Reinforcement Learning for Power Management in Solar Energy Harvesting Sensor Nodes*

Shaswot SHRESTHAMALI, Masaaki KONDO, Hiroshi NAKAMURA

Graduate School of Information Science and Technology, The University of Tokyo

東京大学
THE UNIVERSITY OF TOKYO

INTRODUCTION

Internet of Things (IoT) and Pervasive Computing requires billions and trillions of sensor nodes. Energy Harvesting Wireless Sensor Nodes (EHWSN) play a critical role in forming a **sustainable, maintenance-free network of perpetually communicating autonomous devices** for the IoT infrastructure. Energy autonomy (neutrality) of the sensor nodes should be ensured for perpetual operation. Here we consider the case of maximizing the utility of a solar energy harvesting sensor node while maintaining energy neutrality.

CHALLENGES

1. Energy Neutral Operation

- **Amount of energy harvested should equal the amount of energy consumed** by the node.
- A **minimum duty cycle** should be maintained at all times.
- The battery should never be completely full or depleted.

2. Adaptivity

The node has to be able to adapt to changes in energy harvesting profile due to

- **Diurnal /Seasonal Variations**
- **Climatic Variations**
- **Changes in device parameters**

3. Diversity and Scaling

- The solution should be feasible for sensor nodes regardless of their functionality or working environment.
- Any changes from the original working environment should not compromise the node's performance

REINFORCEMENT LEARNING AS A SOLUTION

- ❑ Hand-engineered solutions will not be practical owing to the diverse scenarios of application and the sheer population of sensor nodes.
- ❑ We propose a one-size-fits-all solution i.e. a power manager for EHWSN that is capable of
  - ✓ **Learning** the best power management strategy through a <u>context aware action-perception learning cycle.</u>
  - ✓ **Adapting** to any changes in its environment
- ❑ We use a Reinforcement Learning (RL), **SARSA(λ)** with a system model shown in Figure 1.

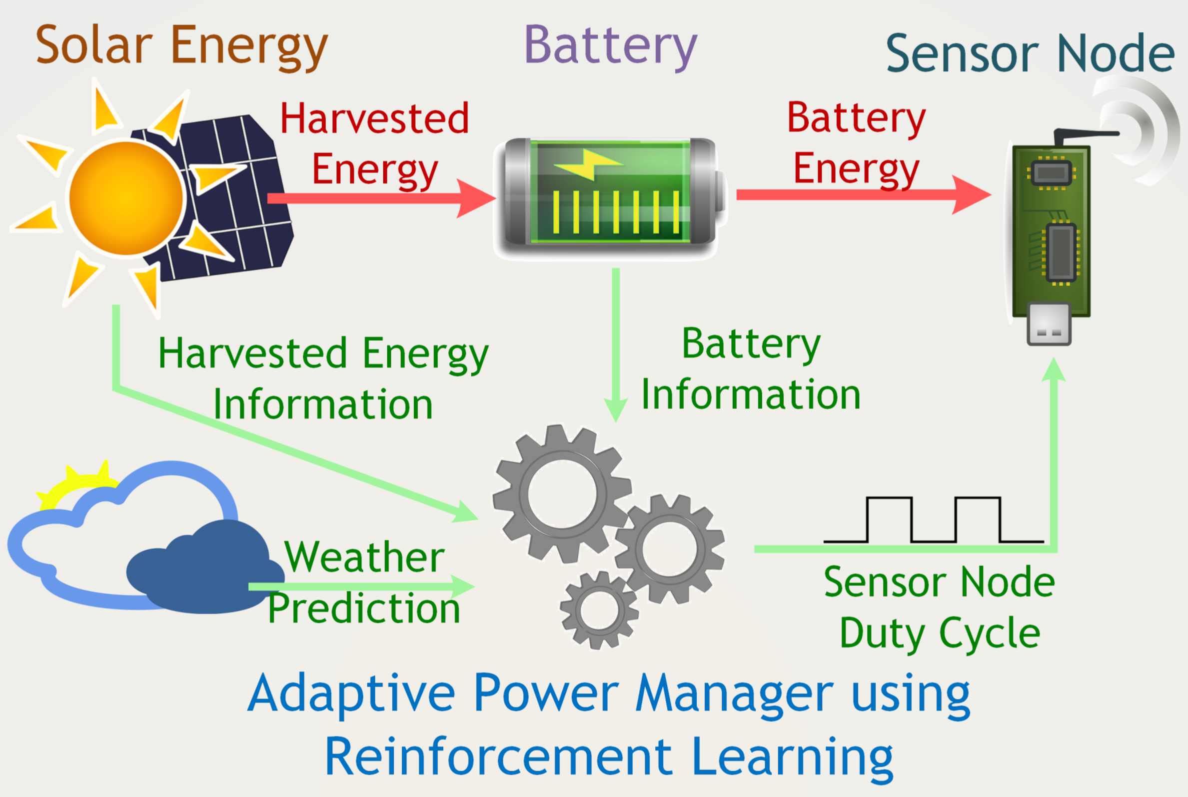
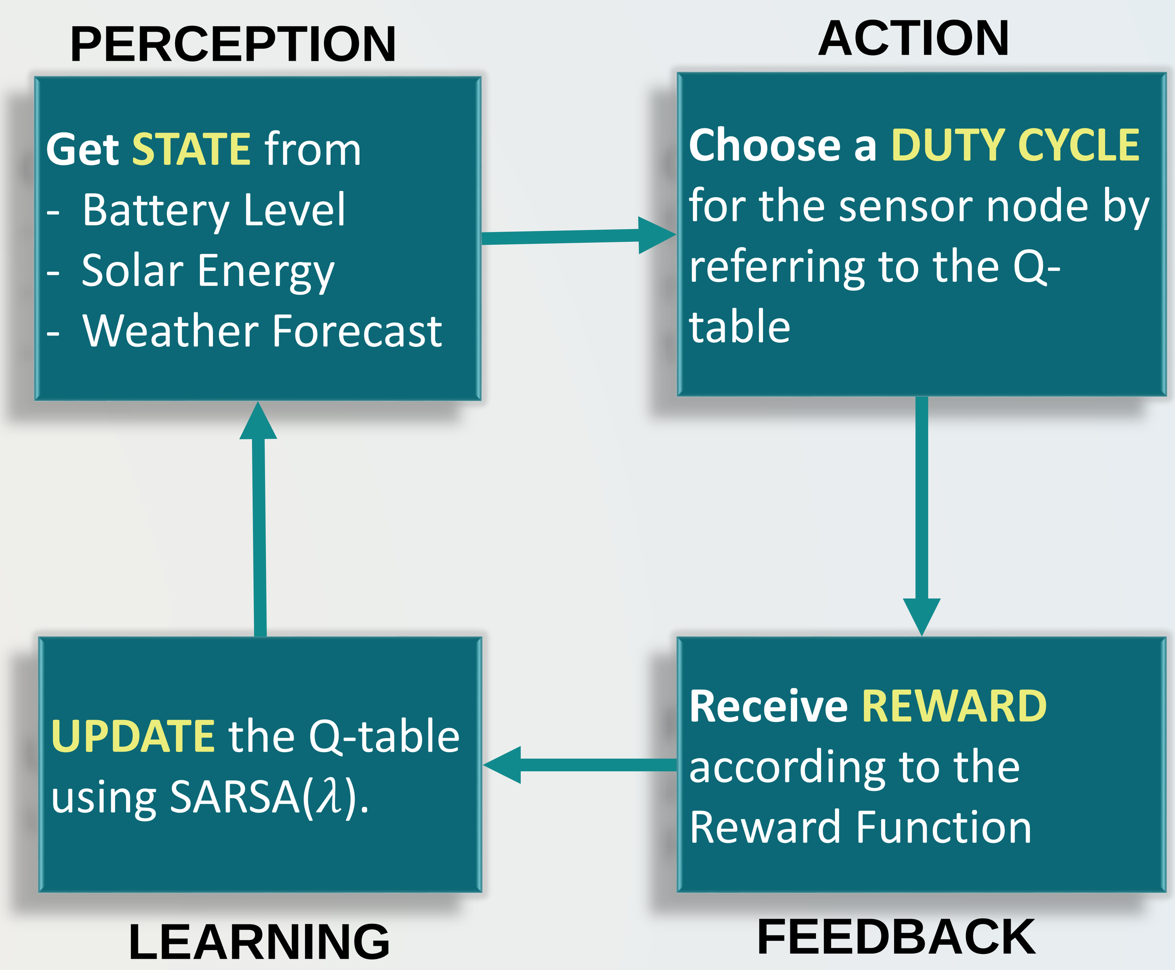


Figure 1: System Model



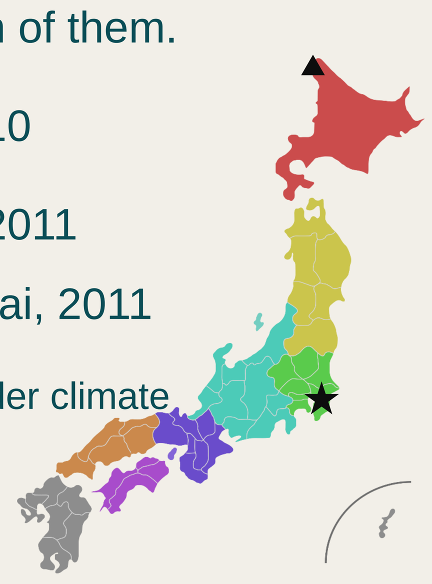
EXPERIMENTAL RESULTS

We ran simulations with historical weather data. We took 12 representative days from 2010 and trained the agent for 10,000 iterations on each of them.

TRAINING: Tokyo, 2010

TESTING: ★ Tokyo, 2011

▲ Wakkanai, 2011



▲ Wakkanai has a much colder climate compared to ★Tokyo

We compare the performance of **SARSA(λ)** with an **Offline Policy**. Offline Policy is calculated using a linear optimizer assuming a perfect omniscient weather predictor.

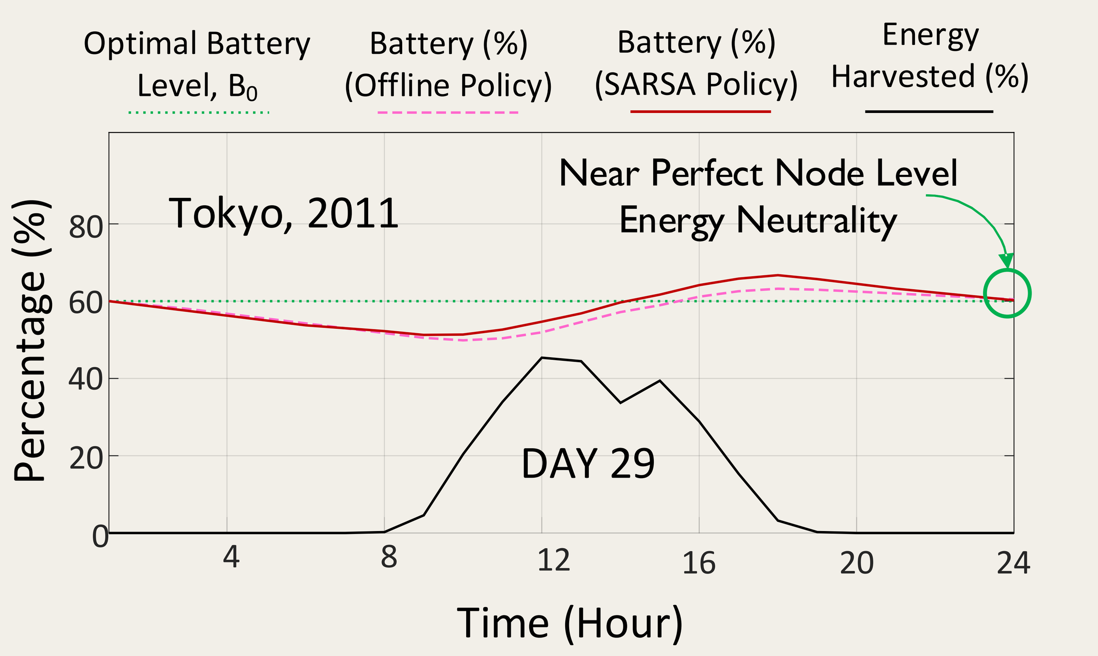


Figure 2: SARSA(λ) achieves near optimal performance

Figure 2 shows the battery profiles for SARSA(λ) and Offline Policy. SARSA(λ) is able to achieve near perfect node level energy neutrality.

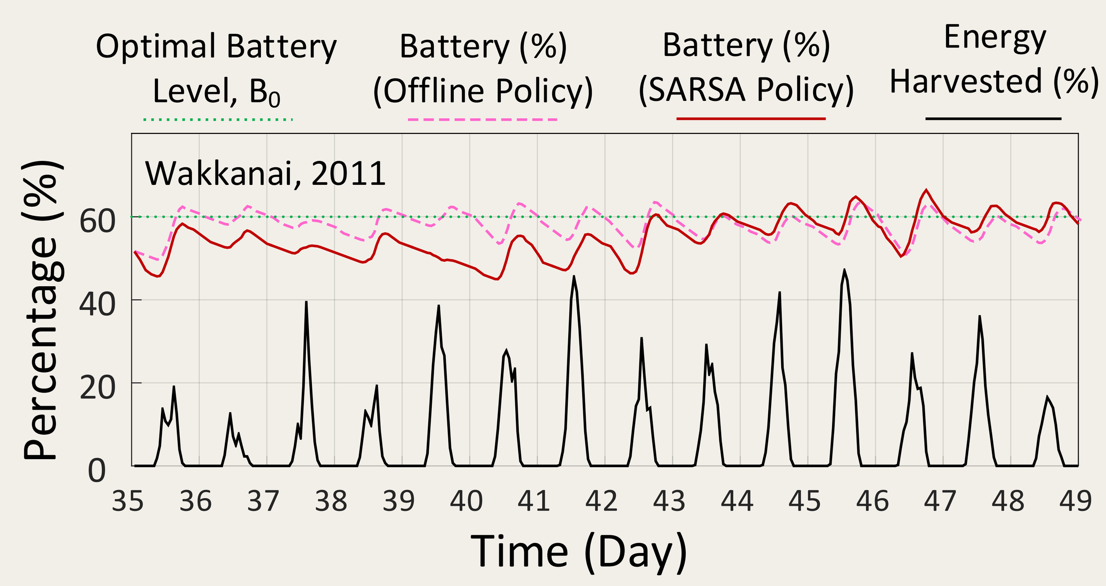


Figure 3: Adaptation to change in location

Figure 3 shows the battery profiles for Wakkanai, 2011. The agent is able to adapt to Wakkanai climate even thought it was trained with Tokyo data.

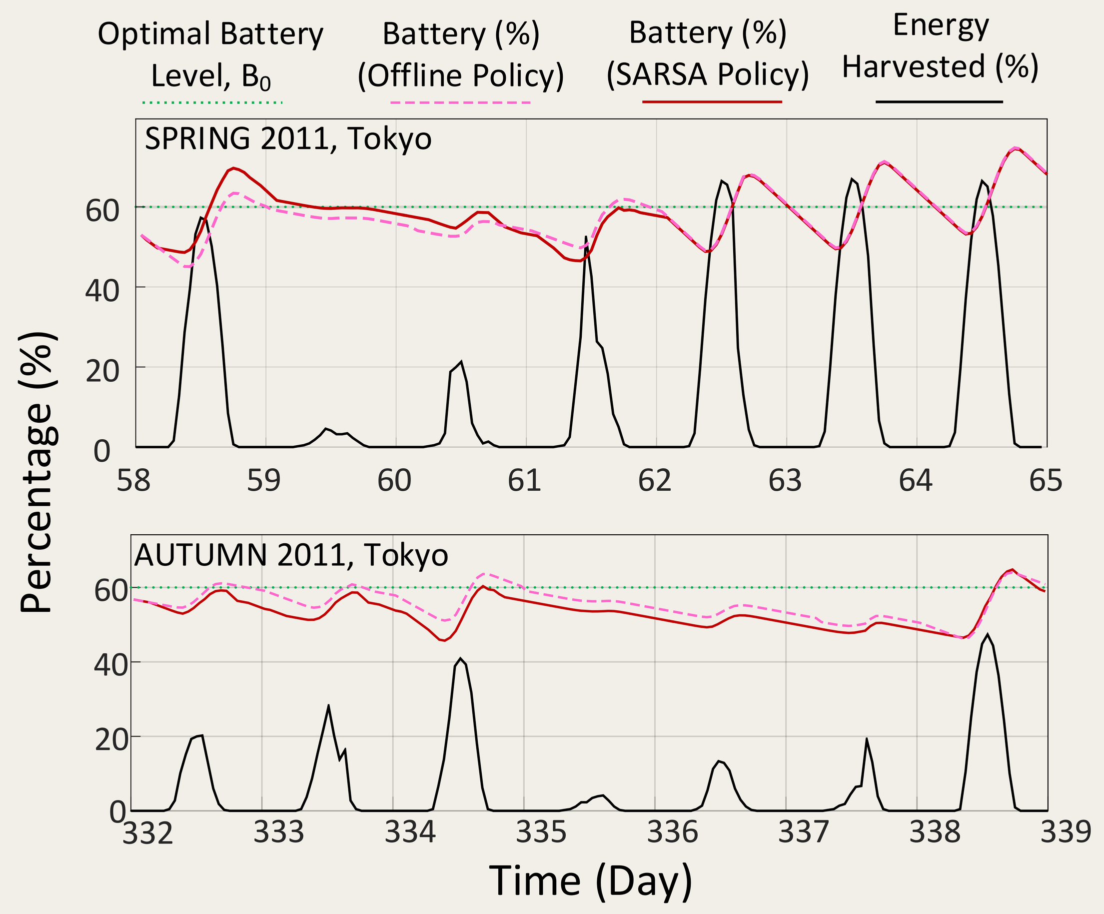


Figure 4: Adaptation to seasonal variation

Figure 4 shows how the agent is able to account for seasonal differences in harvested energy and maintain energy neutrality.

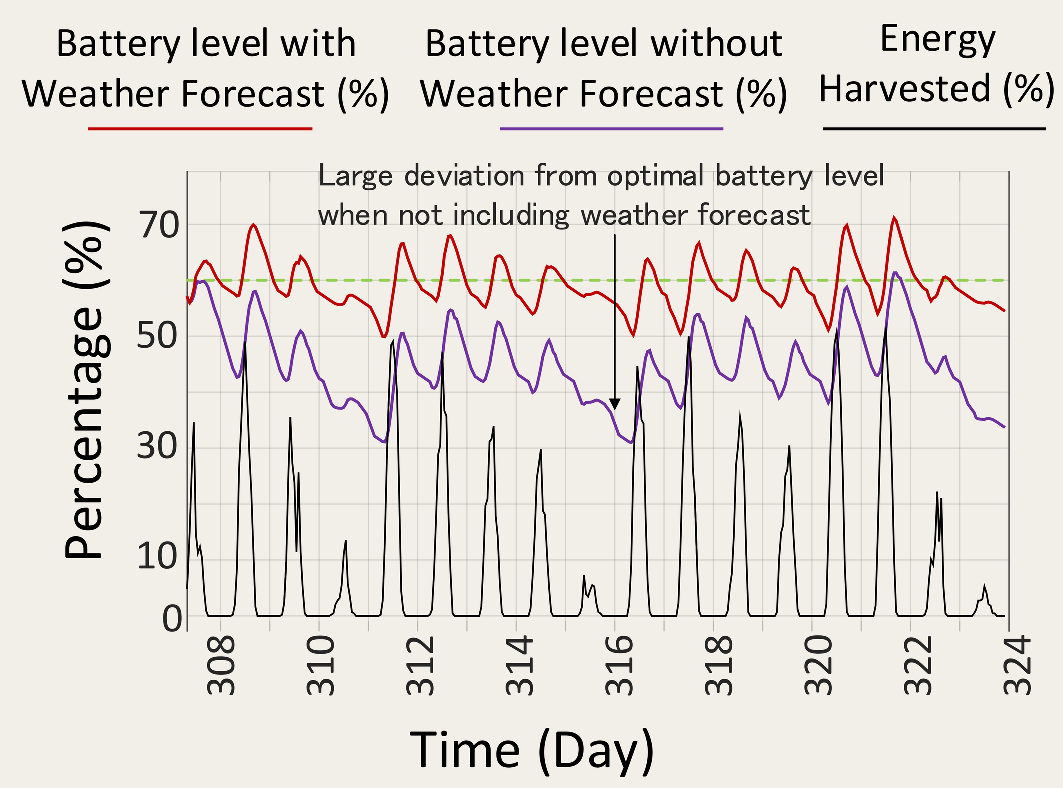


Figure 5: Inclusion of Weather Forecast

Figure 5 shows that inclusion of weather forecast in decision-making results in better performance.

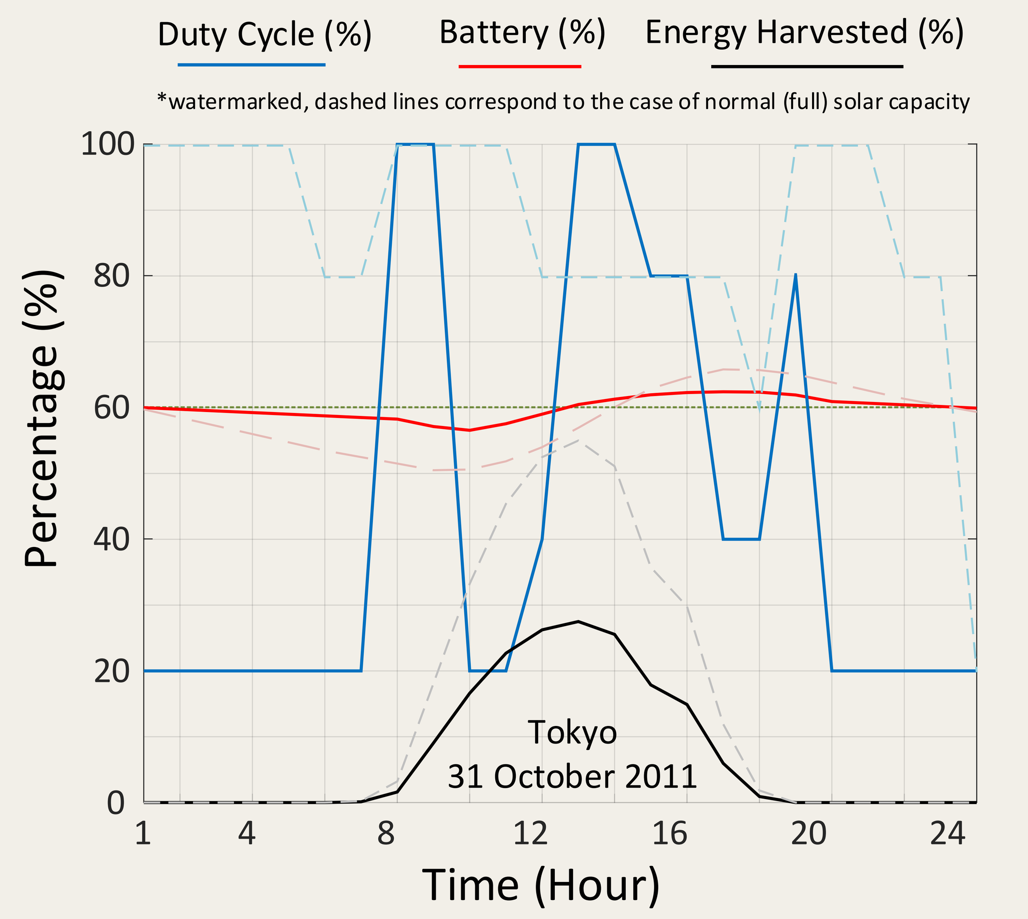


Figure 6: Adaptation to change in device parameters

In Figure 6, the agent is able to achieve energy neutrality even though the solar capacity is halved.

*This work was partially supported by JSPS KAKENHI Grant Number 16K12405.