CISTER
Research Centre in
Real-Time & Embedded
Computing Systems

Journal Paper

Fair Scheduling for Data Collection in Mobile Sensor Networks with Energy Harvesting

Kai Li*
Chau Yuen
Branislav Kusy
Raja Jurdak
Aleksandar Ignjatovic
Salil S. Kanhere
Sanjay Jha

*CISTER Research Centre

CISTER-TR-180902

2018

Fair Scheduling for Data Collection in Mobile Sensor Networks with Energy Harvesting

Kai Li*, Chau Yuen, Branislav Kusy, Raja Jurdak, Aleksandar Ignjatovic, Salil S. Kanhere, Sanjay Jha

*CISTER Research Centre

Polytechnic Institute of Porto (ISEP-IPP)

Rua Dr. António Bernardino de Almeida, 431

4200-072 Porto

Portugal

Tel.: +351.22.8340509, Fax: +351.22.8321159

E-mail: kaili@isep.ipp.pt, yuenchau@sutd.edu.sg, Brano.Kusy@data61.csiro.au, ignjat@cse.unsw.edu.au, sanjay.jha@unsw.edu.au

http://www.cister.isep.ipp.pt

Abstract

We consider the problem of data collection from a network of energy harvesting sensors, applied to tracking mobile assets in rural environments. Our application constraints favor a fair and energy-aware solution, with heavily duty-cycled sensor nodes communicating with powered base stations. We study a novel scheduling optimization problem for energy harvesting mobile sensor network, that maximizes the amount of collected data under the constraints of radio link quality and energy harvesting efficiency, while ensuring a fair data reception. We show that the problem is NP-complete and propose a heuristic algorithm to approximate the optimal scheduling solution in polynomial time. Moreover, our algorithm is flexible in handling progressive energy harvesting events, such as with solar panels, or opportunistic and bursty events, such as with Wireless Power Transfer. We use empirical link quality data, solar energy, and WPT efficiency to evaluate the proposed algorithm in extensive simulations and compare its performance to state-of-the-art. We show that our algorithm achieves high data reception rates, under different fairness and node lifetime constraints.

Fair Scheduling for Data Collection in Mobile Sensor Networks with Energy Harvesting

Kai Li, *Member, IEEE,* Chau Yuen, *Senior Member, IEEE,* Branislav Kusy, *Senior Member, IEEE,* Raja Jurdak, *Senior Member, IEEE,* Aleksandar Ignjatovic, Salil S. Kanhere, *Senior Member, IEEE,* and Sanjay Jha, *Senior Member, IEEE*

***Abstract*—We consider the problem of data collection from a network of energy harvesting sensors, applied to tracking mobile assets in rural environments. Our application constraints favor a fair and energy-aware solution, with heavily duty-cycled sensor nodes communicating with powered base stations. We study a novel scheduling optimization problem for energy harvesting mobile sensor network, that maximizes the amount of collected data under the constraints of radio link quality and energy harvesting efficiency, while ensuring a fair data reception. We show that the problem is NP-complete and propose a heuristic algorithm to approximate the optimal scheduling solution in polynomial time. Moreover, our algorithm is flexible in handling progressive energy harvesting events, such as with solar panels, or opportunistic and bursty events, such as with Wireless Power Transfer. We use empirical link quality data, solar energy, and WPT efficiency to evaluate the proposed algorithm in extensive simulations and compare its performance to state-of-the-art. We show that our algorithm achieves high data reception rates, under different fairness and node lifetime constraints.**

***Index Terms*—Link scheduling, Optimization, Fairness, Energy Harvesting, Mobile Sensor Network**

I. Introduction

RECENT advances in embedded systems and battery technology have enabled a new class of wireless sensing applications [1], [2]. Consider a swarm of micro-aerial vehicles fitted with a variety of sensors that can achieve fine-grained three-dimensional sampling of our physical spaces, enabling exciting new applications such as urban surveillance, disaster recovery and environmental monitoring [3], [4]. It is now possible to monitor individual movement patterns of animals as well as their surrounding environment [5]–[9]. In a typical mobile sensing scenario, sensor nodes mounted on a carrier (e.g., vehicle or animal) collect numerous sensor readings while in transit. The nodes ultimately arrive back at a known rendezvous point (e.g., command center or animal pen), often as a large swarm and remain there for an extended period of time. The data stored on each sensor node is offloaded to a base station (BS) during this time. Moreover, since sensor nodes are typically powered by batteries with limited energy, energy harvesting techniques such as solar panel [10] and Wireless Power Transfer (WPT) [11], [12] have been investigated to extend lifetime of nodes. WPT is implementable by various technologies such as inductive coupling, magnetic resonate coupling, and electromagnetic radiation, for short, medium, and long distance applications, respectively [13], [14]. Presently, the long distance WPT system has been studied to power a large number of devices distributed in a wide area [15]–[17].

A number of considerations make the data collection in animal tracking with energy harvesting non-trivial. First, the number of nodes can be quite large (several hundreds) and while the nodes normally arrive back in large groups, their exact arrival sequence is often unknown. Second, during days with cloudy skies and adverse solar charging weather conditions, the amount of solar energy harvested is reduced. In addition, charging efficiency of WPT becomes very low when the node is far from the WPT transmitter (large-scale channel fading) [18] or encounters antenna orientation bias (shown in Section VI-A). It is thus critical to collect more data from those nodes before their harvested energy is exhausted. Third, the wireless link quality of data transmission between each node and the BS may vary with time. Having a node transmit during instances when the channel quality is poor is likely to result in packet reception errors, which in turn would require retransmissions and thus increased energy expenditure. Fourth, data should be downloaded from the nodes in a fair way, as the social behavior of animals can only be learned when a sufficient amount of data from a large number of nodes is available. In particular, the amount of data collected from each node should be greater than a certain application-specific threshold. This is important to maximize the accuracy of data analysis, e.g., in the context of mobility modeling and population characteristics for animal monitoring [19]–[21].

Conventional scheduling such as the one employed in IEEE 802.15.4 [22], [23] are based on First Come First Served (FCFS), which we refer to as *batch processing*. Batch processing has limited performance in real-world conditions with irregular radio channels and limited bandwidth. Any node with poor link quality occupies the channel due to retransmissions, while the nodes with higher link

An earlier version of this article appeared in the Proceedings of the 11th the European Conference on Wireless Sensor Networks (*EWSN*) [Li et al. 2014]. This article features a new scheduling optimization with energy harvesting, experimental characterization and more complete theoretical analysis.

K. Li is with Real-Time and Embedded Computing Systems Research Centre (CISTER), Portugal. (Email:kaili@isep.ipp.pt)

C. Yuen is with The Singapore University of Technology and Design, Singapore.

B. Kusy and R. Jurdak are with Autonomous Systems Lab, ICT Centre, CSIRO, Australia.

A. Ignjatovic, S. S. Kanhere, and S. Jha are with the School of Computer Science and Engineering, The University of New South Wales, Australia.

quality have to wait. In addition, batch processing does not support data collection fairness, potentially downloading a large amount of data from a small subset of nodes.

As an example, consider the problem of scheduling data transmissions in cattle monitoring application [24]. A sensor collar which contains embedded sensors (e.g., GPS, 3-axis accelerometer and magnetometer) is attached to the cow to record biological data [25]–[27]. The solar panel on the node harvests energy continuously during the sunny daytime. A WPT receiver on the sensor collar harvests energy from the WPT transmitter opportunistically when the animal stays in the charging range. In addition, harvesting energy from both solar and WPT implicitly results in a higher data reception than the one with a single energy harvesting source. The reason is that, firstly, the lifetime of the node is further extended so that more data are collected. Secondly, the sensor node with low energy harvesting from one source also has the chance to be scheduled for data transmission when it harvests a greater amount of energy from the other source. The data is offloaded to a BS which is deployed near a cattle drinking trough. Figure 1 depicts an energy harvesting mobile sensor network (MSN) for data collection.

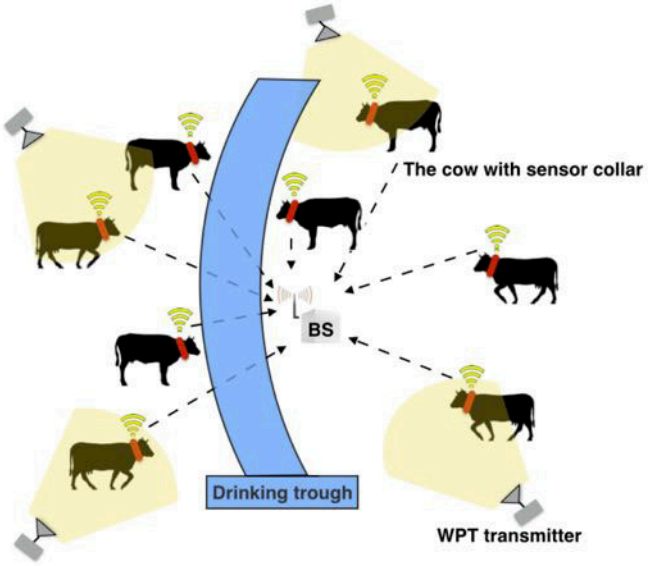


Fig. 1: Motivating Application: the energy harvesting mobile sensor network for cattle monitoring. The sensor collar is equipped with a solar panel and a WPT receiver.

The key to precisely analyzing biological features and social behavior of the animal is to maximize information aggregation from their sensors, besides scheduling the data transmission efficiently and fairly. For that purpose, in this paper, we propose Energy Harvesting Fair Scheduling (EHFS) to optimize data collection in an energy harvesting sensor network. Although energy harvesting sustainability is important, we do not focus on how energy is harvested to maintain a stable battery level on the node, instead, we focus on developing the optimization model to schedule transmissions based on residual energy of the node and link quality, given the energy harvesting uncertainty on solar power and WPT. Then, the optimal scheduling policy becomes that of adaptively changing the time slots' allocation to the nodes according to their instantaneous harvested energy and channel condition. It also ensures fairness by attempting to guarantee a certain application-specific amount of data collected from each node. We first show that this optimization problem is NP-complete. Next, we propose EHFS heuristic algorithm to optimize the scheduling in linear time. The EHFS algorithm prioritizes the nodes for scheduling based on a ratio of the link quality and harvested energy. This enables the nodes with the lowest energy reserves and good communication links to transfer their data first. In addition, we develop a state transition model to address the fairness criterion and maximize overall network goodput. Moreover, a Sensor-WPT testbed is built to characterize the WPT charging efficiency. Specifically, the experimental results show that WPT efficiency is jointly affected by distance between WPT transmitter and receiver, and their antenna orientation. While we use the cattle monitoring application as a case study, the proposed optimization model and EHFS algorithm are application-agnostic and hence applicable to a wide variety of energy harvesting mobile sensing scenarios with delay tolerance. The communication protocol and super frame structure are also able to be applied to a generic mobile sensor network. The node travels nomadically to sense critical environmental data and return to a data processing centre, i.e., the base station. Communication within these networks is often scheduled to conserve power and reduce packet collision.

The rest of paper is organized as follows: Section II presents related work on link scheduling and optimization. We discuss the communication protocol on which EHFS is based in Section III. Section IV formulates system and energy models in data collection. In Section V, we first present the scheduling optimization and constraints. Then we prove that the optimization problem is NP-complete and introduce our suboptimal algorithm. Section VI demonstrates the experiments on Sensor-WPT testbed, and compares the performance of the EHFS algorithm to the state-of-the-art in simulations. Finally, the paper is concluded in Section VII.

II. Related Work

In this section, we review the literature on link scheduling and optimization in wireless networks. To solve different optimization goals, recent work considers throughput, energy consumption or time delay.

Extensive studies have been conducted on link scheduling in cellular networks. In [28], the link quality is predicted by an application framework which tracks the direction of travel of mobile phones at the BS. The authors of [28] develop energy-aware scheduling algorithms for different application workloads such as syncing or streaming. Some scheduling optimizations which consider multicast [29], quality-of-service assurance [30] and fair relaying with multiple antennas [31] are proposed to achieve optimal delay, capacity gain or network utility. The majority of related

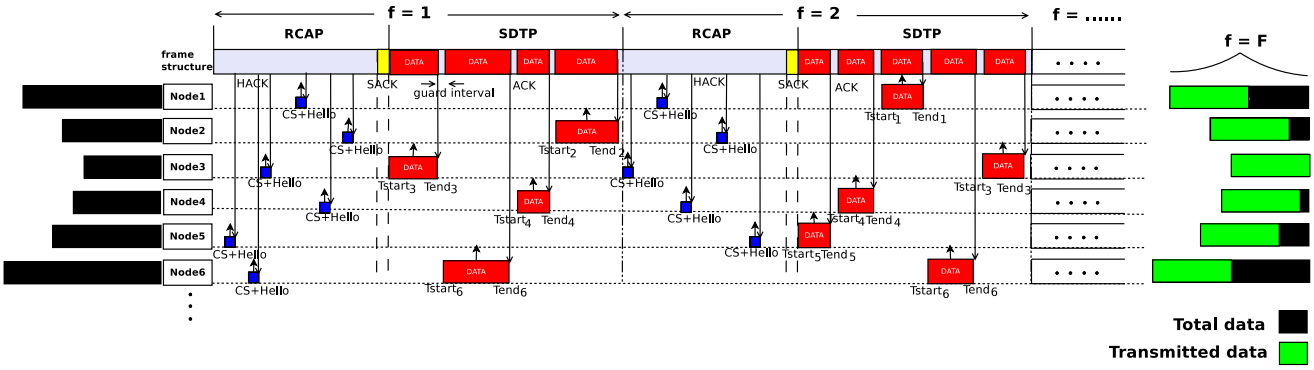


Fig. 2: The timing relationship. $Tstart_i$ and $Tend_i$ stand for the starting and ending time of node i's data transmission respectively. The nodes with different data amounts are scheduled to transmit until $t = F$.

work has focused on addressing the scheduling problem in the context of wireless networks [32]–[35]. However, the notion of fairness in wireless networks focuses on fair allocation, such as channels, tasks among different queues, or time slots among the links in each super frame, which is different from the fairness in data collection of MSN.

A link scheduling for maximum throughput-utility in single-hop networks with the constraint of network delay is presented in [36]. It establishes a delay-based policy for utility optimization. The policy provides deterministic worst-case delay bounds with total throughput-utility guarantee. The author in [37] proposes an opportunistic scheduling algorithm that guarantees a bounded worst case delay in single-hop wireless networks. However, those scheduling algorithms are not applicable in MSNs, because they do not consider the constraints of energy and fairness of collection. In [38], a sensing scheduling among sensor nodes is presented to maximize the overall Quality of Monitoring utility subject to the energy usage. The scheduling algorithm maximizes the overall utility which is to evaluate quality of sensor readings based on the greedy algorithm. For body sensor networks, Sidharth, *et al.* focus on polling-based communication protocols, and address the problem of optimizing the polling schedule to achieve minimal energy consumption and latency [39]. They formulate the problem as a geometric program and solve it by convex optimization.

A recharge model is studied to address energy provisioning problem in WSN [40]. The optimal number of power transmitters is determined to satisfy the energy needs of either static node or mobile node according to a statistical model. Moreover, it estimates a lower bound of the optimal solution and proposes an approximation algorithm to analytically achieve this lower bound. A on-demand mobile charging problem is investigated in [41], which schedules the charging of individual sensor nodes according to their spatial and temporal properties. Nodes send out charging request to mobile chargers when their energy level falls below a certain threshold. By serving a request, the charger moves to the node considering if the new requesting node is of the shortest distance to the mobile charger. The transmit power and the transmission rate of the sensor node are jointly optimized to provide a sustainable and high Quality-of-Service in energy harvesting Body Sensor Networks [42]. The energy harvesting process at each sensor node is formulated as a discrete-time Markov chain, which examines the relationship between data rate and sensor lifetime. In [43], a mobile node with fixed-mobility pattern collects data from the energy harvesting sensor nodes deployed along the trajectory. The network throughput is maximized given that the energy consumption rate of sensor nodes is more than the energy harvesting rate. Moreover, the mobile node determines the maximum number of available time slots allocated to each sensor node. Unfortunately, the optimization model in [43] cannot be applied to the animal monitoring applications since the arrival sequence of the node and the connection time with the BS are unknown.

To the best of our knowledge, there is no research focusing on link scheduling optimization for fair data collection in energy harvesting MSNs. The recent work in the literature is not applicable because they do not optimize the scheduling with the requirements of both energy consumption and data reception fairness. The key difference of our work over previous scheduling optimization is that for a single-hop MSN which includes a large number of energy harvesting nodes, data collection is maximized in a fair way. We formulate the transmission scheduling optimization model in Section V.

III. Communication Protocol

In this section, we present a communication protocol, which extends IEEE 802.15.4 MAC protocol to improve its performance under our specific constraints.

We propose a communication protocol for scheduling optimization in MSN. We utilize a 2-stage communication model, with random channel access period (RCAP) followed by scheduled data transmission period (SDTP) (see Figure 2) [1]. The two periods interchange periodically until all the nodes finish data transmissions.

The purpose of the RCAP is to collect information about sensor nodes, including their current link quality, the amount of available data, and their power resources. This data fits in a single Hello packet and the nodes compete for the channel in a random-access fashion. Nodes check the

radio channel for other data transmissions by using carrier sensing (CS) to avoid packet collisions and the reception of Hello packets is acknowledged by a HACK from the BS, so the nodes can turn off their radios until the end of the RCAP. However, if Hello packets collision happen, the senders have to back off a random time to sense the channel again. Note that a node stops accessing the channel after all its data has been transmitted or a constraint is violated. Consequently, it does not waste energy in RCAP in subsequent super frames. We define φ_i^f as an indicator of node i being active during super frame f during RCAP. If the node i does not compete for the channel in the RCAP of super frame f, φ_i^f is equal to 0.

The BS calculates the transmission schedule at the end of RCAP by running the EHFS algorithm that we illustrate in Section V-B. BS informs all sensor nodes the optimal schedule by broadcasting a SACK packet at the end of the RCAP.

The SDTP is driven by the schedule calculated by the EHFS algorithm. The nodes find their transmission slot (DATA slot) within the super frame and only transmit during their scheduled time to prevent interference. The length of the DATA slots is selected by the scheduler and will typically allow for multiple packet transmissions. We use guard intervals of 2 ms to prevent packet collisions due to time synchronization errors, generally guided by the discussion in [44]. With a large number of nodes, some of them may fail to communicate with the BS during RCAP. However, these nodes consume limited energy due to a long sleeping time during the SDTP.

In addition, an ACK message is used to inform the packets that have been successfully received by the BS. The BS can use one bit to indicate its reception quality of a packet, "1" for success and "0" otherwise. For example, consider a case where the node transmits 40 data packets. The ACK message consists of 5 bytes in total. Due to the small amount of payload, the overhead of this ACK message is negligible. If a packet is not received correctly, the node will retransmit in the future super frame when it is scheduled in SDTP.

IV. System Model

On the basis of Section III, the BS aggregates the nodes and channel information in the RCAP in order to schedule the transmissions. In this section, we explain the basic notations and present an abstract generalizable model of the network, which is used for the optimization model presented in Section V-A. We assume that there are N nodes that can directly communicate with the BS using single-hop communication. Particularly, although the total number of nodes is set to N, the actual number of nodes accessing the channel in RCAP and transmitting data in SDTP in any frame f could be smaller than N since some of the nodes may not arrive at the drinking trough, or residual energy of the node is too low. The residual energy of a node i at the beginning of the first RCAP it participates in is denoted by E_i^0. In order to prevent a node from completely depleting its battery, we assume that a node powers down if the residual energy goes below a certain threshold E_{td}. In this paper, a node in such a state is referred to as a *dead node*. This may happen if any node consumes more energy than it harvests. The wireless channel between each node and the BS is typically influenced by a variety of environmental factors and the transmission noise. The channel variability in turn influences the Packet Reception Rate (PRR) of the node. We estimate the PRR as a function of empirically collected RSSI traces from a real testbed as outlined in Section VI-B.

To illustrate the problem formulation, we list the fundamental variables that have been used in our system model in Table I.

TABLE I: The list of fundamental variables

Notation	Definition
N	total number of sensor nodes
S	total number of slots in SDTP
F	total number of super frames
i	node ID
j	slot sequence number in one super frame
f	frame ID
κ	fairness coefficient
q_i^f	packet reception rate of node i at frame f
γ_f	number of data packets received by the BS in one super frame
α_i^f	data packets collected from node i in frame f
α_i^{total}	total number of data packets collected from node i
λ_i	data payload stored on node i
x_{ij}^f	a binary transmission indicator for node i associated with the slot j in frame f
φ_i^f	an indicator of node i being active during super frame f during RCAP
$\check{E}_A$	energy consumption of a sensor node in RCAP
$\check{E}_{Di}^f$	energy consumption of node i in SDTP of frame f
$\check{E}_{Di}^{total}$	total energy consumption of node i in SDTP
E_{td}	the threshold of node's residual energy
E_i^F	the residual energy of node i when $f = F$
v_{ij}^f	the variable determines slot j in super frame f when node i has finished transmitting all its data payload
e_{tx}	energy consumption of transmitting one data packet
$P_{i,f}^{WPT}$	harvested power of node i in frame f from WPT
$\Delta E_{i,f}$	total harvested energy of node i in frame f

A. Channel Model for Data Transmission

According to the super frame as shown in Figure 2, we divide the SDTP to a number of slots S. Time slot j ($j \in [1, S]$) is allocated by the BS to only one node's transmission for the purpose of avoiding collisions. Let ΔT_i denote the allocated time to node i, where ΔT_i may contain multiple time slots in one super frame. Consequently, we have $S = \sum_{i=1}^{N} \Delta T_i$. EHFS calculates optimal solutions for the nodes in each frame so that the schedule is optimized globally. The sequence number of super frame is denoted as f. We assume the residual energy of node i when it arrives at the data collection centre is $E_i^0 (i \in [1, N])$. The PRR is indicated by q_i^f, where $q_i^f \in [0, 1]$. Additionally,

q_i^f may change from one super frame to the next due to the time-varying channel. We assume q_i^f does not change during the super frame due to block fading. The path loss of the sensor-BS channel can be approximated as free-space path loss [45] and is given by,

$$L(d_i^f) = \frac{(4\pi)^2}{G_{tx}G_{rx}\lambda_0^2}(d_i^f)^\beta, \tag{1}$$

where β indicates the path loss component. d_i^f is the distance between the node i and BS at frame f. G_{tx} and G_{rx} are the antenna gains of the transmitter and receiver, respectively. $\lambda_0 = c/f_0$, which is a ratio of speed of light c and carrier frequency f_0. We define Signal-to-Noise ratio (SNR) for data communication between the node and BS as $\gamma_i^{f\prime}$. Given an additive white Gaussian noise (AWGN) with power N_0,

$$\gamma_i^{f\prime} = \frac{|\hbar|^2 P_i^{tx}}{N_0 L(d_i^f)}, \tag{2}$$

where P_i^{tx} denotes the transmit power of the node i. The small-scale fading is indicated by $\hbar$. Then, the average SNR for the node i is calculated by

$$\overline{\gamma}_i^{f\prime} = \frac{P_i^{tx}}{K_1 N_0 (d_i^f)^\beta}. \tag{3}$$

In this paper, we derive the packet error probability of the channel between the sensor node and the BS based on its outage probability, which provides the lower bound of the packet error probability under an assumption of ideal coding and modulation. For illustration purpose, Rayleigh Block fading is considered [46]. The channel coefficient remains constant within each block, and varies between blocks. At time t, the outage probability at the node i is given by

$$\Pr(\gamma_i^{f\prime} < \gamma_0) = \int_0^{\gamma_0} p(\gamma_i^{f\prime}) d(\gamma_i^{f\prime}) = 1 - \exp(\frac{\gamma_0}{\overline{\gamma}_i^{f\prime}}), \tag{4}$$

where γ_0 is the SNR threshold required for successful reception at the BS.

Substituting Equation (3) into Equation (4), the packet error probability at the BS can be given by

$$\Pr_{i,BS}^f = 1 - \exp(-K_{src} \cdot (d_i^f)^\beta), \tag{5}$$

$$K_{src} = \frac{K_1 N_0 \gamma_0}{P_i^{tx}}. \tag{6}$$

Therefore, the q_i^f can be

$$q_i^f = \exp(-K_{src} \cdot (d_i^f)^\beta). \tag{7}$$

We consider the large-scale path loss which is modeled as free-space propagations in the animal monitoring MSN, as in most cases the sensor node have line-of-sight to the BS. We also consider the small-scale fading which is modeled as independent and identically distributed (i.i.d.) Rayleigh fading. Block fading is assumed on all the wireless links, where the channel gain of a wireless link keeps constant during RCAP and SDTP, but varies between super frames. This assumption is reasonable, because the duration of a frame is typically up to 1000 ms during which the moving distance of animal (e.g., cattle) in the animal pen is negligible.

The data payload stored on each node is represented by λ_i and the fairness coefficient is κ, where $\kappa \in (0, 100\%]$. The data reception fairness ensures that the number of data packets the BS collects from each node is not less than $\kappa \cdot \lambda_i$. Note that nodes do not collect sensor data during data collection, so the data payload of a node is fixed. We define the boolean variable x_{ij}^f as a transmission indicator for node $i \in [1, N]$ associated with the slot $j \in [1, S]$ in the super frame f. Specifically, if $x_{ij}^f = 1$, node i has slot j in SDTP reserved for data transmission in frame f. If $x_{ij}^f = 0$, node i is not scheduled for data transmission at slot j in frame f, hence, it could be possible that the number of scheduled nodes is smaller than N. The number of data packets received by the BS in a super frame is defined as γ_f, where

$$\gamma_f = \sum_{i=1}^{N}\sum_{j=1}^{S} x_{ij}^f \cdot q_i^f. \tag{8}$$

Similarly, for the super frame f, the data received by the BS from any node i is defined as α_i^f, where

$$\alpha_i^f = \sum_{j=1}^{S} x_{ij}^f \cdot q_i^f, (i \in [1, N]). \tag{9}$$

B. Energy Model

The energy consumption of nodes arises from the transmissions in RCAP and SDTP as shown in Figure 2. In this paper, we let $e_{tx-hello}$, $e_{rx-hack}$ and $e_{rx-sack}$ be the energy consumption of transmitting one Hello packet, receiving one HACK and one SACK of the nodes, respectively. The e_{tx} represents energy consumption of transmitting one data packet. Due to the tiny energy consumption of carrier sensing compared to transmitting and receiving packets [47], we neglect the same in our model. The energy consumption of node i in the RCAP is $\check{E}_A$, where

$$\check{E}_A = e_{tx-hello} + e_{rx-hack} + e_{rx-sack}. \tag{10}$$

Therefore, $\sum_{f=1}^{F}(\check{E}_A \cdot \varphi_i^f)$ indicates the energy consumption of the node in the RCAP of all super frames. We next define $\check{E}_{Di}^f$ as the energy that node i consumes on data transmission in the super frame f, where

$$\check{E}_{Di}^f = \sum_{j=1}^{S} x_{ij}^f \cdot e_{tx}, (i \in [1, N]). \tag{11}$$

Note that time slot j is the time to transmit one data packet. Furthermore, we also define $\check{E}_{Di}^{total}$ as total energy consumption on data transmission of node i, where $\check{E}_{Di}^{total} = \sum_{f=1}^{F} \check{E}_{Di}^f$.

For energy harvesting, the node may receive energy input from multiple sources, such as solar, vibration, thermal, or WPT. The total energy input for the node is the sum of energy harvested from these sources over time. In this paper, we focus on two energy harvesting sources, namely,

solar and WPT, and elaborate further on them. The amount of harvested energy from WPT depends on the transmit power, wavelength of the RF signals and the distance between the RF energy source and the harvesting node. We define the transmit power of WPT as P_{tx}^{WPT}. The harvesting power of node i at frame f is $P_{i,f}$. Therefore, the power harvested from the WPT transmitter can be calculated as follows:

$$P_{i,f}^{WPT} = \delta_i(d)\delta_i(\theta)P_{tx}^{WPT}|h_{i,f}|^2, \tag{12}$$

where $\delta_i(d) \in (0,1]$ is a constant indicating WPT efficiency factor given the distance between node i and the charger. The other constant $\delta_i(\theta) \in (0,1]$ denotes WPT efficiency given the antenna alignment between node i and the charger. $h_{i,f}$ is the WPT channel gain between node i and the charger at frame f. Furthermore, we denote the power harvested from solar panel as $P_{i,f}^{solar}$.

Given the time of WPT is τ_i and the solar charging duration is τ_i', the harvested energy of sensor i is given by

$$\Delta E_{i,f} = \delta_i(d)\delta_i(\theta)(P_{tx}^{WPT}\tau_i)|h_{i,f}|^2 + P_{i,f}^{solar}\tau_i'. \tag{13}$$

Let E_i^f denote the residual energy of node i at frame f, and we have

$$E_i^f = E_i^0 - \sum_{f'=1}^{f}(\check{E}_A \cdot \varphi_i^{f'}) + \sum_{f'=1}^{f}\Delta E_{i,f'} - \sum_{f'=1}^{f}\sum_{j=1}^{S} x_{ij}^{f'} \cdot e_{tx}. \tag{14}$$

V. Fair Scheduling with Energy Harvesting

In this section, we first formulate fair scheduling optimization under the constraints of fairness and energy harvesting. We show that the optimization problem is NP-complete. Next, a heuristic algorithm, EHFS is proposed to approximate the optimal solution.

A. Optimization Formulation

Let F denote the total number of super frames needed for all N nodes to finish their data transmissions. The goal of our formulation is to schedule the unique node i to transmit at slot j of each super frame f, which achieves the maximized data reception among the total frames F given the fairness of the nodes. Therefore, the node with the highest PRR in super frame f and the fastest energy drainage speed, i.e., the lowest residual energy in f, is prioritized to transmit in f. Based on the notations in the problem formulation, we formulate the EHFS for finding the optimal schedules as follows. The objective function of the optimization model is to maximize γ_f of all F super frames, which is presented as $maximize \sum_{f=1}^{F}\gamma_f$.

- Constraint ($E_i^F \geq E_{td}$) specifies the minimum remaining energy to be above E_{td}, where E_i^F is given in (14) for $f = F$.
- Constraint ($\alpha_i^{total} \geq \kappa \cdot \lambda_i$) guarantees that the BS receives sufficient data packets to meet the fairness requirement, where $\alpha_i^{total} = \sum_{f=1}^{F}\alpha_i^f$.
- Moreover, it is reasonable that α_i^{total} with any given x_{ij}^f needs to be less than or equal to the total data payload of the node, i.e., λ_i. Therefore, constraint ($\alpha_i^{total} \leq \lambda_i$) limits the value of α_i^{total} by λ_i.
- Constraints ($x_{ij}^f \leq 1$) and ($\sum_{i=1}^{N} x_{ij}^f \leq 1$) specify that at any data transmission time slot only one node communicates with the BS to prevent transmission collisions.

The only unknown is the total number of super frames during which a node is required to transmit. In other words, φ_i^f is not known. To determine φ_i^f, we define a variable v_{ij}^f for node i at any slot j of super frame f. Specifically, v_{ij}^f determines slot j in super frame f when node i has finished transmitting all its data payload (i.e., λ_i), hence, the node will not compete for the channel and transmit data in the future slots and frames. Accordingly, constraint $\lambda_i - (\sum_{g=1}^{f-1}\sum_{w=1}^{S} x_{iw}^g q_i^g + \sum_{w=1}^{j} x_{iw}^f q_i^f) \geq v_{ij}^f$ specifies whether node i has stopped its data transmission at slot j in frame f or not. $\sum_{g=1}^{f-1}\sum_{w=1}^{S} x_{iw}^g q_i^g + \sum_{w=1}^{j} x_{iw}^f q_i^f$ is the total number of received packets until the current slot j in frame f. If the amount of data packets received from node i matches the payload size λ_i, i.e., $\lambda_i = \sum_{g=1}^{f-1}\sum_{w=1}^{S} x_{iw}^g q_i^g + \sum_{w=1}^{j} x_{iw}^f q_i^f$, then v_{ij}^f is equal to 0. As a result, node i will stop channel access and data transmission in the future slots and frames. Otherwise, if node i has not finished all data transmission in f, v_{ij}^f is not fixed regarding (20). Thus, x_{ij}^f is not restrained according to (24). Namely, x_{ij}^f is optimized subject to (15) to (19).

- Constraints ($v_{ij}^f \geq v_{ij'}^f$) and ($v_{ij}^f \geq v_{ij'}^g$) ensure the future slots j' and super frames g have $v_{ij}^f = 0$ if λ_i packets have been received from node i.
- Constraint ($\sum_{a=1}^{F-f} \varphi_i^{f+a} \leq v_{ij}^f$) guarantees all φ_i^f of the future super frames is 0 if $v_{ij}^f = 0$. As a result, the remaining energy of node i which is restricted by the RCAP indicator φ_i^f stops decreasing.
- Constraint ($x_{ij}^f \leq \varphi_i^f$) ensures that the node i stops data transmission if $\varphi_i^f = 0$.

Then, the formulation of the problem is presented from constraints (15) to (24).

B. EHFS Algorithm

Maximizing the collected data presented in Section V-A is a typical 0-1 Multiple Knapsack Problem (MKP) [48]. We reduce an instance of an MKP to our scheduling optimization problem by assigning ΔT_i to each knapsack. Therefore, the capacity of the knapsack is equal to ΔT_i. The items to be put in knapsacks are data packets whose size is prorated by q_i^f. The parameters of the energy and fairness conditions (constraint (15) and (16)) are chosen so that they are satisfied by any placement of items. In this way, optimal placement of items in knapsacks is reduced to such an instance of our scheduling problem. Since the optimal placement of items in knapsacks is an NP problem, this shows that our scheduling optimization problem presented in Section V-A is NP-complete.

$$maximize \quad \sum_{f=1}^{F} \gamma_f$$

$$subject\ to: \quad E_i^F \geq E_{td}, \quad (i \in [1, N]) \tag{15}$$

$$\alpha_i^{total} \geq \kappa \cdot \lambda_i, \quad (i \in [1, N],\ \kappa \in (0, 1]) \tag{16}$$

$$\alpha_i^{total} \leq \lambda_i, \quad (i \in [1, N]) \tag{17}$$

$$x_{ij}^f \leq 1, \quad (i \in [1, N],\ j \in [1, S],\ f \in [1, F]) \tag{18}$$

$$\sum_{i=1}^{N} x_{ij}^f \leq 1, \quad (j \in [1, S],\ f \in [1, F]) \tag{19}$$

$$\lambda_i - (\sum_{g=1}^{f-1} \sum_{w=1}^{S} x_{iw}^g q_i^g + \sum_{w=1}^{j} x_{iw}^f q_i^f) \geq v_{ij}^f, \quad (i \in [1, N],\ j \in [1, S],\ f \in [1, F]) \tag{20}$$

$$v_{ij}^f \geq v_{ij'}^f, \quad (j' \geq j,\ j \in [1, S]) \tag{21}$$

$$v_{ij}^f \geq v_{ij'}^g, \quad (g \geq f, j' \in [1, S], j \in [1, S],\ f \in [1, F]) \tag{22}$$

$$\sum_{a=1}^{F-f} \varphi_i^{f+a} \leq v_{ij}^f, \quad (i \in [1, N],\ j \in [1, S]) \tag{23}$$

$$x_{ij}^f \leq \varphi_i^f, \quad (i \in [1, N],\ j \in [1, S],\ f \in [1, F]) \tag{24}$$

The optimal solutions are inapplicable to practical environments due to the high computational complexity and an assumption of known energy and PRR in all F super frames. Therefore, we propose a EHFS algorithm to approximate the optimal solution. Due to the prominent effect of energy harvesting and link quality variation on the scheduling, we consider the ration η_i^f that is

$$\eta_i^f = \frac{q_i^f}{E_i^f}, \forall i \in [1, N], \forall f \in [1, F]. \tag{25}$$

The motivation of calculating η_i^f is to prioritize the nodes based on both the link quality and harvested energy. The EHFS algorithm gives a high transmission priority to the node with larger η_i^f. Note that the BS is aware of the residual energy and data queue length of the nodes that contend channel in RCAP of each super frame regarding to the proposed communication protocol (in Section III). Therefore, η_i^f is able to be calculated on a frame-by-frame basis. This method achieves large data reception because for the nodes with the same q_i^f, the node with the smallest E_i^f gets higher transmitting priority. Similarly, for the nodes with the same E_i^f, one with higher q_i^f has higher priority. Furthermore, recent work [49] also reveals the prediction of mobility and duration in autonomous long-term wildlife tracking that relies on energy harvesting. The optimal scheduling solutions are aware proactively with the implementation of online individual- or population-based prediction models.

In our algorithm, the node works in three states, Access & Data transmission (AD), NonAccess (NA) and NonData (ND). In AD state, the node competes for the channel in RCAP and eventually transmits data in SDTP as shown in Figure 2. In NA state, the node neither accesses the channel in RCAP nor transmits data in SDTP but only receives the SACK packets for the purpose of saving energy in the super frame. More importantly, none of the nodes, which are in the NA state transmit data given that no time slots are allocated to them. The benefit of NA state is to reduce the channel competition since the number of nodes competing for the channel is decreased, which further helps more nodes achieve fairness. In ND state, the node does not turn on the radio and remains in sleep mode. Note that no matter which state the node works in it still harvests energy from WPT transmitter and solar power.

The EHFS algorithm develops two steps to maximize the data reception with η_i^f. It is implemented as shown in Algorithm 1.

Initially, all nodes are in AD state and the BS schedules the node i ($i \in [1, N]$) which has maximum η_i^f to transmit data. The BS records the number of data packets from the node. Once the node i meets the fairness of data reception (constraint (16)), it transfers to the NA state. Certainly, after the first step, all the nodes have at least $\kappa \cdot \lambda_i$ data packets being transmitted successfully and the fair reception of data is achieved. At the second step, all the nodes change the state from NA to AD. Then, the BS schedules the node with largest η_i^f to transmit first. To maximize data reception, node i remains in AD state until either ($E_i^f \geq E_{td}$) or ($\sum_{f'=1}^{f} \alpha_i^{f'} < \lambda_i$) no longer holds. Moreover, if either constraint is not fulfilled by the node i, it transitions to ND state. By using this approach, the number of data packets collected by the BS is maximized, meanwhile, the energy and fairness requirements are both achieved.

Compared to the optimization model, the proposed EHFS

Algorithm 1 EHFS Algorithm

nodes are in AD state and compete for the channel
The BS calculates η_i^f for the node i
The BS sorts the nodes by η_i^f, then $\eta_i^f \geq \eta_{i'}^f, (i \neq i', i' \in [1, N])$
The BS schedules the node i to transmit
if $\sum_{f'=1}^{f} \alpha_i^{f'} \geq (\kappa \cdot \lambda_i)$ **then**
  The node i goes to NA state
  The BS schedules the next one to transmit
else
  The node i remains in AD state
end if
if every node has $\sum_{f'=1}^{f} \alpha_i^{f'} \geq (\kappa \cdot \lambda_i) \quad \forall i \in [1, N]$ **then**
  All the nodes transfer to AD state
  The BS calculates η_i^f for each node
  The BS sorts the nodes by η_i^f, then $\eta_i^f \geq \eta_{i'}^f, (i \neq i', i' \in [1, N])$
  if $E_i^f \geq E_{td}$ **then**
    The BS schedules the node i to transmit
  else
    The node i changes state to the ND
    The BS schedules the next one to transmit
  end if
  if $\sum_{f'=1}^{f} \alpha_i^{f'} < \lambda_i$ **then**
    The node i remains in AD state
  else
    The node i changes state to the ND
  end if
end if

algorithm is practical with a significantly low computational complexity. Specifically, the calculation of η_i^f in Equation (25) requires 1 division at one super frame, which has a complexity of $O(N)$. Sorting η_i^f in super frame f has linear-time complexity in the number of nodes. The only communication overhead of EHFS algorithm is caused by receiving the SACK packet, compared to the conventional scheduling protocols, e.g., batch processing or distributed scheduling algorithm. The BS uses one bit to indicate whether the node transmits in the specific time slot in SDTP, "1" for transmission and "0" otherwise. Moreover, to notify the working state of the node, the BS uses two bits to present AD, ND and NA states. Consider a case where 300 sensor nodes send Hello packet in RCAP, and accordingly there are 300 time slots in SDTP. Assume no Hello packet is lost, the SACK packet that is broadcasted by the BS to 300 sensor nodes consists of 112.5 bytes in total. For example, consider a case where SDTP has 800 ms and data rate of the node is 250 kbps, one super frame allows the nodes to transmit 25 KB of data payload. Thus, the overhead caused by SACK is less than 1% of the link budget. Moreover, the EHFS algorithm ensures only the nodes with a high η_i^f are scheduled to transmit, which achieves a high data packet reception rate. However, the conventional scheduling protocols consume more energy than EHFS algorithm due to packet collision, which makes the other nodes compete for the channel in RCAP repeatedly and cost energy.

VI. Performance Evaluation

Experiments are conducted on our Sensor-WPT testbed to measure the WPT efficiency as a fusion of distance between WPT transmitter and receiver and their antenna orientation. We then compare the performance of the optimal algorithm that defines AMPL scheduler to our EHFS algorithm that is carried out by Matlab. We utilize empirical link quality, solar and WPT energy harvesting to evaluate the proposed algorithm in extensive simulations and compare its performance to state-of-the-art.

A. Experiments on Sensor-WPT Testbed

We design two experiments to characterize the WPT efficiency factors $\delta_i(d)$ and $\delta_i(\theta)$, on our Sensor-WPT testbed.

In Sensor-WPT, each sensor node is equipped with a rechargeable battery and consumes energy on sensing and data transmission activities. A Powercast [50] wireless charger transmits power to the sensor nodes by WPT. Since WPT charging is carried out in the 915 MHz band while sensor nodes communicate in the 2.4 GHz band, our network achieves simultaneous wireless information and power transfer. The isotropic radiated power of WPT transmitter with 8 dBi integrated antenna gain is 3 W. The sensor node is connected to a P2110 powerharvesting board with a 1 dBi omni-directional antenna [51]. The hardware setup is shown in Figure 3.

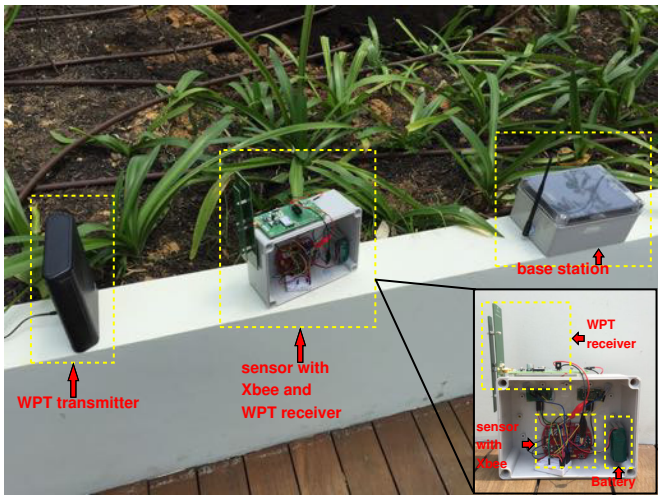


Fig. 3: The hardware setup contains the BS, WPT transmitter and sensor node with RFBee transceiver (for data transmission in 2.4 GHz) and WPT receiver.

In the first experiment, we measure the WPT efficiency factor on distance, $\delta_i(d)$ in Equation (13). As shown in Figure 4, the effective amount of power that can be captured by a sensor node varies with the distance between the node and the BS. We calculated the average value of the received power by WPT in mW, and standard deviation over 250 packets for every distance measured. Moreover, it is also obtained that the received power is inversely proportional to the distance, where the variance is lower than 5 mW. Due to radiation exposure protection, the distance between WPT transmitter and P2110 powerharvesting board has to be further than 20 cm.

In the second experiment, we vary WPT receiver antenna orientation in order to configure $\delta_i(\theta)$ in Equation (13). The distance between WPT transmitter and the sensor node is fixed at 55 cm. Initially, the antenna of WPT receiver on the node and WPT transmitter directly face towards each other. Therefore, the initial orientation is denoted as Zero degree rotation. The orientation increases 45 degrees every 1000 seconds, and the sensor node records 1000 samples at each orientation. The sensor node logs the sequence numbers and RSSI values of received packets in their flash. Figure 5 shows that antenna orientation affects the received power at the WPT receiver. When the antenna orientation is at 90 and 270 degrees (two antennas are orthogonal to each other), the node harvests the lowest energy from the WPT transmitter.

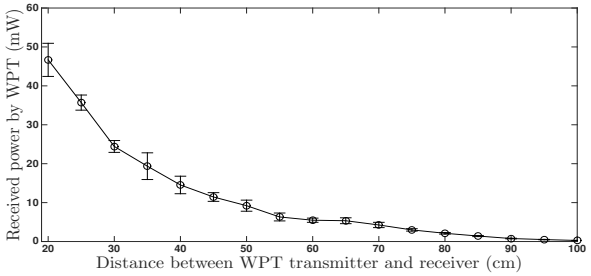


Fig. 4: Received power at the sensor by WPT. The error bars show the standard deviation over 250 packets.

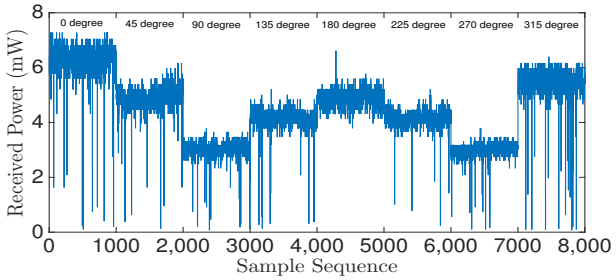


Fig. 5: Received power on the sensor node for varying antenna orientation of the WPT receiver.

Based on the two experiments, we observe that the WPT efficiency is jointly affected by distance between WPT transmitter and receiver, and their antenna orientation. It also indicates that the mobility of the nodes causes unbalanced harvested energy, which results in different η_i^f and further effects the transmission scheduling. Additionally, the parameter $\delta_i(d)$ is given based on the inversely proportional relationship shown in Figure 4, and $\delta_i(\theta)$ is given according to the variation shown in Figure 5. Finally, they are imported to our simulation configuration in the next Section, and we also present their effect in Section VI-D.

B. Simulation Parameters

The data collection network in the simulation utilizes a star topology that contains one BS and N nodes ($N \in [10, 300]$). The BS is placed at the centre of the network so that its radio coverage is able to cover the monitoring area. N nodes are randomly distributed within the open data collection centre. The node communicates with the BS using CC2420 radio in 2.4 GHz. The working temperature is measured as 25°C, therefore, V_{cc}, I_{tx} and I_{rx} is 3 V, 35 mA and 15 mA, respectively [52]. We configure the remaining energy threshold of the sensor, E_{td} to 1.67 mJ.

In our simulator, the length of a super frame is set to 1000 ms, which contains 1000 sub-slots and they are equally divided for RCAP and SDTP. In addition, the number of nodes scheduled transmitting during SDTP is related to the duration of RCAP. Therefore, the duration of RCAP and SDTP can be configured depending on the network scale and requirement of specific applications [22].

Payload of the data packet is 32 bytes. Hello and HACK have the same length of 10 bytes. For our configuration of 26 nodes in the network, the length of SACK packet is 10 bytes. Therefore, we have

$$e_{tx-hello} = V_{cc} \cdot I_{tx} \cdot \frac{10 \times 8}{R_b} = 0.03 \text{ mJ}, \quad (26)$$

$$e_{rx-hack} = e_{rx-sack} = V_{cc} \cdot I_{rx} \cdot \frac{10 \times 8}{R_b} = 0.01 \text{ mJ}, \quad (27)$$

$$e_{tx} = V_{cc} \cdot I_{tx} \cdot \frac{32 \times 8}{R_b} = 0.1 \text{ mJ}, \quad (28)$$

where R_b = 250 kbps. E_i^0 is given by a normal distribution with the mean value of 50 Joules according to the battery capacity of our sensors. The solar charging energy in the simulation makes use of the Camazotz node, which has been developed for wildlife tracking [53]. Camazotz reduces data sampling rate when the solar charge power is low. Figure 6 shows the harvested energy of the two nodes over 43 hours on wild flying foxes. It is observed that solar energy on the different nodes is dynamic due to the mobility of nodes, weather, and landscape. Specifically, the empirical solar energy data (i.e., the power harvested from solar panel at different time) are presented in chronological order, which can be utilized to provide $P_{i,f}^{solar}$ and τ_i' in the energy model (shown in Section IV-B) of our simulator.

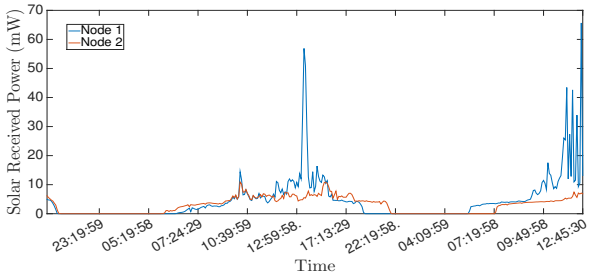


Fig. 6: The solar energy of two Camazotz nodes over 43 hours.

The WPT efficiency parameters, $\delta_i(\theta)$ and $\delta_i(d)$ are 0.5. Additionally, in our simulations, the value of E_i^0, $\delta_i(\theta)$ and $\delta_i(d)$ are given on purpose so that some dead nodes which run out of energy can be observed among different

scheduling algorithms. A random number of nodes are selected to harvest energy from the WPT charger in each super frame, which indicates that the selected animal stays in the charging range of WPT. Moreover, the received power of the node is obtained by Equation (12), which depends on its position and antenna orientation.

In modeling node motion, we adopt a random walk that is a basic mobility model in WSN to generate the simulation scenarios, while our algorithm developed in Section V is general and can support other mobility models. Specifically, a series of waypoints is provided to the node at each super frame, which is represented by the distance to the BS, d_i^f. Therefore, the link quality variation caused by the node's mobility over time is obtained by Equation (7).

Furthermore, the RSSI trace recorded by the sensor in our testbed (shown in Figure 3) is set as the initial link quality of the node, i.e., $q_i^f(f = 1)$. Importing empirical data to our simulator provides an environment to conduct repeatable simulations. In this paper, we convert the RSSI to PRR for the q_i^f by the experimental results of PRR-RSSI relationship [54].

C. Scenarios and Metrics

We simulate the EHFS algorithm in Node On Pasture (NOP) scenario and Node Arriving Pasture (NAP) scenario. In NOP scenario, we assume all the nodes are in the monitoring area from the start of experiment to the end. In NAP scenario, the nodes arrive at the area at different times following a Poisson distribution where the mean arrival rate is 1 node per second. The nodes remain there for an extended period of time. We evaluate three performance metrics: number of data packets received by the BS (data reception), the number of fair nodes and the number of dead nodes. Specifically, the *fair node* denotes the node which fulfills $\alpha_i^{total} \geq \kappa \cdot \lambda_i$ (Constraint (16)). We compare the performance of our EHFS algorithm with optimal solution at first. In NOP scenario, each node carries 80 KB data which is the payload generated by the sensor node. Since the number of nodes communicating with the BS in a short time is small in NAP scenario, we increase the data payload to 300 KB in order to explore the limits of the scheduling algorithms. For this reason, a node occupies the channel longer while more nodes enter the area in NAP scenario.

To evaluate the performance of the EHFS algorithm in the NOP and NAP scenarios, two Greedy scheduling algorithms and FCFS algorithm are constructed in the numerical investigations. Because two basic elements used in the EHFS are the remaining energy represented by E_i^f and link quality q_i^f of node, the Greedy scheduling algorithms are formulated by them. The first Greedy algorithm is called Low Energy (LE) scheduling, namely, the transmission schedule is based solely on the E_i^f of node. Lower E_i^f implies higher priority of transmission at super frame f. High PRR (HP) scheduling is the second algorithm where the node with higher q_i^f has higher priority. We compare them with the EHFS algorithm with κ = 10%, 50% and 90%.

D. Simulation Results

Next, we perform five sets of independent simulations as follows.

1) Comparing to Benchmark: To compare to the optimal schedule shown in Section V-A, we assess the performance of our algorithm when it operates in ten small-scale networks where the number of nodes is increased from 1 to 10. This initial comparison makes us aware of the performance difference between the optimal solution and our algorithm. The node i carries 80 KB data, so $\lambda_i = 2500$. In fact, the comparison is not affected by different κ values, thus we choose $\kappa = 50\%$ for both the optimal schedules and the EHFS algorithm. The optimal schedules achieve a maximum number of received data packets with the fairness and remaining energy constraints. They are constructed using AMPL and a state of the art ILP solver, Cplex 12.5, in a 2.7 GHz Intel core processor with 8 GB of memory.

Table II summarizes running time, and the number of collected data packets. It is also found that there is no dead node in all tests and our algorithm guarantees exactly the same number of fair nodes as optimal schedules. On data reception, the EHFS algorithm and optimal solution have the maximum difference which is 706 when N = 9. On average, the number of packets in our algorithm is less than the AMPL output by around 1.16%. Interestingly, AMPL outperforms EHFS algorithm even in the case of N = 1. We attribute this fact to the randomized PRR of network links in the AMPL and EHFS simulations, which cause a small variance in the number of received packets overall. Moreover, our algorithm is much more efficient than the optimization model on runtime.

2) Node On Pasture Scenario: Figures 7 and 8 show the performance of the aforementioned four scheduling algorithms on the data reception and fairness. When there are only 10 nodes in the network, they have pretty similar performance. However, FCFS, LE and HP collect 75.6%, 45.7% and 41.3% less data packets than our algorithm when N = 300. With WPT energy harvesting, it is observed that more data packets are collected with more nodes. The number of fair nodes of our algorithm is more than the ones of FCFS, LE and HP for 200, 180, 155 nodes when κ = 50% and N = 300. The reason is that LE scheduling fails when the low energy nodes have poor link quality. The nodes with high PRR are not scheduled, however, they still consume energy on channel competitions in RCAP. For HP scheduling, the nodes with high PRR occupy the SDTP for multiple super frames until they finish the transmissions. This leads to a large number of dead nodes. However, those nodes could have potentially gained higher data reception. In contrast, our algorithm makes the schedule based on η_i^f which considers both remaining energy and link quality. Moreover, it also achieves the fairness of data collection.

We find the data reception and fair nodes of FCFS, LE and HP do not vary significantly from N = 150 to 300. The reason is indicated by dead nodes which are shown in Figure 9. It shows FCFS, LE and HP have much more dead nodes than the EHFS algorithm starting from N = 50.

TABLE II: Comparison between the optimal solutions and the EHFS algorithm

Nodes	AMPL (Cplex)		EHFS		Difference	
	Packets	Runtime	Packets	Runtime	Packets	Runtime
1	2499	1 s	2491	0.07 s	8	0.93 s
2	4999	12 s	4981	0.1 s	18	11.9 s
3	7499	28 s	7475	0.04 s	24	27.96 s
4	9998	63 s	9954	0.06 s	44	62.94 s
5	12498	1 m 27 s	12484	0.06 s	14	1 m 26.94 s
6	14998	5 m 15 s	14465	0.06 s	533	5 m 14.94 s
7	17498	1 h 3 m	17353	0.08 s	145	1 h 3 m
8	19997	6 h 53 m	19808	0.08 s	189	1 h 53 m
9	22499	19 h 12 m	21793	0.22 s	706	19 h 12 m
10	24998	36 h 29 m	24583	0.22 s	415	36 h 29 m

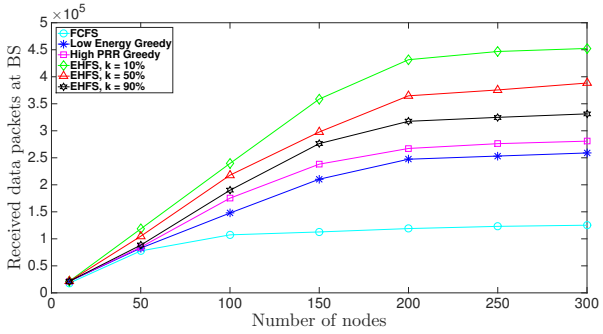


Fig. 7: Data packets collected by the BS, N is from 10 to 300.

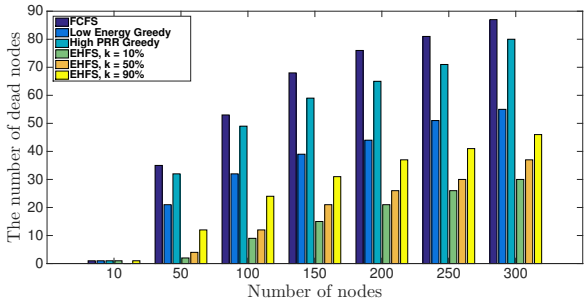


Fig. 9: Number of dead nodes among N, N is from 10 to 300.

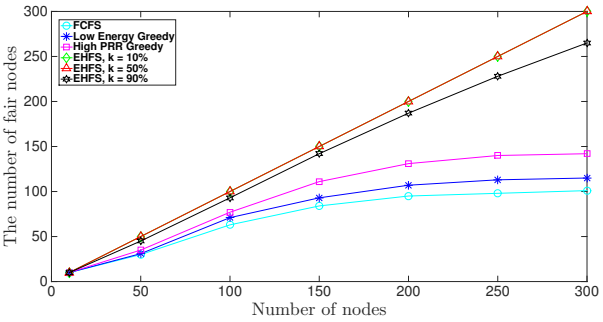


Fig. 8: Number of fair nodes among N, N is from 10 to 300.

According to the EHFS algorithm, we know that κ is a crucial variable which affects the states transition of node i. The performance of our algorithm varies with different κ value. As shown, they are similar for κ = 10%, 50% and 90% when N is 10. From N = 50 to N = 300, κ = 10% performs better than 50% and 90%. The reason is that any node which is scheduled to transmit occupies more super frames when κ is increased due to the fairness constraint (16). It makes the other nodes compete for the channel in RCAP repeatedly and cost energy. However, increasing κ achieves more data collected from the single node, which benefits some application for individual sensor monitoring. Therefore, the configuration of κ depends on the application requirement.

3) Fairness Parameter Effect: Based on the preceding simulations, it is observed that different κ affects the performance of our algorithm. Essentially, the κ decides the fairness level in EHFS. In this experiment, we analyze the impact of κ in the NOP scenario with 300 nodes. Specifically, the κ is varied from 10% to 100%. The performance of data reception, fair nodes and dead nodes are shown in Figure 10.

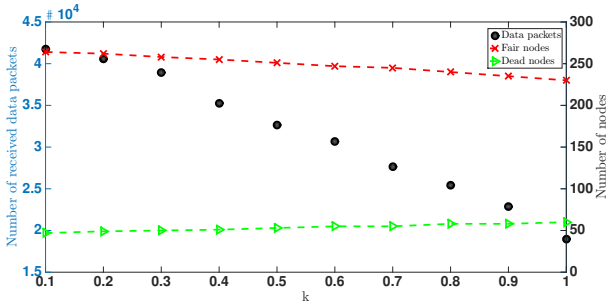


Fig. 10: The effect of κ on the performance of EHFS algorithm. N = 300 and κ is from 10% to 100%.

As shown in Figure 10, data reception rate decreases and the number of dead nodes slowly increases with the increasing κ. This is because the transmission duration of one node is extended when κ is increased. Other nodes with small harvested energy $\Delta E_{i,f}$ deplete their energy due to RCAP if the channel is occupied by someone with high η_i^f for a long time. Their data is not collected by the BS before the nodes exhaust the energy. As observed, energy harvesting can only retard this energy depletion instead of addressing it thoroughly since the charging efficiency is affected by the environmental factors. We also find that the

scheduling with smaller κ achieves larger number of fair nodes.

Therefore, Figure 10 indicates a tradeoff, namely, higher κ guarantees more data packets collected from individual nodes while sacrificing the system throughput; smaller κ achieves a higher system throughput, however, it does not guarantee most of data can be collected from individual node since the BS gives the priority to the one with larger η_i^f after all nodes satisfy the fairness constraint.

4) Node Arriving Pasture Scenario: In this set of experiments, we test the scheduling algorithms when nodes arrive at the data collecting point with a specific arrival rate. We assume the inter-arrival time of nodes is exponentially distributed which is typically used to model situations involving the random time between arrivals to a service facility [55].

From Figure 11 we find that the EHFS algorithm has up to 2.3 times as many collected data packets as FCFS. It outperforms LE and HP by nearly 1.7 times as well. The reason is the newly arrived nodes fail to transmit since the transmitting node have not finished the transmission due to retransmissions. From Figure 12, we observe the difference of fairness which is achieved by different κ is smaller than the one in NOP scenario. That is because the BS schedules a small number of nodes in one super frame in NAP scenario. The first step of EHFS algorithm is completed faster, hence more nodes achieve fairness in NAP scenario. Likewise, the number of dead nodes in our algorithm has small difference in Figure 13. Due to the increase of λ_i in this application, there are 12 dead nodes with the κ = 90% in our algorithm at the maximum. Moreover, in Figure 13, the FCFS, LE and HP also have smaller dead nodes compared with the NOP scenario. The reason is that a small number of nodes is scheduled to transmit at one super frame and they can finish 300 KB data transmission soon. So the newly arrived nodes have small channel competition.

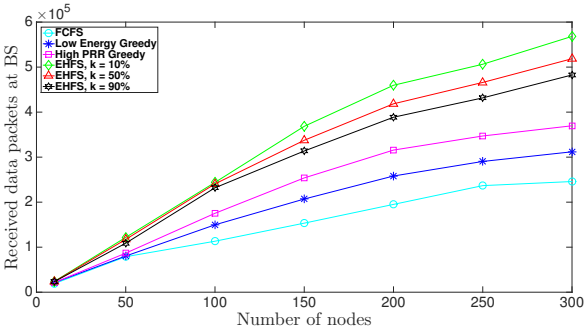


Fig. 11: Data packets collected by the BS. N is from 10 to 300.

5) WPT Efficiency Effect: According to Equation (13) and the experiments in Section VI-A, it is observed that $\delta_i(\theta)$ and $\delta_i(d)$ jointly affect harvested energy $\Delta E_{i,f}$ of the sensor node and the performance of scheduling algorithm. Figure 14 illustrates the impact of the WPT efficiency on the data packets reception of EHFS given that the number of nodes is 50 and κ is 50%. Data reception increases by

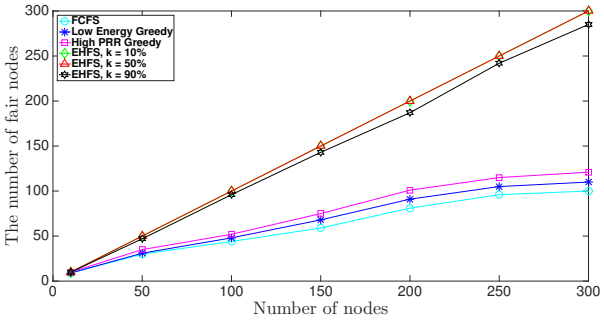


Fig. 12: Number of fair nodes among N, N is from 10 to 300.

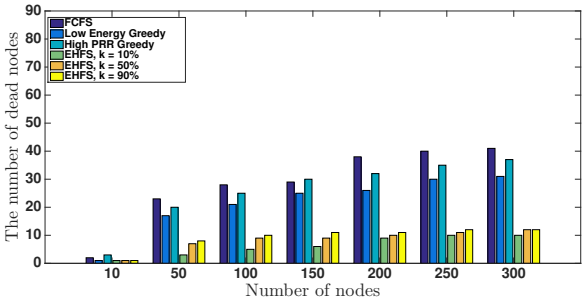


Fig. 13: Number of dead nodes among N, N is from 10 to 300.

increasing $\delta_i(\theta)$ and $\delta_i(d)$ since the nodes harvest more energy via WPT. Specifically, when the nodes are close to the WPT transmitter ($\delta_i(d) = 1$) with WPT receiver antenna alignment ($\delta_i(\theta) = 1$), the data reception has the maximum value which is about 56250 packets. Even in the worst case ($\delta_i(d), \delta_i(\theta) = 0.1$), EHFS algorithm can still achieve the reception of 3251 packets.

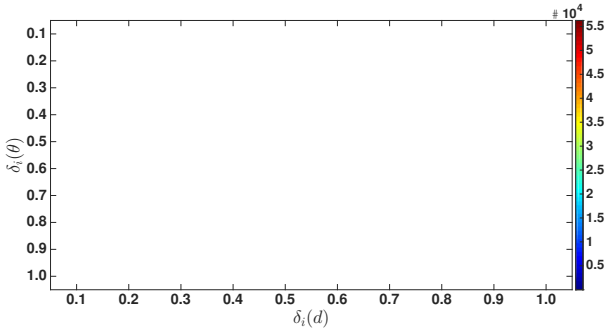


Fig. 14: Data packets reception according to WPT antenna orientation $\delta_i(\theta)$ and distance $\delta_i(d)$. $N = 50$ and $\kappa = 50\%$.

VII. Conclusion and Discussion

In this paper, we have proposed and evaluated a fair link scheduling optimization model with the objective of maximizing the data reception in the data collection of energy harvesting MSN. The super frame structure is developed for the BS to collect data from the sensor nodes. We have

proved that the scheduling optimization is an NP-complete problem. Therefore, the EHFS algorithm is proposed to approximate the optimal solutions in polynomial time. Our algorithm schedules the transmissions of the nodes based on η_i^f and three working states in two steps. With the animal monitoring application and our Sensor-WPT testbed, we have shown the numerical performance of the EHFS algorithm based on the solar energy, WPT charging efficiency, and RSSI. We have compared our algorithm with the optimal schedules of the optimization model and presented extensive simulations incorporating both node on pasture and node arriving pasture scenarios. Specifically, the EHFS algorithm provides a near-optimal scheduling to the data collection in the energy harvesting MSN.

Especially, a single-hop star topology is considered in our network. However, the proposed problem formulation and EHFS algorithm can be further extended to a large-scale MSN with multiple base stations, multiple clusters or small cells [5]. Note that the application only needs to decide which nodes are selected to be scheduled. In addition, the overhead of scheduling acknowledgement messages, i.e., SACK, can be further reduced by optimizing the number of bits for the working states based on the schedule.

ACKNOWLEDGEMENT

This research is funded by National Funds through FCT (Portuguese Foundation for Science and Technology), within the CISTER Research Unit (CEC/04234); the Australian Research Council Discovery Grant DP110104344; NSFC 61750110529. The authors would like to thank the editors and the anonymous reviewers for their constructive comments on the article.

REFERENCES

[1] K. Li, B. Kusy, R. Jurdak, A. Ignjatovic, S. S. Kanhere, and S. Jha, "κ-FSOM: Fair link scheduling optimization for energy-aware data collection in mobile sensor networks." in *EWSN*, 2014, pp. 17–33.

[2] K. Liu, Q. Ma, X. Zhao, and Y. Liu, "Self-diagnosis for large scale wireless sensor networks," in *IEEE INFOCOM*, 2011, pp. 1539–1547.

[3] K. Dantu, B. Kate, J. Waterman, P. Bailis, and M. Welsh, "Programming micro-aerial vehicle swarms with karma," in *ACM SenSys*, 2011, pp. 121–134.

[4] A. Sinha, A. Tsourdos, and B. White, "Multi UAV coordination for tracking the dispersion of a contaminant cloud in an urban region," *European Journal of Control*, vol. 15, no. 3, pp. 441–448, 2009.

[5] P. Sommer, B. Kusy, P. Valencia, R. Dungavell, and R. Jurdak, "Delay-tolerant networking for long-term animal tracking," *IEEE Internet Computing*, vol. 22, no. 1, pp. 62–72, 2018.

[6] V. Dyo, S. A. Ellwood, D. W. Macdonald, A. Markham, N. Trigoni, R. Wohlers, C. Mascolo, B. Pásztor, S. Scellato, and K. Yousef, "WILDSENSING: design and deployment of a sustainable sensor network for wildlife monitoring," *ACM Transactions on Sensor Networks (TOSN)*, vol. 8, no. 4, p. 29, 2012.

[7] M. Israel, "A UAV-based roe deer fawn detection system," in *Proceedings of the International Conference on Unmanned Aerial Vehicle in Geomatics (UAV-g)(H. Eisenbeiss, M. Kunz, and H. Ingensand, eds.)*, vol. 38, 2011, pp. 1–5.

[8] R. Jurdak, P. Corke, D. Dharman, and G. Salagnac, "Adaptive GPS duty cycling and radio ranging for energy-efficient localization," in *ACM SenSys*, 2010, pp. 57–70.

[9] V. Dyo, S. A. Ellwood, D. W. Macdonald, A. Markham, C. Mascolo, B. Pásztor, S. Scellato, N. Trigoni, R. Wohlers, and K. Yousef, "Evolution and sustainability of a wildlife monitoring sensor network," *ACM SenSys*, pp. 127–140, 2010.

[10] P. Sommer, B. Kusy, and R. Jurdak, "Power management for long-term sensing applications with energy harvesting," in *Proceedings of the 1st International Workshop on Energy Neutral Sensing Systems*. ACM, 2013, p. 3.

[11] L. Xie, Y. Shi, Y. T. Hou, and H. D. Sherali, "Making sensor networks immortal: An energy-renewal approach with wireless power transfer," *IEEE/ACM Transactions on Networking (TON)*, vol. 20, no. 6, pp. 1748–1761, 2012.

[12] I. Krikidis, S. Timotheou, and S. Sasaki, "RF energy transfer for cooperative networks: Data relaying or energy harvesting?" *IEEE Communications Letters*, vol. 16, no. 11, pp. 1772–1775, 2012.

[13] C. Park, S. Lee, G.-H. Cho, and C. T. Rim, "Innovative 5-m-off-distance inductive power transfer systems with optimally shaped dipole coils," *IEEE Transactions on Power Electronics*, vol. 30, no. 2, pp. 817–827, 2015.

[14] A. P. Sample, D. Meyer, and J. R. Smith, "Analysis, experimental results, and range adaptation of magnetically coupled resonators for wireless power transfer," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 2, pp. 544–554, 2011.

[15] L. Xiao, P. Wang, D. Niyato, D. Kim, and Z. Han, "Wireless networks with RF energy harvesting: a contemporary survey," *IEEE Communications Surveys & Tutorials*, 2014.

[16] K. Huang and V. K. Lau, "Enabling wireless power transfer in cellular networks: architecture, modeling and deployment," *IEEE Transactions on Wireless Communications (TWC)*, vol. 13, no. 2, pp. 902–912, 2014.

[17] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Transactions on Wireless Communications (TWC)*, vol. 12, no. 5, pp. 1989–2001, 2013.

[18] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," *IEEE Transactions on Wireless Communications (TWC)*, vol. 13, no. 1, pp. 418–428, 2014.

[19] K. Zhao and R. Jurdak, "Understanding the spatiotemporal pattern of grazing cattle movement," *Nature Scientific Reports*, vol. 6, p. 31967, 2016.

[20] K. Zhao, R. Jurdak, J. Liu, D. Westcott, B. Kusy, H. Parry, P. Sommer, and A. McKeown, "Optimal lévy-flight foraging in a finite landscape," *Journal of The Royal Society Interface*, vol. 12, no. 104, p. 20141158, 2015.

[21] R. Jurdak, A. Elfes, B. Kusy, A. Tews, W. Hu, E. Hernandez, N. Kottege, and P. Sikka, "Autonomous surveillance for biosecurity," *Trends in biotechnology*, vol. 33, no. 4, pp. 201–207, 2015.

[22] Y. Liu, C. Yuen, X. Cao, N. Ul Hassan, and J. Chen, "Design of a scalable hybrid mac protocol for heterogeneous M2M networks," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 99–111, 03 2014.

[23] LAN/MAN Standards Committee, "Standard for part 15.4: Wireless medium access control (MAC) and physical layer (PHY) specifications for low rate wireless personal area networks (LR-WPANs)," *ANSI/IEEE 802.15*, vol. 4, 2003.

[24] P. Corke, T. Wark, R. Jurdak, W. Hu, P. Valencia, and D. Moore, "Environmental wireless sensor networks," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1903–1917, 2010.

[25] L. González, G. Bishop-Hurley, R. Handcock, and C. Crossman, "Behavioral classification of data from collars containing motion sensors in grazing cattle," *Computers and Electronics in Agriculture*, vol. 110, pp. 91–102, 2015.

[26] L. González, G. Bishop-Hurley, D. Henry, and E. Charmley, "Wireless sensor networks to study, monitor and manage cattle in grazing systems," *Animal Production Science*, vol. 54, no. 10, pp. 1687–1693, 2014.

[27] R. Dutta, D. Smith, R. Rawnsley, G. Bishop-Hurley, and J. Hills, "Cattle behaviour classification using 3-axis collar sensor and multi-classifier pattern recognition," in *IEEE SENSORS*, 2014, pp. 1272–1275.

[28] A. Schulman, V. Navda, R. Ramjee, N. Spring, P. Deshpande, C. Grunewald, K. Jain, and V. N. Padmanabhan, "Bartendr: a practical approach to energy-aware cellular data scheduling," in *Proceedings of Annual International Conference on Mobile Computing and Networking*. ACM, 2010, pp. 85–96.

[29] T.-P. Low, M.-O. Pun, Y.-W. Hong, and C.-C. Kuo, "Optimized opportunistic multicast scheduling (OMS) over wireless cellular networks," *IEEE Transactions on Wireless Communications (TWC)*, vol. 9, no. 2, pp. 791–801, 2010.

[30] D. Wu and R. Negi, "Downlink scheduling in a cellular network for quality-of-service assurance," *IEEE Transactions on Vehicular Technology (TVT)*, vol. 53, no. 5, pp. 1547–1557, 2004.

[31] Y. Lin and W. Yu, "Fair scheduling and resource allocation for wireless cellular network with shared relays," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 30, no. 8, pp. 1530–1540, 2012.

[32] D. Fooladivanda and C. Rosenberg, "Joint resource allocation and user association for heterogeneous wireless cellular networks," *IEEE Transactions on Wireless Communications (TWC)*, vol. 12, no. 1, pp. 248–257, 2013.

[33] X. Wu, S. Tavildar, S. Shakkottai, T. Richardson, J. Li, R. Laroia, and A. Jovicic, "FlashLinQ: A synchronous distributed scheduler for peer-to-peer ad hoc networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 21, no. 4, pp. 1215–1228, 2013.

[34] L. Zhou, X. Wang, W. Tu, G.-M. Muntean, and B. Geller, "Distributed scheduling scheme for video streaming over multi-channel multi-radio multi-hop wireless networks," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 28, no. 3, pp. 409–419, 2010.

[35] B. Ji, C. Joo, and N. B. Shroff, "Delay-based back-pressure scheduling in multihop wireless networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 21, no. 5, pp. 1539–1552, 2013.

[36] M. J. Neely, "Delay-based network utility maximization," in *IEEE INFOCOM*, 2010, pp. 1–9.

[37] M. Neely, "Opportunistic scheduling with worst case delay guarantees in single and multi-hop networks," in *IEEE INFOCOM*, 2011, pp. 1728–1736.

[38] S. Tang and L. Yang, "Morello: A quality-of-monitoring oriented sensing scheduling protocol in sensor networks," in *IEEE INFOCOM*, 2012, pp. 2676–2680.

[39] S. Nabar, J. Walling, and R. Poovendran, "Minimizing energy consumption in body sensor networks via convex optimization," *International Conference on Body Sensor Networks (BSN)*, pp. 62–67, 2010.

[40] S. He, J. Chen, F. Jiang, D. K. Yau, G. Xing, and Y. Sun, "Energy provisioning in wireless rechargeable sensor networks," *IEEE Transactions on Mobile Computing*, vol. 12, no. 10, pp. 1931–1942, 2013.

[41] L. He, L. Kong, Y. Gu, J. Pan, and T. Zhu, "Evaluating the on-demand mobile charging in wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 14, no. 9, pp. 1861–1875, 2015.

[42] Y. He, W. Zhu, and L. Guan, "Optimal resource allocation for pervasive health monitoring systems with body sensor networks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 11, pp. 1558–1575, 2011.

[43] A. Mehrabi and K. Kim, "Maximizing data collection throughput on a path in energy harvesting sensor networks using a mobile sink," *IEEE Transactions on Mobile Computing (TMC)*, vol. 15, no. 3, pp. 690–704, 2015.

[44] K. Pister and L. Doherty, "TSMP: Time synchronized mesh protocol," *IASTED Distributed Sensor Networks*, pp. 391–398, 2008.

[45] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[46] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[47] S. C. Ergen, "Zigbee/IEEE 802.15. 4 summary," *UC Berkeley, September*, vol. 10, 2004.

[48] S. Martello and P. Toth, *Knapsack problems: algorithms and computer implementations*. John Wiley and Sons, Inc., 1990.

[49] P. Sommer, J. Liu, K. Zhao, B. Kusy, R. Jurdak, A. McKeown, and D. Westcott, "Information bang for the energy buck: Towards energy-and mobility-aware tracking," in *EWSN*, 2016.

[50] (2015, February). [Online]. Available: http://www.powercastco.com/products/powercaster-transmitters/

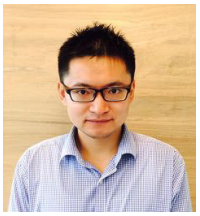
[51] "http://www.powercastco.com/pdf/p2110-evb.pdf."

[52] "http://www.ti.com/product/cc2420."

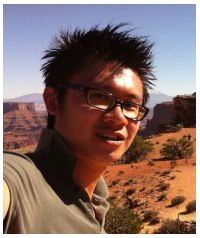
[53] R. Jurdak, P. Sommer, B. Kusy, N. Kottege, C. Crossman, A. Mckeown, and D. Westcott, "Camazotz: multimodal activity-based GPS sampling," *ACM/IEEE IPSN*, pp. 67–78, 2013.

[54] K. Srinivasa and P. Levis, "RSSI is under appreciated," *The Third Workshop on Embedded Networked Sensors (EmNets)*, 2006.

[55] D. Willkomm, S. Machiraju, J. Bolot, and A. Wolisz, "Primary user behavior in cellular networks and implications for dynamic spectrum access," *IEEE Communications Magazine*, vol. 47, no. 3, pp. 88–95, 2009.



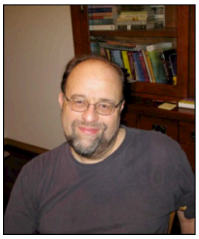
Kai Li (S'09–M'14) received the B.E. degree from Shandong University, China, in 2009, the M.S. degree from The Hong Kong University of Science and Technology, Hong Kong, in 2010, and the Ph.D. degree in Computer Science from The University of New South Wales, Sydney, Australia, in 2014. Currently he is a research scientist and project leader at Real-Time and Embedded Computing Systems Research Centre (CISTER), Portugal. Prior to this, Dr. Li was a postdoctoral research fellow at The SUTD-MIT International Design Centre, The Singapore University of Technology and Design, Singapore (2014-2016). He was a visiting research assistant at ICT Centre, CSIRO, Australia (2012-2013). From 2010 to 2011, he was a research assistant at Mobile Technologies Centre with The Chinese University of Hong Kong. His research interests include vehicular communications and security, resource allocation optimization, Cyber-Physical Systems, Internet of Things (IoT), human sensing systems, sensor networks and UAV networks.



Chau Yuen received the B. Eng and PhD degree from Nanyang Technological University, Singapore in 2000 and 2004 respectively. Dr Yuen was a Post Doc Fellow in Lucent Technologies Bell Labs, Murray Hill during 2005. He was a Visiting Assistant Professor of Hong Kong Polytechnic University in 2008. During the period of 2006–2010, he worked at the Institute for Infocomm Research (Singapore) as a Senior Research Engineer. He joined Singapore University of Technology and Design as an assistant professor from June 2010. He serves as an Editor for IEEE Transactions on Communications and IEEE Transactions on Vehicular Technology. On 2012, he received IEEE Asia-Pacific Outstanding Young Researcher Award. He has held positions on several conference organising committees, and is on technical program committees of various international conferences.

Brano Kusy is a principal research scientist in Autonomous Systems program at CSIRO where he leads the Pervasive Computing team. His research interests include systems topics in distributed systems, such as reliable wireless communication, delay tolerance in sparsely connected networks, and coordination of time and space of individual sensor nodes. He has applied wireless sensing technology in several domains, including energy-aware sensing of airborne animals, high-granularity sensing of occupancy, activity, and user comfort in commercial buildings, and spatio-temporal analysis of environmental parameters in rehabilitated mine environments.

Raja Jurdak is a Senior Principal Research Scientist at CSIRO where he leads the Distributed Sensing Systems Group. He holds a PhD in Information and Computer Science at University of California, Irvine (2005), an MS in Computer Networks and Distributed Computing from the Electrical and Computer Engineering Department at UCI (2001), and a BE in Computer and Communications Engineering from the Faculty of Engineering and Architecture at the American University of Beirut in (2000). His research interests include mobility and energy efficiency in pervasive computing. Jurdak is an Honorary Professor at the University of Queensland, and Adjunct Professor at University of New South Wales, Macquarie University, and James Cook University. He is a Senior Member of the IEEE.



Aleksandar Ignjatovic received the bachelor's and master's degrees both in mathematics from the University of Belgrade, former Yugoslavia, and the PhD degree in mathematical logic from the University of California at Berkeley. After graduation, he was an assistant professor at the Carnegie Mellon University, where he taught for 5 years at the Department of Philosophy and subsequently had a startup in the Silicon Valley. He joined the School of Computer Science and Engineering at the University of New South Wales (UNSW) in 2002, where he teaches algorithms. His research interests include sampling theory and signal processing, applications of mathematical logic to computational complexity theory, algorithms for embedded systems design, and most recently trust-based data aggregation algorithms.

Salil S. Kanhere received the BE degree in electrical engineering from the University of Bombay, India, in 1998, and the MS and PhD degrees in electrical engineering from Drexel University, Philadelphia, in 2001 and 2003, respectively. He is currently a Associate Professor with the School of Computer Science and Engineering at the University of New South Wales, Sydney, Australia. His current research interests include wireless sensor networks, vehicular communication, mobile computing, and network security. He is a member of the ACM and a senior member of the IEEE.

Sanjay Jha is a Professor and Head of the Network Group at the School of Computer Science and Engineering at the University of New South Wales. His research activities cover a wide range of topics in networking including Network and Systems Security, Wireless Sensor Networks, Adhoc/Community wireless networks, Resilience and Multicasting in IP Networks. Sanjay has published over 160 articles in high quality journals and conferences. He is the principal author of the book Engineering Internet QoS and a co-editor of the book Wireless Sensor Networks: A Systems Perspective. He served as an associate editor of the IEEE Transactions on Mobile Computing (TMC). He currently serves on the editorial board of the ACM Computer Communication Review (CCR).