Game-Theoretic Multi-Channel Multi-Access in Energy Harvesting Wireless Sensor Networks

Jianchao Zheng, *Student Member, IEEE*, Honggang Zhang, *Senior Member, IEEE*, Yueming Cai, *Senior Member, IEEE*, Rongpeng Li, and Alagan Anpalagan, *Senior Member, IEEE*

***Abstract*— Energy harvesting (EH) has been proposed as a promising technology to extend the lifetime of wireless sensor networks (WSNs) by continuously harvesting green/renewable energy. However, the intermittent and random EH process as well as the complexity in achieving global network information call for efficient energy management and distributed resource optimization. Considering the complex interactions among individual sensors, we use the game theory to perform distributed optimization for the general multi-channel multi-access problem in an EH-WSN, where strict delay constraints are imposed for the data transmission. Sensors' competition for channel access is formulated as a non-cooperative game, which is proved to be an ordinal potential game that has at least one Nash equilibrium (NE). Furthermore, all the NE of the game is proved to be Pareto optimal, and Jain's fairness index bound of the NE is theoretically derived. Finally, we design a fully distributed, online learning algorithm for the multi-channel multi-access in the EH-WSN, which is proved to converge to the NE of the formulated game. Simulation results validate the effectiveness of the proposed algorithm.**

***Index Terms*— Wireless sensor network, energy harvesting, multi-channel multi-access, potential game, Nash equilibrium, distributed, online learning algorithm.**

I. Introduction

Wireless sensor networks (WSNs) have profound significance towards environmental surveillance and monitoring by placing sensors up close to the phenomena of interest [1]–[5]. However, the limited network lifetime is a major deployment barrier for the traditional battery-operated WSN. In recent years, energy harvesting (EH) has emerged as a promising technology to prolong the lifetime of communication networks by continuously harvesting green/renewable energy from environmental sources, such as sun, wind, vibrations, etc. [6]–[9].

Manuscript received February 3, 2016; accepted February 28, 2016. Date of publication March 8, 2016; date of current version April 26, 2016. This work was supported in part by the Basic Research Program of Jiangsu Province National Natural Science Foundation of China under Grant BK20130067 and in part by the Project of National Natural Science Foundation of China under Grant 61301162 and Grant 61301163. The associate editor coordinating the review of this paper and approving it for publication was Dr. Rosario Morello.

J. Zheng and Y. Cai are with the College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China (e-mail: longxingren.zjc.s@163.com; caiym@vip.sina.com).

H. Zhang is with the Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: honggangzhang@zju.edu.cn).

R. Li is with Huawei Technologies Company, Ltd., Shanghai 210652, China (e-mail: lirongpeng@huawei.com).

A. Anpalagan is with the Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada (e-mail: alagan@ee.ryerson.ca).

Digital Object Identifier 10.1109/JSEN.2016.2539346

Due to the uncertain and dynamically changing environmental conditions, the EH process is intermittent and random in nature, which calls for efficient energy management to guarantee sustainable and reliable network operation [6]–[9]. For convenience of analysis, most of existing works assume that either the transmitter possesses non-causal information on the exact data/energy arrival, or the transmitter knows the statistics of EH and data arrival processes [6]. However, in most practical scenarios, the characteristics of EH and data arrival processes may change with time. Moreover, it may not be possible to know reliable statistical information about these processes before deploying the nodes. Therefore, non-causal information about the data/energy arrival instants and amounts may be infeasible, and thus offline optimization frameworks may not be satisfactory in many practical applications [6], [7].

Besides, existing research on EH mostly focuses on the point-to-point communication system [6], [10], while the network consisting of multiple EH nodes is more challenging to analyze [11]. In general, the complexity in achieving optimal energy utilization policies increases significantly with the number of nodes in the network [11]. Furthermore, the available knowledge of the EH profiles across different sensors is hard to obtain or even unattainable. Therefore, distributed optimization approaches based only on local information should be sought. Game theory is a powerful mathematical tool that models interactive decision-making processes and has been widely used to derive distributed resource allocation in wireless networks [12]–[14]. Due to the complex interactions among distributed sensors, we adopt game theory to exploit the distributed optimization in the autonomous EH-WSN.

In the literature, a lot of work uses game theory to perform distributed optimization for the battery-powered WSN, but only a few study the EH-WSN [1]–[5]. Niyato *et al.* combine queuing theory and bargaining game to study the activation scheduling for solar-powered WSNs [1], but the interactions among sensors are not analyzed. Authors in [2] and [3] design power control games to maximize sensors' throughput, while Michelusi and Zorzi in [4] and [5] consider a multi-access game for maximizing the network utility in the EH-WSN where multiple EH sensors (EHSs) randomly access a shared wireless channel to transmit data packets of random importance to a common fusion center. In practice, the data transmission is usually operated in the multi-channel wireless environment [15], [16]. Thus, we extend the work in [4] and [5] to the case of multiple available channels, and use game theory to study the optimal multi-channel

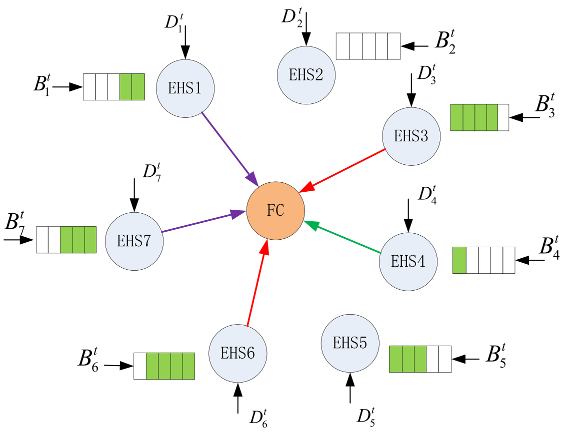


Fig. 1. Energy harvesting wireless sensor network.

multi-access scheme. We also consider strict delay constraints for the data transmission, which is very important in some emergency applications [17], [18] and in the industrial automation and avionics domain [19], [20]. The main contributions of this paper are listed below:

- We formulate the multi-channel multi-access problem in the EH-WSN as a non-cooperative game, which is proved to be an ordinal potential game that has at least one Nash equilibrium (NE). Moreover, all the NE are proved to be Pareto-optimal. The Jain's fairness index (JFI) bound of the NE is also derived.
- Based on the stochastic learning automata, we design a fully distributed, online learning algorithm to find the NE solution to the problem of random multi-channel multi-access in the EH-WSN. The convergence of the algorithm is theoretically analyzed.
- Extensive simulations are conducted to evaluate the performance of the proposed stochastic learning algorithm. Different network scales and different energy harvesting rates are both discussed.

The remainder of this paper is organized as follows. In Section II, we introduce the system model and problem formulation. In Section III, we present a non-cooperative game to analyze the multi-channel multi-access in the EH-WSN. In Section IV, we propose a fully distributed, online adaptive, stochastic learning algorithm to find the NE solution without any information exchange. In Section V, we validate our analysis through simulation. Finally, concluding remarks are given in Section VI.

II. System Model and Problem Formulation

A. System Model

In this work, we consider a network of N EHSs that communicate via M wireless links with a fusion center (FC), as depicted in Fig. 1. The set of EHSs and the set of available channels are denoted by $\mathcal{N} = \{1, 2, \ldots, N\}$, $\mathcal{M} = \{1, 2, \ldots, M\}$, respectively. Each EHS selects one channel for its data transmission. Due to the scarcity of spectrum in practical systems, the number of EHSs N is usually (much) larger than the number of channels M, and thus multiple EHSs need to reuse one channel. We consider a time-slotted fashion, where slot t is the time interval $[tT, tT + T)$, $t \in \mathbb{Z}^+$ and T is the time-slot (TS) duration. At each slot (say t), each EHS (say i) has a new data packet to transmit, and each data packet is of importance $D_i^t > 0$, which is modeled as a continuous random variable with probability density function (pdf) $f_D(x)$, $x \geq 0$.

Each EHS collects ambient energy, which is stored in a rechargeable battery and powers the sensing apparatus and the RF circuitry. We denote the energy harvested in TS t by B_i^t that is modeled as a Bernoulli random process taking values in $\{0, 1\}$. We assume that the transmission of one data packet requires the expenditure of one energy quantum, and use $\{Q_i^t\}$ to denote the action process at EHS i. $Q_i^t = 1$ if the data packet at TS t is transmitted by EHS i, and $Q_i^t = 0$ otherwise. The battery of each EHS is modeled by a buffer with a capacity of $e_{\max}$, and the set of possible energy levels is denoted by $\varepsilon = \{0, 1, \ldots, e_{\max}\}$ because the harvested energy takes values in $\{0, 1\}$. Denoting the amount of energy stored in the battery of EHS i at TS t by E_i^t, the evolution of E_i^t over time t is given by

$$E_i^{t+1} = \min\{E_i^t - Q_i^t + B_i^t, e_{\max}\}. \tag{1}$$

We consider a collision model for packet transmission. The packet of each EHS is successfully delivered to FC if and only if all the other EHSs on the same channel do not transmit. When a collision occurs, the data packet is lost. Moreover, we consider a strict delay constraint for the transmission of data packets; that is, each data packet needs to be transmitted within the TS following its arrival, and it is discarded if not transmitted within the following TS. Obviously, the strict delay constraint could lead to low transmission efficiency for each EHS. However, the strict delay constraint is vital in some emergency situations (e.g., earthquake-tsunami alerting, fire detection in forests, etc.) [17], [18], and in the industrial automation and avionics domain [19], [20].

B. Problem Formulation

It has been shown in [4] and [5] that, for maximizing the average long-term importance of the data reported by each EHS to FC, the following threshold policy is optimal:

$$Q_i^t = \chi(D_i^t \geq \Gamma_i(E_i^t)), \tag{2}$$

where i denotes an arbitrary EHS, $\chi(\cdot)$ is the indicator function, and $\Gamma_i(E_i^t)$ is the importance threshold which is a function of the current energy level E_i^t. According to (2), the EHS only transmits when $D_i^t \geq \Gamma_i(E_i^t)$. We denote the corresponding transmission probability of EHS i in energy level e, induced by the random importance D_i^t, by $\eta_i(e)$. Specifically,

$$\eta_i(e) = \mathbb{E}_{D_i^t}[Q_i^t | E_i^t = e] = \Pr(D_i^t \geq \Gamma_i(e)). \tag{3}$$

The set of admissible policies for EHS i is defined as[1]

$$\Upsilon = \{\eta_i : \eta_i(0) = 0, \eta_i(e_{\max}) \in (0, 1], \eta_i(e) \in (0, 1), \\ e \neq 0, e_{\max}\}. \tag{4}$$

[1]According to (4), $\eta_i(0) = 0$. That is, the EHS will not transmit data when the energy level $e = 0$. Thus, this condition guarantees the non-negative energy state.

If all the other EHSs remain idle (do not transmit), the expected data importance reported by EHS i to FC in energy level e, is given by

$$g(\eta_i(e)) = \mathbb{E}_{D_i^t}[D_i^t Q_i^t | E_i^t = e]$$
$$= \mathbb{E}_{D_i^t}[\chi(D_i^t \geq \Gamma_i(e)) D_i^t] = \int_{\Gamma_i(e)}^{\infty} x f_D(x) dx. \quad (5)$$

Let $\mathbf{E}^t = (E_1^t, E_2^t, \ldots, E_N^t)$ denote the joint state of the energy levels in all EHSs. It has been shown in [5] that the Markov chain $\{\mathbf{E}^t\}$ under the aggregate policy $\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \eta_N) \in \Upsilon^N$ is irreducible, and thus there exists a unique steady-state distribution, $\pi_{\boldsymbol{\eta}}(\mathbf{e})$, $\mathbf{e} \in \varepsilon^N$. Thus, we define the utility function of each EHS as the average long-term aggregate importance of its reported data, given by

$$u_i(a_i, a_{-i}) = \sum_{\mathbf{e} \in \varepsilon^N} \pi_{\boldsymbol{\eta}}(\mathbf{e}) g\left(\eta_i(e_i)\right) \prod_{j \in \mathcal{K}_i} \left(1 - \eta_j(e_j)\right), \quad (6)$$

where a_i denotes the channel selection strategy of EHS i, $a_{-i} = (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_N)$ denotes a channel selection profile of all the EHSs excluding EHS i, $\mathcal{K}_i = \{j \in \mathcal{N} : a_j = a_i\}$ denotes the set of EHSs that choose the same channel as EHS i. Since the utility of EHS i depends not only on its own strategy, but also on the strategies of other EHSs, we denote the utility function as $u_i(a_i, a_{-i})$. By letting $G(\eta_i) = \sum_{e=1}^{e_{\max}} \pi_{\eta_i}(e) g(\eta_i(e))$ and $P(\eta_i) = \sum_{e=1}^{e_{\max}} \pi_{\eta_i}(e) \eta_i(e)$, (6) can be re-written as

$$u_i(a_i, a_{-i}) = G\left(\eta_i\right) \prod_{j \in \mathcal{K}_i} \left(1 - P\left(\eta_j\right)\right). \quad (7)$$

In order to guarantee fairness among the EHSs, we also consider symmetric control policies, i.e., all the EHSs adopt the same transmission policy $\eta_i = \eta$, $\forall i$. Thus,

$$u_i(a_i, a_{-i}) = G\left(\eta\right) \left(1 - P\left(\eta\right)\right)^K, \quad (8)$$

where $K = |\mathcal{K}_i|$ represents the number of EHSs choosing the same channel as EHS i, $|\mathcal{K}_i|$ is the cardinality of the set $\mathcal{K}_i$.

In [4] and [5], the authors have investigated the optimal transmission policy $\boldsymbol{\eta}^*$ in the single-channel system. Given that the data transmission is usually operated in the multi-channel wireless environment [15], [16], we extend the work in [4] and [5] to the multi-channel system, and further study the optimal channel allocation for minimizing transmission collision and improving the network performance. We also consider a fully distributed operation model, in which each EHS possesses only local knowledge about the system state. That is, EHS i's operation at TS t is based only on its own energy level and data importance (D_i^t, E_i^t), and does not need to know the energy level and data importance of other EHSs. There is no central controller for managing the operations of each EHS, which avoids the failure of single point and may demand less information exchange (i.e., signaling overhead) and computational complexity.

III. Channel Access Game

Game theory is a powerful mathematical tool that models and analyzes interactive decision-making processes [12], [13]. Due to the complex transmission collisions among individual EHSs, we adopt game theory to investigate the local information based distributed channel access in the EH-WSN, where each EHS acts as a game player. Mathematically, the game can be formulated as $\mathcal{G} = \left[\mathcal{N}, \{\mathcal{A}_i\}_{i \in \mathcal{N}}, \{u_i\}_{i \in \mathcal{N}}\right]$, where $\mathcal{N}$ is the set of players, $\mathcal{A}_i = \mathcal{M}$ is the set of available actions (channels) for each player i that independently maximizes its individual utility function u_i, i.e.,

$$(\mathcal{G}): \max_{a_i \in \mathcal{A}_i} u_i\left(a_i, a_{-i}\right), \quad \forall i \in \mathcal{N}. \quad (9)$$

Definition 1 (Nash Equilibrium): A channel selection profile $\boldsymbol{a}^* = \left(a_1^*, \ldots, a_N^*\right)$ is a pure-strategy NE if and only if no player can improve its utility by deviating unilaterally, i.e.,

$$u_i\left(a_i^*, a_{-i}^*\right) \geq u_i\left(a_i, a_{-i}^*\right), \quad \forall i \in \mathcal{N}, \quad \forall a_i \in \mathcal{A}_i. \quad (10)$$

Definition 2 (Ordinal Potential Game): A game is an ordinal potential game (OPG) if there exists an ordinal potential function $\Phi : \mathcal{A}_1 \cdots \times \mathcal{A}_N \to R$ such that $\forall i \in \mathcal{N}$, $a_i \in \mathcal{A}_i$, $a_i' \in \mathcal{A}_i$,

$$u_i(a_i', a_{-i}) > u_i(a_i, a_{-i}) \Leftrightarrow \Phi(a_i', a_{-i}) > \Phi(a_i, a_{-i}). \quad (11)$$

It means the variation of the utility function caused by any unilateral strategy update has the same trend with that of the ordinal potential function. Based on the finite improvement property, it has been proved in [21] that every OPG has at least one pure-strategy NE.

Theorem 1: The channel access game $\mathcal{G}$ is an ordinal potential game that has at least one pure-strategy NE point.

Proof: According to (8), the utility u_i is a function of the variable K, η, and thus can be represented as $f(\eta, K)$. According to [4] and [5], the optimal transmission policy η^* in the distributed operation can be uniquely determined by using the symmetric NE solution[2] that maximizes the utility function. That is, $u_i = f(\eta^*, K) = \max_{\eta} f(\eta, K)$.

Suppose EHS i unilaterally deviates from its current channel $a_i = m$ to select another channel $a_i' = n$, and denote the number of EHSs (excluding EHS i) selecting channel m and n by K_m and K_n, respectively. Thus, $u_i(a_i, a_{-i}) = \max_{\eta} f(\eta, K_m)$, $u_i(a_i', a_{-i}) = \max_{\eta} f(\eta, K_n)$. If $K_n < K_m$, $\forall \eta$,

$$\begin{aligned} f(\eta, K_m) &= G\left(\eta\right) \left(1 - P\left(\eta\right)\right)^{K_m}, \\ &= G\left(\eta\right) \left(1 - P\left(\eta\right)\right)^{K_n} \left(1 - P\left(\eta\right)\right)^{K_m - K_n} \\ &< G\left(\eta\right) \left(1 - P\left(\eta\right)\right)^{K_n} \\ &= f(\eta, K_n). \end{aligned} \quad (12)$$

Letting $\eta_n^* = \arg\max_{\eta} f(\eta, K_n)$ and $\eta_m^* = \arg\max_{\eta} f(\eta, K_m)$, we have

$$\begin{aligned} u_i(a_i, a_{-i}) &= f(\eta_m^*, K_m) \overset{(a)}{<} f(\eta_m^*, K_n) \\ &\overset{(b)}{\leq} f(\eta_n^*, K_n) = u_i(a_i', a_{-i}), \end{aligned} \quad (13)$$

where the derivation of step (a) is based on (12), and step (b) holds because $\eta_n^* = \arg\max_{\eta} f(\eta, K_n)$. To summarize,

$$K_n < K_m \Leftrightarrow u_i(a_i, a_{-i}) < u_i(a_i', a_{-i}). \quad (14)$$

[2]In the symmetric NE η^*, all the EHSs employ the same policy that maximizes its utility, and have no incentive to deviate from it. In other words, the symmetric NE η^* is simultaneously optimal for all the EHSs.

In order to prove the existence of NE, we resort to the potential game theory [21], and define a function as

$$\Phi(a_i, a_{-i}) = \frac{1}{\sum_{s=1}^{M} N_s^2}, \tag{15}$$

where $N_s = \sum_{i=1}^{N} \delta(a_i, s)$ denotes the number of EHSs on channel $s \in \mathcal{M}$, $\delta(\cdot)$ is the Kronecker delta function defined as

$$\delta(x, y) = \begin{cases} 1, & \text{if } x = y, \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

For convenience of analysis, we denote the number of EHSs on each channel $s \in \mathcal{M}$ under channel selection profile (a_i, a_{-i}) as N_s, and that under channel selection profile (a_i', a_{-i}) as N_s'. Since $a_i = m$ and $a_i' = n$, it is easy to know $N_m' = N_m - 1$, and $N_n' = N_n + 1$. Thus,

$$N_m' - N_n' = N_m - N_n - 2 < N_m - N_n. \tag{17}$$

According to (12), if $u_i(n, a_{-i}) > u_i(m, a_{-i})$, we have $N_m = K_m + 1 > N_n' = K_n + 1$. Since $N_m' = N_m - 1$, we can get $N_m' - N_n' \geq 0$, which along with (17) yields

$$|N_m' - N_n'| < |N_m - N_n|. \tag{18}$$

Besides, since the strategies of other EHSs do not change, we have

$$\begin{aligned} N_n' + N_m' &= N_n + N_m, \text{ and} \\ N_l' &= N_l, \quad \forall l \neq n, m. \end{aligned} \tag{19}$$

It is well known that $x^2 + y^2 = \frac{1}{2}[(x+y)^2 + |x-y|^2]$, $\forall x, y$. By applying (18) and (19), we can get ${N_n'}^2 + {N_m'}^2 < N_n^2 + N_m^2$, and thus $\sum_{s=1}^{M} {N_s'}^2 < \sum_{s=1}^{M} N_s^2$. Therefore, $\Phi(a_i', a_{-i}) > \Phi(a_i, a_{-i})$. To summarize,

$$u_i(a_i', a_{-i}) > u_i(a_i, a_{-i}) \Leftrightarrow \Phi(a_i', a_{-i}) > \Phi(a_i, a_{-i}). \tag{20}$$

According to Definition 2, we know the game $\mathcal{G}$ is an ordinal potential game with potential function Φ, and thus $\mathcal{G}$ has at least one pure-strategy NE. Therefore, Theorem 1 is proved. ∎

In game theory, efficient resource utilization is determined by the concept of Pareto optimality [12]. A well-known way of characterizing efficiency of an equilibrium is to know whether it is Pareto-optimal, which is defined next.

Definition 3 (Pareto Optimality): A channel selection profile $\boldsymbol{a}^{PO}$ is Pareto-optimal if and only if there exists no other action profile $\tilde{\boldsymbol{a}}$ such that for some $i \in \mathcal{N}$, $u_i(\tilde{\boldsymbol{a}}) > u_i(\boldsymbol{a}^{PO})$, and $\forall j \neq i$, $u_j(\tilde{\boldsymbol{a}}) \geq u_j(\boldsymbol{a}^{PO})$. In other words, $\boldsymbol{a}^{PO}$ is Pareto-optimal if there exists no other feasible allocation that can make one player better without making any other worse.

Theorem 2: All the NE of the game $\mathcal{G}$ are Pareto-optimal.

Proof: We prove this theorem by contradiction. Assume the NE $\boldsymbol{a}^* = (a_1^*, \ldots, a_N^*)$ is not Pareto-optimal. That is, there exists an action profile $\tilde{\boldsymbol{a}}$ such that for some $i \in \mathcal{N}$, $u_i(\tilde{\boldsymbol{a}}) > u_i(\boldsymbol{a}^*)$, and $\forall j \neq i$, $u_j(\tilde{\boldsymbol{a}}) \geq u_j(\boldsymbol{a}^*)$. According to (14), $u_i(\tilde{\boldsymbol{a}}) > u_i(\boldsymbol{a}^*) \Leftrightarrow N_{a_i^*} > N_{\tilde{a}_i}$, and $\forall j \neq i$, $u_j(\tilde{\boldsymbol{a}}) \geq u_j(\boldsymbol{a}^*) \Leftrightarrow N_{a_j^*} \geq N_{\tilde{a}_j}$. Thus,

$$N_{a_i^*} + \sum\nolimits_{j \in \mathcal{N}, j \neq i} N_{a_j^*} > N_{\tilde{a}_i} + \sum\nolimits_{j \in \mathcal{N}, j \neq i} N_{\tilde{a}_j}. \tag{21}$$

However, since the number of total EHSs, N, keeps invariant regardless of the specific channel selection profile, i.e., $N_{a_i^*} + \sum_{j \in \mathcal{N}, j \neq i} N_{a_j^*} = N_{\tilde{a}_i} + \sum_{j \in \mathcal{N}, j \neq i} N_{\tilde{a}_j} = N$, which contradicts with (21). Therefore, the assumption that NE is not Pareto-optimal does not hold. The proof is completed. ∎

Besides, fairness is another important notion for resource allocation in wireless networks. It is believed that the game-theoretic approach can achieve fair solutions, since all players try to select their best strategies to maximize their own utilities and no player is willing to accept low utility. In the following, we analyze the fairness of the NE by using the well-known Jain's fairness index (JFI) [22], [23], which is defined as

$$J_{\mathcal{G}} = \frac{\left(\sum_{i=1}^{N} u_i\right)^2}{N \sum_{i=1}^{N} u_i^2}, \tag{22}$$

where u_i denotes the utility of EHS i. JFI maps a resource allocation vector $\{u_1, u_2, \ldots, u_N\}$ into a score in the interval $[1/N, 1]$, and a higher JFI demonstrates a fairer resource allocation scheme.

Theorem 3: For the formulated game $\mathcal{G}$, the JFI of the NE solution is lower bounded by $\frac{8}{9}$.

Proof: In the NE,

$$|N_m - N_n| \leq 1, \quad \forall m, n, \tag{23}$$

which can be easily proved by contradiction. Assume $\exists m, n$, $N_m - N_n > 1$. According to (14), the EHSs on channel m will definitely deviate their current channel to select channel n in order to improve their own utilities, which contradicts with the definition of NE. This completes the proof of (23).

For convenience of analysis, we divide the problem into two cases. 1) If $N_m = N_n$, according to (8), the utilities of EHSs on channel m and channel n are equal, i.e., $u_i = u_j$, where $a_i = m, a_j = n$; 2) if $N_m - N_n = 1$, according to (14), $u_i < u_j$. Besides, in case 2), due to the motivation of deploying EHSs, more deployed EHSs should bring higher system utility, i.e., $N_m \cdot u_i < N_n \cdot u_j$. Thus, we have

$$1 \leq \frac{u_j}{u_i} < \frac{N_m}{N_n} = \frac{N_n + 1}{N_n} \leq 2, \quad \forall i, j \in \mathcal{N}. \tag{24}$$

We define $u_{\max} = \max\{u_1, u_2, \ldots, u_N\}$, $u_{\min} = \min\{u_1, u_2, \ldots, u_N\}$, and $\alpha = u_{\max}/u_{\min}$. Obviously, $1 \leq \alpha \leq 2$. According to [22], the JFI of $\mathcal{G}$ is lower bounded by:

$$J_{\mathcal{G}} \geq \frac{4\alpha}{(\alpha+1)^2}. \tag{25}$$

Moreover, it is easy to know that $\frac{4\alpha}{(\alpha+1)^2}$ is a decreasing function of α in the interval $[1, \infty)$. Therefore, we can derive the lower bound of $J_{\mathcal{G}}$ as:

$$J_{\mathcal{G}} \geq \frac{4 \cdot 2}{(2+1)^2} = \frac{8}{9}. \tag{26}$$

This concludes the proof. ∎

Algorithm 1 SLA Based Dynamic Channel Access

Initialization: Let $\mathbf{p}_i^t = (p_{i1}^t, \ldots, p_{iM}^t)$ denote EHS i's channel selection probability vector in the t-th time slot, where p_{im}^t denotes the probability for EHS i to select channel m. At the initial time $t = 0$, set $\mathbf{p}_i^t = (\frac{1}{M}, \ldots, \frac{1}{M})$ for all $i \in \mathcal{N}$, based on which each EHS randomly selects a channel.

Loop for $t = 0, 1, 2, \ldots$

1) **Updating transmission strategy:** In the t-th time slot, each EHS, say i, decides its optimal transmission strategy η_i by using [5, Algorithms 1 and 2], which is based on only local information (D_i^t, E_i^t).
2) **Transmitting data packet:** Based on the transmission strategy η_i obtained in Step 1), each EHS transmits its data packet.
3) **Computing utility value:** By detecting the successfully transmitted data packet, each EHS computes its individual utility u_i^t.
4) **Updating channel selection probability:** All the EHSs update their channel selection probabilities according to the following rules:

$$\begin{cases} p_{im}^{t+1} = p_{im}^t + b\, r_i^t (1 - p_{im}^t), & m = a_i^t, \\ p_{im}^{t+1} = p_{im}^t - b\, r_i^t\ p_{im}^t, & m \neq a_i^t, \end{cases} \tag{27}$$

where $0 < b < 1$ is the learning step size, and r_i^t is the normalized received utility defined as $r_i^t = u_i^t / u_i^{\max}$.
5) **Updating channel strategy:** Each EHS updates its channel strategy a_i^{t+1} for the next slot based on its current channel selection probability vector $\mathbf{p}_i^t$.

End loop until one component of $\mathbf{p}_i^t$ approaches 1 (e.g., larger than 0.99), $\forall i \in \mathcal{N}$.

IV. Stochastic Learning for Achieving NE

It has been shown that the NE solution of the formulated game $\mathcal{G}$ achieves both Pareto-optimality and fairness. In this section, based on the stochastic learning automata (SLA) [24], we design a fully distributed, online learning algorithm to find the NE of the channel access game in the EH-WSN. The details of the algorithm is shown in Algorithm 1, and the schematic is described in Fig. 2.

The proposed SLA based dynamic channel access scheme operates iteratively, where each EHS acts as a SLA player. In each iteration period, each SLA player decides its transmission strategy, chooses one of the available channels for transmission, and receives a response (i.e., successfully transmitted data) from the environment. By repeating the above procedure, the SLA player continuously interacts with the random operating environment and finally achieves the optimal strategy among the available channels. Each player operates completely on the basis of its own strategies and the corresponding response from the environment, without any knowledge of the other players in the network, and without prior knowledge of state transition probabilities of the EH and data arrival. Therefore, the proposed algorithm is fully distributed. Moreover, the proposed algorithm is operated online, since each EHS transmits data during each iteration. According to the

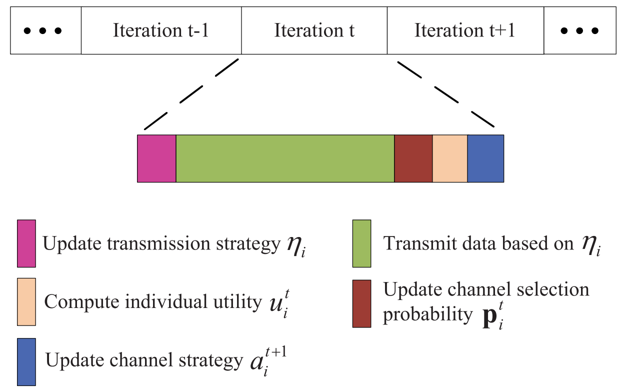


Fig. 2. Schematic of the proposed stochastic learning algorithm.

analysis in [25] and [26], the proposed algorithm runs with computational complexity of $O(T_{it} N)$, where N is the number of nodes, and T_{it} denotes the number of iterations needed for the algorithm convergence.

In order to investigate the convergence of the proposed stochastic learning algorithm, we firstly re-write the updating rule in (27) as

$$\mathbf{p}_i^{t+1} = \mathbf{p}_i^t + b\, r_i^t (\mathbf{I}_{a_i^t} - \mathbf{p}_i^t), \tag{28}$$

where $\mathbf{I}_{a_i^t}$ is a unit vector with the a_i^t-th element being one. Let $\mathbf{P} = (\mathbf{p}_1, \ldots, \mathbf{p}_N)$ denote the mixed strategy profile of all the players, and thus we can achieve the evolution of the mixed strategy profile of the game $\mathcal{G}$ as:

$$\mathbf{P}^{t+1} = \mathbf{P}^t + b\, G(\mathbf{P}^t, \mathbf{r}^t, \mathbf{a}^t), \tag{29}$$

where $\mathbf{r}^t = (r_1^t, \ldots, r_N^t)$, $\mathbf{a}^t = (a_1^t, \ldots, a_N^t)$, and $G(\mathbf{P}^t, \mathbf{r}^t, \mathbf{a}^t)$ represents a function of $(\mathbf{P}^t, \mathbf{r}^t, \mathbf{a}^t)$. Then, according to [24, Th. 3.1], we know the following proposition.

Proposition 1: When the step size b is sufficiently small, i.e., $b \to 0$, the sequence $\{\mathbf{P}^t\}$ will converge to $\mathbf{P}^*$, which is the solution of the following ordinary differential equation (ODE):

$$\frac{d\mathbf{P}}{dt} = f(\mathbf{P}), \tag{30}$$

where $f(\mathbf{P}) = \mathbb{E}\left[G(\mathbf{P}^t, \mathbf{r}^t, \mathbf{a}^t)|\mathbf{P}^t\right]$, and the initial value $\mathbf{P}^0$ is equal to the initial channel selection probability matrix.

Proposition 2: According to [24, Th. 3.2], we have

- *All the stable stationary points of (30) are the Nash equilibria of $\mathcal{G}$.*
- *All the Nash equilibria of $\mathcal{G}$ are the stable stationary points of (30).*

As the proposed game has been shown to be a potential game in Theorem 1, the sequence $\{\mathbf{P}^t\}$ can be proved to converge to a stationary point of the ODE of (30), as proved in [27]. Then, according to Proposition 2, we have:

Theorem 4: When the step size b is sufficiently small, the SLA based dynamic channel access algorithm converges to a pure-strategy NE point of the game $\mathcal{G}$.

V. Simulation Results and Analysis

In this section, we conduct simulations to evaluate the performance of the proposed channel access algorithm. The number of available channels is set to be 10, and each EHS

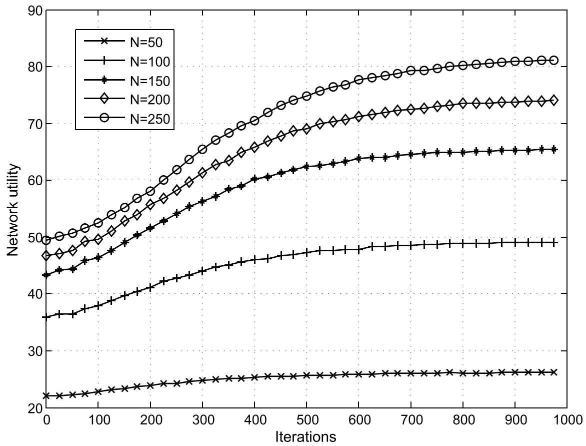


Fig. 3. Convergence behavior.

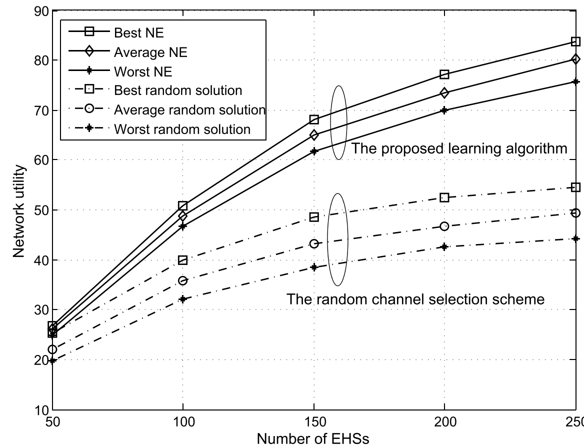


Fig. 4. Performance comparison of the achieved network utility when the number of EHSs varies.

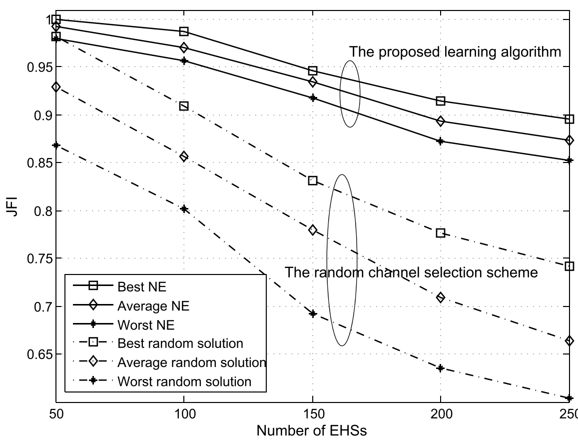


Fig. 5. Performance comparison of the achieved network fairness when the number of EHSs varies.

selects one channel for data transmission. Similar to [5], we model the data packet importance D_i^t as an exponential random variable with pdf $f_D(x) = e^{-x}$, $x \geq 0$. We consider an heterogeneous case that EHS's battery capacity $e_{\max}$ varies in $\{1, 5\}$, and the probability of harvesting one energy quantum $\Pr(B_i^t = 1)$ varies in $\{0.1, 0.5\}$. In addition, the step size b for the channel selection probability updating is set to be 0.1. The network utility $U = \sum_{i \in \mathcal{N}} u_i$, characterizing the average long-term importance of the data reported by all the EHSs to the FC [4], [5], is adopted as the metric for performance evaluation.

A. Performance Evaluation for Different Network Scales

In Fig. 3, we plot the convergence behavior of the proposed algorithm when the number of EHSs is 50, 100, 150, 200, 250, respectively. The curves are obtained by simulating 1000 independent trials and then taking the average results. As shown in the figure, the network utilities increase with the number of iteration steps, and all get largely improved after the algorithm converges. Besides, the convergence speed becomes a little slower when the number of EHSs increases. Although the converging speed is not very fast, it will not weaken the practicality of the proposed algorithm due to its online operation.

Fig. 4 and Fig. 5 show performance comparison for different solutions. To the best of our knowledge, no available algorithms in the literature can be applied to this multi-channel multi-access problem in the EH-WSN,[3] and thus we plot only the random solution for comparison. Due to the dynamic and unknown characteristics of the EH environment, the random selection scheme that each EHS randomly chooses a channel in each time slot is an intuitive solution. Besides, achieving the globally optimal solution for the network-utility maximization is obviously a combinational optimization problem [5]. Moreover, it cannot be found by the exhaustive enumeration due to the huge computational complexity when the number of nodes is large.

Fig. 4 plots the network utilities achieved by different solutions. As the number of EHSs increases from 50 to 250, the network utilities achieved by all the solutions gets significantly improved. It is easy to understand, since more EHSs can transmit more data to the FC if the channels are not heavily occupied. Besides, even the worst NE solution obtains much higher network utility than the best random solution, and the performance gap between the NE and random solutions increases with the number of EHSs. In addition, the performance gap between the best NE and the worst NE is small. It is because all the NE solutions are Pareto-optimal, as proved in Theorem 2.

To evaluate the fairness achieved by the proposed algorithm, we plot the JFI of different solutions in Fig. 5. As more EHSs are deployed in the network, JFI of all solutions decrease. The reason is that the difficulty in achieving fair resource allocation increases with the number of EHSs. Moreover, Fig. 5 shows that the JFI of the NE is higher than that of the random solution, which demonstrates that the NE solution can lead to more fair resource allocation. Besides, when N becomes larger, the JFI gap between the NE and random solutions increases. In addition, it is observed that the JFI of the worst NE is lower than $\frac{8}{9}$ when $N \geq 200$. It is due to the accuracy in convergence of the proposed algorithm. As N is larger, more iterations are needed for convergence to the NE, but the simulation is taken within 1000 iterations, as shown in Fig. 3.

[3]Note that the algorithms in [4] and [5] only apply to determine the distributed transmission policy $\boldsymbol{\eta}$ in the single-channel case.

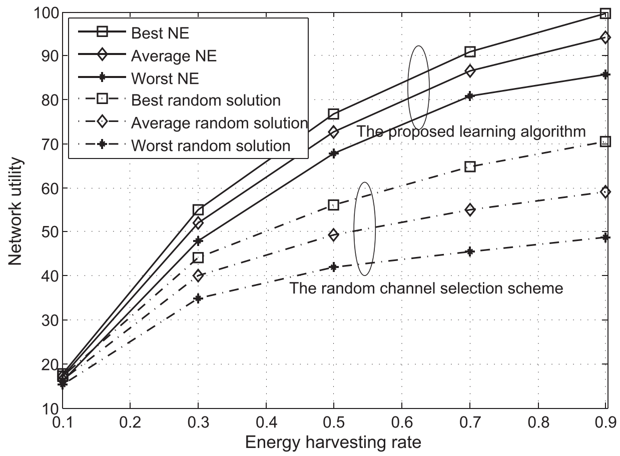


Fig. 6. Performance comparison of the achieved network utility for different energy harvesting rates.

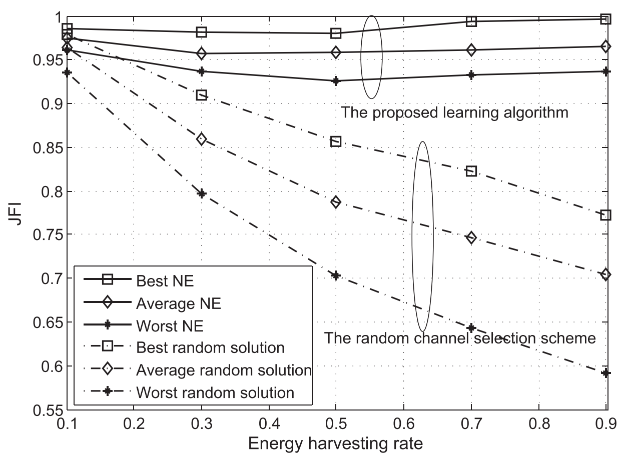


Fig. 7. Performance comparison of the achieved network utility for different energy harvesting rates.

B. Performance Evaluation for Different Energy Harvesting Rates

In this subsection, we evaluate the performance of the proposed algorithm for different energy harvesting rates. Specifically, the probability of harvesting one energy quantum in each time slot is set to be same for all EHSs, and is varied from 0.1 to 0.9. Fig. 6 shows the network utilities achieved by different solutions. The network utilities achieved by all the solutions increase significantly with the energy harvesting rate. However, when the energy harvesting rate is larger, the speed of the utility improvement becomes slower due to the channel congestion. Moreover, the proposed learning algorithm demonstrates much better performance than the random channel selection scheme.

Besides, we plot the JFI of different solutions with respect to different energy harvesting rates in Fig. 7. As shown in the figure, the proposed learning algorithm can achieve very fair solution no matter the energy harvesting rate is low, medium, or high. In contrast, the JFI achieved by the random channel selection algorithm decreases with the increase of the energy harvesting rate. Moreover, when the energy harvesting rate becomes larger, the JFI gap between the NE and random solutions increases.

VI. Conclusion

In this paper, we have studied the distributed multi-channel multi-access in an energy harvesting wireless sensor network (EH-WSN) under strict delay constraints for data transmission. Each EHS autonomously optimizes its channel selection strategy based on only local information, which has been formulated as a non-cooperative game. The utility of each EHS is defined as the average long-term aggregate importance of its reported data. We have proved the game to be an ordinal potential game that has at least one NE. Moreover, all the NE have been shown to be Pareto-optimal. The Jain's fairness index bound of the NE has been also derived. Besides, we have designed a fully distributed, online learning algorithm to converge to the NE of the channel access game. Finally, simulations have been conducted for different network scales and different energy harvesting rates, and the results have validated the effectiveness of the proposed algorithm in terms of network utility improvement and resource allocation fairness.

For future work, we will study energy cooperation among nodes. Since the EH process is intermittent and random in nature, it is better to enhance cooperation among nodes to increase the energy robustness. Besides, due to the dynamic EH process, the typical problems (multihop routing, transmission power control, topology control, etc.) that are inherent in the traditional battery-operated WSN become quite different for the EH-WSN, which thus needs further investigation.

References

[1] D. Niyato, E. Hossain, and A. Fallahi, "Sleep and wakeup strategies in solar-powered wireless sensor/mesh networks: Performance analysis and optimization," *IEEE Trans. Mobile Comput.*, vol. 6, no. 2, pp. 221–236, Feb. 2007.

[2] I. Menache and E. Altman, "Battery-state dependent power control as a dynamic game," in *Proc. WiOPT*, 2008, pp. 242–250.

[3] F.-Y. Tsuo, H.-P. Tan, Y. H. Chew, and H.-Y. Wei, "Energy-aware transmission control for wireless sensor networks powered by ambient energy harvesting: A game-theoretic approach," in *Proc. IEEE ICC*, Jun. 2011, pp. 1–5.

[4] N. Michelusi and M. Zorzi, "Optimal random multiaccess in energy harvesting wireless sensor networks," in *Proc. IEEE ICC*, Jun. 2013, pp. 463–468.

[5] N. Michelusi and M. Zorzi, "Optimal adaptive random multiaccess in energy harvesting wireless sensor networks," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1355–1372, Apr. 2015.

[6] P. Blasco, D. Gunduz, and M. Dohler, "A learning theoretic approach to energy harvesting communication system optimization," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1872–1882, Apr. 2013.

[7] J. Zheng, Y. Cai, X. Shen, Z. Zheng, and W. Yang, "Green energy optimization in energy harvesting wireless sensor networks," *IEEE Commun. Mag.*, vol. 53, no. 11, pp. 150–157, Nov. 2015.

[8] 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, pp. 3872–3882, Aug. 2013.

[9] S. Zhang, "Energy-aware traffic offloading for green heterogeneous networks," *IEEE J. Sel. Areas Commun.*, to be published, doi: 10.1109/JSAC.2016.2520244.

[10] T. Zhang, W. Chen, Z. Han, and Z. Cao, "A cross-layer perspective on energy-harvesting-aided green communications over fading channels," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1519–1534, Apr. 2015.

[11] D. Gündüz, K. Stamatiou, N. Michelusi, and M. Zorzi, "Designing intelligent energy harvesting communication systems," *IEEE Commun. Mag.*, vol. 52, no. 1, pp. 210–216, Jan. 2014.

[12] G. Owen, *Game Theory*, 3rd ed. San Diego, CA, USA: Academic, 2001.

[13] J. Zheng, Y. Cai, Y. Xu, and A. Anpalagan, "Distributed channel selection for interference mitigation in dynamic environment: A game-theoretic stochastic learning solution," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4757–4762, Nov. 2014.

[14] N. Zhang, H. Liang, N. Cheng, Y. Tang, J. W. Mark, and X. Shen, "Dynamic spectrum access in multi-channel cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 11, pp. 2053–2064, Nov. 2014.

[15] Y. Bi, K.-H. Liu, L. X. Cai, X. Shen, and H. Zhao, "A multi-channel token ring protocol for QoS provisioning in inter-vehicle communications," *IEEE Trans. Wireless Commun.*, vol. 8, no. 11, pp. 5521–5631, Nov. 2009.

[16] W. Guo, X. Sheng, H. Liu, and X. Zhu, "Development of a multi-channel compact-size wireless hybrid SEMG/NIRS sensor system for prosthetic manipulation," *IEEE Sensors J.*, vol. 16, no. 2, pp. 447–456, Jan. 2016.

[17] P. Pinto, A. Pinto, and M. Ricardo, "Cross-layer admission control to enhance the support of real-time applications in WSN," *IEEE Sensors J.*, vol. 15, no. 12, pp. 6945–6953, Dec. 2015.

[18] Y. Bi, H. Shan, X. S. Shen, N. Wang, and H. Zhao, "A multi-hop broadcast protocol for emergency message dissemination in urban vehicular ad hoc networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 3, pp. 736–750, Mar. 2016.

[19] T. Zheng, M. Gidlund, and J. Akerberg, "WirArb: A new MAC protocol for time critical industrial wireless sensor network applications," *IEEE Sensors J.*, vol. 16, no. 7, pp. 2127–2139, Apr. 2016.

[20] P. Suriyachai, U. Roedig, and A. Scott, "A survey of MAC protocols for mission-critical applications in wireless sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 2, pp. 240–264, Second Quarter 2012.

[21] D. Monderer and L. S. Shapley, "Potential games," *Games Econ. Behavior*, vol. 14, no. 1, pp. 124–143, 1996.

[22] R. K. Jain, D.-M. W. Chiu, and W. R. Haws, "A quantitative measure of fairness and discrimination for resource allocation in shared computer system," Eastern Res. Lab., Digit. Equip. Corp., Hudson, MA, USA, Tech. Rep., 1984.

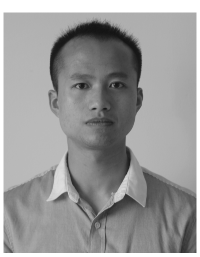
[23] J. Zheng, Y. Cai, Y. Liu, Y. Xu, B. Duan, and X. Shen, "Optimal power allocation and user scheduling in multicell networks: Base station cooperation using a game-theoretic approach," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 6928–6942, Dec. 2014.

[24] P. S. Sastry, V. V. Phansalkar, and M. Thathachar, "Decentralized learning of Nash equilibria in multi-person stochastic games with incomplete information," *IEEE Trans. Syst., Man, Cybern.*, vol. 24, no. 5, pp. 769–777, May 1994.

[25] Y. Xu, Q. Wu, L. Shen, J. Wang, and A. Anpalagan, "Opportunistic spectrum access with spatial reuse: Graphical game and uncoupled learning solutions," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 4814–4826, Oct. 2013.

[26] J. Zheng, Y. Cai, W. Yang, Y. Xu, and A. Anpalagan, "A game-theoretic approach to exploit partially overlapping channels in dynamic and distributed networks," *IEEE Commun. Lett.*, vol. 18, no. 12, pp. 2201–2204, Dec. 2014.

[27] J. Zheng, Y. Cai, N. Lu, Y. Xu, and X. Shen, "Stochastic game-theoretic spectrum access in distributed and dynamic environment," *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4807–4820, Oct. 2015.



Jianchao Zheng received the B.S. degree in electronics engineering from the College of Communications Engineering, PLA University of Science and Technology, Nanjing, China, in 2010, where he is currently pursuing the Ph.D. degree in communications and information system with the College of Communications Engineering. His research interests focus on interference mitigation techniques, green communications, game theory, learning theory, and optimization techniques.

Honggang Zhang was the International Chair Professor of Excellence for Université Européenne de Bretagne and Supélec, Rennes, France, from 2012 to 2014. He is currently a Full Professor with the Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China. He is also an Honorary Visiting Professor with the University of York, York, U.K. He served as the Chair of the Technical Committee on Cognitive Networks of the IEEE Communications Society from 2011 to 2012. He is currently active in the research on green communications and was the Leading Guest Editor of the *IEEE Communications Magazine* special issues on Green Communications. He was the Co-Author and an Editor of two books *Cognitive Communications-Distributed Artificial Intelligence*, *Regulatory Policy and Economics, Implementation* (Wiley), and *Green Communications: Theoretical Fundamentals, Algorithms and Applications* (CRC Press).

Yueming Cai received the B.S. degree in physics from Xiamen University, Xiamen, China, in 1982, and the M.S. degree in microelectronics engineering and the Ph.D. degree in communications and information systems from Southeast University, Nanjing, China, in 1988 and 1996, respectively. His current research interest includes cooperative communications, signal processing in communications, wireless sensor networks, and physical layer security.

Rongpeng Li received the B.E. degree from Xidian University, Xi'an, China, in 2010, and the Ph.D. degrees from Zhejiang University, Hangzhou, China, in 2015, both as Excellent Graduates. He was a Visiting Doctoral Student with Supélec, Rennes, France, in 2013, and an Intern Researcher with the China Mobile Research Institute, Beijing, China, in 2014. He is currently a Researcher with Huawei Technologies Company, Ltd., Shanghai, China. He has authored or co-authored several papers in the related fields. His research interests currently focus on resource allocation of cellular networks (especially full duplex networks), applications of reinforcement learning, and analysis of cellular network data. He served as the Web Design Chair of the IEEE OnlineGreenComm in 2015, and the Web and System Chair of the IEEE ISCIT in 2011.

Alagan Anpalagan (SM'–) is currently a Professor with the Department of Electrical and Computer Engineering, Ryerson University, Canada, where he directs a research group working on radio resource management and radio access and networking areas within the WINCORE Laboratory. He served as Editor of the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS (2012-14), the IEEE COMMUNICATIONS LETTERS (2010-13), *Wireless Personal Communications* (Springer) (2011-13), and the *EURASIP Journal of Wireless Communications and Networking* (2004-2009). He co-authored three edited books entitled *Design and Deployment of Small Cell Networks*, (Cambridge University Press, 2014), *Routing in Opportunistic Networks* (Springer, 2013), and *Handbook on Green Information and Communication Systems* (Academic Press, 2012).

Dr. Anpalagan received the B.A.Sc., M.A.Sc., and Ph.D. degrees from the University of Toronto, all in electrical engineering. He is currently a registered Professional Engineer in the province of Ontario, Canada, and a fellow of the Institution of Engineering and Technology. He was a recipient of Dean's Teaching Award (2011), the Faculty Scholastic, Research and Creativity Award (2010, 2014), and the Faculty Service Award (2011, 2013) at Ryerson University. He received the Exemplary Editor Award from the IEEE ComSoc (2013) and the Editor-in-Chief of the Top10 Choice Award in the *Transactions on Emerging Telecommunications Technology* (2012). He served as the TPC Co-Chair of the IEEE Globecom'15: SAC Green Communication and Computing, the IEEE WPMC'12 Wireless Networks, the IEEE PIMRC'11 Cognitive Radio and Spectrum Management, and the IEEE CCECE'04/08. He has served as the Vice Chair of the IEEE SIG on Green and Sustainable Networking and Computing with Cognition and Cooperation (2015-), the IEEE Canada Central Area Chair (2012-14), the IEEE Toronto Section Chair (2006-07), the ComSoc Toronto Chapter Chair (2004-05), and the IEEE Canada Professional Activities Committee Chair (2009-11).