling, Reinforcement Learning, Wireless Sensor Networks

1. INTRODUCTION

Wireless Sensor Networks (WSNs), the backbone of 'Internet of Things' is being empowered by autonomous and intelligent devices [1, 2]. However, network lifetime is a major concern in the development of resilient and reliable sensor networks [3, 4]. In general, the energy-efficiency in WSNs can be accomplished using various approaches like clustering of nodes, mobile sink deployment, in-network data aggregation, routing, adaptive duty cycling and many more [5–7]. These schemes are also used in cooperative ways to improve the network performance further [8]. However, scheduling the mobile sink in an asynchronous WSN is very challenging because the number of active nodes vary with time. In this situation, energy-efficient data aggregation and mobile sink scheduling becomes very challenging [9]. Therefore, scheduling of mobile sink should be adaptive in accordance with asynchronous behaviour of the network. For ease of mobile sink scheduling, the network can be partitioned into a number of clusters [8]. Mobile sink traverses the whole network via clusters and associated way-points to collect the data from sensor nodes. The halt-time or time for mobile sink to stay in a cluster depends on total number of nodes in that cluster. In a well synchronized and clustered WSN, mobile sink can be provided with a fixed halt-time per cluster. However, in asynchronous and clustered network, fixed halt-time of mobile sink in a cluster leads to information loss, buffer overflow and poor quality of services. Hence, an intelligent mobile sink scheduling is required in this type of network scenario.

In this work, a method for joint mobile sink scheduling and data aggregation is proposed using reinforcement learning. In the proposed method, a reinforcement learning based framework is used for mobile sink scheduling. Mobile sink will learn the asynchronous behaviour of the network while simultaneously performing the data aggregation process. The traversing order of clusters is decided based on minimum tour length of mobile sink in the network. Therefore, with a limited energy consumption, mobile sink covers the whole network assuring minimum information loss in every cluster. The adaptive halt-time of mobile sink in each cluster plays a key role here. The variation in halt-time in a cluster depends on the number of active nodes in that cluster, residual energy of the mobile sink and the information loss in the network. Based on the current state of the network, mobile sink decides whether to move to next cluster or stay in present cluster.

The proposed method utilizes various advantages of clustering, mobile sink and asynchronous behaviour of nodes in maximizing the energy-efficiency. Hence, the proposed joint mobile sink scheduling and data aggregation prolongs the network lifetime while minimizing the information loss. The contributions of this paper are as follows. A deep reinforcement learning framework is implemented to model an intelligent mobile sink for wireless sensor networks. A cooperative and energy-efficient data gathering approach is proposed utilizing the strengths of clustering, mobile sink and asynchronous behaviour of the sensor nodes. Intelligent mobile sink scheduling is used to minimize the information loss in the network.

The rest of the paper is organized as follows. Section 2 presents the network model and problem definition. Design and implementation of intelligent mobile sink is presented in Section 3. The performance evaluation is presented in Section 4. Finally, the work is concluded in Section 5.

2. NETWORK MODEL AND PROBLEM DEFINITION

The network consists of a large number of sensor nodes that function based on an adaptive duty cycling algorithm [7]. The network is divided into a number of clusters. In the presence of mobile sink, all nodes of a cluster can transmit data to it. The network assumed with uneven distribution of nodes on the field may have different number of nodes in each cluster. A mobile sink is deployed in this clustered and asynchronous network for energy-efficient data gathering. Mobile sink traverses the whole network via clusters and their way-points. All the nodes in a cluster are single-hop away from the mobile sink present on the way-point of that cluster. An example of energy-efficient data collection scheme utilizing intelligent mobile sink is shown in Figure 1. Mobile sink stays in every cluster with

This work was supported by TCS Research Scholarship Program under project TCS/EE/2011191K.

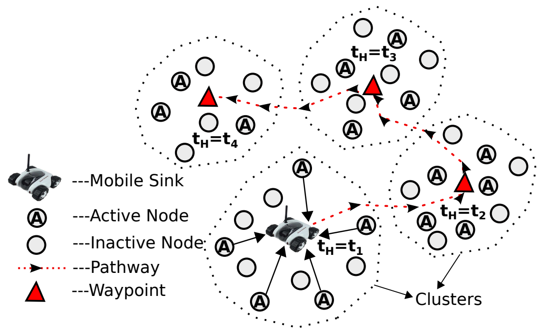

Fig. 1. Illustration of energy-efficient data collection scheme over clustered, asynchronous WSN. An adaptive halt-time (t_H) is assigned to each cluster.

different halt-time(t_H). The adaptive halt-time of mobile sink at any way-point depends on the active number of nodes in the present cluster.

2.1. Problem Definition

Given a wireless sensor network $W(N, C)$ with total number of nodes N and partitioned into C number of clusters. A mobile sink is deployed in the clustered and unknown asynchronous behavioured network with total energy $\mathcal{E}_{ms}^T$. Hence, mobile sink budgets its residual energy among all clusters depending on the number of active nodes in the present cluster and the active nodes expected to be encountered in the future. The problem is how to jointly schedule the mobile sink and aggregate data so that mobile sink can decide the halt-time adaptively to each cluster according to the network conditions. Hence, the problem can be formulated as

$$\begin{aligned}
&\min \ f\left(\mathcal{EC}_{network}^T, \mathcal{I}_{loss}^T\right) \\
&\text{subject to} \\
&(\mathcal{N}_c, \mathcal{W}_c) \in \Gamma_w \quad \text{for} \quad c = 1, 2, 3, ..., C \\
&\{n_i \mid \mathcal{H}_l(\mathcal{W}_c, n_i) = 1\} \in \mathcal{N}_c \qquad (i) \\
&\mathcal{E}_{tour} \leq \mathcal{E}_{ms}^T \qquad (ii) \\
&\mathcal{EC}(c) \leq \mathcal{B} \cdot \mathcal{A}_c \cdot \mathcal{E}_{elec} \qquad (iii) \\
&t_H(c) = \frac{\mathcal{B} \cdot \mathcal{A}_c}{\mathcal{R}_r} \qquad (iv)
\end{aligned} \tag{1}$$

where $\mathcal{EC}_{network}^T$ is total energy consumption in the network. $\mathcal{I}_{loss}^T$ is the total information loss in the network. Each cluster $\mathcal{N}_c$ with a way-point $\mathcal{W}_c$ is a part of the trajectory Γ_w for a mobile sink. Clusters are maintained in such a way that its way-point is one-hop away from every node n_i of that cluster. $\mathcal{H}$ defines the hop-length between two nodes on a graph or network. Energy consumed in a tour or round $\mathcal{E}_{tour}$ by the mobile sink is constrained to be less than the total available energy $\mathcal{E}_{ms}^T$. The energy consumption $\mathcal{EC}(c)$ in a c^{th} cluster depends on the the memory buffer size of each sensor node $\mathcal{B}$ and the number of active nodes $\mathcal{A}_c$ in that cluster.. $\mathcal{E}_{elec}$ is the energy consumed in receiving one packet of information. The halt-time t_H for each cluster is adaptive depending on the number of active nodes. Hence, $t_H(c)$ of a cluster is ratio of the amount of information to be received and the reception rate of the mobile sink $\mathcal{R}_r$. The various constraints given in equation 1 helps in developing the intelligent mobile sink. The constraint (i) bounds the movement of mobile sink to the fixed way-points. The energy consumption in every tour is limited by constraint (ii). Adaptive halt-time and managing with residual energy of the mobile sink is dealt with constrains (iii) and (iv). Energy consumption and information loss for the proposed network model are defined as follows.

Energy Consumption, $\mathcal{EC}_{network}^T$: It is the sum of energy consumed by mobile sink and the sensor nodes. Hence, the minimization of $\mathcal{EC}_{network}$ results in efficient utilization of mobile sink as well as prolonged network lifetime.

$$\mathcal{EC}_{network}^T = \mathcal{EC}_{ms} + \mathcal{EC}_{nodes} \tag{2}$$

where $\mathcal{EC}_{ms}$, $\mathcal{EC}_{nodes}$ are the energy consumption by mobile sink and sensor nodes respectively.

Information Loss, $\mathcal{I}_{loss}^T$: In a mobile sink equipped network, it is the amount of information that was available but not collected by mobile sink because of various issues like delay and the unavailability of residual energy in mobile sink. Total information loss in a clustered network is sum of information loss in an individual cluster and is given as

$$\mathcal{I}_{loss}^T = \sum_{c=1}^{C} I_{loss}(c) \tag{3}$$

where $I_{loss}(c)$ is the information loss in c^{th} cluster.

3. JOINT MOBILE SINK SCHEDULING AND DATA AGGREGATION USING Q-LEARNING

The presence of asynchronous behaviour which is caused by an unknown adaptive duty cycling protocol or inaccuracy in synchronization protocols needs an adaptive scheduling for the mobile sink. Hence, mobile sink is required to stop adaptively in each cluster for collecting the information. In this way, an intelligent mobile sink needs to be modelled. A reinforcement learning framework helps in modelling an agent that learns from the environment and acts smartly [10]. A deep reinforcement learning framework is modelled for continuous controlling of robot in [11]. The state-action-reward framework of reinforcement learning in modelling the intelligent mobile sink is framed as follows.

State, S_t: For the given network model, the state S_t is defined in terms of network conditions at time t which are: current cluster details $\mathcal{N}_c$, $\mathcal{I}_{left}^c (= \mathcal{I}_{total}^c - \mathcal{I}_{collected}^c)$ and the residual energy of the mobile sink $\mathcal{E}_{ms}^c$ at that cluster. The information available initially at c^{th} cluster is $\mathcal{I}_{total}^c = \mathcal{B} \times \mathcal{A}_c$ with $\mathcal{B}$ as the buffer size of the node and $\mathcal{A}_c$ as the active nodes in the cluster. Therefore, each way-point is associated with the state of its cluster, $S_t = [\mathcal{N}_c, \mathcal{I}_{left}^c, \mathcal{E}_{ms}^c]$.

Action, A_t: Based on the state $S_t(c)$, a change in pre-associated halt-time $t_H(c)$ of the c^{th} cluster is decided. Therefore, the available actions are, $A_t = \{move, halt\}$. Whenever a *halt* action is executed, the mobile sink collects data from the current cluster for a pre-decided time interval t_{unit} at the end of which, a new state is reached. Thus the total halt-time $t_H(c) = n \times t_{unit}$ where n is the number of *halt* actions taken at the cluster.

Reward, R_{t+1}: The reward which is the function of consequences to an action taken, is defined in terms of the objectives or goals of the problem. Therefore, it is defined as a function of energy consumption and information loss. An increase in the reward should indicate either a fall in the information loss or a fall in the energy consumed or both. The formulation of this function is explained in next subsection.

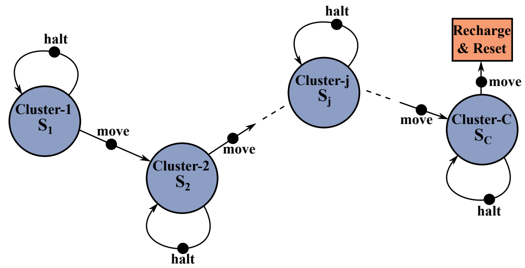

Fig. 2. A state-action space model for joint mobile sink scheduling and data aggregation.

3.1. Q-Learning Approach

The joint mobile sink scheduling and data aggregation decides the optimal traversal of the graph formed by the clusters and stop time at each waypoint according to the network conditions. To model the adaptive scheduling of mobile sink, a Q-learning approach of reinforcement learning is developed herein. In Q-learning, the value $Q(S_t, A_t)$ of an action A_t in S_t state at current time t is updated based on the value of next state $Q(S_{t+1}, A_{t+1})$. The updated action value can be given as

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha[R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)] \tag{4}$$

where α is the step-size which assures the convergence of estimates of future value function. γ is discount rate parameter $(0 \leq \gamma \leq 1)$ by which it determines the worth of future rewards with respect to the present. However, since the number of possible states is considerable (it is a 3-tuple), this approach needs to be augmented with neural networks and deep learning in order to be implemented. Therefore, for appropriate modelling, the proposed method uses the deep Q-network (DQN) algorithm presented in [12].

3.1.1. Motivating the Reward

The reinforcement learning framework operates on the notion of the *favourable* condition which is represented by a positive reward and the *unfavourable* conditon that warrants a negative reward. In our context, a *favourable* condition is when the usage of energy is optimized with the collection of information. Before formulating the reward, a function that will model the desire of proposed model behaviour is defined as

$$f = b_1 \cdot \underbrace{(I^c_{left}/I^c_{total})}_{\text{NIL}} + b_2 \cdot \underbrace{(\mathcal{E}^c_{ms}/\mathcal{E}^T_{ms})}_{\text{NRE}} \tag{5}$$

where b_1, b_2 are scaling parameters. The first and second bracketed terms respectively represent the normalized valued of the information loss (NIL) (should the sink were to move) and the normalized residual energy (NRE). We proceed to model the desired behaviour by following an intuitive policy presented in Table 1.

3.1.2. Reward Formulation

Table 1 shows that a high value of the function f is good for *halt* actions but is bad for *move* actions. Therefore, the instantaneous

Table 1. Intuitive Policy for Modelling Desired Behaviour

Action	NIL	NRE	Function 'f'	Condition
move	$\uparrow$	$\uparrow$	Very High	*unfavourable*
	$\downarrow$	$\uparrow$	High	*less unfavourable*
	$\uparrow$	$\downarrow$	Low	*slightly favourable*
	$\downarrow$	$\downarrow$	Very Low	*favourable*
halt	$\uparrow$	$\uparrow$	Very High	*favourable*
	$\downarrow$	$\uparrow$	High	*slightly favourable*
	$\uparrow$	$\downarrow$	Low	*unfavourable*
	$\downarrow$	$\downarrow$	Very Low	*unfavourable*

reward as a function of the state and the action taken is defined as:

$$R_{t+1}(S_t, A_t) = \begin{cases} +f - \lambda; & A_t = halt \\ -f + \lambda; & A_t = move \end{cases} \tag{6}$$

where λ is a tuning constant s.t. $0 \leq \lambda \leq (b_1 + b_2)$ and represents the turning point on the function below which values start getting *favourable* for *move* but *unfavourable* for *halt*. Hence, the reward function formulated above, gives a positive value whenever the condition grows favourable and returns negative values otherwise.

Consequently, maximizing the cumulative reward $(\sum_t R_{t+1}.)$ will result in solving the problem defined in Section 2.

4. PERFORMANCE EVALUATION

In this section, performance evaluation of the proposed method of joint mobile sink scheduling and data aggregation is presented. Improvements in energy-efficiency and information loss in an asynchronous WSN are studied in detail.

4.1. Network Model and Parameters

A wireless sensor network with 520 nodes is deployed in the environment. The network is assumed to be divided into 7 clusters with different number of nodes, $\{60, 70, 80, 75, 70, 85, 80\}$. As the behaviour of the network is asynchronous, the active number of nodes varies with time in each cluster. The asynchronous behaviour level (uncertainty) is the measure of variance of active number of nodes in the cluster which are sampled using a truncated Gaussian distribution. The uncertainty bounds the upper and lower limits on active number of nodes in a cluster. The list of hyperparameters used in mobile sink scheduling is given in Table 2.

4.2. Mobile Sink Scheduling Analysis

The performance analysis of the modelled intelligent mobile sink is presented in Figure 3. The reinforcement learning algorithm tries to maximize the reward and in doing so minimizes information loss $\mathcal{I}^T_{loss}$ and the energy consumed $\mathcal{EC}_{ms}$ as given in the problem definition. It can be clearly noted that after a certain number of episodes, the mobile sink starts converging on a scheduling policy. The convergence of cumulative reward is shown in Figure 3 (a). The total energy consumption of mobile sink $\mathcal{EC}_{ms}$ and information loss $\mathcal{I}^T_{loss}$ of the network with uncertainty less than 10% is presented in Figure 3 (b) and (c) respectively. With the increase in uncertainty in the network behaviour, total information loss increases and it takes more time to converge as shown in Figure 3. However, mobile sink is still capable of learning the environment with a penalty for information loss.

Table 2. Hyperparameters Values for Mobile Sink Scheduling

Parameters	**Value**
Learning rate parameter, α	0.0001
Discount rate, γ	0.99
Batch size	20
Initial exploration	1
Exploration decay rate	0.0007
Final exploration	0.2

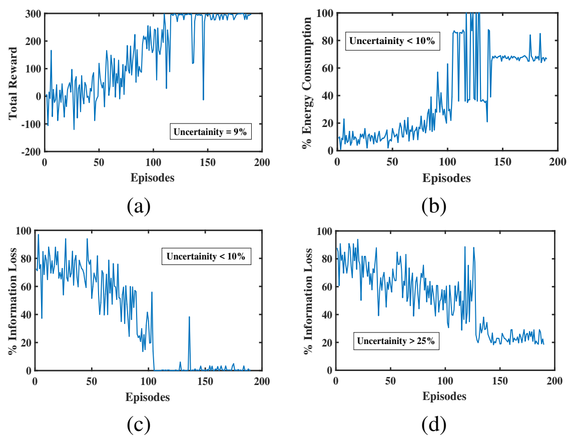

Fig. 3. Mobile sink scheduling analysis: convergence of (a) Total reward obtained by mobile sink after every episode (tour), (b) Energy consumption (c) Information loss after each episode with the asynchronous (uncertainty) behaviour level $< 10\%$. The information loss with high uncertainty ($> 25\%$) is shown in fig (d).

4.3. Data Aggregation and Energy-Efficiency Analysis

In the proposed method, data aggregation and energy-efficiency is improved using the advantages of single-hop clustering, intelligent mobile sink scheduling and adaptive duty cycling of the network. With an autonomous and intelligent mobile sink, active nodes of a cluster transmit their data to the mobile sink when it reaches a particular cluster. Mobile sink covers the network with a limited energy consumption and minimum information loss. For an asynchronous WSN of 520 nodes, the information loss using a fixed amount of energy for a trained mobile sink is presented in Figure 4. Figure 4 (a) shows the comparison of information loss using intelligent mobile sink scheduling and conventional scheduling. The upper limit on energy consumption of mobile sink is considered 36%. As this limit is becoming stronger (28%) in Figure 4 (b), more information loss occurs in conventional mobile sink sink scheduling. In conventional sink scheduling scheme, halt-time of mobile sink is fixed for each waypoint [13] which results in more information loss. The energy consumption analysis to achieve less than 5% information loss in an asynchronous network is presented in Figure 5. It can be noted that after an uncertainty level of 40%, mobile sink fails to achieve an information loss less than 5%. However, it is successfully achieved using the proposed intelligent mobile sink scheduling scheme.

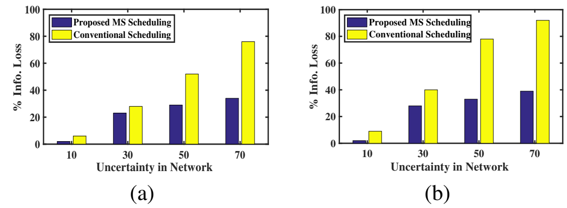

Fig. 4. Information loss analysis. (a) Comparison of information loss using the proposed intelligent scheduling and conventional scheduling with a maximum of 36% energy consumption of mobile sink, (b) With a maximum of 28% of energy consumption in an asynchronous network.

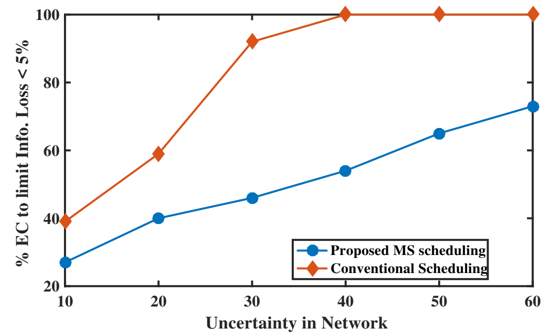

Fig. 5. Energy consumption analysis. Comparison of energy consumption using proposed sink scheduling and conventional scheduling to achieve less than 5% information loss in an asynchronous network.

5. CONCLUSION

In this work, we have proposed a method to develop a joint mobile sink scheduling and data aggregation which learns the network behaviour of asynchronous WSNs over the time. For this, a deep Q-learning approach is used in the framework of asynchronous WSN. It is noted that mobile sink reduces the information loss and improves the network lifetime. Significant improvements are obtained for various uncertainty level of asynchronous behaviour. Performance evaluation shows that mobile sink is converging with the network behaviour after a fixed number of episodes. The obtained results are motivating enough to bring the proposed method into practice of autonomous 'Internet of Things'. Future work will focus on developing the learning framework for multiple mobile sink to improve the energy-efficiency further.

6. REFERENCES

[1] Cesare Alippi, Romano Fantacci, Dania Marabissi, and Manuel Roveri, "A cloud to the ground: The new frontier of intelligent and autonomous networks of things," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 14–20, 2016.

[2] Oladayo Bello and Sherali Zeadally, "Intelligent device-to-device communication in the internet of things," *IEEE Systems Journal*, vol. 10, no. 3, pp. 1172–1182, 2016.

[3] Ian F Akyildiz, Weilian Su, Yogesh Sankarasubramaniam, and Erdal Cayirci, "Wireless sensor networks: a survey," *Computer networks*, vol. 38, no. 4, pp. 393–422, 2002.

[4] Yunxia Chen and Qing Zhao, "On the lifetime of wireless sensor networks," *IEEE communications letters*, vol. 9, no. 11, pp. 976–978, 2005.

[5] Wendi Rabiner Heinzelman, Anantha Chandrakasan, and Hari Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *System sciences, 2000. Proceedings of the 33rd annual Hawaii international conference on*. IEEE, 2000, pp. 10–pp.

[6] Z Maria Wang, Stefano Basagni, Emanuel Melachrinoudis, and Chiara Petrioli, "Exploiting sink mobility for maximizing sensor networks lifetime," in *Proceedings of the 38th annual Hawaii international conference on system sciences*. IEEE, 2005, pp. 287a–287a.

[7] Lei Tang, Yanjun Sun, Omer Gurewitz, and David B Johnson, "Pw-mac: An energy-efficient predictive-wakeup mac protocol for wireless sensor networks," in *INFOCOM, 2011 Proceedings IEEE*. IEEE, 2011, pp. 1305–1313.

[8] S. Redhu and R. M. Hegde, "Energy-efficient landmark tracing in wsns using random walks on network graphs," in *2017 9th International Conference on Communication Systems and Networks (COMSNETS)*, Jan 2017, pp. 399–400.

[9] Shouling Ji and Zhipeng Cai, "Distributed data collection and its capacity in asynchronous wireless sensor networks," in *INFOCOM, 2012 Proceedings IEEE*. IEEE, 2012, pp. 2113–2121.

[10] Richard S Sutton and Andrew G Barto, *Reinforcement learning: An introduction*, vol. 1, MIT press Cambridge, 1998.

[11] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

[12] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[13] Hamidreza Salarian, Kwan-Wu Chin, and Fazel Naghdy, "An energy-efficient mobile-sink path selection strategy for wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 5, pp. 2407–2419, 2014.