

# Wireless Power Transfer and Data Collection in Wireless Sensor Networks

Kai Li<sup>1</sup>, Member, IEEE, Wei Ni<sup>2</sup>, Senior Member, IEEE, Lingjie Duan<sup>1</sup>, Member, IEEE, Mehran Abolhasan<sup>3</sup>, Senior Member, IEEE, and Jianwei Niu<sup>4</sup>, Senior Member, IEEE

**Abstract**—In a rechargeable wireless sensor network, the data packets are generated by sensor nodes at a specific data rate, and transmitted to a base station. Moreover, the base station transfers power to the nodes by using wireless power transfer (WPT) to extend their battery life. However, inadequately scheduling WPT and data collection causes some of the nodes to drain their battery and have their data buffer overflow, whereas the other nodes waste their harvested energy, which is more than they need to transmit their packets. In this paper, we investigate a novel optimal scheduling strategy, called EHMDP, aiming to minimize data packet loss from a network of sensor nodes in terms of the nodes' energy consumption and data queue state information. The scheduling problem is first formulated by a centralized MDP model, assuming that the complete states of each node are well known by the base station. This presents the upper bound of the data that can be collected in a rechargeable wireless sensor network. Next, we relax the assumption of the availability of full state information so that the data transmission and WPT can be semidecentralized. The simulation results show that, in terms of network throughput and packet loss rate, the proposed algorithm significantly improves the network performance.

**Index Terms**—Wireless sensor network, wireless power transfer, Markov decision process, scheduling, optimization.

## I. INTRODUCTION

**I**NEXPENSIVE sensors capable of significant computation and wireless communications are becoming available. However, sensor nodes are severely constrained by the amount of battery power, limiting the network lifetime and quality of service. Rechargeable Wireless Sensor Network (RWSN) has been extensively studied, where energy is harvested by Wireless Power

Manuscript received January 2, 2017; revised July 26, 2017; accepted November 2, 2017. Date of publication November 13, 2017; date of current version March 15, 2018. This work was supported in part by the SUTD-ZJU Joint Collaboration Grant (Project SUTDZJU/RES/03/2014) and in part by the Singapore Ministry of Education Academic Research Fund Tier 2 under Project MOE2016-T2-1-173. The review of this paper was coordinated by Dr. C. Chen. (Corresponding author: Lingjie Duan.)

K. Li and L. Duan are with the Singapore University of Technology and Design, Singapore 487372 (e-mail: kai\_li@sutd.edu.sg; lingjie\_duan@sutd.edu.sg).

W. Ni is with the Commonwealth Scientific and Industrial Research Organization, Sydney, NSW 2122, Australia (e-mail: wei.ni@csiro.au).

M. Abolhasan is with the University of Technology Sydney, Ultimo, NSW 2007, Australia (e-mail: mehran.abolhasan@uts.edu.au).

J. Niu is with the State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China (e-mail: niujianwei@buaa.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVT.2017.2772895

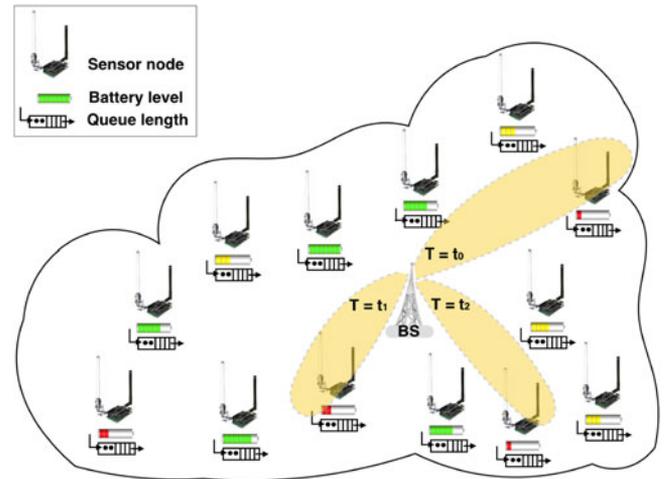


Fig. 1. Data collection and WPT in wireless powered sensor network.

Transfer (WPT) to recharge battery of a sensor node [1]–[4]. Fig. 1 depicts that a number of fixed sensor nodes serve as data sources with sensing ability, for example, monitoring the environment and detecting abnormal events. The nodes, equipped with a data communication antenna and a wireless power receiver, generate data packets at an application-specific sampling rate. A base station (BS) is deployed to transfer power to sensor nodes via WPT, and collect sensory data from the nodes [5]–[8]. Beamforming is used at the BS, either electronically or mechanically. The use of beamforming allows for the concentrated transfer of energy towards the intended node, which overcomes the broadcast nature of radio, and avoids the dispersion and waste of energy [9]–[11]. During each time slot (or epoch), one of the nodes is scheduled to transmit data to and harvest energy from the BS. To this end, the BS only generates one beam per epoch for both energy transfer and data collection, thereby reducing the overhead of beamforming. Especially, the network in Fig. 1 can be generically extended to a scalable RWSN that is composed of multiple cells, where each cell is covered by one BS. The working mode of WPT includes power splitting, and time switching in terms of the WPT receiver structure [12]. With time switching, a receiver switches its operation between the power transfer and data transmission over time, while with power splitting, the received signal is split into two streams with one stream used for transferring power and the other stream for transmitting data. In particular, we consider a time-switching

WPT network in this work. The WPT and data transmission are designed to be in the same radio frequency band, and different time slots so that the nodes only need a single RF chain, which reduces the hardware cost. It is critical to schedule the WPT and data collection to extend the battery life of the nodes and to minimize the packet loss. Inadequately scheduling WPT and data collection causes some of the nodes to drain their battery and have their data buffer overflow, while the other nodes waste their harvested energy, which is more than they need to transmit their packets.

In this paper, we aim to optimize the scheduling of WPT and data collection in RWSNs. The scheduling problem is first formulated as an off-line finite state Markov Decision Process (MDP), called EHMDP, where transmission order, modulation level, and WPT duration of the nodes are jointly optimized. As a result, the packet loss of entire network can be minimized and the battery lifetime of all the nodes are leveraged. Note that in EHMDP, the BS needs to have real-time knowledge on the packet arrivals and battery levels of the nodes. This requires the nodes to report prior to every time slot. Extending from the centralized scheme, we further propose a semi-decentralized approach that allows the nodes to spontaneously self-nominate for data transmission and energy harvesting. Each sensor node makes its own decision on whether it turns on transmission in a time slot, as well as the modulation level if the transmission is turned on. We proposed a transmission probability to impose urgency to the nodes with built-up queues, avoid their transmissions being withheld due to insufficient energy harvested, and in turn, avoid data queue overflows at the nodes and reduce packet losses. Furthermore, an Energy-Queue Aware Transmission (E-QAT) algorithm is proposed in terms of the transition probability of the node. In addition, while this paper is structured around WSNs, our framework can be applied to any wireless networks with energy harvesting capabilities.

The rest of this paper is organized as follows. The related work is summarized in Section II. Section III presents network structure of RWSN. The centralized MDP model with global information for WPT and data collection is investigated in Section IV. In Section V, a semi-decentralized algorithm based on individual energy and data queue is proposed. Simulations and numerical results are presented in Section VI. Conclusions and future work are given in Section VII.

## II. RELATED WORK

Earlier energy-efficient transmission schemes were focused on a single point-to-point link [13]–[17] or in cellular frameworks [18]. In [13], a string tautening algorithm was proposed to produce the most energy-efficient schedule for delay-limited traffic, first under the assumption of negligible circuit power, and then extended to energy-harvesting powered transmissions with non-negligible transmitter circuit powers [14], [15]. Later, simple network topologies, such as three-party two-way relay, were considered for the maximization of energy efficiency, first in a halfduplex transmission mode [16] and then in a full-duplex mode [17]. Several studies have considered to integrate WPT

technologies into communication networks to achieve simultaneous data transmission and power transfer [19]–[29]. In [19], [20], a mobile charger is employed to power wireless identification and sensing nodes which support sensing, computing, and communicating. The movement strategy of the mobile charger is formulated, such that the charging latency of all nodes can be minimized while the continuous operation of the sensor nodes can be guaranteed. In [21], the energy harvested by a sensor node is allocated for sensing and data transmitting in a balanced fashion. The sensing rate and routing are optimized, given the dynamic features of network topology. In [22], a data sender harvests energy from a power transmitter via WPT in the downlink before transmitting information to a data receiver in the uplink. It is found that data reception performance degrades when the time for energy harvesting and data transmission is unbalanced. In [23], power transfer weight and power allocation at the BS are designed to balance the transferred power to the receivers given different SINR requirements. Moreover, the problem of joint information and energy transmission is formulated to give the upper bound of the sum of transferred power. Zhou *et al.* present a resource allocation problem for simultaneous data and WPT in downlink multiuser OFDM system, where the users harvest energy and decode data using the same signals received from the BS [24]. A tradeoff between the weighted sum-rate of all users and transferred energy is obtained given lower bound of harvested energy on each user and upper bound of total transmission power. Moreover, an energy harvesting resource allocation scheme has been investigated with imperfect channel state information in [25]. In [26], WPT is considered to support multiple users concurrently, where the BS broadcasts the energy in the network. A frame structure is proposed to allocate time for WPT and data transmission, which achieves the maximum of sum-throughput and minimum of charging and transmission time.

To achieve a balance between throughput and data collection fairness, a scheduling strategy is proposed to allocate the power transfer among users and the proportion of the time between energy harvests [27]. In [28], a data rate selection in RWSN is studied, which considers fairness of data rate and WPT duration among the nodes. Moreover, a set of algorithms are developed to obtain a optimal rate assignment with the Water-Filling-Framework under different routing types. A WPT system is studied to balance a tradeoff between spectral and energy efficiencies by jointly designing beamforming for WPT and data transmission [29]. Moreover, both WPT and data transmission beamforming are adaptive to channel fading by exploiting the benefits of channel state information.

However, most of the work in simultaneous WPT and data transmission only focus on improving energy efficiency of WPT. The data lost caused by buffer overflow is not correlatively considered.

MAC protocol design for scheduling data transmission and power transfer is given in the literature [30]–[32]. It was noted that the power transfer varies depending on surrounding environments. A MAC protocol based on fair polling scheme is studied to achieve a data collection fairness [30]. Moreover, each node contends channel adaptively to their energy harvesting rate.

TABLE I  
THE LIST OF FUNDAMENTAL VARIABLES DEFINED IN SYSTEM MODEL

Notation	Definition
$N$	number of sensor nodes
$i$	node ID
$\mathbf{h}$	channel gain between the BS and the node
$P^E$	power that is transferred to the node
$P_e$	output power of the BS
$L$	number of bits of the sensory data packet
$\epsilon$	required BER of the node
$e_i$	battery level of node $i$
$E$	battery capacity of the node
$K$	the highest battery level of the node
$q_i$	queue length of node $i$
$Q$	maximum queue length of the node
$\rho$	modulation scheme of the node
$M$	the highest modulation order
$\lambda$	arrival probability of the data packet
$\mathcal{A}$	action set of MDP
$T_\lambda$	time when a new incoming data packet is put into the queue
$\widehat{T}$	time interval of the proposed discrete-time Markov decision process.

In [31], a RF-MAC protocol is studied to jointly schedule power transmitters and energy harvesting rates in RWSN. RF-MAC focuses on the amount of energy delivery to the nodes, while eliminating disruption to data communication. In [32], multiple WPT transmitters are grouped into different sets based on an estimate of their separation distance from the node to reduce the impact of destructive interference. Moreover, a MAC protocol is presented to transfer power to the nodes on request. Unfortunately, the MAC protocols in literature schedule WPT and data communication with the objective of interference elimination, which is different from the problem in this work.

### III. SYSTEM DESIGN AND STRUCTURE

Notations used in this paper are listed in Table I. The network under consideration consists of a BS and  $N$  geographically distributed energy-harvesting powered nodes, as illustrated in Fig. 1. The BS, connected to persistent power supply, is responsible for remotely charging the nodes using WPT. Equipped with  $N_c$  antennas ( $N_c \gg 1$ ), the BS can exploit transmit beamforming techniques to produce a narrow beam to each node. As a result, energy is transferred with improved focus and transfer efficiency. The BS is also responsible for collecting sensory data. Additionally, receive beamforming techniques enable the BS to enhance the received signal strength (RSS) and reduce bit error rate (BER). In particular, other advanced multi-user beamforming techniques, e.g., zero-forcing beamforming, are not considered in this work since they achieve spatial multiplexing or diversity, but would require real-time feedback on channel state information in most cases.

Each of the energy-harvesting powered nodes, e.g., node  $i$  ( $i = 1, \dots, N$ ), harvests energy from the BS to power its operations, e.g., sensing and communicating. The rechargeable battery of the node is finite with the capacity of  $E_i$  Joules, and it overflows if overcharged. The complex coefficient of the reciprocal wireless channel between the BS and the node is  $\mathbf{h}_i$ ,

which is precisely known at both the BS and the node through channel reciprocity. Considering the non-persistent power supply and subsequently limited signal processing capability of the nodes, we let each node be equipped with a single antenna.  $\mathbf{h}_i \in \mathcal{C}^{N_c \times 1}$ . Suppose that the BS employs maximal ratio combining (MRC) [33] to maximize the beamforming gain. The power transferred to node  $i$  can be given by

$$P_i^E = P_e \|\mathbf{h}_i\|^2, \quad (1)$$

where  $P_e$  is the constant output power of the BS, and  $\|\cdot\|$  stands for norm.

Node  $i$  also keeps sensing its ambient environment, packetizes and queues (in a first-in-first-out (FIFO) fashion) the sensory data in packets of  $L_i$  bits, and transmits the packets to the BS through the wireless channel. The arrival/queueing process of sensory data at node  $i$  is modeled as a random process, where a new packet is put into the FIFO queue at a probability  $\lambda_i$  within every  $T_\lambda$  seconds. The node has a finite buffer to accommodate the maximum queue length of  $Q_i$  packets or  $Q_i L_i$  bits. The buffer starts overflowing, once the maximum queue length is reached while the transmission of the sensor is withheld due to insufficient energy harvested. In other words, the arrival rate of sensory data exceeds the departure rate, due to the insufficient energy harvested. To this end, the scheduling algorithm needs to be appropriately designed to assign the sensor sufficient channel and energy resources.

The modulation scheme that node  $i$  uses to transmit packets is denoted by  $\rho_i$ .  $\rho_i \in \{1, 2, \dots, M\}$ , where  $\rho_i = 1, 2$ , and 3 indicates binary phase-shift keying (BPSK), quadrature-phase shift keying (QPSK), and 8 phase-shift keying (8PSK), respectively, and  $\rho_i \geq 4$  corresponds to  $2^{\rho_i}$  quadrature amplitude modulation (QAM).  $M$  is the highest modulation order.

Suppose that the bit error rate (BER) requirement of node  $i$  is  $\epsilon_i$ , and the MRC is conducted at the BS to maximize the signal-to-noise ratio (SNR). The required transmit power of node  $i$  depends on  $\rho_i$ ,  $\epsilon_i$ , and  $\mathbf{h}_i$ , and can be given by [34], [35]

$$P_i^D(\rho_i) \approx \frac{\kappa_2^{-1} \ln \frac{\kappa_1}{\epsilon_i}}{\|\mathbf{h}_i\|^2} (2^{\rho_i} - 1), \quad (2)$$

where  $\kappa_1$  and  $\kappa_2$  are channel related constants.

For illustration convenience, we consider a homogeneous network, where all the nodes have the same battery size, buffer size, packet length, packet arrival probability, wireless channel, and the BER requirement. The subscript “ $i$ ” is suppressed in  $E_i$ ,  $Q_i$ ,  $L_i$ ,  $\lambda_i$ ,  $\mathbf{h}_i$ , and  $\epsilon_i$ . However, the proposed scheduling protocols can be extended to a heterogeneous network, where the complexity of the scheduling problem may grow, as the result of an increased number of states. Moreover, all the channels involved are assumed to be block-fading, i.e., the channels remain unchanged during each transmission block, and may change from block to block.

### IV. CENTRALIZED SCHEDULING AND OFF-LINE OPTIMIZATION

#### A. Centralized Scheduling Protocol

One time slot, also known as a scheduling interval, lasts  $\widehat{T}$  mini-slots, during which the node sends a packet to the BS,

followed by WPT from the BS to the node. In this case, the BS takes the action of choosing a node and deciding the associated modulation order for every time slot. After receiving the packet of the node, the BS starts to transfer energy and it can acknowledge its successful or failed reception of the packet by changing the phase or waveform of the energy signal [36]. The node discards the packet, if successfully delivered, from its queue, based on the acknowledgement. Given such centralized coordination, the nodes are activated in a TDMA fashion to transmit sensory data and harvest energy, where one node per time slot. We note that there is no collision between the data transmissions of the nodes in the case of centralized scheduling. The receive and transmit beams can also be produced ahead of the actual data transmission or energy transfer.

We note that the BS's actions are not only based on the current network state (i.e., the battery level and queue length of every node  $i$ , denoted by  $e_i$  and  $q_i$ , respectively,  $i = 1, \dots, N$ ), but take into account potential influence on future evolutions of the network as well. Particularly, the current action that the BS takes can affect the future battery level and queue length of every node, and in turn, influences the future actions to be taken. Such action taking is a discrete time stochastic control process which is partly random (due to the random and independent arrival/queueing process of sensory data at every node) and partly under the control of the decision-making BS. The action of selecting node and modulation for every time slot can be optimized in a sense that the optimality in regards of a specific metric, e.g., packet loss, can be achieved in long term over the entire stochastic control process (rather than in an individual time slot).

### B. Off-Line EHMDP Modeling

As noted earlier, the optimization of the action, i.e., the selection of node and modulation, for every time slot needs to be conducted over the entire stochastic control process of centralized scheduling. The correlation between actions taken in different time slots needs to be captured, and to validate the long-term optimality of the actions.

We consider optimizing the actions to minimize the overall packet loss of the entire WPT powered sensor network. The packet loss can result from buffer overflows at the nodes where data transmissions are withheld due to insufficient energy harvested. The packet loss can also result from unsuccessful data transmissions over wireless fading channels.

It has been known that battery readings are continuous variables with variance difficult to be traced in real-time. The proposed quantization that discretizes the continuous battery readings to  $K$  levels can facilitate generating MDP states in EHMDP. Moreover, the accuracy of the quantization can be improved by reducing the quantization interval, but can result in an increased number of MDP states, and hence an increasing complexity of MDP. Therefore, to improve the mathematical tractability of the problem and illustration convenience, the continuous battery is discretized into  $K$  levels, as  $0 < \mathcal{E} < 2\mathcal{E} < \dots < K\mathcal{E} = E$ .  $e_i \in \{0, \mathcal{E}, 2\mathcal{E}, \dots, K\mathcal{E}\}$  [29]. In other words, the battery level of a node is lower rounded to the closest discrete level. The queue length  $q_i \in \{0, 1, \dots, Q\}$  is discrete.

The nodes use a small control message to update the BS with their battery level and queue length at the beginning of every time slot. For example, consider a network of 40 nodes, battery level of 100 and queue length of 100 packets, the overhead of one node takes 12 bits, and total overhead is 480 bits, which is much smaller than the size of a data packet. Therefore, we assume that both the transmission time and the energy consumption are negligible. We also assume that the BS receives all the necessary information timely and correctly. In other words, the information required to take optimal actions is assumed to be globally available in the case of centralized scheduling. This assumption is reasonable, because the control messages typically adopt reliable, low-order modulation schemes, e.g., BPSK (sometimes with repetition coding).

Such centralized scheduling can be formulated as a discrete-time Markov decision process (MDP) with time interval of  $\widehat{T}$ , where each state, denoted by  $\mathcal{S}_\alpha$ , represents the battery levels and queue lengths of all the nodes in the network, i.e.,  $\{(e_{\alpha,n}, q_{\alpha,n}), n = 1, \dots, N\}$ . The size of the state space (i.e., the number of such states) is  $(K(Q+1))^N$ . The action to be taken during a transition, denoted by  $\mathcal{A}$ , is to select the node to be activated and specify its modulation.  $\mathcal{A} \in \{(i, \rho_i) : i = 1, \dots, N, \rho_i \in \{1, \dots, M\}\}$ , where  $M$  denotes the highest modulation level. The size of the action set is  $NM$ .

Furthermore,  $\mathcal{A}$  can be reduced to only consist of the selected node, i.e., the  $n$ -th action  $\mathcal{A}_n = \{n\}$ . The maximum energy that can be harvested into the battery of the selected node  $n$  can be given by

$$\Delta\mathcal{E}_n = \left[ \left( \widehat{T} - \frac{L}{\rho_n^* W} \right) \frac{P_e \|\mathbf{h}_n\|^2}{\mathcal{E}} - \frac{L\kappa_2^{-1} \ln\left(\frac{\kappa_1}{\epsilon}\right)}{\|\mathbf{h}_n\|^2 \rho_n^* W \mathcal{E}} (2^{\rho_n^*} - 1) \right] \mathcal{E} \quad (3)$$

where  $\rho_n^*$  denotes the optimal modulation scheme of node  $n$ , which is obtained by (18) (See Appendix). The actions of the MDP can be optimized by judiciously deriving the transition probabilities and formulating a dynamic programming problem, and solved by using dynamic programming techniques [37]. Details will be provided shortly in this section.

Note that the centralized scheduling is ideal and can only be conducted off-line, under the assumptions of negligible signaling time and energy, as well as the resultant global availability of the required information. The optimal actions of the MDP are inapplicable to practical environments with non-negligible signaling time and energy consumption. Nevertheless, the MDP (or its optimal actions) provides the lower bound of packet loss to any scheduling designs of the WPT powered sensor networks. Moreover, the MDP also lays foundations to and eases the understanding of the on-line optimization of semi-decentralized scheduling, as will be described in detail in the next section.

### C. Transition Probability and Packet Loss

Given the action  $\mathcal{A}_k = \{k\}$  ( $1 \leq k \leq N$ ), the transition probability of the MDP, from state  $\mathcal{S}_\alpha$  to  $\mathcal{S}_\beta$  ( $1 \leq (\alpha, \beta) \leq$

$K^N (Q + 1)^N$ ), can be given by

$$\Pr \left\{ \mathcal{S}_\beta \mid \mathcal{S}_\alpha, \mathcal{A}_k \right\} = \Pr \left\{ (e_{\beta,k}, q_{\beta,k}) \mid (e_{\alpha,k}, q_{\alpha,k}), k \in \mathcal{A}_k \right\} \\ \times \prod_{n=1, n \notin \mathcal{A}_k}^N \Pr \left\{ (e_{\beta,n}, q_{\beta,n}) \mid (e_{\alpha,n}, q_{\alpha,n}), n \notin \mathcal{A}_k \right\}, \quad (4)$$

where the two parts of the RHS are specified in (5).

Specifically, (5a) corresponds to the selected node  $k$ , where the first case is that the queue of the node increases due to the failed transmission of a packet and the arrival of a new sensory packet. The second case is that the queue decreases due to a successful transmission and no new packet arrival. The third case is that the queue does not change, which is due to either a successful transmission and a new packet arrival, or a failed transmission and no new packet arrival. The battery level of the selected node all increases by  $\Delta \mathcal{E}_k$  in the three cases, given the optimized modulation order  $\rho_k$ .

(5b) corresponds to the unselected nodes  $n \neq k$ , where the first case is that the queue of the node increases due to a new sensory packet arrival. The second case is that the queue does not change, without a new packet arrival. The battery level of the node does not change, since the node is not selected to harvest energy. (5) shown at the bottom of this page.

The packet loss, resulting from buffer overflow during the transition, can be given by

$$R \left\{ \mathcal{S}_\beta \mid \mathcal{S}_\alpha, \mathcal{A}_k \right\} = R \left\{ (e_{\beta,k}, q_{\beta,k}) \mid (e_{\alpha,k}, q_{\alpha,k}), k \in \mathcal{A}_k \right\} \\ + \sum_{n=1, n \notin \mathcal{A}_k}^N R \left\{ (e_{\beta,n}, q_{\beta,n}) \mid (e_{\alpha,n}, q_{\alpha,n}), n \notin \mathcal{A}_k \right\}; \quad (6)$$

$$R \left\{ (e_{\beta,k}, q_{\beta,k}) \mid (e_{\alpha,k}, q_{\alpha,k}), k \in \mathcal{A}_k \right\} = \\ \begin{cases} (1 - (1 - \epsilon)^L) \lambda, & \text{if } q_{\alpha,k} = q_{\beta,k} = Q; \\ 0, & \text{otherwise.} \end{cases} \quad (7a)$$

$$R \left\{ (e_{\beta,n}, q_{\beta,n}) \mid (e_{\alpha,n}, q_{\alpha,n}), n \notin \mathcal{A}_k \right\} = \\ \begin{cases} \lambda, & \text{if } q_{\alpha,n} = q_{\beta,n} = Q; \\ 0, & \text{otherwise.} \end{cases} \quad (7b)$$

The first case of (7a) is the probability that the buffer of the selected node  $k$  overflows. Specifically, the new generated packet at State  $\mathcal{S}_\beta$  with the arrival probability  $\lambda$  has to be dropped since the node  $k$  failed to transmit the packet at State  $\mathcal{S}_\alpha$  given  $Q$  packets in the queue. Similarly, the first case of (7b) is the probability that the buffer of an unselected node  $n \neq k$  overflows.

#### D. MDP Solver

Generally, value iterations, policy iterations, and linear programming are three major methods for computing optimal actions for MDP with expected total discounted rewards. The actions of the MDP are optimized by using dynamic programming techniques [37]. Particularly, in this paper, the value iteration algorithm [38], [39] is applied. The value iteration algorithm computes the optimal cost function by assuming first a one-stage finite horizon, then a two-stage finite horizon, and so on. The cost functions are computed to converge in the limit to the optimal cost function. Therefore, the policy associated with the successive cost functions converges to the optimal policy in a finite number of iterations. In addition, the reason is that policy iteration requires policy evaluation at each of iterations, which is a protracted iterative computation involving multiple sweeps through the state set. Linear programming requires the MDP problem can be solved in a number of arithmetic operations polynomial in  $N$  and  $M$ , which run extremely slowly in practice. However, note that policy iteration and linear programming are also adoptable to obtain the optimal actions of the MDP problem.

We define that a policy  $\pi$  is a mapping from states to actions, and the set of all the policies is given as  $\Pi$ . Moreover, if the policy is independent of the current stage, it is said to be stationary. The goal of the algorithm is to minimize the expected total packet loss which is denoted as  $v(S_\beta)$ ,

$$v(S_\beta) = \min_{\pi \in \Pi} \left\{ E_S^\pi \sum_{t=1}^{\infty} \omega^{t-1} R \left\{ \mathcal{S}_\beta \mid \mathcal{S}_\alpha, \mathcal{A}_k \right\} \right\} \quad (8)$$

where  $\omega (\omega \in [0, 1])$  denotes the discount factor for future states. To find the action  $k$  to calculate  $v(S_\beta)$ , we evaluate for each action  $k \in \mathcal{A}_k$  and select the actions that achieve (8). Thus, we

$$\Pr \left\{ (e_{\beta,k}, q_{\beta,k}) \mid (e_{\alpha,k}, q_{\alpha,k}), k \in \mathcal{A}_k \right\} = \begin{cases} (1 - (1 - \epsilon)^L) \lambda, & \text{if } e_{\beta,k} = e_{\alpha,k} + \Delta \mathcal{E}_k \text{ and } q_{\beta,n} = q_{\alpha,n} + 1; \\ (1 - \epsilon)^L (1 - \lambda), & \text{if } e_{\beta,k} = e_{\alpha,k} + \Delta \mathcal{E}_k \text{ and } q_{\beta,n} = q_{\alpha,n} - 1; \\ (1 - (1 - \epsilon)^L) (1 - \lambda) + (1 - \epsilon)^L \lambda, & \text{if } e_{\beta,k} = e_{\alpha,k} + \Delta \mathcal{E}_k \text{ and } q_{\beta,n} = q_{\alpha,n}; \\ 0, & \text{otherwise.} \end{cases} \quad (5a)$$

$$\Pr \left\{ (e_{\beta,n}, q_{\beta,n}) \mid (e_{\alpha,n}, q_{\alpha,n}), n \notin \mathcal{A}_k \right\} = \begin{cases} \lambda, & \text{if } e_{\beta,n} = e_{\alpha,n} \text{ and } q_{\beta,n} = q_{\alpha,n} + 1; \\ 1 - \lambda, & \text{if } e_{\beta,n} = e_{\alpha,n} \text{ and } q_{\beta,n} = q_{\alpha,n}; \\ 0, & \text{otherwise.} \end{cases} \quad (5b)$$

**Algorithm 1:** Value Iteration Algorithm.

---

```

1: 1. Initialize:
2:  $v^0(S_\alpha) = 0$  for the state  $S_\alpha$ ,  $t \in [0, \infty]$ , specify  $\varepsilon > 0$ ,
   and set  $w = 0$ 
3: 2. Iteration:
4: while  $S_\alpha \in \mathcal{S}$  do
5:    $v^{w+1}(S_\alpha) = \min_{k \in \mathcal{A}_k} \{R\{S_\beta | S_\alpha, \mathcal{A}_k\}$ 
      $+ \sum_{S_\beta \in \mathcal{S}} \omega Pr\{S_\beta | S_\alpha, \mathcal{A}_k\} v(S_\beta)\}$ 
6:   if  $\|v^{w+1} - v^w\| < \varepsilon(1 - \omega)/(2\omega)$  then
7:     for  $S_\alpha \in \mathcal{S}$  do
8:        $k = \arg \min_{k \in \mathcal{A}_k} \{R\{S_\beta | S_\alpha, \mathcal{A}_k\}$ 
          $+ \sum_{S_\beta \in \mathcal{S}} \omega Pr\{S_\beta | S_\alpha, \mathcal{A}_k\} v(S_\beta)\}$ 
9:     end for
10:  end if
11: end while

```

---

have

$$v(S_\alpha) = \min_{k \in \mathcal{A}_k} \{R\{S_\beta | S_\alpha, \mathcal{A}_k\} + \sum_{S_\beta \in \mathcal{S}} \omega Pr\{S_\beta | S_\alpha, \mathcal{A}_k\} v(S_\beta)\} \quad (9)$$

Therefore, the optimal action  $k$  which satisfies (9) can be given by

$$k = \arg \min_{k \in \mathcal{A}_k} \left\{ R\{S_\beta | S_\alpha, \mathcal{A}_k\} + \sum_{S_\beta \in \mathcal{S}} \omega Pr\{S_\beta | S_\alpha, \mathcal{A}_k\} v(S_\beta) \right\} \quad (10)$$

Note that  $k$  may not be unique, but at least one minimizing action exists due to a finite  $\mathcal{A}_k$ . Specifically, if  $k$  has an unique value, it is the optimal decision in state  $S_\alpha$ . If not, choosing any minimizing action can achieve the minimum expected packet loss.

The Value Iteration Algorithm is summarized in Algorithm 1.

Since the running time for each iteration is  $O(MN^2)$ , the presented value iteration algorithm is polynomial given the polynomial total number of required iterations.

Note that the centralized scheduling is ideal and can only be conducted off-line, under the assumptions of negligible signaling time and energy, as well as the resultant global availability of the required information. The optimal actions of the MDP are inapplicable to practical environments with non-negligible signaling time and energy consumption. Nevertheless, the MDP (or its optimal actions) can provide the lower bound of packet loss to any scheduling designs of the WPT powered sensor networks. The MDP also lays foundations to and eases the understanding of the on-line optimization of semi-decentralized scheduling, as will be described in detail in the next section.

## V. SEMI-DECENTRALIZED ENERGY-QUEUE AWARE DATA TRANSMISSION

Note that in EHMDP, the BS needs to have real-time knowledge on the packet arrivals and battery levels of the nodes. This requires the nodes to report prior to every time slot. Extend-

ing from the centralized scheme, we further propose a semi-decentralized approach that allows the nodes to spontaneously self-nominate for data transmission and energy harvesting. Each sensor node makes its own decision on whether it turns on transmission in a time slot, as well as the modulation level if the transmission is turned on. The decisions/actions that node  $i$  takes are based on the current battery level and queue length of its own, i.e.,  $e_i$  and  $q_i$ , as other nodes' battery and queue statuses (i.e.,  $e_j$  and  $q_j$ , for any  $j \neq i$ ) are unavailable to node  $i$  in distributed networks. Collisions occur under such distributed scheduling, due to the independent actions that each node takes.

### A. Protocol Design

To alleviate the collisions, for every node  $i = 1, \dots, N$ , we propose to design a transmission probability, denoted by  $p_i \in [0, 1]$ , at which the node is activated to transmit a packet and then harvest energy in the current time slot. Apparently,  $p_i$  needs to depend on  $e_i$  and  $q_i$ , which is given as  $f(e_i, q_i)$ . Our design philosophy is to design  $p_i$  to monotonically increase with  $q_i$  and decrease with  $e_i$ . The purpose of such design is to impose urgency to the nodes with built-up queues, avoid their transmissions being withheld due to insufficient energy harvested, and in turn, avoid buffer overflows at the nodes and reduce packet losses. Note that the definition of  $p_i$  is generic, which can be given by any classical probability distributions. Especially, Exponential, Sigmoid, and Gamma distributions are considered as three design examples of  $p_i = f(e_i, q_i)$  in this paper. The reason is that they present variant  $p_i$  regarding to the  $e_i$  and  $q_i$  so that the impact of  $p_i$  on the proposed protocol can be observed. Specifically, the three examples of  $p_i$  are given by

$$\text{Exponential design : } p_i = (1 - e^{-\kappa_e e_i}) e^{-\kappa_q q_i}$$

$$\text{Sigmoid design : } p_i = \sin\left(\frac{\pi q_i}{2Q}\right) \cos\left(\frac{\pi e_i}{2E}\right)$$

$$\text{Gamma design : } p_i = \frac{1}{\Gamma(\varrho)} \gamma\left(\varrho, \frac{1}{\theta} \frac{q_i}{e_i}\right) \quad (11)$$

where for exponential distribution,  $\lambda$  denotes the rate parameter; for Gamma distribution,  $k > 0$  is the shape value, and  $\theta > 0$  is the scale value.  $\Gamma(\varrho)$  is the Gamma function, and  $\gamma(\varrho, \frac{1}{\theta} \frac{q_i}{e_i})$  is the lower incomplete gamma function.

Other techniques to alleviate collisions in distributed scheduling include channel-sensing multiple access/collision-avoidance (CSMA/CA) and exponential back-off retransmission which are widely used in IEEE 802.11/802.15 distributed networks [40], [41]. An extended CSMA-based protocol, namely, Q-CSMA, is able to schedule nodes based on their queue lengths [42]. However, these techniques are not suitable for energy-restrained sensor nodes, because CSMA/CA requires non-negligible energy to continuously sense the availability of the channel and would quickly drain the batteries.

Under distributed scheduling, the BS becomes the passive data collector. Consider an ideal case, in a time slot where only a single node transmits a packet (collision-free), the BS starts by detecting the PHY/MAC header (i.e., the pilot signal) of the packet, then generates a receive beam to the node, and continues to receive the sensory data part of the packet. Once the node

completes transmitting the packet, the BS forms a transmit beam to transfer energy to the sensor node over the remaining part of the time slot. The BS can acknowledge its successful or failed reception of the packet by changing the phase or waveform of the energy signal. The node discards the packet, if successfully delivered, from its queue, based on the acknowledgement.

Consider multiple nodes have their transmissions collided in a time slot, the BS stays idle, since it can by no means identify the activated nodes. We denote  $\Pr_k^c$  as the packet collision probability of the selected node  $k$  by our MDP model, and it is given by [43]

$$\Pr_k^c = 1 - \prod_{n \neq k} (1 - p_n) \quad (12)$$

Equation (13) depicts the probability of node  $k$  that transmits data with packet collision probability and harvested energy. In the first case, the battery level of selected node  $k$  increases by  $\Delta\mathcal{E}_k$ , and the queue length also increases, which indicates that a new packet arrives and the packet of State  $\alpha$  is transmitted successfully with no collision. The second case is that the queue length of selected node  $k$  decreases since there is no new packet arrival and the packet of State  $\alpha$  is transmitted successfully, and the battery level increases via energy harvesting of WPT. The third case is that the queue length of the node  $k$  does not change. The battery level increases at State  $\beta$ , which denotes that a new packet arrives while the sensory packet of State  $\alpha$  is transmitted successfully. The fourth case is that the queue length of selected node  $k$  remains the same. However, the battery level decreases by  $\Delta\nu_k$ . Therefore, in this case, there is no new packet arrival, moreover, the transmission at State  $\alpha$  has packet collision, which requests a retransmission. In the last case, the queue length of selected node  $k$  decreases, and the battery level also drops by  $\Delta\nu_k$ , which indicates that no new packet arrival and the transmitted packet at State  $\alpha$  collides.

Note that a key difference between the probability of the selected node  $k$  in (13) and the one in (5a) and (5b) is due to the potential packet collision in the semi-decentralized approach. As a result, the probability formulated in (13) comprehensively considers transmit probability  $p_n$  ( $n \neq k$ ) and the packet collision probability  $\Pr_k^c$ .

Given a slot-based data transmission and energy harvesting, the BS needs to broadcast a beacon periodically to all the nodes

---

**Algorithm 2:** E-QAT algorithm.

---

**Input:**  $\Pr \left\{ (e_{\beta,k}, q_{\beta,k}) \middle| (e_{\alpha,k}, q_{\alpha,k}), k \in \mathcal{N} \right\}_{col}, p_k$   
1: **Repeat:**  
2: **if**  $\Pr \left\{ (e_{\beta,k}, q_{\beta,k}) \middle| (e_{\alpha,k}, q_{\alpha,k}), k \in \mathcal{N} \right\}_{col} < \Pr_{thresh}^0$   
**then**  
3: node  $k$  does not transmit.  
4:  $p_k^f \leftarrow \min\{1, (1 + \alpha)^f p_k\}$ .  
5: **else**  
6: node  $k$  transmits the data packet.  
7: **if** the packet is collided **then**  
8:  $p_k^f \leftarrow \min\{1, (1 + \alpha)^f p_k\}$ .  
9: back off a random time to retransmit.  
10: **else**  
11:  $p_k \leftarrow f(e_k, q_k)$   
12: **end if**  
13: **end if**

---

for network synchronization.  $p_n$  ( $n \in [1, N]$  and  $n \neq k$ ) can also be broadcast along with the beacon. As a result, the collision probability can be evaluated at each individual node  $k$  by using (12). (13) shown at the bottom of this page.

### B. Energy-Queue Aware Data Transmission

We propose an Energy-Queue Aware Transmission (E-QAT) algorithm. Specifically, node  $k$  first calculates the transition probability that is given by (13). In particular, the transition probability is formulated based on the node's packet collision probability, variation of its battery level, harvested energy, and data queue state information. If  $\Pr \left\{ (e_{\beta,k}, q_{\beta,k}) \middle| (e_{\alpha,k}, q_{\alpha,k}), k \in \mathcal{N} \right\}_{col}$  is lower than a given threshold,  $\Pr_{thresh}^0$ , the node does not transmit the packet. In particular,  $\Pr_{thresh}^0$  can be determined using collision-based medium access techniques, e.g., the one presented in [44]. Then, the transmission probability of node  $k$  is updated by  $p_k^f = \min\{1, (1 + \alpha)^f p_k\}$ , where  $f$  is the number of frames during which the node failed in data transmission and  $\alpha$  is the indicator parameter that can be given by [40]. Node  $k$  transmits its data packet only when  $\Pr \left\{ (e_{\beta,k}, q_{\beta,k}) \middle| (e_{\alpha,k}, q_{\alpha,k}), k \in \mathcal{N} \right\}_{col} \geq \Pr_{thresh}^0$ . If the packet is successfully received by the BS, the algorithm continues to process the next data transmission, and  $p_k$  returns

$$\Pr \left\{ (e_{\beta,k}, q_{\beta,k}) \middle| (e_{\alpha,k}, q_{\alpha,k}), k \in \mathcal{N} \right\}_{col} = \begin{cases} (1 - (1 - \epsilon)^L) \lambda \prod_{n \neq k} (1 - p_n), & \text{if } e_{\beta,k} = e_{\alpha,k} + \Delta\mathcal{E}_k \text{ and } q_{\beta,n} = q_{\alpha,n} + 1; \\ (1 - \epsilon)^L (1 - \lambda) \prod_{n \neq k} (1 - p_n), & \text{if } e_{\beta,k} = e_{\alpha,k} + \Delta\mathcal{E}_k \text{ and } q_{\beta,n} = q_{\alpha,n} - 1; \\ \left( (1 - (1 - \epsilon)^L) (1 - \lambda) + (1 - \epsilon)^L \lambda \right) \prod_{n \neq k} (1 - p_n), & \text{if } e_{\beta,k} = e_{\alpha,k} + \Delta\mathcal{E}_k \text{ and } q_{\beta,n} = q_{\alpha,n}; \\ \left( (1 - (1 - \epsilon)^L) (1 - \lambda) + (1 - \epsilon)^L \lambda \right) \Pr_k^c, & \text{if } e_{\beta,k} = e_{\alpha,k} - \Delta\nu_k \text{ and } q_{\beta,n} = q_{\alpha,n}; \\ (1 - \epsilon)^L (1 - \lambda) \Pr_k^c, & \text{if } e_{\beta,k} = e_{\alpha,k} - \Delta\nu_k \text{ and } q_{\beta,n} = q_{\alpha,n} - 1; \\ 0, & \text{otherwise.} \end{cases} \quad (13)$$

to  $f(e_k, q_k)$ . However, if the data packet collision happens, node  $k$  has to back off a random time to retransmit, meanwhile, the transmission probability is correspondingly increased by  $p_k^f = \min\{1, (1 + \alpha)^f p_k\}$ .

The detail is shown in Algorithm 2. Moreover, note that the modulation scheme  $\rho_k$  has been decided by (14) in Section IV. Since the sensor nodes have independent  $e_k$  and  $q_k$ , the  $p_k^f$  is independently updated in E-QAT, which significantly reduces the collision probability.

As discussed, the off-line EHMDP protocol and the E-QAT algorithm provide solutions to scheduling data transmission and WPT in centralized and semi-decentralized networks, respectively. In terms of performance, EHMDP provides a lower bound for the overall packet loss and an upper bound for throughput, at the cost of real-time signaling and centralized coordination. Moreover, the transition probability formulation in EHMDP, i.e., (5a) and (5b), jointly considers the battery level and queue length of the node. This inspires the design of transmit probability with collision in the E-QAT algorithm, i.e., (13), where the nodes can self-nominate for data transmission and energy harvesting based on the current battery level and queue length of its own.

## VI. NUMERICAL EVALUATION

The simulations prototyping our proposed scheduling are investigated in this section. We compare the performance between our proposed EHMDP and E-QAT, and existing scheduling strategies. We present extensive results showing the performance of E-QAT with different packet collision probability of the scheduled node  $k$ . Moreover, we explain how  $Pr_k^c$  varies with the MDP states given different probability distribution of  $p_i$  ( $i \in [1, N]$ ).

### A. Simulation Configuration

The sensor nodes ( $N \in [10, 40]$ ) are deployed in the range of 50 meters, which are all one-hop away from the BS. The node has the maximum discretized battery level  $E = 5$  and queue length  $Q = 6$ , and the highest modulation  $M$  is 5. Each data packet has a payload of 32 bytes, i.e.,  $L = 256$ . Moreover, the data transmission period is given as a sequence of time slots where each node generates one data packet and put into the queue. For the  $P_i^D(\rho_i)$  in (2), the two constants,  $\kappa_1$  and  $\kappa_2$  are 0.2 and 3, respectively. We set the target  $\epsilon = 0.05\%$  for the numerical results, i.e., the transmitted bits have an error no more than 0.05, however, this value can be configured depending on the traffic type and quality-of-service (QoS) requirement of the sensory data. Furthermore, the transmit power of WPT transmitter at BS is 3 watts, and power transfer efficiency,  $\delta$ , is set to 0.4.

We simulate the proposed EHMDP in a centralized network, and E-QAT in a distributed network. In either of the scenarios, a random scheduling strategy and a greedy algorithm are simulated for performance comparison. In centralized RWSN, the first algorithm, referred to as “Full Queue (FQ)”, is a greedy algorithm, where the scheduling is based on the data queue length of node. The node with full queue has a higher priority to transmit data and harvest power. The second algorithm is named as “Random Selection (RS)”, where the BS randomly selects one

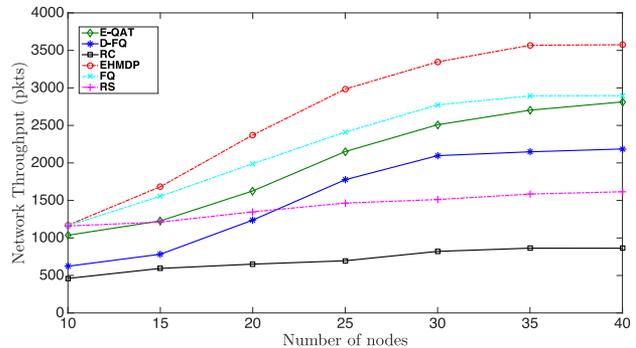


Fig. 2. A comparison of network throughput using different scheduling strategies in the centralized and distributed RWSN.

node to transmit data and transfer power. In the network with distributed scheduling, the node transmits data and harvests energy based on the local MDP states. The first strategy, referred to as “Decentralized FQ (D-FQ)”, is a distributed algorithm based on the data queue, where the node with a full queue transmits. The second strategy is named as “Random Contention (RC)”, where the node transmits data in a random-contention fashion. Note that the BS only transfers power to the node that wins the contention and completes data transmission in the slot. However, due to the distributed scheduling, the packets collide when multiple nodes are in the same MDP state.

### B. Performance Evaluation

1) *Centralized and Distributed Scheduling*: The network throughput of aforementioned six scheduling strategies are presented in Fig. 2. For centralized scheduling algorithms, EHMDP, FQ, and FS, they have similar throughput when there are 10 nodes in the network. However, from  $N = 15$  to  $N = 40$ , EHMDP outperforms the other three existing centralized scheduling algorithms. EHMDP achieves 17%, and 52% higher throughput than FQ and RS when  $N = 40$ . For semi-decentralized scheduling,  $p_i$  ( $i \in [1, N]$ ) is given by Sigmoid distribution in (11). It is also observed that the throughput of E-QAT is higher than D-FQ and RC by 21% and 68%, respectively.

Furthermore, the throughput of EHMDP is more than E-QAT by around 20%. The reason is that the state information of each node is known by the BS, which enables the centralized scheduling of the node and modulation for every time slot.

In general, the throughput of distributed scheduling algorithms is lower than the centralized ones, which is caused by the packets collision. This issue is also observed by Fig. 3, which presents the network packet loss rate. The packet loss rate of the centralized algorithms is generally lower than the distributed ones. Specifically, EHMDP achieves 42%, 57%, and 72% less packet loss than E-QAT, D-FQ, and RC when  $N = 10$ . Moreover, the packet loss rate of E-QAT is less than D-FQ and RC output by 16% and 33%. We also see that the packet loss rate of the distributed algorithms grows to 100% with an increase of number of nodes since more time slot contentions and higher  $Pr_k^c$ . Moreover, the network throughput and packet loss achieved by E-QAT are much closer to those of the optimal EHMDP, as compared to the other distributed protocols.

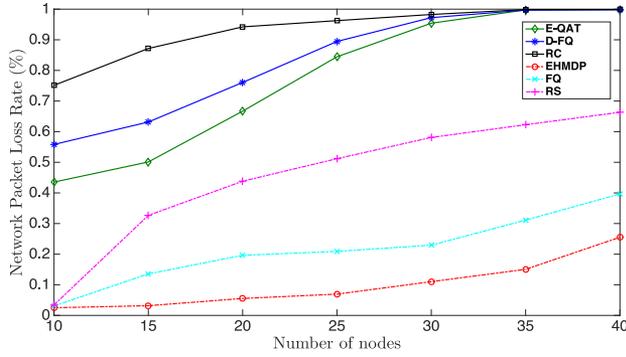


Fig. 3. A comparison of network packet loss rate by our EHMDP and E-QAT, and the typical scheduling strategies.

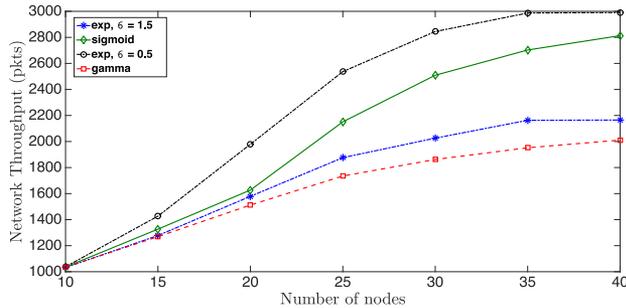


Fig. 4. Network throughput with different the packet transmission probability in terms of  $p_i (i \in [1, N])$ .

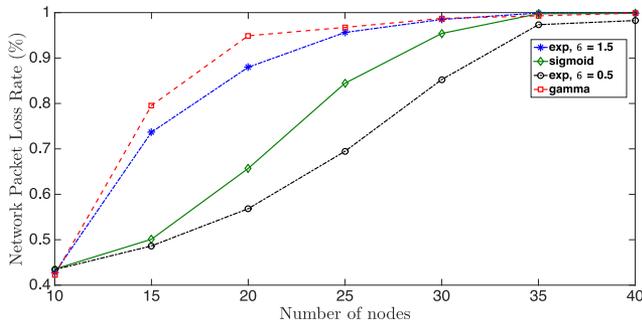


Fig. 5. Network packet loss with different  $p_i (i \in [1, N])$ .

Note that in EHMDP, the BS needs to have real-time knowledge on the packet arrivals and battery levels of the nodes. This requires the nodes to report prior to every time slot. However, the semi-decentralized approach allows the nodes to spontaneously self-nominate for data transmission and energy harvesting. Each sensor node makes its own decision on whether it turns on transmission in a time slot, as well as the modulation level if the transmission is turned on.

2) *Impact of  $p_i$  to E-QAT*: Figs. 4 and 5 depict the network throughput and packet loss rate of E-QAT with respect to  $p_i (i \in [1, N])$  in (11). In Fig. 4,  $p_i$  following Exponential distribution with  $\lambda = 0.5$  (shown as exp,  $\lambda = 0.5$ ) achieves the highest network throughput, about 3000 packets when  $N = 40$ . Moreover, the Gamma distribution of  $p_i$  gives the lowest network throughput, which is about 2000 packets. In Fig. 5, when  $N = 10$ , the four distributions perform similar packet loss rate, which is around 43%. With an increase of number of nodes, Gamma

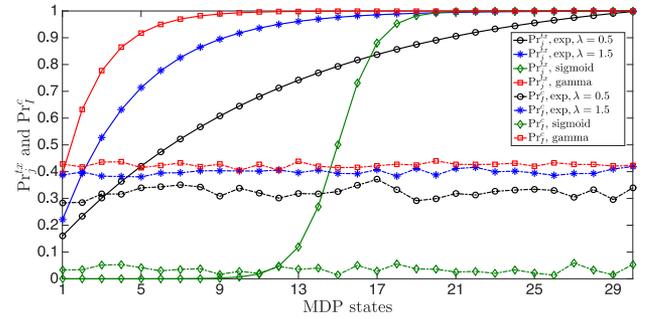


Fig. 6.  $p_i (i \in [1, N])$  and  $\Pr_k^c$  with different probability distributions.

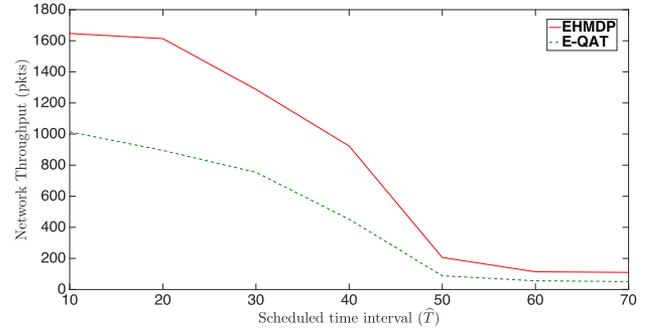


Fig. 7. Network throughput regarding to different  $\hat{T}$ .

distribution performs 7% 30%, and 40% more packet loss than (exp,  $\lambda = 1.5$ ), Sigmoid, and (exp,  $\lambda = 0.5$ ), respectively.

To further explain the reason of performance difference, Fig. 6 plots the impact of different probability distributions to the  $\Pr_j^{t,x}$  and  $\Pr_k^c$  over all MDP states. In total, the node has 30 MDP states since the maximum discretized energy level  $E$  is 5 and queue length  $Q$  is 6. Given the four probability distributions,  $\Pr_j^{t,x}$  generally increases with the growth of the state, because the node with more energy or longer data queue has a higher transmission probability. We also observe in Fig. 6 that  $\Pr_k^c$  stabilizes regarding to the states. Moreover, the Sigmoid distribution performs lower  $\Pr_k^c$  than the other three. The reason is that the  $\Pr_j^{t,x}$  of Sigmoid distribution is lower than the others from state 1 to 16. Thus, when the nodes ( $\neq k$ ) are in low MDP state, the  $\Pr_k^c$  has low value due to the definition of (12).

3) *Impact of  $\hat{T}$* : In this case, we study impact of the scheduling interval that lasts  $\hat{T}$  mini-slots. The number of nodes is fixed to 15, and the other settings are as configured in Section VI-A. Figs. 7 and 8 present the network throughput and packet loss rate of EHMDP and E-QAT, where  $\hat{T} \in [10, 40]$ . It is observed that the network throughput of EHMDP and E-QAT generally drops with  $\hat{T}$  while the packet loss rates of EHMDP and E-QAT grow. It confirms the fact that an increasing number of packets is generated and injected into the data queues of nodes with a longer  $\hat{T}$ . However, only one node is scheduled to transmit data and harvest energy during each  $\hat{T}$ , though the node can harvest more energy with a longer  $\hat{T}$  due to (3). As a result, the nodes that are not scheduled suffer from more packet loss caused by data buffer overflow, compared to those with shorter  $\hat{T}$ . Moreover, EHMDP achieves higher network throughput and lower packet loss than E-QAT, since the former has state information of each node and can accordingly achieve the optimal schedules.

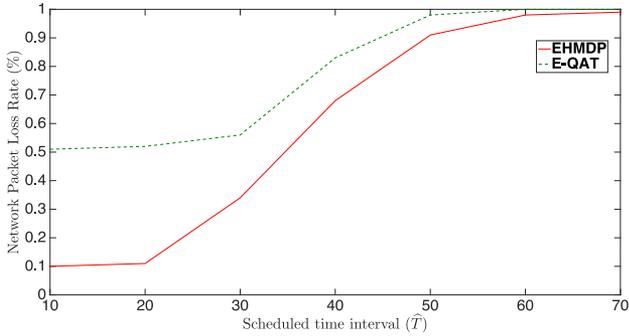


Fig. 8. Network packet loss regarding to different  $\hat{T}$ .

## VII. CONCLUSION

In this paper, we address a scheduling problem on the data collection and WPT in RWSN. The problem is formulated as a finite state MDP with the objective of minimizing the packet loss. In terms of the scheduling mechanism of MAC protocol, two different MDP formulations have been studied, i.e., collision-free centralized scheduling, EHMDP, and contention-based semi-decentralized scheduling, E-QAT. Moreover, an off-line optimization of actions in centralized scheduling is investigated to provide the lower bound of packet loss to the scheduling algorithms of the WPT powered sensor networks. For the semi-decentralized scheduling, we develop a solution to obtain the optimal scheduling policy with different packet collision probability distributions. Numerical results have shown that our proposed MDP formulation and algorithm outperform existing algorithms with substantial gains on throughput and packet loss.

## APPENDIX

Given the goal of the MDP to minimize the packet loss stemming from insufficient energy,  $\rho_i$  is to be chosen to maximize the energy gained during each scheduled time interval, or “epoch”, with a duration of  $\hat{T}$ . The optimal modulation of node  $i$ ,  $\rho_i$ , is independent of the battery level and the queue length of the node  $i$ . This is because  $\rho_i$  is selected to maximize the increase of the battery level at node  $i$ , not the battery level itself, under the bit error rate requirement  $\epsilon_i$  for the packet transmitted. As a result,  $\rho_i$  is decoupled from  $\mathcal{A}$ , and optimized a priori by

$$\rho_i = \arg \max_{\rho=1, \dots, M} \left\{ \left( \hat{T} - \frac{L}{\rho W} \right) P_i^E - \frac{L}{\rho W} P_i^D(\rho) \right\}, \quad (14)$$

the right-hand side (RHS) of which, by substituting (1) and (2), can be rewritten as

$$\max_{\rho=1, \dots, M} \left\{ \left( \hat{T} - \frac{L}{\rho W} \right) P_e \|\mathbf{h}_i\|^2 - \frac{L \kappa_2^{-1} \ln(\frac{\kappa_1}{\epsilon})}{\|\mathbf{h}_i\|^2 \rho W} (2^\rho - 1) \right\} \quad (15)$$

where  $W$  is the bandwidth of the uplink data transmission,  $\frac{1}{W}$  is the duration of an uplink symbol,  $\frac{L}{\rho W}$  is the duration of uplink data transmission, and  $(\hat{T} - \frac{L}{\rho W})$  is the rest of the epoch used for downlink WPT.

By using the first-order necessary condition of the optimal solution, we have

$$\frac{d}{d\rho} \left( \left( \hat{T} - \frac{L}{\rho W} \right) P_e \|\mathbf{h}_i\|^2 - \frac{L \kappa_2^{-1} \ln(\frac{\kappa_1}{\epsilon})}{\|\mathbf{h}_i\|^2 \rho W} (2^\rho - 1) \right) = 0 \quad (16)$$

$$\begin{aligned} \rho^{-2} \frac{L}{W} P_e \|\mathbf{h}_i\|^2 - \frac{L \kappa_2^{-1} \ln(\frac{\kappa_1}{\epsilon})}{\|\mathbf{h}_i\|^2 W} (\rho^{-1} 2^\rho \ln 2 - \rho^{-2} 2^\rho) \\ - \frac{L \kappa_2^{-1} \ln(\frac{\kappa_1}{\epsilon})}{\|\mathbf{h}_i\|^2 W} \rho^{-2} = 0 \end{aligned} \quad (17)$$

The  $\rho$  values are then given as follows:

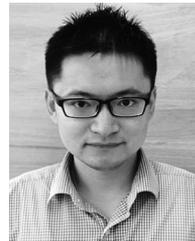
$$\rho 2^\rho \ln 2 - 2^\rho = \frac{L}{W} P_e \|\mathbf{h}_i\|^2 \frac{\|\mathbf{h}_i\|^2 W}{L \kappa_2^{-1} \ln(\frac{\kappa_1}{\epsilon})} - 1 \quad (18)$$

Since the left-hand side of (18) monotonically increases with  $\rho$ , the optimal value  $\rho^*$  is able to be obtained by applying a bisectional method, and evaluating the two closest integers about the fixed point of the bisectional method.

## REFERENCES

- [1] L. Fu, L. He, P. Cheng, Y. Gu, J. Pan, and J. Chen, “ESync: Energy synchronized mobile charging in rechargeable wireless sensor networks,” *IEEE Trans. Veh. Technol.*, vol. 65, no. 9, pp. 7415–7431, Sep. 2016.
- [2] Y. L. Che, L. Duan, and R. Zhang, “Spatial throughput maximization of wireless powered communication networks,” *IEEE J. Sel. Areas Commun.*, vol. 33, no. 8, pp. 1534–1548, 2015.
- [3] L. Xiao, P. Wang, D. Niyato, D. Kim, and Z. Han, “Wireless networks with RF energy harvesting: A contemporary survey,” *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 757–789, Secondquarter 2014.
- [4] S. Sudevalayam and P. Kulkarni, “Energy harvesting sensor nodes: Survey and implications,” *IEEE Commun. Surveys Tuts.*, vol. 13, no. 3, pp. 443–461, Third Quarter 2011.
- [5] Y. Zhang, D. Niyato, P. Wang, and D. I. Kim, “Optimal energy management policy of mobile energy gateway,” *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3685–3699, May 2016.
- [6] K. Li, C. Yuen, B. Kusy, R. Jurdak, A. Ignjatovic, S. S. Kanhere, and S. Jha, “Fair scheduling for data collection in mobile sensor networks with energy harvesting,” *arXiv preprint arXiv:1603.02476*, 2016.
- [7] Y. Shu *et al.*, “Near-optimal velocity control for mobile charging in wireless rechargeable sensor networks,” *IEEE Trans. Mobile Comput.*, vol. 15, no. 7, pp. 1699–1713, Jul. 2016.
- [8] K. Li, C. Yuen, and S. Jha, “Fair scheduling for energy harvesting WSN in smart city,” in *Sensys. ACM*, 2015, pp. 419–420.
- [9] Z. Wang, L. Duan, and R. Zhang, “Adaptively directional wireless power transfer for large-scale sensor networks,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1785–1800, 2016.
- [10] S. Gong, L. Duan, and N. Gautam, “Optimal scheduling and beamforming in relay networks with energy harvesting constraints,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1226–1238, 2016.
- [11] Y. Che, J. Xu, L. Duan, and R. Zhang, “Multiantenna wireless powered communication with cochannel energy and information transfer,” *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2266–2269, 2015.
- [12] Z. Ding *et al.*, “Application of smart antenna technologies in simultaneous wireless information and power transfer,” *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 86–93, Apr. 2015.
- [13] Z. Nan, T. Chen, X. Wang, and W. Ni, “Energy-efficient transmission schedule for delay-limited bursty data arrivals under nonideal circuit power consumption,” *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6588–6600, 2016.
- [14] X. Chen, W. Ni, X. Wang, and Y. Sun, “Provisioning quality-of-service to energy harvesting wireless communications,” *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 102–109, 2015.
- [15] X. Chen, W. Ni, X. Wang, and Y. Sun, “Optimal quality-of-service scheduling for energy-harvesting powered wireless communications,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3269–3280, 2016.

- [16] Q. Cui, T. Yuan, and W. Ni, "Energy-efficient two-way relaying under non-ideal power amplifiers," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1257–1270, 2017.
- [17] Q. Cui, Y. Zhang, W. Ni, M. Valkama, and R. Jäntti, "Energy efficiency maximization of full-duplex two-way relay with nonideal power amplifiers and non-negligible circuit power," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6264–6278, 2017.
- [18] W. Ni and I. B. Collings, "A new adaptive small-cell architecture," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 829–839, 2013.
- [19] S. He, J. Chen, F. Jiang, D. K. Yau, G. Xing, and Y. Sun, "Energy provisioning in wireless rechargeable sensor networks," *IEEE Trans. Mobile Comput.*, vol. 12, no. 10, pp. 1931–1942, Oct. 2013.
- [20] L. Fu, P. Cheng, Y. Gu, J. Chen, and T. He, "Optimal charging in wireless rechargeable sensor networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 278–291, Jan. 2016.
- [21] Y. Zhang, S. He, and J. Chen, "Data gathering optimization by dynamic sensing and routing in rechargeable sensor networks," *IEEE/ACM Trans. Netw.*, vol. 24, no. 3, pp. 1632–1646, Jun. 2016.
- [22] F. Zhao, L. Wei, and H. Chen, "Optimal time allocation for wireless information and power transfer in wireless powered communication systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1830–1835, Mar. 2016.
- [23] J. Xu, L. Liu, and R. Zhang, "Multiuser MISO beamforming for simultaneous wireless information and power transfer," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4798–4810, Sep. 2014.
- [24] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer in multiuser OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 2282–2294, Apr. 2014.
- [25] K. Lee and J.-P. Hong, "Energy-efficient resource allocation for simultaneous information and energy transfer with imperfect channel estimation," *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, pp. 2775–2780, Apr. 2016.
- [26] X. Kang, C. K. Ho, and S. Sun, "Full-duplex wireless-powered communication network with energy causality," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5539–5551, Oct. 2015.
- [27] N. Tekbiyik, T. Girici, E. Uysal-Biyikoglu, and K. Leblebicioglu, "Proportional fair resource allocation on an energy harvesting downlink," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1699–1711, Apr. 2013.
- [28] J. Marañević, C. Stein, and G. Zussman, "Max-min fair rate allocation and routing in energy harvesting networks: Algorithmic analysis," *Algorithmica J.*, vol. 78, no. 2, pp. 521–557, 2017.
- [29] C. Qin, W. Ni, H. Tian, R. P. Liu, and Y. J. Guo, "Joint beamforming and user selection in multiuser collaborative MIMO SWIPT systems with non-negligible circuit energy consumption," *IEEE Trans. Veh. Technol.*, 2017.
- [30] M. Kunikawa, H. Yomo, K. Abe, and T. Ito, "A fair polling scheme for energy harvesting wireless sensor networks," in *Proc. IEEE 81st Veh. Technol. Conf.*, 2015, pp. 1–5.
- [31] M. Y. Naderi, P. Nintanavongsa, and K. R. Chowdhury, "RF-MAC: A medium access control protocol for re-chargeable sensor networks powered by wireless energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3926–3937, Jul. 2014.
- [32] P. Nintanavongsa, M. Y. Naderi, and K. R. Chowdhury, "Medium access control protocol design for sensors powered by wireless energy transfer," in *Proc. IEEE Conf. Comput. Commun.*, 2013, pp. 150–154.
- [33] X. Chen, X. Wang, and X. Chen, "Energy-efficient optimization for wireless information and power transfer in large-scale MIMO systems employing energy beamforming," *IEEE Wireless Commun. Lett.*, vol. 2, no. 6, pp. 667–670, Dec. 2013.
- [34] K. Li, W. Ni, X. Wang, R. P. Liu, S. S. Kanhere, and S. Jha, "Energy-efficient cooperative relaying for unmanned aerial vehicles," *IEEE Trans. Mobile Comput.*, vol. 15, no. 6, pp. 1377–1386, 2016.
- [35] T. He, X. Wang, and W. Ni, "Optimal chunk-based resource allocation for OFDMA systems with multiple BER requirements," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4292–4301, Nov. 2014.
- [36] A. Kadrolkar, R. X. Gao, R. Yan, and W. Gong, "Variable-word-length coding for energy-aware signal transmission," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 4, pp. 850–864, Apr. 2012.
- [37] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Hoboken, NJ, USA: Wiley, 2014.
- [38] E. F. Arruda, F. Ourique, and A. Almudevar, "Toward an optimized value iteration algorithm for average cost Markov decision processes," in *Proc. IEEE Conf. Decis. Control*, 2010, pp. 930–934.
- [39] C. Sun, E. Stevens-Navarro, and V. W. Wong, "A constrained MDP-based vertical handoff decision algorithm for 4G wireless networks," in *Proc. IEEE Int. Conf. Commun.*, 2008, pp. 2169–2174.
- [40] T. Chaturvedi, K. Li, C. Yuen, A. Sharma, L. Dai, and M. Zhang, "On the design of MAC protocol and transmission scheduling for Internet of Things," in *IEEE Region 10 Conf. (TENCON)*, 2016, pp. 2000–2003.
- [41] K. Li, B. Kusy, R. Jurdak, A. Ignjatovic, S. S. Kanhere, and S. Jha, " $\kappa$ -FSOM: Fair link scheduling optimization for energy-aware data collection in mobile sensor networks," in *European Conference on Wireless Sensor Networks*. New York, NY, USA: Springer, 2014, pp. 17–33.
- [42] J. Ni, B. Tan, and R. Srikant, "Q-CSMA: Queue-length-based CSMA/CA algorithms for achieving maximum throughput and low delay in wireless networks," *IEEE/ACM Trans. Netw.*, vol. 20, no. 3, pp. 825–836, Jun. 2012.
- [43] B. Shrestha, E. Hossain, K. W. Choi, and S. Camorlinga, "A Markov decision process (MDP)-based congestion-aware medium access strategy for IEEE 802.15.4," in *Proc. Global Telecommun. Conf.*, 2011, pp. 1–5.
- [44] Y. Tay, K. Jamieson, and H. Balakrishnan, "Collision-minimizing CSMA and its applications to wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1048–1057, Aug. 2004.



**Kai Li** (S'09–M'14) received the B.E. degree from Shandong University, Jinan, China, in 2009, the M.S. degree from The Hong Kong University of Science and Technology, Hong Kong, in 2010, and the Ph.D. degree in computer science from the University of New South Wales, Sydney, NSW, Australia, in 2014. He is currently a Research Scholar with the Real-Time and Embedded Computing Systems Research Centre (CISTER), Porto, Portugal. From 2014 to 2016, he was a Postdoctoral Research Fellow with the SUTD-MIT International Design Centre, Singapore University of Technology and Design, Singapore. He was a Visiting Research Assistant with the ICT Centre, Commonwealth Scientific and Industrial Research Organisation, Australia, from 2012 to 2013. From 2010 to 2011, he was a Research Assistant with the Mobile Technologies Centre, The Chinese University of Hong Kong. His research interests include vehicular communications and security, cyber-physical systems, human sensing systems, resource allocation optimization, wireless sensor networks, and UAV networks.



**Wei Ni** (M'09–SM'15) received the B.E. and Ph.D. degrees in electronic engineering from Fudan University, Shanghai, China, in 2000 and 2005, respectively. He is currently a Team Leader and a Project Leader with Cyber Physical System, Commonwealth Scientific and Industrial Research Organisation, Sydney, NSW, Australia. He also holds adjunct positions at the University of New South Wales, Sydney, NSW, Australia, Macquarie University, Sydney NSW, USA, and the University of Technology Sydney, Ultimo, NSW, USA. Prior to this, he was a Postdoctoral Research Fellow with Shanghai Jiaotong University (2005–2008), a Deputy Project Manager with Bell Labs R&I Center, Alcatel/Alcatel-Lucent (2005–2008), and a senior researcher with Devices R&D, Nokia (2008–2009). His research interests include stochastic optimization, game theory, graph theory, as well as their applications to communication, network and security. He served as an Editor for *Hindawi Journal of Engineering* from 2012 to 2014, a Secretary of IEEE NSW VTS Chapter from 2015 to 2017, the Track Chair for the VTC-Spring 2017, the Track Chair for the IEEE VTC-Spring 2016, and the Publication Chair for BodyNet 2015. He also served as the Student Travel Grant Chair for WPMC 2014, a Program Committee member for CHINACOM 2014, and a TPC member for IEEE ICC'14, ICC'15, EICE'14, and WCNC'10.



**Lingjie Duan** (S'09–M'12) received the Ph.D. degree from The Chinese University of Hong Kong, Hong Kong, in 2012. In 2011, he was a Visiting Scholar at the University of California at Berkeley, Berkeley, CA, USA. He is currently an Assistant Professor with the Engineering Systems and Design Pillar, Singapore University of Technology and Design, Singapore. His research interests include network economics and game theory, cognitive communications and cooperative networking, energy-harvesting wireless communications, and mobile crowdsourcing.

He was a recipient of the 10th IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2015, the Hong Kong Young Scientist Award (Finalist in Engineering Science track) in 2014, and the CUHK Global Scholarship for Research Excellence in 2011. In 2016, he was a Guest Editor for the *IEEE Journal on Selected Areas in Communications* for special issue on human-in-the-loop mobile networks and the IEEE WIRELESS COMMUNICATIONS MAGAZINE for feature topic sustainable green networking and computing in 5G systems. He served as the Program Co-Chair for the IEEE INFOCOM 2014 GCCCN Workshop, the ICCS 2014 special session on economic theory and communication networks, the Wireless Communication Systems Symposium of the IEEE ICC 2015, the GCNC Symposium of the IEEE ICNC 2016, and the IEEE INFOCOM 2016 GSNC Workshop. He also served as a Technical Program Committee member for many leading conferences in communications and networking, such as ACM MobiHoc, the IEEE INFOCOM, SECON, WiOPT, ICC, GLOBECOM, WCNC, and NetEcon. He is an Editor for the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS and a SWAT member of the Editorial Board of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.



**Mehran Abolhasan** (M'01–SM'11) received the B.E. degree in computer engineering and the Ph.D. degree in telecommunications from the University of Wollongong, Wollongong, NSW, Australia, in 1999 and 2003, respectively. He is currently an Associate Professor and the Deputy Head of the School for Research, School of Computing and Communications, University of Technology Sydney, Ultimo, NSW, Australia. He has led several major research and development projects in various areas of wireless communications since 2003 and authored more than

60 international publications. He was a recipient of more than \$1 million in research funding. His current research interests include wireless mesh, fifth-generation cooperative networks, and body-area/sensor networks.



**Jianwei Niu** (SM'12) received the M.S. and Ph.D. degrees in computer science from Beihang University, Beijing, China, in 1998 and 2002, respectively. He was a Visiting Scholar with the School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA, from 2010 to 2011. He is currently a Professor with the School of Computer Science and Engineering, Beihang University. He has authored or coauthored more than 150 referred papers and has filed more than 30 patents. His current research interests include wireless sensor networks and mobile

and pervasive computing. He served as a TPC member for InfoCom, PerCom, ICC, WCNC, GLOBECOM, and LCN. He was a recipient of the New Century Excellent Researcher Award from the Ministry of Education of China in 2009, the first prize of the Technical Invention Award from the Ministry of Education of China in 2012, the Innovation Award from the Nokia Research Center, and the Best Paper Award at JNCA 2015, IEEE ICC 2013, WCNC 2013, ICACT 2013, CWSN 2012, and GreenCom 2010. He served as the DySON Workshop Co-Chair for INFOCOM 2014, the Program Chair for the IEEE SEC 2008, and the Executive Co-Chair for the TPC of CPSCOM 2013. He has served as an Associate Editor for the *International Journal of Ad Hoc and Ubiquitous Computing*, *Journal of Network and Computer Applications*, and *Mobile Networks and Applications*.