Green Energy Optimization in Energy Harvesting Wireless Sensor Networks

Jianchao Zheng, Yueming Cai, Xuemin (Sherman) Shen, Zhongming Zheng, and Weiwei Yang

Jianchao Zheng, Yueming Cai, and Weiwei Yang are with PLA University of Science and Technology.

Xuemin (Sherman) Shen and Zhongming Zheng are with the University of Waterloo.

ABSTRACT

This article studies the sensor activation control for the optimization of green energy utilization in an EH-WSN, where both energy generation and target distribution exhibit temporal and spatial diversities. Decentralized operation is considered for the green energy optimization in the EH-WSN. The optimization is achieved in two dimensions: dynamic (activation) mode adaptation in the temporal dimension and energy balancing in the spatial dimension. Due to the interactions among autonomous distributed sensors, game theory is applied to the local information based decentralized optimization for the spatial energy balancing problem. In addition, reinforcement learning techniques are proposed to address the temporal mode adaptation in the dynamic and unknown environment. Simulation results are provided to demonstrate the effectiveness of the proposed approaches.

INTRODUCTION

Wireless sensor networks (WSNs) have profound significance toward environmental surveillance and monitoring by spreading throughout factories, forests, oceans, battlefields, and so on [1–3]. However, the limited network lifetime constrained by the battery capacity is a major deployment barrier for traditional WSNs. Recently, energy harvesting (EH) has emerged as a promising technology to extend the lifetime of communication networks by continuously harvesting green energy from environmental sources, such as the sun, wind, and vibrations [4, 5].

Due to the uncertain and dynamically changing environmental conditions, the intermittent and random nature is the most typical characteristics of the EH process. Thus, efficient energy management becomes critical to ensure continuous and reliable network operation [6, 7]. Most existing works assume that either the transmitter has non-causal information on the exact data/energy arrival instants and amounts, or the transmitter knows the statistics of underlying EH and data arrival processes [6]. Nonetheless, in many practical scenarios, the characteristics of EH and data arrival processes may change over time. Moreover, it may not be possible to have reliable statistical information about these processes before deploying the nodes. Hence, non-causal information about the data/energy arrival instants and amounts may be infeasible, so offline optimization frameworks may not be satisfactory in most practical scenarios. Besides, existing research on EH mainly focuses on a point-to-point communication system [6], while the network of multiple EH nodes is more challenging to study [8].

This article studies a general EH-WSN, where multiple energy harvesting sensors (EHSs) are deployed to monitor a target area, as depicted in Fig. 1. Energy harvested by sensors at different locations at different times are usually varying, which reflects both temporal and spatial diversities. Besides, due to the moving characteristics and random (usually not uniform) scattering of targets (e.g., in the battlefield environment), the target distribution also exhibits both temporal and spatial diversities. We focus on the dynamic and online sensor activation (activate/sleep) scheduling for green energy management.

According to characteristics of energy generation and target distribution, green energy optimization in the EH-WSN is a challenging problem that involves optimization in two dimensions: dynamic (activation) mode adaptation in the temporal dimension and energy balancing in the spatial dimension. Specifically, dynamic mode adaptation aims to optimize the green energy usage in multiple time slots to adapt to temporal dynamics of green energy generation and target mobility, while spatial energy balancing maximizes the utilization of green energy in each time slot by balancing the energy consumption among sensors to adapt to the spatial diversity of the energy generation and target distribution.

In general, the complexity of centralized schemes to achieve the optimal energy utilization increases significantly with the number of nodes in the network [8]. Moreover, the optimal solutions depend heavily on the available knowledge of the EH profiles across different sensors, which is difficult to obtain or even unattainable. Therefore, decentralized opti-

mization based only on local information has drawn more attention [2, 3]. However, the existing works either consider a restrictive energy generation/event occurrence model, or simplify the study of the interactions among distributed sensors [9, 10].

Due to the complex interactions among individual sensors, we adopt game theory [11–13] to investigate the local-information-based decentralized green energy optimization in the spatial dimension. By carefully designing the utility function for each sensor, the network nodes can be made to exhibit desired behaviors while individual nodes simply perform local tasks. Moreover, to address the dynamic mode adaptation problem in temporal dimension, reinforcement learning techniques are proposed. After multiple iterative learning, self-regulating sensors can adapt their behaviors to the dynamic and unknown environment, and thus obtain a satisfactory solution that maximizes the green energy utilization.

The rest of this article is organized as follows. In the next section, we present an overview of the studied EH-WSN. Then we discuss key issues and technical challenges of green energy optimization. Following that, reinforcement learning techniques are incorporated into game theory to deal with the complex optimization for green energy utilization. Research directions are then discussed, followed by the conclusion.

OVERVIEW OF THE EH-WSN

In the EH-WSN, multiple EHSs are deployed for target monitoring, as shown in Fig. 1. Generally, both the energy generation and target distribution exhibits temporal and spatial diversities. The temporal diversity of target distribution indicates that the target distribution in the network varies in different time slots due to the moving characteristics of targets and probably the joining of new targets. Moreover, targets are randomly distributed in the area; thus, sensors at different locations may experience different target distribution intensities, which reflects the spatial diversity. Besides, green energy generation also possesses both temporal and spatial diversities. For example, solar energy generation depends on many factors such as temperature, sunlight intensity, the geographical location of the solar panel, and so on [7]. Therefore, energy generation by sensors at different locations is different. Moreover, the daily solar energy generation in a given area exhibits temporal dynamics that peak around noon and bottom out during the night.

In order to keep continuous and sustaining target monitoring, green energy utilization should be optimized by coping with the temporal and spatial diversities of green energy generation and target distribution. The characteristics of the energy arrival and target distribution in the current time slots as well as in future time slots need to be considered. The key problems for green energy optimization in the EH-WSN include medium access control (MAC) [1], power control [2], topology control, activation scheduling [3], and so on. We focus on the activation scheduling problem in this article.
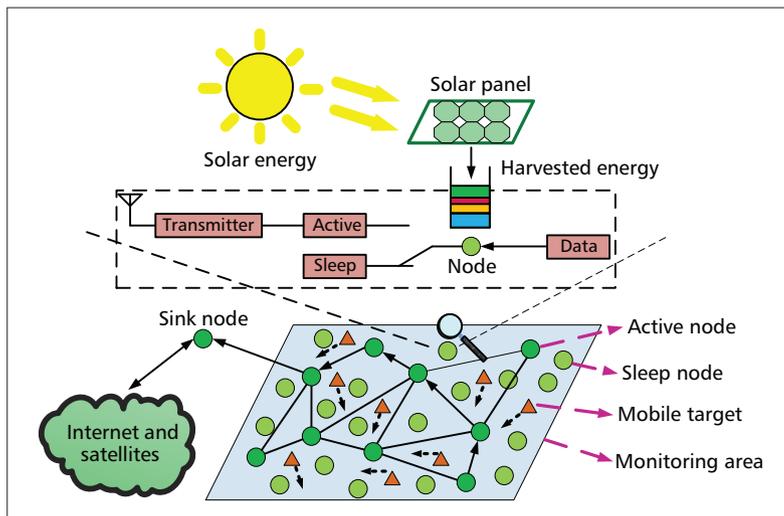

Figure 1. Energy harvesting wireless sensor network.

GREEN ENERGY OPTIMIZATION IN THE EH-WSN

In this section, we discuss the motivation for activation scheduling based green energy optimization, and then introduce its key issues and technical challenges in the EH-WSN.

MOTIVATION OF GREEN ENERGY OPTIMIZATION

Due to the seamless deployment of sensors, a target is often covered by multiple sensors. In the traditional battery-operated WSN, optimally turning some of these sensors to sleep mode will prolong the lifetime of the network while maintaining complete target coverage. In essence, the energy efficiency is improved only by optimizing the energy consumption. However, to improve the energy efficiency in the EH-WSN, not only energy consumption but also green EH should be taken into consideration. In other words, we should minimize the energy consumption and maximize the green energy collection at the same time.

Without loss of generality, we take an example with three sensors and three targets, as shown in Fig. 2a. The node sensing area is the disk centered at the sensor, with the radius equal to the sensing range. A sensor covers a target if the Euclidean distance between the sensor and the target is smaller than or equal to a predefined sensing range [14].

Assume each sensor has two units of energy in storage, and one unit can supply each sensor to be active for one time slot. Thus, if all sensors are active continuously, the network lifetime is two time slots. To prolong the network lifetime, we can permit each sensor to sleep alternately to save energy while ensuring all targets are monitored continuously by at least one sensor. In order to cover all the targets, at least two sensors need to keep active at any time. Therefore, we can divide the sensors into three cover sets: $C_1 = \{s_1, s_2\}$, $C_2 = \{s_2, s_3\}$, $C_3 = \{s_1, s_3\}$, and let each cover set be active for one time slot. This scheme will achieve a longer lifetime (i.e., $1 \times 3 = 3$ time slots) irrespective of the activation order of the cover sets.
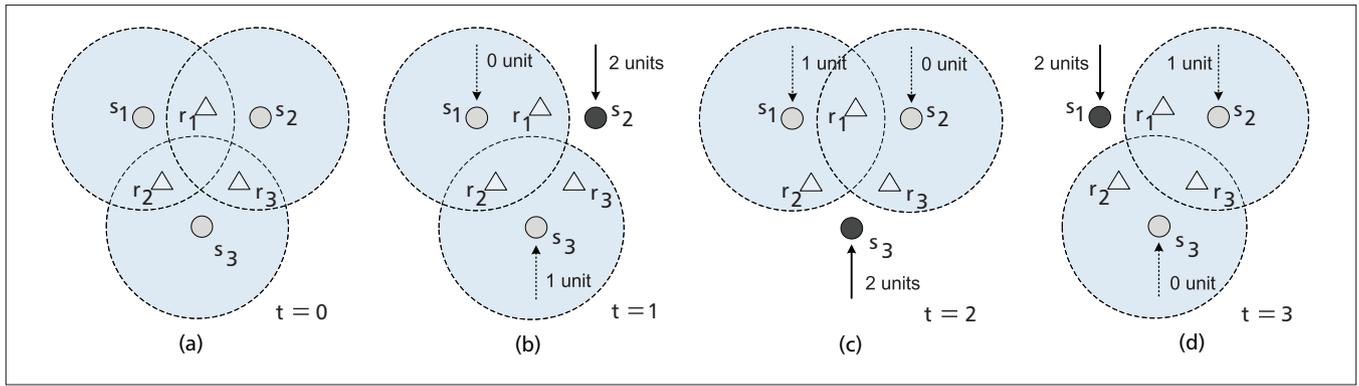
Figure 2. An example with three sensors $C = \{s_1, s_2, s_3\}$ and three targets $R = \{r_1, r_2, r_3\}$.

However, if sensors are equipped with the EH ability, they can get the recharging opportunity for energy supplementation.[1] Take 3 time slots for study, and assume the energy arrival processes for sensor 1, sensor 2, and sensor 3 are (0, 1, 2),[2] (2, 0, 1) and (1, 2, 0), respectively. By exhaustive searching, we can derive the optimal activation scheduling scheme as (C_3, C_1, C_2), as depicted in Figs. 2b–d. That is, C_3, C_1, and C_2 are active in the first, second, and third time slots, respectively. We can see that each sensor enters sleep state to collect energy when its energy arrival reaches the maximum. Although each sensor consumes two units of energy to activate for two time slots, it also harvests two units of energy from the environment. Therefore, through green energy optimization, not only is the energy consumption minimized, but the green energy collection is also maximized. Each sensor obtains sustaining supplementation for its energy consumption, which further enhances the energy efficiency and prolongs the lifetime of the network.

KEY ISSUES AND TECHNICAL CHALLENGES

According to the above example, the optimal usage of green energy depends on characteristics of green energy generation and target distribution, both of which exhibit temporal and spatial diversity.[3] Therefore, the green energy optimization is a challenging problem that involves optimization in two dimensions: the temporal dimension and the spatial dimension.

Dynamic Mode Adaptation: Since mobile targets show temporal dynamics, sensors' energy demands change over time. Moreover, green energy supplements vary along the time horizon. Thus, in order to optimize their performance, sensors should adjust their activation modes adaptively, that is, determine when to activate and when to sleep. If a sensor stays active for more time at the current stage, it can provide better coverage, but more energy is utilized, and it may suffer from tracking discontinuity due to energy shortages in future stages. To solve the temporal mode adaptation problem, parameters such as the current energy arrival and consumption, and estimations of future energy

arrival and consumption should be considered. However, characteristics of energy generation and target distribution change over time, and it is usually not possible to have statistical information before deploying the nodes. Thus, offline optimization frameworks cannot apply to EH-WSNs.

Spatial Energy Balancing: Due to the seamless deployment of sensors, a sensor can enter sleep mode by offloading its covering target to neighboring sensors. In this way, sensors' power consumption is adapted while ensuring complete target coverage. In order to maximize network sustainability, green energy utilization should be optimized by balancing the power consumption among sensors according to the availability of green energy. The power consumption of sensors is balanced by properly deciding the sleep time of each sensor. Sensors that have more harvested energy can keep working longer while letting sensors that are energy deficient enter sleep mode for energy conservation and new energy collection.

In practice, it is difficult to collect global information for centralized operation due to the time-varying characteristics of energy generation and target distribution. Alternatively, decentralized schemes for spatial energy balancing can be considered, which can achieve robust, scalable, and energy-efficient operation.

In summary, the main challenges of green energy optimization in the EH-WSN are listed below:
- Due to the complex coupling of the optimization in temporal and spatial dimensions, it is challenging to achieve optimal green energy utilization.
- The existing static and offline optimization schemes cannot adapt to the dynamics of energy generation and target distribution, which need dynamic and online optimization techniques.
- Energy arrival and target moving are stochastic, which means the information about future energy arrival and target distribution is nondeterministic and unknown.
- Self-organizing sensors perform decentralized information processing for proper operation only based on the local observed information.

[1] It is assumed that sensors cannot harvest green energy when they are active for target monitoring, since monitoring targets and collecting green energy simultaneously complicates the hardware design of the sensors.

[2] The ith element denotes the amount of energy arrival in the ith time slot.

[3] For simplicity, the temporal and spatial diversity of the target distribution are not studied in the above example.
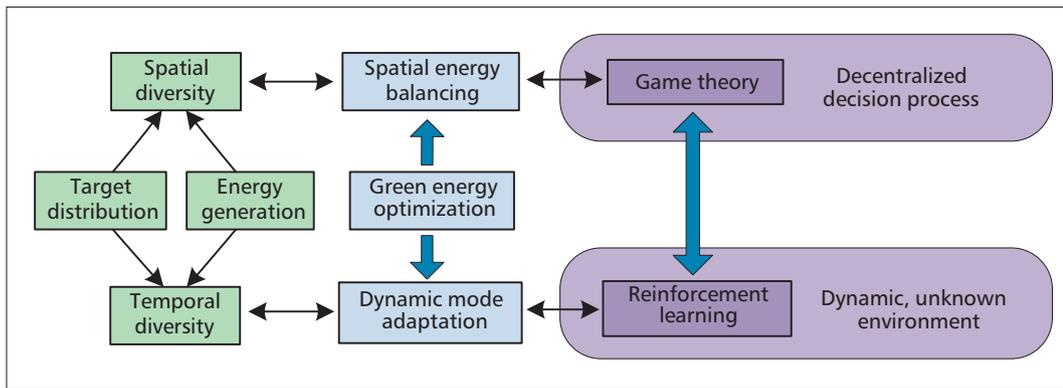
Figure 3. Schematic of green energy optimization in the EH-WSN.

SPATIO-TEMPORAL OPTIMIZATION IN A DECENTRALIZED, DYNAMIC, AND UNKNOWN ENVIRONMENT

In this section, we incorporate reinforcement learning techniques into game theory to deal with the complex coupling of optimization in the temporal and spatial dimensions. Specifically, game theory is adopted to investigate local information based decentralized optimization for the spatial energy balancing problem, while reinforcement learning techniques are employed to address the temporal mode adaptation problem in a dynamic and unknown environment, as shown in Fig. 3.

DECENTRALIZED OPTIMIZATION USING GAME THEORY

Game theory is a mathematical tool applied to model and analyze interactive decision making processes [11–13]. Recently, there has been a great deal of interest in using game theory for analyzing communication networks. It is driven by the need to develop autonomous and flexible network structures as well as design low-complexity distributed algorithms. Generally, a game model consists of three components: a set of players, a set of available actions for each player, and a set of utility functions mapping the action profiles into real numbers. By using game theory, the interactions among multiple interdependent decision makers can be well modeled and analyzed, and the outcome of complex interactions is predictable, and thus can be improved by properly designing the utility function and action update rule of each player.

In WSNs, the channel quality, required packet transmission energy, and acquired data value of each sensor all depend on the activity of other sensors. Due to the complex interactions among individual sensors, game theory becomes an attractive tool to investigate decentralized green energy optimization in the spatial dimension. By formulating a game, each sensor acts as an autonomous game player that observes and reacts to other sensors' behavior in an optimal fashion. This sets up a dynamic system wherein each sensor and its environment (i.e., other sensors) continuously self-adjust to adapt to each other, instead of treating other sensors as static entities [12]. This reactive behavior is also the reason game-theoretic optimization works better than a deterministic scheme for greedy optimization. In the literature, a lot of works use game theory to perform distributed optimization for battery-powered WSNs, but only a few study EH-WSNs.

Michelusi and Zorzi in [1] consider a multi-access game to design an optimal MAC protocol for maximizing the network utility in the EH-WSN, while [2] designs a power control game to maximize sensors' throughput. As for activation scheduling, Niyato *et al.* combine queuing theory and bargaining game to formulate a model for solar-powered WSNs[3], but the interactions among sensors are not analyzed.

For the game design, there are both similarities and differences between the battery-powered WSN and the EH-WSN. In both networks, each sensor's utility depends on its acquired data value and the associated energy cost, which are both affected by the activities of its neighboring sensors. Taking sensor i, for instance: if too many of sensor i's neighbors activate simultaneously, excessive energy is consumed due to the spatio-temporal correlation of sensors' measurements, and the value of data collected by sensor i decreases. Moreover, the probability of successful transmission drops due to channel congestion, which means more energy for packet transmission is required to keep success rate fixed. Therefore, sensors are motivated to activate when the majority of neighbors are in sleep mode and/or its measurement is far from the local aggregated parameter [13]. However, the activation strategy of each sensor in the EH-WSN also depends on its time-varying energy state. If the residual energy is sufficient, it is natural for the sensor to take a positive activation strategy for monitoring targets, which takes the burden off the shoulders of other sensors that are short of energy. On the other hand, when the sensor has less energy in storage, it should be able to enter sleep mode for energy saving as well as new energy collection.

Besides, due to the temporal diversities of the green energy generation and target distribution in the EH-WSN, sensors' energy states, energy consumption, and acquired data values all vary dynamically. Thus, unlike the battery-powered WSN, mainly adopting static and deterministic game models, the dynamic and Bayesian games
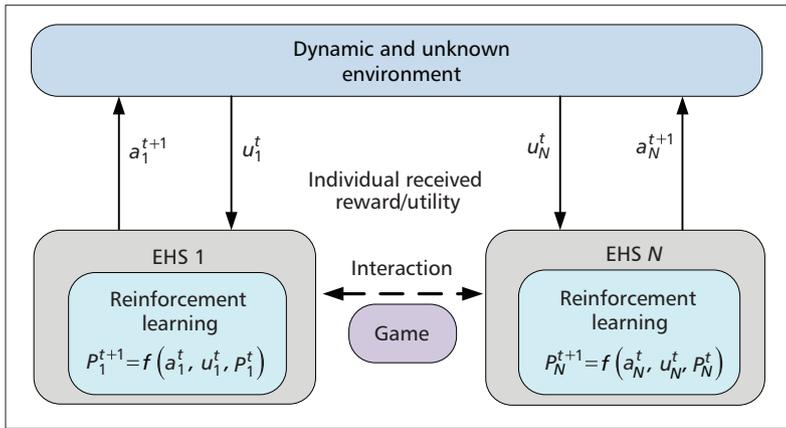
Figure 4. Diagram of the reinforcement learning in the dynamic and unknown EH-WSN. $P_i^t = (P_i^t(0), P_i^t(1))$ is the player i's probability vector for activation strategy at time t, and $P_i^{t+1} = f(a_i^t, u_i^t, P_i^t)$ represents the probability updating rule that depends on the current strategy a_i^t and the received utility u_i^t.

are more suitable to the EH-WSN, which can be formally given by $\mathcal{G} = [\mathcal{N}, \{\mathcal{A}_i\}_{i \in \mathcal{N}}, \mathbf{X}, \{\bar{u}_i\}_{i \in \mathcal{N}}]$. Here, $\mathcal{N} = \{1, 2, \ldots, N\}$ is the set of players[4] (i.e., sensors), $\mathcal{A}_i = \{0, 1\}$ is the set of activation strategies (0 denotes sleep and 1 represents activate) for each player, $\mathbf{X}$ is a random variable characterizing the dynamic and unknown environment, and $\bar{u}_i = \mathbb{E}_{\mathbf{X}}[u_i(\mathbf{X})]$ is the mathematical expectation of the state-based utility function u_i, which is designed to trade off the energy cost of acquiring data against its value, based on the current energy state, that is,

$$u_i = \begin{cases} D_i^t - \gamma \dfrac{E_i^t}{\phi_i^t}, & \text{Activate,} \\ 0, & \text{Sleep,} \end{cases} \quad (1)$$

where D_i^t denotes the value of data collected by sensor i at time t, E_i^t is the amount of energy consumption for activation, ϕ_i^t represents the current energy state of sensor i, and γ is a parameter that weighs the energy cost against its performance. Each player repeatedly plays a game where the actions are whether to activate or sleep and accordingly receives a reward/utility u_i. No pre-computed strategy is given, and players learn their activation strategies through repeated play, continuously adapting their strategies to maximize the expected utility $\bar{u}_i$.

REINFORCEMENT LEARNING TECHNIQUES

Adapting to various unknown and time-varying characteristics in EH systems is a challenging research topic. The results available so far are few and limited. In [6, 10], the authors introduce the Markov decision problem (MDP) to address the dynamic process of data and energy arrival. However, the MDP requires the data/energy state transition to follow the Markov model and the state transition probability to be known. In this subsection, we propose two reinforcement learning techniques for a completely unknown dynamic environment. Reinforcement learning techniques are artificial intelligence tools that provide a system with the necessary information

in order to plan its actions to maximize the reward it receives from the environment [15]. By adapting to the dynamic and stochastic characteristics of the wireless environment, reinforcement learning can result in a significant improvement in network performance.

No-Regret Reinforcement Learning: The no-regret procedure [11] is a regret-based reinforcement learning approach for optimization in a dynamic and unknown environment. In each time period, a player may either continue playing the same activation strategy as in the previous period or switch to another strategy, with probabilities that are proportional to the difference in the accumulated payoff caused by the strategy change, which we call the regret value. Taking time period t, for instance, each player first calculates the utility of the current strategy $a_i \in \mathcal{A}_i$ and the utility of choosing the other strategy $a_i' \in \mathcal{A}_i$, and then updates its regret value R_i^t by

$$R_i^t(a_i, a_i') = \frac{1}{t} \sum_{\tau \leqslant t:\, a_i^{(\tau)} = a_i} \left[u_i^\tau\left(a_i', a_{-i}^{(\tau)}\right) - u_i^\tau\left(a_i^{(\tau)}, a_{-i}^{(\tau)}\right) \right], \quad (2)$$

where $a_{-i}^{(\tau)}$ is the joint strategy of all the players excluding i at time τ. By regret, the player compares its average utility to that of the other activation strategy, and then makes an intelligent decision on the optimal strategy for the next period according to the probability $P_i^{t+1}(a_i') = 1/\mu [R_i^t(a_i, a_i')]^+$, where $[R_i^t(a_i, a_i')]^+ = \max\{R_i^t(a_i, a_i'), 0\}$, and μ is a application-dependent normalization coefficient ensuring the probability in interval $[0, 1]$. It was proven in [11] that the average regret vanishes at the rate of $O(T^{-1/2})$, where T is the number of time periods. Having no regret means that no other strategies would significantly improve the player's utility.

However, existing no regret procedures mainly focus on the *full information model*, in which the utility of every action is observed at each time period, and all the history information needs to be exchanged among neighboring players by necessary communication [12]. This would create heavy signaling overhead for energy-hungry sensor networks. Therefore, we are more concerned with the *partial information model*, where at each time period only the utility/reward of the selected action is observed, as shown in Fig. 4. Each player may not even know the number of players participating in the game, let alone the actions the other players choose. Learning automata (LA) [15] are powerful learning tools that can be used toward this goal.

Reward-Based Learning Automata: At each period, a LA player that resides in a certain state chooses one of the available actions, performs it, and receives a new state from the environment as well as the environmental response. By repeating the above procedure, the LA player continuously interacts with the random operating environment in order to find the optimal strategy among the available actions. Specifically, the probability vector for activation strategy is updated according to the following rules:

---

[4] We will use sensor and game player interchangeably in this article.

$$\begin{cases} P_i^{t+1}(j) = P_i^t(j) + b\tilde{u}_i^t(1 - P_i^t(j)), & \text{if } j = a_i^t, \\ P_i^{t+1}(j) = P_i^t(j) - b\tilde{u}_i^t P_i^t(j), & \text{otherwise,} \end{cases} \quad (3)$$

where $a_i^t \in \mathcal{A}_i$ denotes the activation strategy played by player i at time t, $\tilde{u}_i^t$ is the normalized utility value for choosing a_i^t, $P_i^t(j)$ is the probability of choosing strategy $j \in \mathcal{A}_i$ for player i at time t, and b is the step size controlling the learning rate. The player operates entirely on the basis of their own strategies and the corresponding response from the environment, without any knowledge of other players in the network, and without prior knowledge of state transition probabilities or rewards. Therefore, it is particularly attractive to apply LA to address the dynamic and unknown characteristics in the EH-WSN.

Besides, although game theory copes with the spatial optimization while reinforcement learning deals with the temporal adaptation, the two techniques are executed at the same time to address the complex coupling of the optimization in the temporal and spatial domains, as shown in Fig. 4. The game is played repeatedly in the time horizon, and the player's reward in the dynamic learning process is essentially the game-theoretic utility. In this way, the proposed game-theoretic learning approach effectively solves the two-dimensional optimization of green energy utilization in the EH-WSN.

PERFORMANCE EVALUATION

Unlike the traditional battery-operated WSN, wherein both the target monitoring performance and network lifetime should be considered, the performance evaluation for an EH-WSN is mainly from the perspective of target monitoring, since the capability of harvesting green energy promises a potentially infinite lifetime [8]. In simulations, we consider an EH-WSN where multiple EHSs and 10 targets are randomly scattered in a 100 m² area. The system operates in a time-slotted fashion, where slot t is the time interval $[t\Delta_{TS}, t\Delta_{TS} + \Delta_{TS})$, $t \in \mathbb{Z}^+$, and $\Delta_{TS} = 10$ ms is the time slot duration. The target's location evolves according to a slow Markov process. Specifically, in every slot, each of the 10 targets randomly jumps to a new location with probability $\rho = 0.01$. Besides, we denote the energy harvested in slot t by B_i^t, which is modeled as a Bernoulli random process taking values in {0, 1}, and the probability of harvesting one energy quantum $\Pr(B_i^t = 1)$ varies in {0.1, 0.5}. In addition, the normalization parameter of the no regret procedure is set to be $\mu = 10$, and the step size for the reward-based LA is set to be $b = 0.1$. Similar to [1, 12], the network utility quantifying the performance of target monitoring is defined as the aggregate value of data collected by all sensors, that is, $U = \Sigma_i D_i$, where D_i denotes the value/importance of data reported by sensor i.

Figure 5 plots the convergence behavior of the two reinforcement learning algorithms when 300 EHSs are deployed. It can be seen that the no regret learning procedure converges faster to a better equilibrium, while the reward-based LA converges with more fluctuations to an inferior
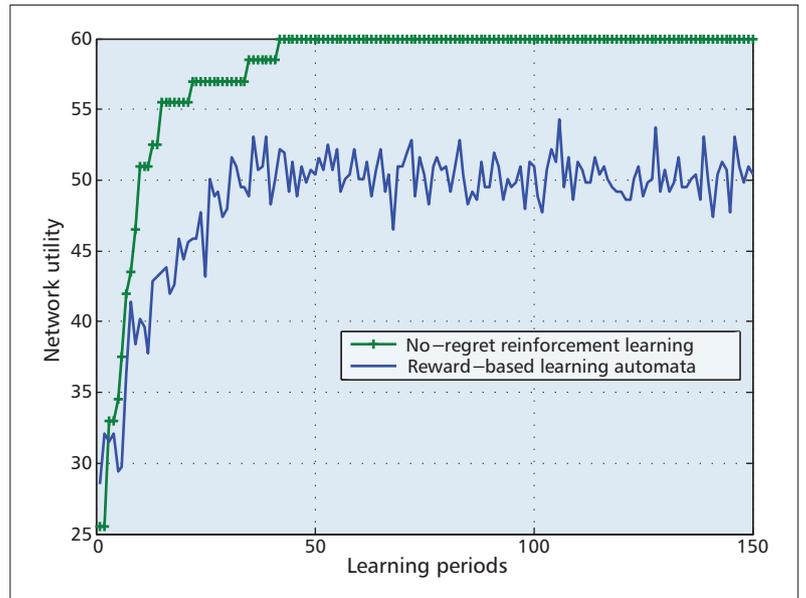

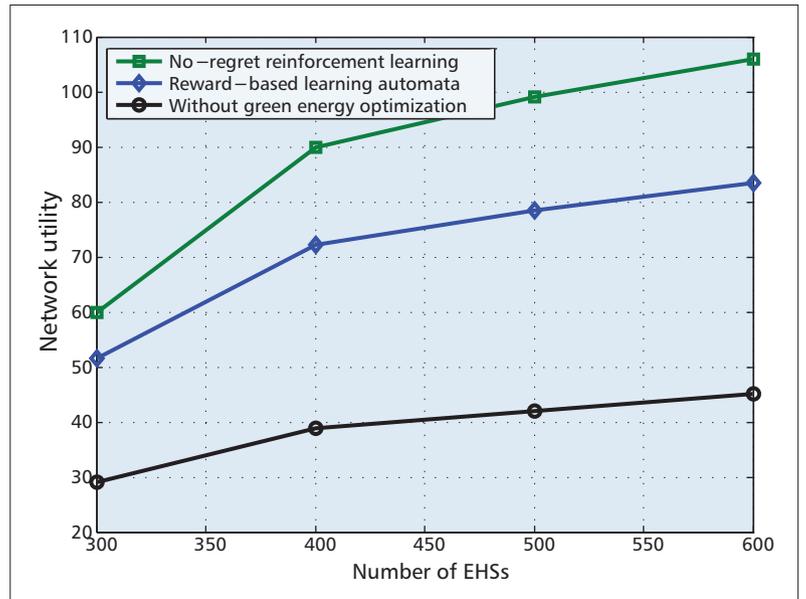**Figure 5.** Convergence behavior of the reinforcement learning algorithms.

Figure 6. Utility performance comparison.

solution. However, the no regret learning procedure requires more information exchange for strategy updating to achieve the *no regret* play, while the LA does not need inter-node communication for information exchange. Besides, according to [11, 15], the computational complexity for both algorithms is $O(|\mathcal{A}_i|)$, where $|\mathcal{A}_i|$ is the cardinality of $\mathcal{A}_i$.

Figure 6 presents a performance comparison for different solutions in terms of the obtained network utility when the number of EHSs varies from 300 to 600. As the number of EHSs increases, the target monitoring performance obtained by all the solutions gets improved, but the proposed reinforcement learning algorithms are much better than the scheme in [12] without green energy optimization. The reinforcement learning algorithms achieve significant perfor-

Simulation results have demonstrated the effectiveness of the proposed approaches. In addition, several research directions have been identified and discussed, aiming to provide insights and guidelines for the researchers in this field.

mance improvement due to their advantages of dealing with the dynamic and stochastic EH environment. Moreover, since the regret/reward value for the algorithm implementation is based on the game-theoretic utility, the green energy utilization is also optimized/balanced among sensors in the spatial dimension.

RESEARCH DIRECTIONS

To better utilize the green energy in EH-WSNs and further improve the network performance, the following potential research topics can be studied.

Designing More Efficient Game Models: The efficiency of the game-theoretic solution depends largely on the design of utility function for each player. This article only considers a simple and intuitive utility function to study the key problem. Further research can improve the game efficiency by carefully designing each player's utility function. Furthermore, different from the non-cooperative game models adopted in this article, cooperative game models can be applied to decrease competition among interactive players and improve energy cooperation among individual players.

Investigating the Trade-off between the Performance and the Cost of the Learning Techniques: The no regret algorithm achieves better performance at the expense of heavier information exchange overhead than the LA algorithm. As can be expected, the green energy optimization performance can be improved at the cost of convergence speed, computational complexity, signaling overhead, and so on. Therefore, it is important to investigate the trade-off between the performance and the cost as well as design proper algorithms based on specific applications.

Studying Energy Cooperation among Different Energy Sources: This article only studies solar-powered systems. However, due to the intermittent and random nature of the EH process, it is better to incorporated multiple energy sources into the network to increase the energy robustness. For example, at night, solar energy may not be available, but there can be wind for energy generation. Therefore, how to properly deploy sensors with different energy supplies and promote energy cooperation among them is a critical and interesting topic.

Studying Multihop Routing, Medium Access Control, Transmission Power Control, and Topology Control in the EH-WSN: Due to the dynamic EH process, these typical problems that are inherent in the traditional battery-operated WSN become quite different for the EH-WSN. However, existing works have not addressed these problems well; hence, they need further investigation.

CONCLUSION

In this article, we have investigated the green energy optimization in an EH-WSN, which involves two subproblems: the dynamic mode adaptation problem in the temporal dimension and the energy balancing problem in the spatial dimension. We have proposed game-theoretic methods with reinforcement learning techniques to deal with the complex coupling of the two-dimensional optimization in the EH-WSN. Simulation results have demonstrated the effectiveness of the proposed approaches. In addition, several research directions have been identified and discussed, aiming to provide insights and guidelines for researchers in this field.

REFERENCES

[1] N. Michelusi and M. Zorzi, "Optimal Random Multiaccess in Energy Harvesting Wireless Sensor Networks," *Proc. IEEE ICC*, 2013.
[2] F. Tsuo *et al.*, "Energy-Aware Transmission Control for Wireless Sensor Networks Powered by Ambient Energy Harvesting: A Game-Theoretic Approach," *Proc. IEEE ICC*, 2011.
[3] D. Niyato et al., "Sleep and Wakeup Strategies in Solar-Powered Wireless Sensor/Mesh Networks: Performance Analysis and Optimization," *IEEE Trans. Mobile Computing*, vol. 6, no. 2, Feb. 2007, pp. 221–36.
[4] Z. Zheng et al., "Sustainable Communication and Networking in Two-tier Green Cellular Networks," *IEEE Wireless Commun.*, Aug. 2014, pp. 47–53.
[5] R. Zhang *et al.*, "MIMO Broadcasting for Simultaneous Wireless Information and Power Transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, May 2013, pp. 1989–2001.
[6] P. Blasco *et al.*, "A learning Theoretic Approach to Energy Harvesting Communication System Optimization," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, Apr. 2013, pp. 1872–82.
[7] T. Han and N. Ansari, "On Optimizing Green Energy Utilization for Cellular Networks with Hybrid Energy Supplies," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, Aug. 2013, pp. 3872–82.
[8] D. Gündüz *et al.*, "Designing Intelligent Energy Harvesting Communications Systems," *IEEE Commun. Mag.*, Jan. 2014, pp. 210–16.
[9] K. Kar *et al.*, "Dynamic Node Activation in Networks of Rechargeable Sensors," *IEEE/ACM Trans. Networking*, vol. 14, no. 1, Feb. 2006, pp. 15–26.
[10] Z. Ren *et al.*, "Dynamic Activation Policies for Event Capture in Rechargeable Sensor Network," *IEEE Trans. Parallel and Distrib. Sys.*, vol. 25, no. 12, Dec. 2014, pp. 3124–34.
[11] J. Zheng *et al.*, "Distributed Channel Selection for Interference Mitigation in Dynamic Environment: A Game-Theoretic Stochastic Learning Solution," *IEEE Trans. Vehic. Tech.*, vol. 63, no. 9, Nov. 2014, pp. 4757–62.
[12] V. Krishnamurthy *et al.*, "Decentralized Adaptive Filtering Algorithms for Sensor Activation in an Unattended Ground Sensor Network," *IEEE Trans. Signal Process.*, vol. 56, no. 12, Dec. 2008, pp. 6086–6101.
[13] O. Gharehshiran *et al.*, "Distributed Energy-Aware Diffusion Least Mean Squares: Game-Theoretic Learning," *IEEE J. Sel.Topics Signal Processing*, vol. 7, no. 5, Oct. 2013, pp. 821–36.
[14] M. Cardei *et al.*, "Energy-Efficient Target Coverage in Wireless Sensor Networks," *Proc. IEEE INFOCOM*, 2005.
[15] P. Nicopolitidis *et al.*, "Adaptive Wireless Networks Using Learning Automata," *IEEE Wireless Commun.*, Apr. 2011, pp. 75–81.

BIOGRAPHY

JIANCHAO ZHENG [S'12] (longxingren.zjc.s@163.com) received a B.S. degree in electronic engineering from the College of Communications Engineering, PLA University of Science and Technology, Nanjing, China, in 2010. He is currently pursuing a Ph.D. degree in communications and information systems in the College of Communications Engineering, PLA University of Science and Technology. His research interests focus on interference mitigation techniques, learning theory, game theory, and optimization techniques.

Yueming Cai [M'05,SM'12] (caiym@vip.sina.com) received a B.S. degree in physics from Xiamen University, China, in 1982, and an M.S. degree in micro-electronics engineering and a Ph.D. degree in communications and information systems, both from Southeast University, Nanjing, China, in 1988 and 1996, respectively. His current research interests include cooperative communications, signal processing in communications, wireless sensor networks, and physical layer security.

Xuemin (Sherman) Shen [M'97, SM'02, F'09] (sshen@uwaterloo.ca) received a B.Sc.(1982) degree from Dalian Maritime University, China, and M.Sc. (1987) and Ph.D. (1990) degrees from Rutgers University, New Jersey, all in electrical engineering. He is a professor and University Research Chair, Department of Electrical and Computer Engineering, University of Waterloo, Canada. He was the associate chair for graduate studies from 2004 to 2008. His research focuses on resource management in interconnected wireless/wired networks, wireless network security, social networks, smart grid, and vehicular ad hoc and sensor networks. He is a co-author/editor of six books, and has published more than 600 papers and book chapters in wireless communications and networks, control, and filtering. He is an elected member of the IEEE ComSoc Board of Governors and Chair of the Distinguished Lecturers Selection Committee. He served as TPC Chair/Co-Chair for IEEE INFOCOM '14, IEEE VTC-Fall '10, and IEEE GLOBECOM '07; Symposia Chair for IEEE ICC '10; Tutorial Chair for IEEE VTC-Spring '11 and IEEE ICC '08, General Co-Chair for Chinacom '07 and QShine '06; and as Chair of the IEEE Communications Society Technical Committees on Wireless Communications, and P2P Communications and Networking. He also serves/has served as Editor-in-Chief of *IEEE Network*, *Peer-to-Peer Networking and Application*, and *IET Communications*; a Founding Area Editor for *IEEE Transactions on Wireless Communications*; an Associate Editor for *IEEE Transactions on Vehicular Technology*, *Computer Networks*, *ACM/Wireless Networks*, and others; and Guest Editor for *IEEE JSAC*, *IEEE Wireless Communications*, *IEEE Communications Magazine*, *ACM Mobile Networks and Applications*, and so on. He received the Excellent Graduate Supervision Award in 2006, and the Outstanding Performance Award in 2004, 2007, and 2010 from the University of Waterloo; the Premier's Research Excellence Award (PREA) in 2003 from the Province of Ontario, Canada; and the Distinguished Performance Award in 2002 and 2007 from the Faculty of Engineering, University of Waterloo. He is a registered Professional Engineer of Ontario, Canada, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, and a Distinguished Lecturer of the IEEE Vehicular Technology and Communications Societies.

Zhongming Zheng (forezero@gmail.com) received B.Eng. (2007) and M.Sc. (2010) degrees from the City University of Hong Kong. Currently, he is pursuing his Ph.D degree in electrical and computer engineering at the University of Waterloo in the Broadband Communication Research Group. His research focuses on green wireless communication, smart grid, and wireless sensor networks.

Weiwei Yang (wwyang1981@163.com) received B.S., M.S. and Ph.D. degrees from the Institute of Communications Engineering, PLA University of Science and Technology, in 2003, 2006, and 2011, respectively. He is assistant professor with the College of Communications Engineering, PLA University of Science and Technology. His research interests are mainly in OFDM systems, signal processing in communications, cooperative communications, and physical layer security.