Predictive Energy-Aware Adaptive Sampling with Deep Reinforcement Learning

Author(s):
Heo, Seonyeong; Mayer, Philipp; Magno, Michele

Publication date:
2022

Permanent link:
https://doi.org/10.3929/ethz-b-000587335

Originally published in:
https://doi.org/10.1109/ICECS202256217.2022.9971120

This article has been accepted for publication in the proceedings of the 2022 29th IEEE International Conference on Electronics, Circuits and Systems (ICECS). Content may change prior to final publication. DOI: 10.1109/ICECS202256217.2022.9971120

Predictive Energy-Aware Adaptive Sampling with Deep Reinforcement Learning

Seonyeong Heo, Philipp Mayer, Michele Magno
Dept. of Information Technology and Electrical Engineering, *ETH Zürich*, Switzerland
seonyeong.heo@pbl.ee.ethz.ch, mayerph@iis.ee.ethz.ch, michele.magno@pbl.ee.ethz.ch

***Abstract*—Energy harvesting can enable wireless smart sensors to be self-sustainable by allowing them to gather energy from the environment. However, since the energy availability changes dynamically depending on the environment, it is difficult to find an optimal energy management strategy at design time. One existing approach to reflecting dynamic energy availability is energy-aware adaptive sampling, which changes the sampling rate of a sensor according to the energy state. This work proposes deep reinforcement learning-based predictive adaptive sampling for a wireless sensor node. The proposed approach applies deep reinforcement learning to find an effective adaptive sampling strategy based on the harvesting power and energy level. In addition, the proposed approach enables predictive adaptive sampling by designing adaptive sampling models that consider the trend of energy state. The evaluation results show that the predictive models can successfully manage the energy budget reflecting dynamic energy availability, maintaining a stable energy state for a up to 11.5% longer time.**

***Index Terms*—Adaptive sampling, energy harvesting, energy management, wireless smart sensors, reinforcement learning**

I. Introduction

Energy harvesting enables wireless smart sensors to operate with a low maintenance cost by removing the necessity of frequent battery replacement [1]. Harvesting energy from ambient energy sources such as solar light and wind, wireless sensor nodes can charge their batteries and keep them from draining. With the energy, wireless sensor nodes can acquire and transfer sensor data for an intelligent service. While global attention to environmental issues is growing, energy harvesting is becoming a more popular option for wireless sensor networks [2].

Although energy harvesting wireless sensor nodes should manage energy consumption well to prevent service disruption, finding an optimal energy management strategy at design time is difficult because the amount of energy available for harvesting changes over time [3]. For example, in the case of solar energy, the solar irradiance may change dramatically depending on the weather [4]. Since the weather is barely predictable far in advance, precise solar energy prediction is hardly possible at design time. Therefore, wireless sensor nodes should reflect dynamic energy availability to enhance energy utilization [5].

One existing approach to dynamic energy management is energy-aware adaptive sampling, which changes the sampling rate of a wireless sensor node according to the energy state [6]–[10]. In wireless sensor nodes, each sampling task includes sensor reading, data processing, and wireless communication. Too frequent sensor sampling may drain the battery very quickly, especially in the presence of power-hungry sensors [6], [11]. Therefore, energy-aware adaptive sampling can help increase the lifetime of wireless sensor nodes by reducing the sampling rate when the energy availability is low [12].

Previous work proposes various approaches to energy-aware adaptive sampling, either manually designing an adaptive sampling algorithm [6]–[8] or using machine learning to find an effective adaptive sampling strategy [9], [10]. Especially, Fraternali et al. [10] and Aoudia et al. [9] apply reinforcement learning to find an optimal sensing rate, showing the potential of reinforcement learning for energy-aware adaptive sampling. However, the existing approaches consider the current energy state only, unaware of the trend of the energy state. For this reason, their adaptive sampling models are likely to be reactive to the current energy state.

This paper proposes a predictive energy-aware adaptive sampling method for wireless sensor networks. More in detail, this work applies deep actor-critic reinforcement learning to find an effective adaptive sampling strategy based on the history of harvesting power and battery level. The proposed method formulates a reinforcement learning problem and designs a simulation environment for an energy harvesting wireless sensor node. In addition, this work presents adaptive sampling models that consider the trend of energy states to determine the sampling rate.

To be effective, this work implements the reinforcement learning environment for predictive energy-aware adaptive sampling on top of existing reinforcement learning frameworks [13], [14]. To evaluate the performance of the adaptive sampling models, we train the adaptive sampling models with existing energy harvesting datasets and evaluate the models in terms of energy management. The evaluation results show that the predictive models can successfully manage the energy budget reflecting dynamic energy availability, maintaining a stable energy state for a up to 11.5% longer time.

This paper is organized as follows. Section II discusses related work on energy-aware adaptive sampling. Section III describes the proposed predictive adaptive sampling method, presenting the problem formulation for reinforcement learning and the adaptive sampling algorithm. Section IV presents the evaluation results of the proposed method. Section V concludes this work.

II. Related Work

Previous work has proposed various approaches to energy-aware adaptive sampling for optimizing the energy management of wireless sensor nodes [6]–[10].

Traditional Adaptive Sampling Algorithms: Extending an existing adaptive sampling algorithm [15], [16], Srbinovski et al. [6] propose the energy-aware energy adaptive sampling algorithm (EASA), which monitors the battery level and adjusts the sampling frequency when the battery level is too low. Similarly, Lee and Lee [7] introduce two different adaptive sampling algorithms called RASA and CASA. Similar to EASA, RASA adjusts the sampling frequency when the battery level is too low. CASA reduces the sampling frequency to save extra energy when the quality of energy harvesting is good enough. Unlike the previous algorithms [7], [15], Lee and Kim [8] propose an adaptive sampling algorithm that uses a trend estimation model to predict changes in energy sources based on the previous data.

Although the energy-aware adaptive sampling algorithms can enhance the lifetime of a wireless sensor node at a low cost [6]–[8], the traditional algorithms require manual adjustment of parameters for the target environment.

Machine Learning-based Adaptive Sampling: Fraternali et al. [10] and Aoudia et al. [9] use reinforcement learning to find an optimal sensing rate for wireless sensor networks. Fraternali et al. apply Q-learning [17] to automatically choose one of the four performance states with different sampling rates. Fraternali et al. define the energy state as a tuple of light intensity level, energy level, and a boolean indicating whether it is a weekday or the weekend. On the other hand, Aoudia et al. apply actor-critic reinforcement learning [18] to predict the target sampling (packet) rate directly. Aoudia et al. simply define the energy state as an energy level.

Although the existing approaches [9], [10] show the potential of reinforcement learning for energy-aware adaptive sampling, they only consider the current energy state to determine the sampling rate. Then, an adaptive sampling strategy is likely to be reactive to the current energy state.

III. Predictive Energy-Aware Adaptive Sampling

This section explains the system model and formulates a reinforcement learning problem to train energy-aware adaptive sampling models.

A. System Model

This work assumes a system model where a wireless sensor node connects to the base station, similar to previous work [6], [7]. As illustrated in Fig. 1, the wireless sensor node consists of a smart sensor, an energy harvester, and a battery. At the time t, the sensor node sends a sensor value s_t, the harvesting power P_t^h, and the battery level E_t^b to the base station. After getting the data from the sensor node, the base station determines the sampling frequency f for the sensor node and transfers it to the sensor node. After the period of f^{-1}, the sensor node sends the set of values again to the base station.

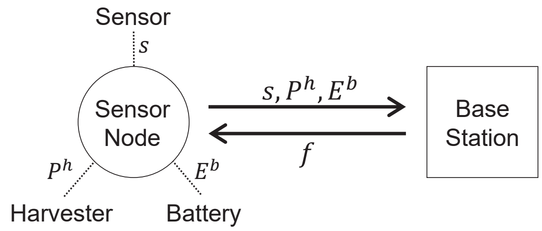


Fig. 1: System model of this work

B. Reinforcement Learning

The main goal of this work is to find an effective adaptive sampling strategy using deep reinforcement learning. In general, a reinforcement learning problem consists of an environment and an agent. At every step, an agent observes a *state* from the environment and takes an *action* based on the observation. After the action is taken, the environment gives a *reward* for the action to the agent. In reinforcement learning, an episode is a sequence of steps in which an agent takes actions in sequence and terminates in a certain condition.

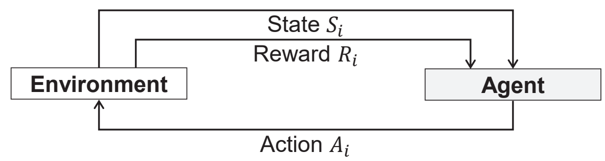


Fig. 2: General concept of reinforcement learning

1) Formulation: This work first defines the state, action, and reward for energy-aware adaptive sampling as follows.

- The state S_i is defined as a tuple of the current time t_i, the last sampling frequency f_{i-1}, the harvesting power $P_{t_i}^h$, and the battery level $E_{t_i}^b$ at step i. Note that in the system model, the sensor node transfers the state to the base station which determines the sampling rate.

$$S_i := (t_i, f_{i-1}, P_{t_i}^h, E_{t_i}^b)$$

- The action A_i can be one of the three actions: `Up`, `Stay`, and `Down`. Based on the recent history of the states, the agent may increase the sampling frequency (`Up`), keep the same sampling frequency (`Stay`), or decrease the sampling frequency (`Down`).

$$A_i := \texttt{Up} \mid \texttt{Stay} \mid \texttt{Down}$$

- The reward R_i is defined as the amount of time spent for the step plus one. Here, the additional reward of 1 is given for a single sensor acquisition.

$$R_i := 1 + (t_i - t_{i-1})$$

In our setting, an episode terminates when the average battery level becomes lower than the threshold. Therefore, if an agent maintains a stable energy state for a longer time, it will obtain a larger total reward.

Algorithm 1 Energy Harvesting Environment Simulation

Initialize k_0 to k_{max}
Initialize the first w states $S_0, ..., S_{w-1}$
while not $done$ **do**
    $A_i \leftarrow \texttt{Agent}(\{S_{i-w+1}, ..., S_i\})$
    if $A_i = \texttt{Up}$ **then**
        $k_i \leftarrow \max(k_{min}, k_{i-1} - 1)$
    else if $A_i = \texttt{Down}$ **then**
        $k_i \leftarrow \min(k_{i-1} + 1, k_{max})$
    end if
    $f_i \leftarrow f_{base}/k_i$
    while $t < t_{i-1} + 1/f_i$ **do**
        $t \leftarrow t + 1$
        $P_t^h \leftarrow$ Get a value from the dataset
        $E_t^b \leftarrow E_{t-1}^b + P_{t-1}^h - E_{t-1}^c$
        if $t = t_{max}$ or $\overline{E^b} < E^h$ **then**
            $done \leftarrow true$
            break
        end if
    end while
    $t_i \leftarrow t$
    $R_i \leftarrow 1 + (t_i - t_{i-1})$
    $i \leftarrow i + 1$
end while

2) Environment: For reinforcement learning, an energy harvesting environment is designed and simulated based on the system model. Algorithm 1 summarizes the simulation process of a single episode. First, it initializes the states and the divisor for the sampling frequency. After the initialization, it starts iterating through the main loop. At each iteration (or step), the agent chooses an action based on the recent history of states. Here, w indicates the length of window. According to the action, it adjusts the divisor and updates the sampling frequency. After adjusting the sampling frequency, it simulates the interval updating the energy state. If the simulation time is over or the average battery level falls below the threshold E^h during the interval, it terminates the episode. In this paper, we set the threshold to the initial energy level E_0^b.

For simplicity, the energy consumption of a sensor node is estimated assuming that the average power consumption of sensor sampling is constant. Then, the energy consumption at the time t, E_t^c is calculated as

$$E_t^c := P^a \cdot T_t^a + P^s \cdot (1 - T_t^a)$$

where P^a is the average power when the node acquires and transfers the sensor value, P^s is the average power when the node is in a sleep mode, and T_t^a is the amount of time taken to acquire and transfer the sensor value within $[t, t+1)$. This work chooses the parameters based on the energy characterization of a wireless sensor node.

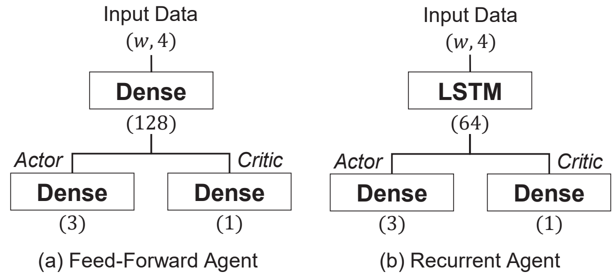


Fig. 3: Two different agent designs

3) Agent: This work designs two types of agents that determine the proper action for energy-aware adaptive sampling based on the recent history of states: (i) a feed-forward agent and (ii) a recurrent agent. Each reinforcement learning agent consists of small actor and critic models as illustrated in Fig. 3. The actor models take the normalized states as input and return the probabilities of the actions as output. Note that the output dimension of the actor models is three because the number of possible actions is three. The critic models take the same input and return the value of the state as output.

IV. Evaluation

A. Experimental Setup

To evaluate the reinforcement learning-based models, this work implements the reinforcement learning environment on top of OpenAI Gym [13]. In addition, this work implements the feed-forward and recurrent agents with Tensorflow [14]. For training, this work uses indoor solar harvesting datasets [19] and splits each dataset for training and testing. This work randomly chooses a two-week period from a dataset and simulates the period as one episode. This work trains each agent with 150 episodes on a desktop computer.

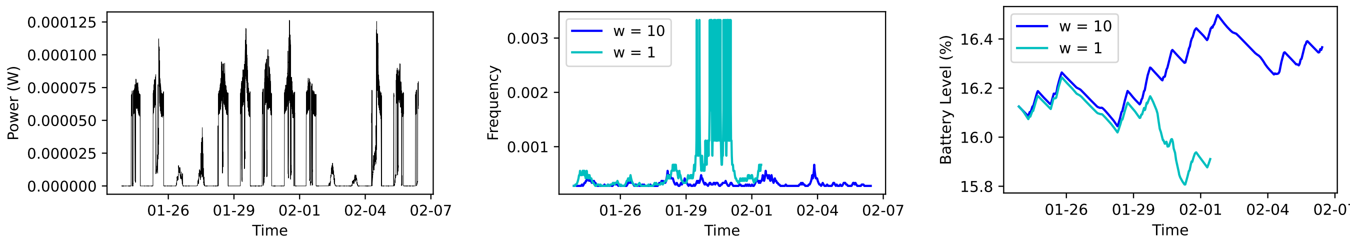
To estimate the energy consumption based on a real deployment of wireless sensor nodes with power hungry sensors, the energy characterization of previous work [6] has been used. Note that this work reduces the sampling time to 1 second instead of 10 seconds considering the amount of energy obtainable from indoor solar harvesting. Table I summarizes the parameter values used in evaluation.

TABLE I: Simulation Parameters

Param.	Value	Description
k_{min}	1	Minimum divisor value
k_{max}	12	Maximum divisor value
f_{base}	1/300 Hz	Base sampling frequency
P^a	36.858 mW	Power for sensor sampling
P^s	0.0055 mW	Power in a sleep mode

B. Results

This work first evaluates the performance of the adaptive sampling models with the average total reward obtained in the test episodes. To show the effectiveness of predictive adaptive sampling, this work compares the adaptive sampling models with different window lengths ($w = 10$ or 1).



(a) Solar Power (b) Frequency (c) Battery Level

Fig. 4: Comparison of feed-forward agents with different window lengths

Table II summarizes the average total reward of each agent for two datasets that have a similar range of daily energy. In the table, each percentage in parenthesis indicates the percentage of the simulation time (i.e., two weeks in total) for which the agent maintains a stable energy state. The result shows that the feed-forward agent with $w = 10$ obtained the highest average reward, maintaining a stable energy state for a up to 11.5% longer time than the feed-forward agent with $w = 1$.

TABLE II: Average Total Reward

Agent	w	#06	#13
Feed-Forward	10	**791,134 (65.6%)**	**626,823 (52.0%)**
Feed-Forward	1	656,668 (54.4%)	488,638 (40.5%)
Recurrent	10	615,161 (51.0%)	406,914 (33.7%)
Recurrent	1	608,019 (50.4%)	399,013 (33.1%)

Overall, using a longer window results in a higher average reward than using the current energy state only like previous work [9], [10]. Fig. 4 shows sample simulation results of feed-forward agents with different window lengths. The agent with the shorter window tends to be reactive to the current energy state. Therefore, it is more likely to become unstable if the energy availability changes dynamically.

V. Conclusion

This paper proposes deep reinforcement learning-based predictive adaptive sampling for a wireless sensor node. The main goal of this work is to enable predictive adaptive sampling by designing adaptive sampling models that consider the trend of energy state. The evaluation results show that the models with the larger window can more safely manage the energy budget reflecting dynamic energy availability.

References

[1] A. J. Williams, M. F. Torquato, I. M. Cameron, A. A. Fahmy, and J. Sienz, “Survey of energy harvesting technologies for wireless sensor networks,” *IEEE Access*, vol. 9, pp. 77 493–77 510, 2021.

[2] J. Singh, R. Kaur, and D. Singh, “Energy harvesting in wireless sensor networks: A taxonomic survey,” *International Journal of Energy Research*, vol. 45, no. 1, pp. 118–140, 2021.

[3] N. Stricker, R. Da Forno, and L. Thiele, “Robustness of predictive energy harvesting systems: Analysis and adaptive prediction scaling,” *IET Computers & Digital Techniques*, vol. 16, no. 4, pp. 106–124, 2022.

[4] N. Yamin and G. Bhat, “Online solar energy prediction for energy-harvesting internet of things devices,” in *2021 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED)*, 2021, pp. 1–6.

[5] B. J. Fletcher, D. Balsamo, and G. V. Merrett, “Power neutral performance scaling for energy harvesting mp-socs,” in *2017 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2017, pp. 1516–1521.

[6] B. Srbinovski, M. Magno, B. O’Flynn, V. Pakrashi, and E. Popovici, “Energy aware adaptive sampling algorithm for energy harvesting wireless sensor networks,” in *2015 IEEE Sensors Applications Symposium (SAS)*, 2015, pp. 1–6.

[7] C. Lee and J. Lee, “Harvesting and energy aware adaptive sampling algorithm for guaranteeing self-sustainability in wireless sensor networks,” in *2017 International Conference on Information Networking (ICOIN)*, 2017, pp. 57–62.

[8] C. Lee and S.-L. Kim, “The sampling period estimation based adaptive sampling algorithm for a self-sustainable disaster monitoring system,” in *2020 European Conference on Networks and Communications (EuCNC)*, 2020, pp. 165–169.

[9] F. Ait Aoudia, M. Gautier, and O. Berder, “RLMan: An energy manager based on reinforcement learning for energy harvesting wireless sensor networks,” *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 2, pp. 408–417, 2018.

[10] F. Fraternali, B. Balaji, and R. Gupta, “Scaling configuration of energy harvesting sensors with reinforcement learning,” in *Proceedings of the 6th International Workshop on Energy Harvesting & Energy-Neutral Sensing Systems*, ser. ENSsys ’18. New York, NY, USA: Association for Computing Machinery, 2018, p. 7–13.

[11] T. Kannan and H. Hoffmann, “Protecting adaptive sampling from information leakage on low-power sensors,” in *Proceedings of the 27th ACM International Conference on Architectural Support for Programming Languages and Operating Systems*, 2022, pp. 240–254.

[12] M. Mehrani, I. Attarzadeh, and M. Hosseinzadeh, “Sampling rate prediction of biosensors in wireless body area networks using deep-learning methods,” *Simulation Modelling Practice and Theory*, vol. 105, p. 102101, 2020.

[13] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, “OpenAI Gym,” 2016.

[14] M. Abadi et al., “TensorFlow: Large-scale machine learning on heterogeneous systems,” 2015, software available from tensorflow.org. [Online]. Available: https://www.tensorflow.org/

[15] C. Alippi, G. Anastasi, C. Galperti, F. Mancini, and M. Roveri, “Adaptive sampling for energy conservation in wireless sensor networks for snow monitoring applications,” in *2007 IEEE International Conference on Mobile Adhoc and Sensor Systems*, 2007, pp. 1–6.

[16] C. Alippi, G. Anastasi, M. Di Francesco, and M. Roveri, “An adaptive sampling algorithm for effective energy management in wireless sensor networks with energy-hungry sensors,” *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 2, pp. 335–344, 2010.

[17] C. J. Watkins and P. Dayan, “Q-learning,” *Machine learning*, vol. 8, no. 3, pp. 279–292, 1992.

[18] V. Konda and J. Tsitsiklis, “Actor-critic algorithms,” in *Advances in Neural Information Processing Systems*, S. Solla, T. Leen, and K. Müller, Eds., vol. 12. MIT Press, 1999.

[19] L. Sigrist, A. Gomez, and L. Thiele, *Long-Term Tracing of Indoor Solar Harvesting*. Zenodo, sep 2019, Appears in the Proceedings of the 2nd Workshop on Data Acquisition To Analysis (DATA ’19).