

論文の内容の要旨

論文題目 Reinforcement Learning-Based Optimization in Energy Harvesting Wireless
Sensor Nodes
 (強化学習を用いた環境発電駆動センサーノードの最適化に関する研究)

氏 名 シュレスタマリ サソット

Energy Harvesting Wireless Sensor Nodes (EHWSNs) constitute a major proportion of the rapidly growing number of connected devices in the Internet of Things (IoT). Given this scale of growth, autonomous operation of EHWSNs is now more of a requirement rather than a preferred feature. Since they harvest energy from their working environment, they do not require periodic energy replenishment and could therefore theoretically operate forever. With intelligent and adaptive policies, it is possible to extract maximum utility and uninterrupted operation, even when the working environment is complex and unpredictable.

Our work investigates Reinforcement Learning (RL) optimization methods for EHWSNs to learn to self-regulate its resources – especially with regards to its energy as this is the most critical resource in EHWSNs. This entails first - balancing the device energy consumption with the unpredictable energy supply, i.e., Energy Neutral Operation (ENO) and second - ensuring that the energy used has maximum utility for the user. Traditionally, analytic methods were used on account of their theoretical guarantees and low compute requirements. However, these methods have limited applications, are not adaptive, and require manual calibrations. More recently, RL methods have gained popularity mainly because of their adaptivity and general scope of application. However, RL methods suffer from high learning costs with respect to time, risk and computation which prevents them from being used in the real world. This is mainly due to non-optimal formulation of the RL problem (especially the reward function) and inefficient exploration. With this work, we tackle these limitations in three different orthogonal directions and contribute solutions to overcome them.

For our first contribution, we develop tabular RL solutions for a general EHWSN that can generate near-optimal energy management policies and adapt to the working environment. We propose a general ENO reward function that abstracts the problem formulation from application-specific details. We demonstrate that even with the most basic RL methods and our novel reward function, we can expect adaptive behavior and superior policies.

Our second contribution proposes an efficient exploration strategy to minimize the learning costs that are required by Deep RL (DRL) to learn energy-neutral policies. DRL replaces the look-up tables of tabular RL with more powerful neural network-based function approximators. In doing so, the problem formulation can accommodate continuous states and actions which extends its applicability. We propose a distributed learning solution that coordinates the many sensor nodes of a network to efficiently explore the large state space concurrently. Our method significantly minimizes learning risks and accelerates the learning process.

In the final part of our contribution, we introduce a general Multi-Objective RL (MORL) framework for EHWSN optimization. Modern EHWSNs are quite sophisticated and can use its energy to for many different tasks (e.g., sensing, communication, and processing). Thus, it is necessary not only to ensure ENO but also to allocate energy between the different tasks within the constraints of a limited budget to maximize node utility. Current methods are unable to consolidate between the multiple objectives and tradeoff between them. Our MORL framework is more general and we show that it provides solutions that can learn more intelligent policies with minimal learning costs compared to traditional RL methods. More importantly, with this framework, one can dynamically adjust the energy management policy to tradeoff during runtime between different performance objectives without additional training.

Thus, in our work we show that RL methods are very promising alternatives for optimization of EHWSNs. We show that with our novel problem formulation, they can learn to adapt to unknown environments for long-term uninterrupted operation. We also show that it is possible to minimize the learning costs by leveraging the distributed structures of nodes in the sensor network. Finally, we demonstrate that with MORL, we can further use the RL approach to not only ensure ENO but also optimize the node utility by allowing runtime tradeoffs. With our contributions, EHWSNs are a step closer to autonomous and perpetual operation. As we inch closer to full autonomy, we can expect IoT to be more seamlessly integrated into our world and enrich our lives.