Reinforcement Learning Based on Energy Harvesting in Wireless Sensor Networks

P.M. Allwyn Ebenezer, Francis Xavier Engineering College, Department of CSE, Allwyn93ebezer@gmail.com

P.C.Arumugam, Francis Xavier Engineering College, Department of CSE, arumugam28ajay@gmail.com

Ms.N.RajaPriya, Assistant Professor(CSE), Francis Xavier Engineering College, nrp.priya10@gmail.com

Abstract

A promising solution to achieve autonomous wireless sensor networks is to enable each node to harvest energy in its environment. A novel energy management algorithm based on reinforcement learning, named RLMan. By continuously exploring the environment, RLMan adapts its energy management policy to time-varying environment, regarding both the harvested energy and the energy consumption of the node.Linear function approximations are used to achieve very low computational and memory footprint, making RLMan suitable for resource-constrained systems such as wireless sensor nodes.RLMan is more robust to variability of the node energy consumption.

Keywords: Reinforcement Learning · Indoor Localisation · SARSA ·Energy Efficiency · Pervasive Health.

1. Introduction

Many applications, such as smart cities, precision agriculture and plant monitoring, rely on the deployment of a large number of individual sensors forming Wireless Sensor Networks(WSNs). These individual nodes must be able to operate for long periods of time, up to several years or decades, while being highly autonomous to reduce maintenance costs. As refilling the batteries of each device can be expensive or impossible if the network is dense or if the nodes are deployed in a harsh environment, maximizing the lifetime of typical sensors powered by individual batteries of limited capacity is a perennial issue. Therefore, important efforts were devoted in the last decades to develop low power consumption devices as well as energy efficient algorithms and communication schemes to maximize the lifetime of WSNs. Typically, a node quality of service (sensing rate, packet rate) is set at deployment to a value that guarantees the required lifetime. However, as batteries can only store a finite amount of energy, the network is doomed to die. A promising solution to increase the lifetime of WSNs is to enable each node to harvest energy in its environment. In this scenario, each node is equipped with one or more energy harvesters, as well as an energy buffer (battery or capacitor) to allow storing part of the harvested energy for future use during periods of energy scarcity. Various energy sources are possible, such as light, wind, motion, fuel cells. As the energy sources are typically dynamic and uncontrolled, it is required to dynamically adapt the power consumption of the nodes, by adjusting their quality of service, in order to avoid power failure while maximizing the energy efficiency and ensuring the fulfilment of application requirements. This task is done by a software module called Energy Manager (EM). For each node, the EM is responsible for maintaining the node in Energy Neutral Operation (ENO)state, the amount of consumed energy never exceeds the amount of harvested energy over a long period of time. Ideally, the amount of harvested energy equals the amount of consumed energy over a long period of time, which means that no energy is wasted by saturation of the energy storage device.

Many energy management schemes were proposed in the last years to address the nontrivial challenge of designing efficient adaptation algorithms, suitable for the limited resources provided by sensor nodes in terms of memory, computation power, and energy storage. They can be classified based on their requirement of predicted information about the amount of energy that can be harvested in the future, prediction-based and prediction-free. Prediction-based Ems rely on prediction, of the future amount of harvested energy over a finite time horizon to take decision on the node energy consumption. In contrast with prediction-based energy management schemes, prediction-free approaches do not rely on forecasts of the harvested energy. These approaches were motivated by the significant errors from which energy predictors can suffer, which incur overuse or underuse of the harvested energy, and the fact that energy prediction requires the ability to measure the harvested energy, which incur a significant overhead. Indeed, most of the EMs require an accurate control of the spent energy, as well as detailed tracking of the previously harvested and previously consumed energies to operate properly. Considering these practical issues, we propose RLMan, a novel EM scheme based on Reinforcement Learning (RL) theory. RLMan is a prediction-free EM, whose objective is to maximize the quality of service, defined as the packet rate, the frequency at which packets are generated (e.g., by performing measurements) and sent, while avoiding power failure.

It is assumed that the node is in the sleep state between two consecutive packet generations and sendings, in order to allow energy saving. RLMan requires only the state of charge of the energy storage device to operate and aims to set the packet rate by both exploiting the current knowledge of the environment and exploring it, in order to improve the policy. Continuous exploration of the environment enables adaptation to varying energy harvesting and energy consumed dynamics, making this approach suitable for the uncertain environment of Energy Harvesting WSNs (EH-WSNs). The problem of maximizing the quality of service in EH-WSNs is formulated as a Markovian Decision Process (MDP), and a novel EM scheme based on RL theory is introduced.

A formulation of the problem of maximizing the quality of service in energy harvesting WSNs using the RL framework. A novel EM scheme based on RL theory named RLMan, which requires only the state of charge of the energy storage device, and which uses function approximations to minimize the memory footprint and computational overhead. An exploration of the impact of the parameters of RLMan using extensive simulations, and real measurements of both indoor light and outdoor wind. The comparison of RLMan to three state-of-the-art Ems (P-FREEN, Fuzzyman and LQ-Tracker) that aim to maximize the quality of service, regarding the capacitance of the energy storage device and the variability of the node task energy cost. The rest of this paper is organized as follows: Section II presents the related work focusing on prediction-free EMs, and Section III provides the relevant background on RL theory. In Section IV, the problem of maximizing the packet rate in energy harvesting WSNs is formulated using the RL framework, and RLMan is derived based on this formulation. In Section V, RLMan is evaluated. First, the simulation setup is presented. Next, exploration of the parameters of RLMan is performed. Finally, the results of the comparisons of RLMan to three other EMs are shown. Section VI concludes this paper.

2.Related Work

RL is a framework for optimizing the behavior of an agent, or controller, that interacts with its environment. Many energy management schemes were proposed in the last years to address the nontrivial challenge of designing efficient adaptation algorithms, suitable for the limited resources provided by sensor nodes in terms of memory, computation power, and energy storage. The first prediction-free EM was LQ-Tracker, proposed in 2007 by Vigorito. This scheme relies on linear quadratic tracking, a technique from control theory, to adapt the duty-cycle according to the current state of charge of the energy buffer. In this approach, the energy management problem is modeled as a first order discrete time linear dynamical system with colored noise, in which the system state is the state of charge of the energy buffer, the controller output is the duty-cycle and the colored noise is the moving average of state of charge increments produced by the harvested energy. The objective is to minimize the average squared tracking error between the current state of charge and a target residual energy level. The authors used classical control theory results to get the optimal control, which does not depend on the colored noise, and the control law coefficients are learned online using gradient descent. Finally, the outputs of the control system are smoothed by an exponential weighting scheme to reduce the duty-cycle variance. Peng et al. proposed P-FREEN, an EM that maximizes the duty-cycle of a sensor node in the presence of battery storage inefficiencies, As solving this kind of problem directly is computationally intense, they proposed a set of budget assigning principles that maximizes the duty-cycle by only using the current observed energy harvesting rate and the residual energy.

The proposed algorithm requires the current state of charge of the energy buffer as well as the harvested energy to take decision about the energy budget. If the state of charge of the energy buffer is below a fixed threshold or if the amount of energy harvested at the previous time slot is below the minimum required energy budget, then the node is operating with the minimum energy budget. Otherwise, the energy budget is set to a value that is a function of both the amount of energy harvested at the previous time slot and the energy storage efficiency. In the authors proposed to use fuzzy control theory to dynamically adjust the energy consumption of the nodes. Fuzzy control theory proposes to extend conventional control techniques to ill-modeled systems, such as EH-WSNs. With this approach, named Fuzzyman, an intuitive strategy is formally expressed as a set of fuzzy IF-THEN rules. The algorithm requires as an input both the residual energy and the amount of harvested energy since the previous execution of the EM.

The first step of Fuzzyman is to convert the crisp entries into fuzzy entries. The so-obtained fuzzy entries are then used by an inference engine which applies the rule base to generate fuzzy outputs. The outputs are finally mapped to a single crisp energy budget that can be used by the node. The main drawbacks of Fuzzyman are that it requires the amount of energy harvested which can be unpractical to measure, and the lack of systematic way to set the parameters. RL theory was used in the context of energy harvesting communication systems. The authors addressed the problem of maximizing the throughput at the receiver in energy harvesting two-hop communications. The problem was first formulated as a convex optimization problem, and then reformulated using the reinforcement learning framework, so it can be solved using only local causal knowledge.

The well-known SARSA algorithm was used to solve the obtained problem. With RLMan, we focus on maximizing the packet rate in point-to-point communication systems. considered energy harvesting WSNs with packet rate requirement, and used Q-Learning, a well-known RL algorithm, to meet the packet rate constraints. In this approach, states and actions are discretized, and a reward is defined according to the satisfaction of the packet rate constraints. The aim of the algorithm is to maximize the overall rewards, by learning the Q-values, the accumulative reward associated with a given state-action pair. The proposed EM requires the tracking of the harvested energy and the energy consumed by the node in addition to the state of charge. Moreover it uses two dimensional look-up tables to store the Q values, which incurs significant memory footprint. At each execution of the EM, the action is chosen according to the Q-values using soft-max function, and the Q-value of the last state-action pair observed is updated using the corresponding last observed reward. With RLMan, we propose an approach that requires only the state of charge of the energy storage device in order to maximize the packet rate. Indeed, measuring the amount of harvested energy or consumed energy requires additional hardware, which incurs additional cost, and increases the form factor of the node. Moreover, linear function approximators are used to minimize the memory footprint and computation overhead, which are critical when focusing on constrained systems such as WSN nodes.

3. Problem Statement

RL theory was used in the context of energy harvesting communication systems. Its maximizing the throughput at the receiver in energy harvesting two-hop communications. The problem was first formulated as a convex optimization problem, and then reformulated using the reinforcement learning framework so it can be solved using only local causal knowledge. The well-known SARSA(State action reward state action) algorithm was used to solve the obtained problem. With RLMan, we focus on maximizing the packet rate in point-to-point communication systems. Suitable for point to point communication systems. Less quality of service in energy communications

4.System Architecture

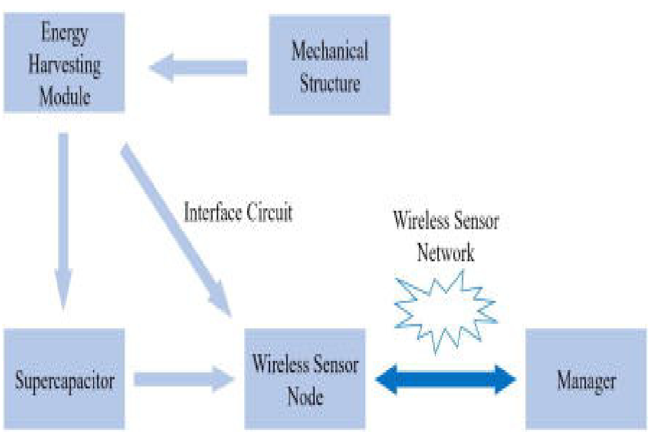


Fig. System Architecture

It is assumed that time is divided into equal length time slots of duration Ts, and that the EM is executed at the beginning of every time slot. The amount of residual energy, the amount of energy stored in the energy storage device, is denoted by er and the energy storage device is assumed to have a finite capacity denoted by Emaxr . The hardware failure threshold is denoted by Efailr , and corresponds to the minimum amount of residual energy required by the node to operate, i.e., if the residual energy drops below this threshold, a power failure arises. It is assumed that the job of the node is to periodically send a packet at a packet rate denoted by P rXmin;Xmaxs, and that the goal of the EM is to dynamically adjust the performance of the node by setting. Choosing a continuous action space enables more accurate control of the consumed energy, as a continuum of packet rates is available. The goal of the EM is to maximize the packet rate while keeping the node sustainable, avoiding power failure.

5.Modules

- **Tuning RLMan**
- **LQ-Tracker**
- **Energy Storage Device**
- **Impact Of The Variability**
- **Evaluation Of RLMan**

Tuning RLMan

Simulations were performed for both indoor light and outdoor wind energy traces, but as the obtained results are similar for both cases. RLMan requires three parameters to be set: the discount factor , the trace decay parameter and the policy standard deviation . Using the setup previously introduced, the impact of these parameters on the performance of the proposed scheme was explored. For each value of a parameter, one hundred simulations were run, each performed using different seeds for the random number generators. The capacitance was set to 1 F, as it is the value used by the PowWow platform that was simulated, leading to Efail r _ 3:9 J and Emax r _ 13:5 J.. Each cross represents the result of one simulation run, while the dots show the average of all the simulation runs for one value of a parameter. The results show that does not significantly impact the performance of RLMan, and therefore the results regarding this parameter are not exposed. Simulations were performed for both indoor light and outdoor wind energy traces, but as the obtained results are similar for both cases, only the results of indoor light are shown.

LQ-Tracker

LQ-Tracker were only fed with the value of the residual energy. Both the indoor light and the wind energy traces were considered. LQTracker achieve more than 99:9 % efficiency, for indoor light and outdoor wind, for all capacitance values, and despite the fact that they require only the residual energy as an input. In addition, when the node is powered by outdoor wind, RLMan always outperforms the other Ems in terms of average packet rate for all capacitance values. When the node is powered by indoor light, RLMan also outperforms all the other Ems, except LQ-Tracker when the values of the capacitance are higher than 2:8 F. Moreover, the advantage of RLMan over the other Ems is more significant for small values of the capacitance. Especially, the average packet rate is more than 20 % higher compared to LQ-Tracker in the case of indoor light, and almost 70 % higher in the case of outdoor wind, when the capacitance value is set to 0:5F. This is encouraging as using small capacitance leads to lower cost and lower form factor.

Energy Storage Device

RLMan over the other Ems is more significant for small values of the capacitance. This is encouraging as using small capacitance leads to lower cost and lower form factor. The Ems were evaluated for capacitance sizes ranging from 0:5 F to 3:5 F, and for a value of _ of 0:16. All the Ems successfully avoid power failure when powered by indoor light or outdoor wind.

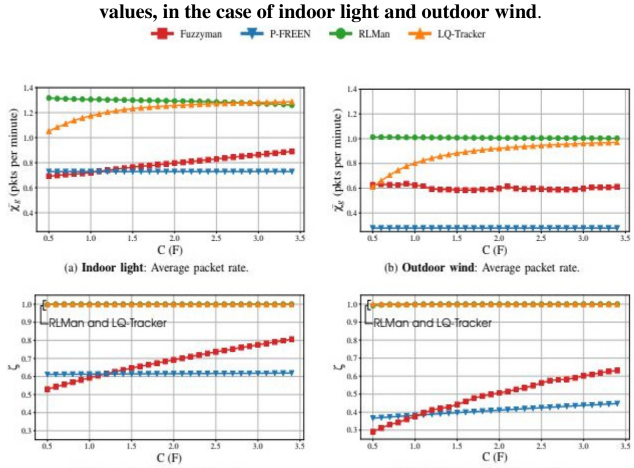
Impact Of The Variability

One can imagine that the performance of the other schemes could be increased by tracking the energy consumption of the node, and using an adaptive algorithm to compute the packet rate from the output of the EM. The Ems were evaluated for different values of _. Etyp C and Emax C were kept constant, while _C was changed by varying the parameters of the beta distribution. The capacitance of the energy storage device was set to 1:0 F, as it is the value used by PowWow. Results. It can be seen from that in both the cases of indoor light and outdoor wind, increasing the variability of EC does not affect the efficiency of RLMan and LQ-Tracker, both of which achieve an efficiency of 1. Concerning Fuzzyman and P-FREEN, increasing _ enables higher efficiencies. Indeed, P-FREEN and Fuzzyman outputs are an energy budget, from which a packet rate must be calculated. The calculation of the packet rate was performed based on a typical cost of a task execution Etyp C . However, as _ increases, EC tends to be higher than Etyp C more often, which causes less energy waste as more energy is required to achieve the same quality of service.

Evaluation Of RLMan

RLMan was evaluated using exhaustive simulations. This section starts by presenting the simulation setup. Then, the impact of the different parameters required by RLMan is explored. RLMan does not require any knowledge on the energy consumption of the node. It successfully adapts to variations of the energy consumption of the node task (e.g., due to channel noise or sensors variability), and ensures high energy efficiency and packet rate. One can imagine that the performance of the other schemes could be increased by tracking the energy consumption of the node, and using an adaptive algorithm to compute the packet rate from the output of the EM, but this would require additional resources and increase the complexity, memory and computation footprint of the energy manager. RLMan inherently adapts to the dynamics of both the harvested energy and the consumed energy, and therefore does not require this additional complexity.

Average packet rate and energy efficiency for different capacitance values, in the case of indoor light and outdoor wind.



(a) **Indoor light**: Average packet rate. (b) **Outdoor wind**: Average packet rate.

(c) **Indoor light**: Energy efficiency. (d) **Outdoor wind**: Energy efficiency.

Conclusion& Future Enhancement

RLMan uses function approximations to achieve low memory and computational overhead, and only requires the state of charge of the energy storage device as an input. RLMan enables significant gains regarding packet rate compared to state of the art approaches, both in the case of indoor light and outdoor wind. RLMan successfully adapts to variations of the consumed energy, and does not require additional energy consumption tracking schemes to keep high efficiency and quality of service.

References:

1. Adaptive Control of Duty Cycling in Energy-Harvesting Wireless Sensor Networks -June 2007 C. M. Vigorito, D. Ganesan
2. Prediction free energy neutral power management for energy harvesting wireless sensor nodes – 2014, S. Peng and C. P. Low
3. Fuzzy power management for energy harvesting Wireless Sensor Nodes – 2016, F. Ait Aoudia, M. Gautier, and O. Berder

4. Reinforcement Learning for Energy Harvesting Decode-and-Forward Two-Hop Communications – 2017, A. Ortiz, H. Al-Shatri
5. Multi objective Reinforcement Learning: A Comprehensive Overview – 2015, C. Liu, X. Xu, and D. Hu
6. C. M. Vigorito, D. Ganesan, and A. G. Barto, "Adaptive Control of Duty Cycling in Energy-Harvesting Wireless Sensor Networks," in 4th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), June 2007.
7. [8] S. Peng and C. P. Low, "Prediction free energy neutral power management for energy harvesting wireless sensor nodes," Ad Hoc Networks, vol. 13, Part B, 2014.
8. [9] F. Ait Aoudia, M. Gautier, and O. Berder, "Fuzzy power management for energy harvesting Wireless Sensor Nodes," in IEEE International Conference on Communications (ICC), May 2016.
9. [10] A. Ortiz, H. Al-Shatri, X. Li, T. Weber, and A. Klein, "Reinforcement Learning for Energy Harvesting Decode-and-Forward Two-Hop Communications,"IEEE Transactions on Green Communications and Networking, vol. 1, no. 3, pp. 309–319, September 2017.
10. M.Gorlatova, A.Wallwater, and G.Zussman, "Networking Low-Power Energy Harvesting Devices: Measurements and Algorithums",in IEEE INFOCOM,April 2011.