Distributed decision making policy for frequency band selection boosting RF energy harvesting rate in wireless sensor nodes

S. Darak, Christophe Moy, Jacques Palicot

Distributed Decision Making Policy for Frequency Band Selection Boosting RF Energy Harvesting Rate in Wireless Sensor Nodes

S. J. Darak · Christophe Moy · Jacques Palicot

Abstract Emerging paradigms such as smart cities and Internet of Things are expected to be an intrinsic part of next generation communication standards. To bring these paradigms to life, self-sustainable wireless sensor network (WSN) nodes capable of seamless and maintenance free operation at remote locations are desired. Recently, radio frequency energy harvesting (RFEH) circuits capable of harvesting RF power transmitted by base stations, TV towers and other ambient RF sources have been developed. Low power requirements and architectural compatibility between WSN nodes and RFEH circuits make RFEH a promising and feasible solution for WSN nodes. In this paper, a novel multi-stage decision-making policy (DMP) for RFEH enabled WSN nodes has been proposed. It offers an intelligence, via online learning algorithm, for characterization and selection of frequency bands based on their RF potential especially in the dynamic spectrum environment. Furthermore, proposed DMP supports multi-antenna multi-band harvesting capabilities of the RFEH circuits. The final contribution includes tunable RFEH duration that leads to significant improvement in the harvested energy and fewer number of frequency band switchings (FBS). Derived theoretical performance bounds and simulation results validate the superiority of proposed DMP in terms of the harvested RF energy and

S. J. Darak
ECE Dept., IIIT-Delhi, India-110020
Tel.: +91-11-26907427
E-mail: sumit@iiitd.ac.in

Christophe Moy
SCEE, CentraleSupélec/IETR, Rennes, France
E-mail: Christophe.Moy@centralesupelec.fr

Jacques Palicot
SCEE, CentraleSupélec/IETR, Rennes, France
E-mail: Jacques.Palicot@centralesupelec.fr

throughput of the WSN nodes. Furthermore, the fewer number of FBS makes the proposed DMP suitable for resource-constrained WSN nodes.

Keywords RF Energy Harvesting · Decision Making Policy · Wireless Sensor Nodes · Multi-armed Bandit

1 Introduction

Wireless sensor network (WSN) nodes are small-sized transceivers, designed to sense and transmit the sensed information to the desired receiver [1, 2]. WSN nodes have generated significant interest in the research community because of their applications in emerging paradigms such as smart cities, smart grids and Internet of Things (IoTs) [3, 4]. Expectations from these paradigms are huge because of their potential to enable data exchange anywhere and anytime between a wide range of devices. However, to bring the smart cities, smart grids and IoTs to life, efficient solutions to the power requirements, maintenance-free operations and remote deployments of WSN nodes need to be investigated [3, 4, 6]. One such solution is an ambient energy harvesting making WSN nodes self–sustainable and environment-friendly, i.e., green [5, 6].

Recently, radio frequency energy harvesting (RFEH) circuits capable of harvesting energy from signals transmitted by RF sources have been developed [3–8]. RF sources can be dedicated sources or ambient sources such as cellular base stations, TV towers, and WiFi access points. This is exciting because, in addition to enabling the data transmission at farther distances, the broadcast nature of the RF signals can simultaneously charge multiple WSN nodes. Since WSN nodes are usually distributed over a large area and deployed at remote locations, RFEH from the ambient RF sources,

instead of the dedicated RF sources, seems to be a promising and practical solution. Recent advances in the design of reconfigurable antennas (i.e., antennas with tunable center frequency) also make RFEH from ambient RF sources feasible [9–12]. However, the limited bandwidth of reconfigurable antennas, as well as dynamic spectrum environment, demands RFEH circuit to be capable of giving higher preference to the frequency bands having higher RF potential over other bands [9–16]. Furthermore, recent RFEH circuits offer RFEH simultaneously from multiple frequency bands via multiple reconfigurable antennas or single multiband reconfigurable antenna [9–16]. Such RFEH circuits need an intelligence and hence, the decision making policy (DMP) to characterize the frequency bands based on their RF potential and minimize the time spent on harvesting from sub-optimal bands.

The RF potential of the frequency band is location dependent which means that an intelligence (i.e. information about the RF potential of frequency bands) gained by WSN node may not be useful for other WSN nodes. In the case of distributed DMPs, no such information exchange takes place between WSN nodes. Furthermore, distributed DMPs have the advantages of robustness to link or node failures and zero communication overhead. The design and development of such distributed DMP is the focus of the work presented in this paper.

In addition to characterization of the frequency band(s) based on the RF potential, the number of frequency band switchings (FBS) should be as small as possible. Here, FBS refers to switching of RFEH circuit from one frequency band to another in the successive time slot. Higher the number of FBS, the higher is the penalty in terms of the dynamic power consumption, RFEH duration and reconfiguration overhead for antenna adjustments and impedance matching circuit. Thus, DMP should have a fewer number of FBS. Note that reduction in the number of FBS must be achieved without compromising on the total harvested RF energy.

In this paper, a new multi-stage DMP for RFEH enabled WSN nodes has been proposed. The proposed work is an extension of [17] and significantly novel due to following contributions.

1. The first contribution is to offer an intelligence to WSN nodes for the characterization and selection of the frequency bands based on their RF potential. This is a non-trivial task due to the dynamic nature of RF potential of the different frequency bands occupied by ambient sources. We have formulated such characterization and selection problem in a multi-armed bandit (MAB) framework and the solution is obtained using online learning based Bayes Upper Confidence Bound (BUCB) algorithm.

2. The second contribution is to improve the harvested RF energy by minimizing the time spent on the suboptimal frequency bands by exploiting multiple reconfigurable antennas of the RFEH circuits.

3. Third contribution includes the tunable RFEH duration that leads to significant improvement in the harvested energy as well as a number of FBS.

4. Final contribution deals with the derivation of theoretical performance bounds as well as simulation results to compare the performance of the proposed DMP over existing DMPs in terms of average harvested RF energy, the number of FBS and data throughput.

5. This paper is a significant extension of conference paper in [17]. In [17], the DMP for single antenna RFEH circuit is presented with a focus on the formulation of RFEH problem in MAB framework. The work presented in this paper deals with the design of the DMP for advanced RFEH circuits with multiple fixed antennas. For such RFEH circuits, DMP in [17] is sub-optimal. In addition, we have derived the performance bounds for the proposed DMP along with extensive simulation results including the comparisons of various DMPs with respect to the parameters such as harvested RF energy, FBS as well as data throughput of the WSN nodes.

The paper is organized as follows. The literature review of the recent advances in RFEH is done in Section 2. The detailed design of the proposed DMP is presented in Section 3 followed by the simulation results in Section 4. Section 5 concludes the paper.

2 Literature Review

Recently, there has been a significant surge of interests in the RFEH enabled WSN nodes from the academia as well as industry [3, 4, 6–8, 13–15]. For example, researchers successfully demonstrated the RFEH using WiFi signal (referred to as PoWi-Fi) [15]. In PoWi-Fi, WSN nodes and even small cameras, are designed to harvest the RF energy transmitted by WiFi access points. This is achieved with simple modifications to existing access point architecture and data transmission protocol without compromising on the quality of service for WiFi users. Similarly, WSN nodes capable of harvesting ambient RF energy have been demonstrated by Drayson Technologies and Intel [7]. Furthermore, exponential increase in the data traffic and number of cellular users guarantee the availability of sufficient ambient RF energy in the environment at any time of the day.

From an architecture perspective, RFEH circuit has advantages of smaller dimensions and easier integration with the analog front-end of WSN node terminals thereby making it a preferred choice for WSN nodes compared to solar, thermal, vibrational etc. based energy harvesting approaches [14]. These examples signify the feasibility and commercial interests in RFEH capable WSN nodes.

Research efforts relevant to the RFEH in WSN nodes mainly deal with the design of RFEH circuits and DMPs. Discussion on the former which includes the design of multi-band reconfigurable antenna based RFEH circuit, multiple fixed antennas based RFEH circuit, impedance matching and multi-stage RFEH circuits is out of the scope of this paper and reader may refer to [9–12, 14–16] for more information. To summarize, an efficiency of the RFEH circuits depend on the amount of incident RF power and hence, harvesting bandwidth. The higher the received RF power or the bandwidth, the higher is the harvested RF energy but the relationship is non-linear [12, 14–16]. Furthermore, though existing commercial RFEH circuits have a single fixed antenna, it is expected that the RFEH circuits with multiple reconfigurable antennas or single multi-band reconfigurable antenna capable of harvesting from multiple non-contiguous frequency bands will be available in near future. The harvesting bandwidth of such circuit is expected to be narrow (i.e. few MHz) making wideband RFEH infeasible. Hence, RFEH circuits need intelligence to characterize the frequency bands based on their RF potential and tune the RFEH circuit to the optimum frequency bands for faster RFEH.

Comprehensive reviews of the recent advances in wireless communication network design related to RFEH are given in [6, 13]. In [18, 19], DMPs for scheduling dedicated RF sources to serve multiple WSN nodes has been proposed while the focus of the proposed work is on RFEH from ambient RF sources.

In [14], relay-based cooperative RFEH using the dedicated RF source with the possibility of ambient RF harvesting has been discussed. The DMPs in [6, 15] deal with the switching actions between data communication and RFEH modes and the goal is to maximize the total network throughput. However, it is assumed that the frequency band statistics are known and only single band with fixed bandwidth is considered. This means that the challenge of frequency band selection for the decentralized networks has not been considered in [6, 15]. The optimization, as well as learning based DMPs for RFEH, enabled the secondary user in the cognitive radio network has been proposed in [6] and later extended to the multiuser cognitive radio network in [20, 21]. An optimization based DMPs are

computationally complex and may not be feasible for the battery-operated resource-constrained WSN nodes. Hence, online learning access based DMPs need to be investigated.

The DMPs for frequency band selection in RFEH enabled secondary users have been recently proposed in [22, 23]. In [22, 23], we have introduced RFEH in the context of D2D communications for the cognitive radio network. The DMP in [22, 23] improves the average spectrum utilization by minimizing collisions between users and dynamically switching between the D2D and RFEH modes. However, the usefulness and performance analysis of Bayesian online learning algorithms for frequency band characterization has not been explored yet in the literature. Furthermore, the design of DMP for RFEH circuits with multiple antennas and/or multi-band reconfigurable antenna has not been discussed in the literature yet. In the next section, the design of the proposed DMP for such RFEH circuits is presented.

3 Proposed Decision Making Policy

Existing DMPs [13–15, 17–23] consider RFEH from single RF source which uninterruptedly transmits in a fixed and known frequency band. In dynamic spectrum environment consisting of multiple ambient RF sources and limited bandwidth RFEH circuits, WSN nodes need intelligence, i.e. DMP, to characterize frequency bands based on their RF potential. The design of the DMP which exploits multiple antennas of RFEH circuit to increase the energy harvested as well as data throughput of WSN nodes is presented in this section. To begin with, the network model considered for the proposed DMP is discussed in the next sub-section 3.1 followed by the proposed decision-making framework in 3.2.

3.1 Network Model

Consider wideband spectrum consisting of N frequency bands of uniform bandwidth, B_s $(=1/N)$ and n denotes the frequency band index, i.e., $n \in \{1, 2, .., N\}$. Hereafter, frequency specifications are normalized with respect to half the sampling frequency. The RF potential of the nth frequency band is denoted as Ω_n and equal to $\mu_n \cdot \chi_n$ where μ_n is the probability of the band n being occupied and χ_n is the average normalized incident RF energy when the frequency band n is chosen for RFEH. It assumed that Ω_n evolves as an independent and identically distributed (i.i.d.) random process with variance σ_n^2, stationary and unknown to WSN nodes.

The RFEH circuit, consisting of L reconfigurable antennas, harvests RF energy from the frequency bands

chosen by the DMP in each time slot. Each time slot duration is fixed and equal to Δt and k denotes the time slot index with horizon, K, i.e., $k \in \{1, 2, .., K\}$. The kth time slot begins at $[\Delta t \cdot (k-1)]$ units and we have

$$\hat{t}_{l,k} = \Delta t - M_{l,k} \cdot \bar{t}_{l,k} \qquad (1)$$

where $M_{l,k}$ denotes the number of stages of frequency band selection for antenna $l \in \{1, 2, .., L\}$, $\bar{t}_{l,k}$ is the time required for frequency band selection, front-end reconfiguration and antenna adjustments and $\hat{t}_{l,k}$ is the time available for RFEH or idle time depending on the status of the chosen frequency band. The value of $M_{l,k}$ depends on the status of the chosen frequency band. For example, $M_{l,k}=2$ indicates that the first frequency band chosen by antenna l in time slot k does not meet required RF energy constraints and hence, the second frequency band is chosen in the same time slot.

In any given time slot, DMP decides the bandwidth and center frequency of the frequency band(s) which are then filtered by the analog front-end of WSN node. Due to the reconfigurability constraints of the analog front-end, the bandwidth of the chosen frequency band(s) is assumed as fixed and equal to B_{afe}. Here, B_{afe} is the bandwidth of the analog front-end of WSN nodes and is assumed to be an integer multiple of B_s for simplicity of analysis. Then, the number of frequency band choices, N_{rf}, of bandwidth B_{afe} are given by

$$N_{rf} = N + 1 - \frac{B_{afe}}{B_s} \qquad (2)$$

Let the actual harvested power from the lth antenna of RFEH circuit in the kth time slot be $p_{l,k}$. Let P_t^* and P_t denote the total RF power harvested using the genie-aided DMP (i.e. the DMP where the frequency band statistics i.e., $\mu_{\bar{n}}$ and $\chi_{\bar{n}}, \forall \bar{n} \in \{1, 2, .. N_{rf}\}$, are known a priori) and the distributed DMP, respectively. Then, the total expected loss in terms of harvested RF energy, U_t, is given by

$$U_t = P_t^* - P_t = \left\lfloor \frac{t}{\Delta t} \right\rfloor \cdot L \cdot p^* - \sum_{k=1}^{K=\left\lfloor \frac{t}{\Delta t} \right\rfloor} \sum_{l=1}^{L} \mathbb{E}[p_{l,k}] \quad (3)$$

In Eq. 3, p^* denotes the RF energy harvested by using single antenna in one time slot using genie-aided DMP and is given by,

$$p^* = \chi_{\bar{n}^*} \cdot \eta_{rf}(\chi_{\bar{n}^*}) \cdot \Delta t \qquad (4)$$

Here $\eta_{rf}(\cdot)$ is the efficiency of the RFEH circuit for a given incident RF power and $\bar{n}^*$ denotes the frequency band index having maximum RF potential i.e.,

$$\bar{n}^* = \underset{\bar{n} \in \{1,2,..,N_{rf}\}}{\arg\max} \chi_{\bar{n}} \cdot \mu_{\bar{n}} \qquad (5)$$

Similarly, in Eq. 3, $p_{l,k}$ denotes the RF energy harvested by antenna l in the kth time slot over the frequency band chosen by the DMP and is given by,

$$p_{l,k} = \chi_{\bar{n}_{l,k,M_{l,k}}} \cdot \eta_{rf}(\chi_{\bar{n}_{l,k,M_{l,k}}}) \cdot (\Delta t - M_{l,k} \cdot \bar{t}_{l,k}) \quad (6)$$

where $\bar{n}_{l,k,m_{l,k}}$ denotes the frequency band chosen by the lth antenna in the stage, $m_{l,k}$, of the kth time slot, $\bar{n}_{l,k,m_{l,k}} \in \{1, 2, ..., N_{rf}\}$ and $m_{l,k} \in \{1, 2, ..M_{l,k}\}$. Total number of FBS of the DMP, FBS_t, is given by

$$FBS_t = \sum_{k=2}^{K} \sum_{l=1}^{L} \prod_{\bar{l}=1}^{L} \mathbb{E}[\mathbf{1}_{\{\bar{n}_{l,k,1} \neq \bar{n}_{\bar{l},k-1,M_{\bar{l},k-1}}\}}] +$$

$$\sum_{k=1}^{K} \sum_{l=1}^{L} \sum_{m_{l,k}=2}^{M_{l,k}} \mathbb{E}[\mathbf{1}_{\{\bar{n}_{l,k,m_{l,k}} \neq \bar{n}_{l,k,m_{l,k}-1}\}}] \qquad (7)$$

where the first term indicates the FBS from the one-time slot to another while the second term indicates the FBS in a given time slot due to multi-stage DMP. An indicator function[1] is used in the calculation of FBS_t.

From Eq. 3-7, it can be observed that the harvesting time in the distributed DMP is lower compared to the genie-aided DMP. This is because distributed DMP needs to explore various frequency bands before settling down to the optimum band in addition to the time required for hardware reconfiguration when frequency band switching occurs. In genie-aided DMP, such switching does not exist since it has prior knowledge of optimum frequency bands. Hence, the main objective of the proposed DMP is to have a lower loss in terms of harvested RF energy, i.e. U_t, and minimize the number of FBS by minimizing the number of selections of sub-optimal bands. In Table 1, all the notations used in this paper are given along with their definitions.

3.2 Proposed Decision Making Framework

The proposed decision-making framework enumerating the different decision-making stages in each time slot is shown in Fig. 1. The first decision to be taken at the beginning of each time slot, say $k(> 1)$, is whether to continue RFEH over the frequency band chosen for the last time slot, i.e., $(k-1)$. The advantage of an affirmative decision (i.e., $skip\_BS = 1$) is that the total time available for the RFEH is same as genie-aided DMP since $M_{l,k} = 0$ thereby leading to the zero penalties in terms of RFEH time and FBS. However, if the same frequency band is chosen each time, DMP may not be able to characterize the frequency bands accurately due

---

[1] Indicator function: $\mathbf{1}_{\{logical\ expression\}}=1$ if logical expression=true; else 0.

Table 1 Notations and definitions

Notations	Definitions
N	No. of frequency bands of bandwidth B_s
Bs	Minimum frequency band bandwidth
μ_n	Vacancy statistics of the nth band of bandwidth B_s
χ_n	Average RF energy of the nth band when occupied
B_{afe}	Bandwidth of the analog front-end of WSN nodes
N_{rf}	No. of bands of bandwidth B_{afe}
L	No. of antennas
δt	Time slot duration
k	Time slot index
$\bar{t}_{l,k}$	Reconfiguration time for antenna l in slot k
$\hat{t}_{l,k}$	Time available for RFEH for antenna l in slot k
$M_{l,k}$	No. of stages of frequency band selection for antenna l in slot k
$p_{l,k}$	Actual harvested power from the lth antenna in the kth time slot
P_t^*	Total RF power harvested using the genie-aided DMP
P_t	Total RF power harvested using the distributed DMP
U_t	Total expected harvested RF energy loss of the distributed DMP
$skip\_BS$	Flag to decide whether to skip decision making or not
ms_RFEH	Flag to decide whether another stage of RFEH is possible or not
$\eta_{rf}(\cdot)$	Efficiency of the RFEH circuit for a given incident RF power
$\bar{n}^*$	Frequency band index having maximum RF potential
$\bar{n}_{l,k,m_{l,k}}$	Frequency band chosen by the lth antenna in the stage, $m_{l,k}$, of the kth time slot
FBS_t	Total number of FBS of the DMP
$T_{\bar{n},k}$	Number of time slots up to k slots in which the band $\bar{n}$ is chosen by WSN node
$X_{\bar{n},k}$	Average normalized harvested RF energy from the frequency band $\bar{n}$ till slots k
I	Learning algorithm specific constant
$K(p,q)$	Kullback-Leibler divergence factor
$Q(x)$	Probability that normal random variable gets a value larger than x standard deviations above the mean
$Beta$	Complete beta function
ϕ	RF thereshold for ms_RFEH
$\Omega_n = \mu_n \cdot \chi_n$	RF potential of the nth band of bandwidth B_s
$\Omega_{\bar{n}}$	Observed RF potential of the $\bar{n}^{th}$ band
$\bar{\chi}_{\bar{n}}$	Estimated value of $\chi_{\bar{n}}$
$\bar{\mu}_{\bar{n}}$	Estimated value of $\mu_{\bar{n}}$
$\hat{\mu}_{\bar{n},k}$	Learned occupancy statistic of the $\bar{n}^{th}$ frequency band till the kth time slot
$\hat{\chi}_{\bar{n},k}$	Learned normalized RF potential statistic of the $\bar{n}^{th}$ frequency band till the kth time slot
$\delta_{W\mu}(k)$	Average difference between learned occupancy statistics during a window period W
$\delta_{W\chi}(k)$	Average difference between learned normalized RF potentials during a window period W
$\bar{\delta}_{W\mu}(k)$	Cumulative moving average of $\delta_{W\mu}(k)$
$\bar{\delta}_{W\chi}(k)$	Cumulative moving average of $\delta_{W\chi}(k)$
ϵ	Error between the actual and learned frequency band statistic
ε	Online algorithm convergence threshold

to exploration-exploitation trade-off and this may lead to high loss, U_t, due to the selection of sub-optimal bands. Thus, $skip\_BS$ needs to be chosen wisely and is one of the important contributions of the proposed DMP. When $skip\_BS = 0$, the frequency band is chosen using a MAB algorithm.

After the frequency band selection, short time RFEH is carried out. If the harvested RF energy is above a certain threshold, harvesting is continued over the same band for the rest of the time slot. Else, if multi-stage sensing is possible, i.e., $ms\_RFEH = 1$, another band is chosen followed by the short time RFEH. If $ms\_RFEH = 0$, the RFEH circuit remains idle until the beginning of the next time slot. Note that the feed-back about the RF potential of the chosen frequency band is given by RFEH circuit or battery unit and there is no need of additional spectrum detector.

At the end, the parameters of an online learning algorithm are updated based on the feedback received from the RFEH circuit. Similar decision making needs to be carried out L times i.e. for each antenna or each band of a reconfigurable antenna. One of the contributions of the proposed DMP is to exploit the feedback received from L antennas or L bands of a reconfigurable antenna to quickly and accurately characterize the frequency bands. This means that the frequency band selection is not carried out independently. Instead, all L feedbacks are taken while selecting the frequency band
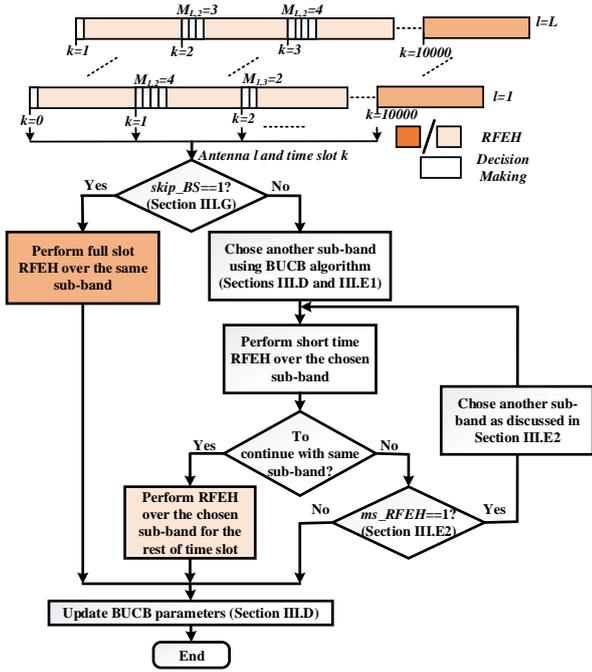
Fig. 1 Proposed decision making framework

for a given antenna or any band of a reconfigurable antenna.

3.3 Harvesting Bandwidth Constraints

Ideally, the harvesting bandwidth of RFEH circuit should be as high as possible. This is because:

1. Total harvested RF power increases with an increase in the harvesting bandwidth.
2. Efficiency of the RFEH circuit increases with an increase in the RF input power which in turn depends on the harvesting bandwidth.
3. From the frequency band characterization perspective, the number of frequency band choices increases with the decrease in harvesting bandwidth [6, 14]. This, in turn, leads to the higher amount of time spent on the sub-optimal bands.

For example, when $N = 4$, the number of frequency band choices are 4 for harvesting bandwidth equal to B_s compared to 2 and 1 for harvesting bandwidth equal to $3B_s$ and $4B_s$, respectively. The higher the number of choices, the higher is the time spent on the sub-optimal frequency bands which lead to the penalty in terms of the harvested RF power. Also, the number of FBS increases with the increase in the number of frequency bands. Thus, for the proposed DMP, the harvesting bandwidth is fixed and equal to the maximum that RFEH circuit can have i.e. B_{afe}.

For a given B_{afe}, the number of frequency bands, N_{rf}, are given by Eq. 2. Ideally, the RFEH circuit should be able to harvest from the entire wideband input signal of bandwidth, i.e., $B_{afe} = N \cdot B_s$. To have $B_{afe} = N \cdot B_s$, wideband reconfigurable antennas are desired. The design of such antennas is a challenging and non-trivial problem and hence, in practice, we have narrow but multi-band reconfigurable antenna with $B_{afe} \ll N \cdot B_s$ [9–12, 14–16]. The frequency band characterization is discussed next.

3.4 Frequency Band Characterization

The size and power constraints of WNS nodes limit the value of L for RFEH circuits. For instance, in existing RFEH circuits with reconfigurable antenna, it can be safely assumed that $N_{rf} \gg L$ [9–12]. As a result, RFEH enabled WSN nodes need an intelligence to choose frequency bands for RFEH in each time slot to minimize the loss given by Eq. 3. Random selection approach would lead to a linear increase in the loss while the greedy approach may not guarantee accurate characterization and hence, convergence to the optimal frequency bands leading to inferior performance.

In the proposed DMP, we have formulated the frequency band characterization and selection problem into parallel MAB framework [24–26]. Here, each arm is analogous to frequency bands and L arms are chosen in each time slot out of N_{rf} arms. Selection of bands in each time slot is done by online learning algorithm which needs to balance between exploration of N_{rf} bands and exploitation of top L bands. For instance, online learning algorithm should meet following criteria [26]

$$\liminf_{k \to \infty} \frac{\mathbb{E}[T_{\bar{n},k}]}{L \cdot \ln k} = \frac{I}{K(\chi_{\bar{n}} \cdot \mu_{\bar{n}}, \chi_{\bar{n}^*} \cdot \mu_{\bar{n}^*})}, \ \forall \bar{n} \qquad (8)$$

where

$$\bar{n}^* = \underset{\bar{n} \in \{1,2,..,N_{rf}\}}{\arg\max} \chi_{\bar{n}} \cdot \mu_{\bar{n}} \qquad (9)$$

Here $T_{\bar{n},k}$ denotes the number of time slots up to k slots in which the band $\bar{n}$ is chosen by WSN node and I is algorithm specific constant. Also, $K(p,q)$ denotes the Kullback-Leibler divergence factor and is the measure of the difference between the probabilities, p and q. Eq. 8 guarantees accurate characterization of the frequency bands via Kullback-Leibler divergence factor and the algorithm for which $I=1$ is said to be optimal as it guarantees the high value of $T_{\bar{n}^*,k}$, which is the number of times the optimal frequency band has been selected.

Online learning algorithms include frequentist approach based Upper Confidence Bound (UCB) algorithm and its extensions [24], optimization based KL-UCB [25] algorithm, Bayes-UCB (BUCB) [27] and Thompson Sampling (TS) algorithms. Though all these algorithms are asymptotically optimal, it has been recently proved that the parameter I of the BUCB and TS algorithms is lower than others [26, 27]. Also, both of them have lower computational complexity than others [26, 27]. Empirically, we observed that the BUCB algorithm offers slightly better performance than TS algorithm and more importantly, it leads to a fewer number of the FBS. These advantages make the BUCB algorithm a preferred choice for the proposed DMP.

The main idea behind BUCB algorithm is to assume some prior distribution on the probability statistics of each band (e.g. uniform prior) and at any time slot, sample the frequency bands according to its posterior probability of being optimum [27]. In the proposed DMP, BUCB algorithm is employed to rank the frequency bands based on their RF potential as follows

$$G(\bar{n}, k) = Q\left\{1 - \frac{1}{k}; Beta[X_{\bar{n},k} + 1, T_{\bar{n},k} - X_{\bar{n},k} + 1]\right\} \quad (10)$$

where $X_{\bar{n},k}$ is the average normalized harvested RF energy from the frequency band $\bar{n}$ which is chosen for $T_{\bar{n},k}$ time slots out of k time slots, $Q(x)$ is the probability that any normal random variable gets a value larger than x standard deviations above the mean and $Beta$ represents the complete beta function, i.e., Euler integral of the first kind. The higher the value of $G(\bar{n}, k)$, the higher is the RF potential of the corresponding band, $\bar{n}$. To the best of our knowledge, the usefulness of the BUCB algorithm for RFEH application has not been studied yet in the literature. Furthermore, existing BUCB algorithm, designed for conventional sequential MAB framework, needs to be adapted for the multi-stage multi-band RFEH circuits via parallel MAB framework. After characterization, next step is frequency band selection which is discussed below.

3.5 Frequency Band Selection for Multi-stage DMP

In the proposed DMP, frequency band selection depends on the values of $skip\_BS$ and ms_RFEH. When $skip\_BS=1$, the frequency band is same as that of the previous time slot. The frequency band selection when $skip\_BS=0$ is discussed next.

3.5.1 First stage frequency band selection

When $skip\_BS=0$, the proposed DMP calculates the quality index of all frequency bands using BUCB algorithm given by Eq. 10 and selects the frequency band with highest quality index. The chosen frequency band is denoted as $\bar{n}_{l,k,m_{l,k}}$ (i.e., $\bar{n}_{l,k,1}$) where $m_{l,k}$ denotes the frequency band selection stage for antenna l in time slot k and hence, it is equal to 1 in the first stage. After short time RFEH on the chosen frequency band, if harvested RF energy is above the threshold, ϕ, RFEH is continued over the same band for the rest of time slot as shown in Fig. 1. In this case, $ms\_RFEH=0$ and $M_{l,k}=1$ where $M_{l,k}$ is the number of frequency band selection stages for antenna l in time slot k.

The value of the threshold, ϕ, affects the number of FBS since high value of ϕ makes the DMP greedy towards optimal action leading to higher number of FBS. Small ϕ may even lead to drainage of battery power if harvested energy is not sufficient to support RFEH operations. However, the threshold, ϕ, depends on the sensitivity of the RFEH circuit and is equal to the minimum input RF power for which the harvested energy is non-zero. The value of ϕ is taken from the product specification or data sheet of the RFEH circuit.

3.5.2 Subsequent frequency band selection

If the harvested RF power from the frequency band $\bar{n}_{l,k,1}$ is below ϕ, then the proposed DMP starts the multi-stage frequency band selection process. In each stage, the feasibility of RFEH, i.e., whether available time i.e. $(\Delta t - m_{l,k} \cdot \bar{t}_{l,k})$, is more than the wake-up time of the RFEH circuit is checked. If yes, then $ms\_RFEH=1$ and $m_{l,k}$ is incremented by 1. When $ms\_RFEH=1$, short term RFEH over the chosen frequency band is carried out. If the harvested RF energy is above ϕ, the RFEH is continued over the same band for the rest of time slot as shown in Fig. 1. Otherwise, the above process of feasibility check and subsequent RFEH is repeated. When $ms\_RFEH=0$, the RFEH circuit remains idle until the beginning of next time slot.

One possible drawback of the multi-stage approach is that it may lead to battery drainage if poor bands are chosen for $m_{l,k} > 1$. Such probability is very high in BUCB algorithm as frequency band ranking is not accurate except optimal band since it was originally designed for sequential MAB framework. Hence, proposed DMP follows greedy frequency band selection when $m_{l,k} > 1$. This means that, instead of calculating quality index again, the frequency bands are chosen based on their observed RF potential, $\Omega_{\bar{n}}$, which is

given by

$$\Omega_{\bar{n}} = \bar{\chi}_{\bar{n}} \cdot \bar{\mu}_{\bar{n}} = \frac{X_{\bar{n},k}}{T_{\bar{n},k}} \tag{11}$$

Here, $\bar{\chi}_{\bar{n}}$ and $\bar{\mu}_{\bar{n}}$ indicate the estimated values of $\chi_{\bar{n}}$ and $\mu_{\bar{n}}$, respectively. For example, when $m_{l,k} = 2$ and $m_{l,k} = 3$, frequency bands with the highest and second highest $\Omega_{\bar{n}}$ values are chosen, respectively. Another advantage of greedy approach is fewer number of FBS due to lower $M_{l,k}$.

3.6 Multi-antenna RFEH Circuit

Multi-antenna or multi-band RFEH circuits can harvest higher RF energy than single antenna RFEH circuit at the expense of increased circuit complexity. Note that the conversion efficiency remains the same and is independent of the number of antennas, L. Our existing DMP in [17] was designed for single antenna RFEH circuit and its direct extension for multi-antenna RFEH circuit leads to sub-optimal DMP since all antennas must be tuned to the same frequency band. Such approach is suited only for genie-aided DMP since it has prior knowledge of the frequency band statistics. In the proposed DMP, antennas are tuned to L frequency bands chosen in each time slot by the decision-making framework discussed in previous sub-sections. After the convergence of the learning algorithm, i.e. accurate characterization of RF energy potential and occupancy statistics of all the frequency bands, all antennas are tuned to the same band chosen by the proposed decision-making framework.

The convergence of an online learning algorithm depends on the Eq. 8 which indicates the minimum number of times each frequency band should be chosen. Since the frequency band statistics are unknown, a new approach to calculate desired minimum value of $T_{\bar{n},k}, \forall \bar{n}$ using frequency band statistics estimated based on RFEH events up to time slots, k as well as recent gradient changes in these statistics, is presented below.

Let $\hat{\mu}_{\bar{n},k}$ and $\hat{\chi}_{\bar{n},k}$ be the occupancy and normalized RF potential statistics of the $\bar{n}^{th}$ frequency band, respectively, learned until the kth time slot. Let $\delta_{W\mu}(k)$ and $\delta_{W\chi}(k)$ be the average difference between learned occupancy probability and normalized RF potentials during a window period W and is given by

$$\delta_{W\mu}(k) = \frac{\sum\limits_{\bar{n}=1}^{N_{rf}} \sum\limits_{v=1}^{W} |\hat{\mu}_{\bar{n},k} - \hat{\mu}_{\bar{n},k-v}|}{N_{rf} \cdot W} \tag{12}$$

$$\delta_{W\chi}(k) = \frac{\sum\limits_{\bar{n}=1}^{N_{rf}} \sum\limits_{v=1}^{W} |\hat{\chi}_{\bar{n},k} - \hat{\chi}_{\bar{n},k-v}|}{N_{rf} \cdot W} \tag{13}$$

where $\delta_{W\mu}(k)$ and $\delta_{W\chi}(k)$ are the positive values less than or equal to 1 and $k \geq W$. As the value of k increases, $\delta_{W\mu}(k)$ and $\delta_{W\chi}(k)$ decreases. Let $\bar{\delta}_{W\mu}(k)$ and $\bar{\delta}_{W\chi}(k)$ be the cumulative moving average of $\delta_{W\mu}(k)$ and $\delta_{W\chi}(k)$, respectively. Then,

$$\bar{\delta}_{W\mu}(k) = \frac{\sum\limits_{v=W}^{k} \delta_{W\mu}(v)}{k - W + 1}, \quad \bar{\delta}_{W\chi}(k) = \frac{\sum\limits_{v=W}^{k} \delta_{W\chi}(v)}{k - W + 1} \tag{14}$$

When $\bar{\delta}_{W\mu}(k)$ and $\bar{\delta}_{W\chi}(k)$ are less than ε, algorithm can be said be close to convergence. For small value ε, it can be concluded that the algorithm has converged and satisfy the property of optimality at time slot k where following two conditions are satisfied.

$$T_{\bar{n},k} \geq \frac{ln(k)}{K(\mu^*_{\ k}, \hat{\mu}_{\bar{n},k})}, \forall \bar{n} \tag{15}$$

$$\bar{\delta}_{W\mu}(k) \leq \varepsilon, \quad \bar{\delta}_{W\chi}(k) \leq \varepsilon \tag{16}$$

After convergence, all antennas are tuned to the same frequency band. Based on $\hat{\mu}_{\bar{n},k}$ and $\hat{\chi}_{\bar{n},k}, \forall \bar{n}$ values at time slot k, each antenna is tuned to the optimum frequency band chosen by the proposed decision-making framework discussed in previous sub-sections. The proposed approach leads to at the most L time faster convergence to the optimum frequency band than previous approach [17]. Next, the proposed method to determine the value of $skip\_BS$ is presented.

3.7 Higher RFEH Rate

In the conventional cellular networks, data transmission and reception are done in a time-slotted manner due to the hardware and network synchronization constraints. The advantage of the ambient RFEH is that RFEH circuits are not constrained to have perfect synchronization with other entities in the network since they neither intend to receive error-free data nor they are worried about the collisions with other WSN nodes transmitting on the non-orthogonal frequency bands. The only constraint is that the WSN nodes should be aware of the start of each time slot for accurate characterization of the frequency bands.

In the proposed DMP, RF potential of each frequency band observed via online learning algorithm is exploited to skip the need of the decision making stage in the appropriately chosen time slots. This leads to increase in the RFEH duration by $\bar{t}_{l,k}$, in corresponding time slots leading to higher RFEH rate. Furthermore, it also leads to the reduction in the number of FBS since the frequency band remains the same. The selection of such time slots is discussed below.

According to Chebyshev inequality, there is an upper bound on the probability that the difference between the learned and actual frequency band statistic is more than some constant, ϵ, and it is given by

$$P\left\{\left|\Omega_{\bar{n}} - (\hat{\chi}_{\bar{n},k} \cdot \hat{\mu}_{\bar{n},k})\right| > \epsilon\right\} \leq \frac{\sigma_{\bar{n}}^2}{T_{\bar{n},k} \cdot \epsilon^2} \tag{17}$$

According Eq. 17, for small ϵ, $T_{\bar{n},k}$ should be sufficiently high. Small ϵ also indicates that the magnitude change in statistics, $\Delta_{\bar{n},k}$, over consecutive time slot is small where

$$\Delta_{\bar{n},k} = \hat{\chi}_{\bar{n},k} \cdot \hat{\mu}_{\bar{n},k} - \hat{\chi}_{\bar{n},k-1} \cdot \hat{\mu}_{\bar{n},k-1} \tag{18}$$

In the proposed DMP, if $\Delta_{\bar{n},k}$ is sufficiently small for all frequency bands and the frequency band chosen in the previous time slots has highest observed RF potential, then $skip\_BS=1$. When $skip\_BS$ is set to 1 in time slot k, it remains 1 for subsequent $log(k)$ number of time slots. This means that the higher the value of k, the higher is the number time slots for which $skip\_BS=1$ and hence, higher is the amount of harvesting RF energy compared to the case where $skip\_BS$ is always 0.

3.8 Bounds on Loss, U_t

In this section, we derived the upper and lower bounds on the loss, U_t, of the proposed DMP. For simplicity of analysis, we assume that the frequency bands are arranged in the decreasing order of their RF potential, i.e., $\Omega_1 \geq \Omega_2 \geq \Omega_2 ... \geq \Omega_{N_{rf}}$. To begin with, upper and lower bounds on the loss, U_t, for single stage DMP and single antenna RFEH circuit are given by

$$R_{st}(N_{rf}, k) \geq \sum_{\bar{n}=2}^{N_{rf}} \Delta(1, \bar{n}) \cdot \mathbb{E}[T_{\bar{n},k}] \tag{19}$$

$$R_{st}(N_{rf}, k) \leq \Omega_{N_{rf}} \left[\sum_{\bar{n}=2}^{N_{rf}} \mathbb{E}[T_{\bar{n},k}]\right] \tag{20}$$

where $\Delta(1, 2) = (\Omega_1 - \Omega_2)$. The bounds on the $T_{\bar{n},k}$ for online learning algorithms has been discussed in the literature and reader may refer to [24, 25, 27] for more information. Furthermore, it has been shown that BUCB and TS algorithms have superior bounds on $T_{\bar{n},k}$ than others. Based on this observation, we can conclude that proposed DMP using BUCB algorithm lead to a fewer number of selection of sub-optimal frequency bands and hence, lower loss (and higher amount of harvested RF energy).

Next, the bounds for the proposed multi-stage DMP for single antenna RFEH circuit, i.e. $L=1$, are given by

$$R_{mt}(N_{rf}, k) \geq R_{st}(N_{rf}, k) -$$

$$\sum_{\bar{n}=2}^{N_{rf}} \left\{ (\Omega_1 - \Omega_{\bar{n}}) \cdot \mathbb{E}[T_{\bar{n},k}] \left[\sum_{v=1}^{\bar{n}-1} \left(\prod_{s=0, \Omega_0 = \Omega_{\bar{n}}}^{v-1} (1 - \Omega_s) \right) \Omega_v \right] \right\} \tag{21}$$

$$R_{mt}(N_{rf}, k) \leq R_{st}(N_{rf}, k) -$$

$$\sum_{\bar{n}=2}^{N_{rf}} \left\{ \Omega_{N_{rf}} \cdot \mathbb{E}[T_{\bar{n},k}] \left[\sum_{v=1}^{N_{rf}-1} \left(\prod_{s=0, \Omega_0 = \Omega_{N_{rf}}}^{v-1} (1 - \Omega_s) \right) \Omega_v \right] \right\} \tag{22}$$

It can be observed that the proposed multi-stage DMP offers lower loss than single stage DMP when $N_{rf} > 1$ and $\Omega_{\bar{n}}$ is non-zero for every $\bar{n} > 1$. For $N_{rf} = 1$, the loss of single and multi-stage DMP is same since $L=1$.

The loss of conventional multi-stage DMP with L-antenna RFEH circuit is L times the loss of DMP with single-antenna multi-stage DMP since the same frequency band is chosen by all antennas. In the case of proposed DMP, it is hard to find the actual value of such loss. Instead, we provide the upper bound of difference between loss of conventional DMP and proposed DMP, C_{ub}, which is given by

$$C_{ub} = \sum_{\bar{n}=2}^{N_{rf}} \left\lfloor \frac{\mathbb{E}[T_{\bar{n},k}]}{L} \right\rfloor (\Omega_1 - \Omega_{\bar{n}}) \tag{23}$$

It can be observed that $C_{ub} \geq 0$. Numerically, the proposed DMP exploits L observations per time slots to reduce the time spend on sub-optimal bands. Thus, the number of times the optimal frequency band is chosen and hence, the total harvested RF energy is higher (i.e., lower loss) in the proposed DMP.

4 Simulation Results

In this section, simulation results are presented to evaluate the performance of single-stage DMP (ST-DMP), multi-stage DMP (MT-DMP) and the proposed DMP with respect to genie-aided DMP. Here, genie-aided DMP is the DMP which has complete knowledge of the status of all N frequency bands in each time slot. In ST-DMP, only one frequency band is chosen in each time slot and all antennas choose the same band. If the chosen band is vacant, RFEH circuit remains idle until the beginning of next time slot. In MT-DMP, antennas can switch to

another frequency band if the chosen frequency band is vacant. The difference between the proposed DMP and MT-DMP are, 1) All antennas choose identical frequency band in MT-DMP which may not true in the proposed DMP, and 2) In MT-DMP, frequency bands are always chosen using an underlying online learning algorithm in each time slot. On the other hand, in the proposed DMP, frequency bands are chosen using an online learning algorithm only in the first stage in each time slot. Three versions of ST-DMP and MT-DMP, designed using UCB algorithm, TS algorithm, and BUCB algorithm respectively, are considered in order to evaluate the performance of these algorithms for RFEH application.

Consider WSN node with L-antenna RFEH circuit capable of harvesting RF energy from wideband spectrum consisting of $N(=9)$ frequency bands of bandwidth, B_s. The available choices for number of antennas, L, and the harvesting bandwidth, B_{afe}, are $\{2, 4\}$ and $\{B_s, 2B_s, 4B_s\}$, respectively. For illustration, three different types of frequency band statistics are considered which are given below:

Case 1 : $\Omega_s=$ [.05 .08 .10 .15 .25 .40 .50 .65 .85]
Case 2 : $\Omega_s=$ [.01 .08 .85 .05 .25 .40 .15 .65 .10]
Case 3 : $\Omega_s=$ [.46 .47 .48 .49 .50 .51 .52 .53 .54]

For example, $\Omega_s(9) =0.85$ in Case 1 indicates that the average normalized RF energy of the 9th frequency band is 0.85. Case 3 considers approximately identical RF potential for all frequency bands making characterization challenging for the learning algorithm based DMPs. Each numerical result reported hereafter is the average of the values obtained over 50 independent experiments and simulations consider a time horizon of 10000 iterations. All simulations are carried out using MATLAB.

4.1 Harvesting Throughput

To begin with, we compare total harvested RF energy in % w.r.t. genie-aided DMP at the end of the horizon, i.e. 10000 time slots, for different values of parameters, L and B_{afe}. The plots are shown in Fig. 2 (a), Fig. 2 (b), and Fig. 2 (c) for frequency band statistics in Case 1, Case 2, and Case 3, respectively. Note that we have considered ST-DMP using three algorithms while MT-DMP using BUCB algorithm since ST-DMP using BUCB algorithm is shown to offer superior performance over ST-DMPs using UCB and TS algorithms. MT-DMP leads to the higher amount of harvested RF energy than ST-DMP which validates bounds given in Section 3.8. The proposed DMP harvest 10-15% higher RF energy than that of ST-DMPs and MT-DMP for

Case 1 and Case 2. In addition, the difference between the performance of the proposed DMP and other DMPs further increases as the value of B_{afe} decreases. This is because, for smaller B_{afe} and hence higher N_{rf}, DMP needs more time to characterize the frequency bands. In ST-DMPs and MT-DMP, all antennas choose the same frequency bands leading to slower exploration while the proposed DMP exploits L antennas using parallel MAB framework, discussed in Section 3.6, leading to faster exploration and hence, a higher number of selection of optimal frequency bands.

In Case 3, the performance of all DMPs is nearly same since all frequency bands have approximately identical RF potential. This was as expected since the difference in the harvested energy when DMP choose optimal band and second or third optimal band is quite less. Even in such case, proposed DMP offers around 2% improvement in the performance since it successfully identifies optimal band earlier than other DMPs because of accurate characterization by the Bayesian learning algorithm.

Next, Fig. 3 shows total harvested RF energy in % w.r.t. genie-aided DMPs at different time instants of the horizon. Each sub-figure in Fig. 3 corresponds to particular value of L and B_s is equal to B_{afe}. All plots in Fig. 3 become flat after some time which indicates that DMP repetitively chooses optimal frequency band identified by an underlying learning algorithm. It can be observed that the proposed DMP offers superior performance at all time instants of the horizon. Numerically, the proposed DMP leads to more than 10% improvement in the harvested RF energy over other DMPs. Harvesting throughput results in Fig. 2 and Fig. 3 also show that the BUCB algorithm offers significant improvement over UCB and TS algorithms. Since the plots of various DMPs for Case 3 are overlapping due to identical RF potential and hence, difficult to distinguish, they are not shown here.

Simulation results in Fig. 2 and Fig. 3 show that total harvested energy increases by increasing L and B_{afe}. To understand the trade-off between L and B_{afe}, we compare actual number of units of harvested RF energy in Table I. Note that single antenna can harvest at the most one, two and four units of RF energy from bandwidth, B_s, $2B_s$ and $4B_s$, respectively, in single time slot. Table I shows that the harvested RF energy do not increase linearly with the harvesting bandwidth unlike the number of antennas, L. This is because, when harvesting bandwidth is increased, the RFEH circuit needs to select contiguous frequency bands which may not be an optimal choice. For example, when $B_{afe}=2B_s$ and $L=1$, RFEH circuit needs to select two adjacent frequency bands each of bandwidth
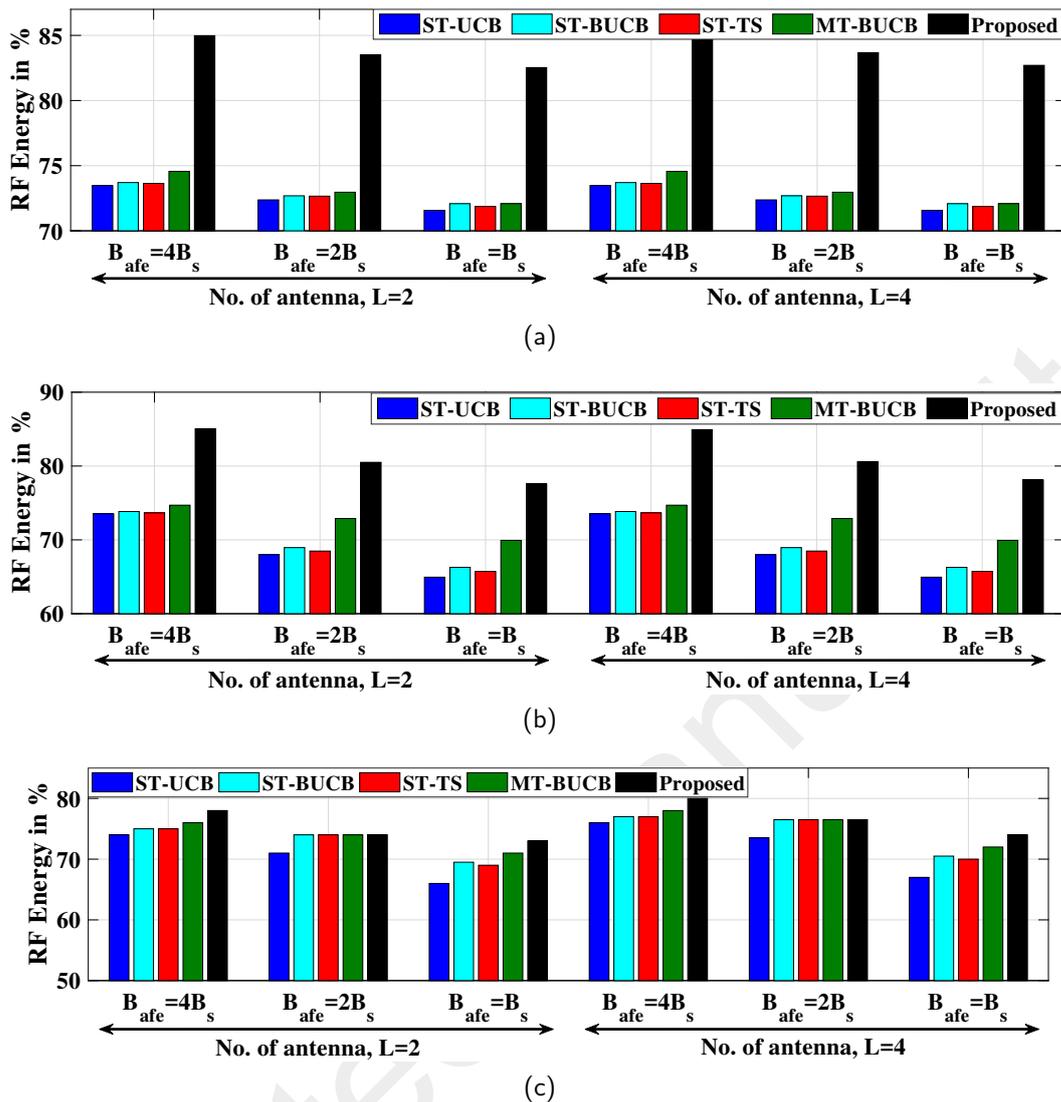
Fig. 2 Average harvested RF power in % w.r.t. genie-aided DMP for frequency band statistics (a) Case 1, (b) Case 2 and (c) Case 3.

B_s (i.e. Bands 2 and 3 in Case 2). On the other hand, when $B_{afe}=B_s$ and $L=2$, RFEH circuit can choose any two frequency bands each of bandwidth, B_s (i.e. Bands 8 and 9 in Case 2). Similarly, the harvested energy when $L, B_{afe} = \{2, 2B_s\}$ is much lower than the same when $L, B_{afe} = \{4, B_s\}$ though in both the cases, four frequency bands of the bandwidth, B_s, are chosen. Thus, increasing the number of antennas is the preferred option than increasing the bandwidth, B_{afe}. However, from the architecture perspective, increasing L leads to higher implementation complexity due to the need of L RF chains and multi-band RFEH circuit compared to the latter where only one RF chain and wide band RFEH circuit are needed. However, the design of RFEH circuit with wider harvesting bandwidth is also a challenging problem.

4.2 Data Throughput of WSN Node

The harvested RF energy is used by WSN nodes for the data pre-processing and transmission tasks. Thus, WSN nodes should have sufficient energy every time they intend to transmit data. This means that the DMP should not just harvest more energy but it should harvest continuously over the horizon in order to avoid the battery outage. Let θ denotes the number of units of the harvested RF energy required for transmission over the bandwidth, B_s using a single antenna. For example, RFEH units with two antennas need 2θ and 4θ units of RF energy, respectively, for transmission over the bandwidth, B_s and $2B_s$. For simplicity of analysis, it is assumed that all transmissions are successful and transmission power is fixed. In the case of exist-
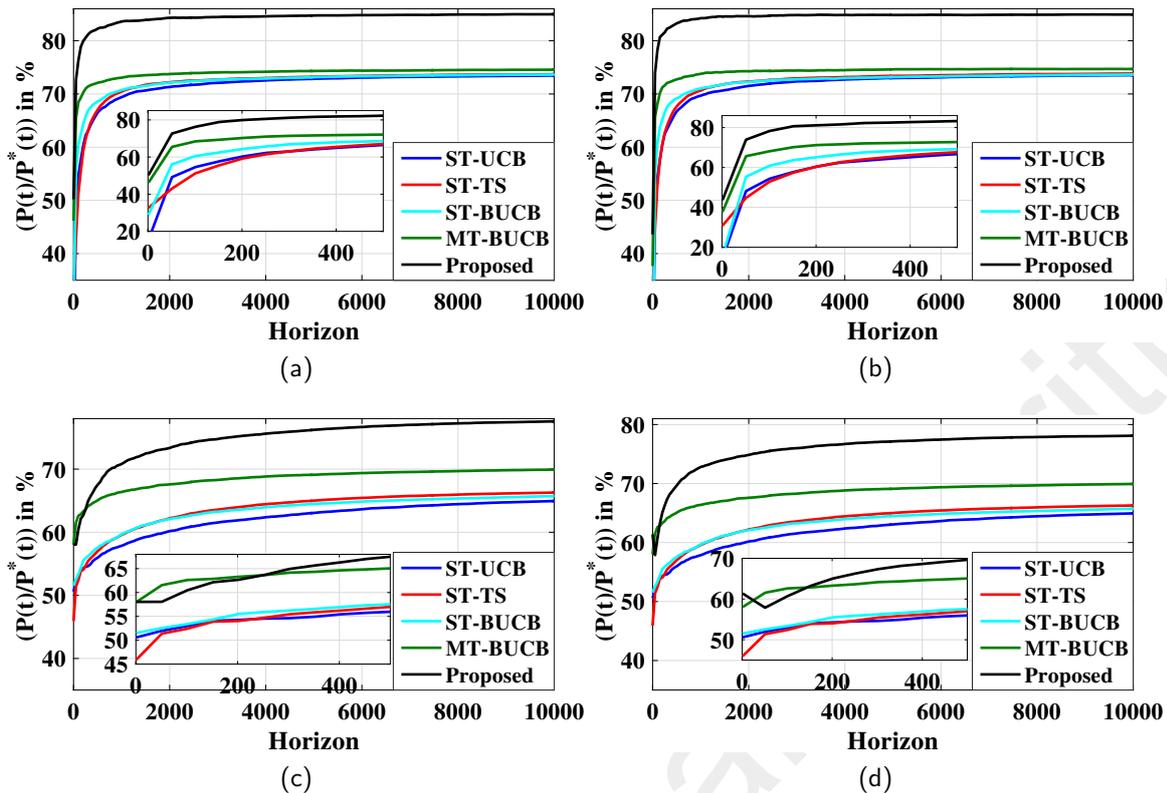
Fig. 3 Average harvested RF power in %w.r.t the genie-aided DMP at different stages of the horizon for (a) Case 1, $L=2$, and $B_{afe} = B_s$, (b) Case 1, $L=4$, and $B_{afe} = B_s$, (c) Case 2, $L=2$, and $B_{afe} = B_s$, (d) Case 2, $L=4$, and $B_{afe} = B_s$.

ing RFEH circuits and WSN node architectures, we can safely assume that $\theta \geq 4$. In Fig. 4, we compare the data throughput, $S(t)$ in % for different value of θ. It can be observed that the proposed DMP offers 2-13% higher throughput than existing DMPs. This is because the proposed DMP grantees selection of optimal frequency band with higher RF potential leading to higher harvested RF energy.
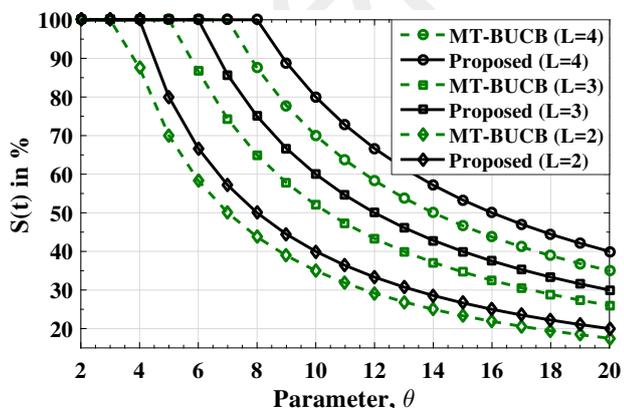

**Fig. 4** Throughput of a WSN node, $S(t)$, in % vs. θ.

4.3 Number of FBSs

As discussed in Section 1, the number of FBS should be as small as possible for energy-efficient WSN nodes. In Fig. 5, the number of FBS for different values of L and B_{afe} at the end of horizon (i.e.,$k=10000$) for frequency band statistics in (a) Case 1, (b) Case 2, and (c) Case 3 are shown. It can be observed that the number of FBS in MT-BUCB are higher than the ST-DMPs except for ST-UCB. However, in spite of being multistage DMP, the number of FBS in the proposed DMP are significantly lower than other DMPs. Such improvement is achieved due to the proposed schemes discussed in Sections 3.5 and 3.7. For Case 3, we observed that the proposed DMP offers only 2% improvement in total harvested RF energy over other DMPs. However, it can be observed from Fig. 5(c) that proposed DMP offers significantly fewer number of FBS than others for Case 3 validating the accurate and faster characterization capability of the proposed DMP compared to others.

As expected, the number of FBS reduces as B_{afe} increases. This is because, increase in B_{afe} leads to decrease in N_{rf} which means fewer number of frequency band choices. It can be observed that the number of FBS in DMPs using the BUCB and TS algorithms are

Table 2 Total Harvested RF Energy Units

DMPs	Case 1						Case 2					
	No. of Antenna=2 & B_{afe}=			No. of Antenna=4 & B_{afe}=			No. of Antenna=2 & B_{afe}=			No. of Antenna=4 & B_{afe}=		
	B_s	$2B_s$	$4B_s$	B_s	$2B_s$	$4B_s$	B_s	$2B_s$	$4B_s$	B_s	$2B_s$	$4B_s$
ST-UCB	14209	25663	34569	28417	51326	69140	14224	14895	20147	28448	29790	40294
ST-TS	14254	25774	34816	28507	51549	69632	14283	15105	20561	28566	30210	41121
ST-BUCB	14241	25764	34711	28482	51528	69423	14247	15002	20388	28494	30005	40775
MT-BUCB	14420	25871	34816	28840	51742	69633	14449	15969	21693	28897	31938	43386
Proposed	**16434**	**29610**	**39861**	**32806**	**59340**	**79887**	**16448**	**17635**	**24062**	**32856**	**35301**	**48456**

(a)

(b)

(c)
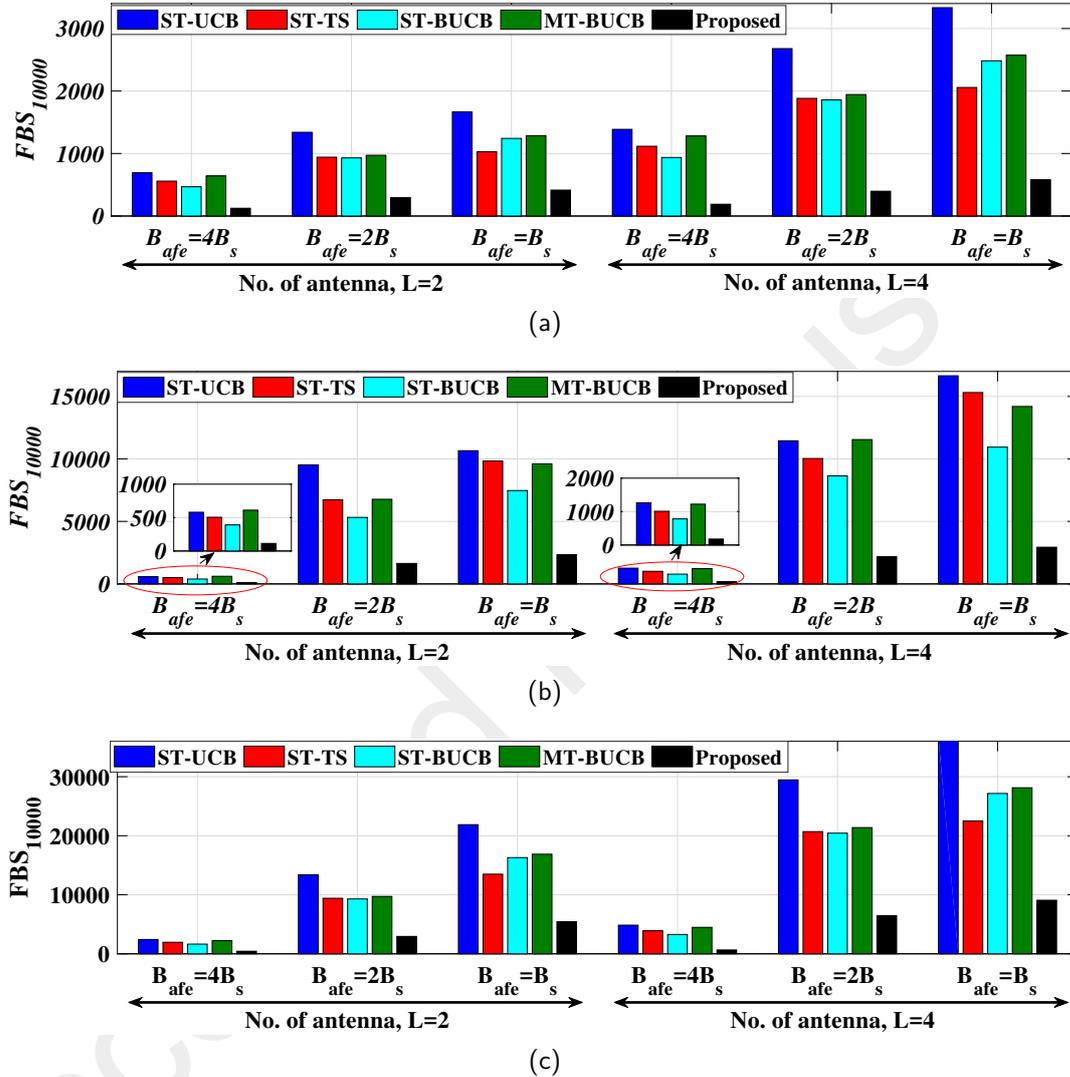
Fig. 5 Total no. of FBS for different values of L and B_s for frequency band statistics in (a) Case 1, (b) Case 2, and (c) Case 3.

lower compared to the DMP using the UCB algorithm. This validates the superiority of the Bayesian online learning algorithms based DMP for wireless communication applications.

5 Conclusions and Future Works

In this paper, a new decision-making policy (DMP) for the frequency band selection in multi-antenna RF energy harvesting (RFEH) circuits is presented. The proposed multi-stage DMP, designed using the Bayesian multi-armed bandit (MAB) algorithm, leads to superior performance in terms of total harvested RF power and number of frequency band switchings over existing

DMPs. Furthermore, the proposed DMP offers higher data throughput for RFEH based WSN nodes than other DMPs. To the best of our knowledge, the proposed DMP is the first one which uses MAB framework for RFEH. Future works include validation of simulation results in real radio environment via integration of the proposed DMP with the multi-antenna RFEH circuit from Powercast [28].

The MAB-based DMP have been used for opportunistic spectrum access in cognitive radio [29–31] and can be extended further using channel bonding techniques within MAB framework [32]. The proposed MAB-based DMP can be extended to other applications such as heterogeneous networks [33], full duplex radios [34], demand response in smart grid [35], electric vehicle charging [36] and relay selection in D2D communications. For instance, full duplex radios in wireless networks promise significant improvement in throughput provided that they transmit over optimal bands and avoid jammers. Such problem can be formulated in MAB framework. Similarly, proposed DMP can be used for scheduling the electric vehicles for charging in order to avoid overload on charging stations and minimize the electricity cost for the consumers.

References

1. I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirco, "A Survey on Sensor Networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, Nov. 2002.

2. B. Rashid and M. H. Rehmani, "Applications of Wireless Sensor Networks for Urban Areas: A Survey," *Journal of Network and Computer Applications*, vol. 60, no. 1, pp. 192–219, Jan. 2016.

3. P. Kamalinejad, C. Mahapatra, Z. Sheng, S. Mirabbasi, V. C. M. Leung and Y. L. Guan, "Wireless Energy Harvesting for the Internet of Things," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 102–108, June 2015.

4. Y. He, X. Cheng, W. Peng and G. Stuber, "A Survey of Energy Harvesting Communications: Models and Offline Optimal Policies," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 79–85, June 2015.

5. F. Akhtar and M. H. Rehmani, "Energy Replenishment Using Renewable and Traditional Energy Resources for Sustainable Wireless Sensor Networks: A Review," *Renewable and Sustainable Energy Reviews*, vol. 45, no. 1, pp. 769–784, May 2015.

6. S. Ulukus, A. Yener, E. Erkip, O. Simeone, M. Zorzi, P. Grover and K. Huang, "Energy Harvesting Wireless Communications: A Review of Recent Advances," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 3, pp. 360–381, Mar. 2015.

7. Drayson Technologies, "RF Energy Harvesting for the Low Energy Internet of Things, " *white paper*, 2015.

8. X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless Networks with RF Energy Harvesting: A Contemporary Survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 757-789, May 2015.

9. J. Costantine, Y. Tawk, S. Barbin, and C. G. Christodoulou, "Reconfigurable Antennas: Design and Applications," *IEEE Proceedings*, vol. 103, no. 2, pp. 424-437, Mar. 2015.

10. P. K. Li, Z. H. Shao, Q. Wang and Y. J. Cheng, "Frequency and Pattern Reconfigurable Antenna for Multistandard Wireless Applications," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, no. 1, pp. 333-336, Feb. 2015.

11. M. Pinuela, P. D. Mitcheson and S. Lucyszyn, "Ambient RF Energy Harvesting in Urban and Semi-Urban Environments," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 7, pp. 2715-2726, July 2013.

12. T. Soyata, L. Copeland, and W. Heinzelman, "RF Energy Harvesting for Embedded Systems: A Survey of Tradeoffs and Methodology," *IEEE Circuits and Systems Magazine*, vol. 16, no. 1, pp. 22–57, Feb. 2016.

13. L. Mohjazi, M. Dianati, G. K. Karagiannidis, S. Muhaidat and M. Al–Qutayri, "RF-Powered Cognitive Radio Networks: Technical Challenges and Limitations," *IEEE Communications Magazine*, vol. 53, no. 4, pp. 94–100, April 2015.

14. D. Mishra, S. De, S. Jana, S. Basagni, K. Chowdhury and W. Heinzelman, "Smart RF Energy Harvesting Communications: Challenges and Opportunities," *IEEE Communications Magazine*, vol. 53, no. 4, pp. 70–78, April 2015.

15. V. Talla, B. Kellogg, B. Ransford, S. Naderiparizi, J. R. Smith and S. Gollakota, "Powering the Next Billion Devices with Wi-Fi," *Communications of the ACM*, vol. 60, no. 3, pp. 83–91, Mar. 2017.

16. C. R. Valenta and G. D. Durgin, "Harvesting Wireless Power: Survey of Energy-Harvester Conversion Efficiency in Far-Field, Wireless Power Transfer Systems," *IEEE Microwave Magazine*, vol. 15, no. 4, pp. 108–120, June 2014.

17. S. J. Darak, C. Moy and J. Palicot, "Smart Decision Making Policy for Faster Harvesting From Ambient RF Sources in Wireless Sensor Nodes," in 13th *International Symposium on Wireless Communication Systems (ISWCS)*, pp. 148-152, Poznan, Poland, Sept. 2016.

18. A. Baknina and S. Ulukus, "Online Scheduling for Energy Harvesting Broadcast Channels with Finite Battery," in *IEEE International Symposium on Information Theory*, pp. 1984-1988, Barcelona, Spain, July 2016.

19. D. Shaviv, A. OZGUR and H. H. Permuter, "Capacity of Remotely Powered Communication, " in *IEEE International Symposium on Information Theory*, pp. 1979-1983, Barcelona, Spain, July 2016.

20. D. T. Hoang, D. Niyato, P. Wing and D. I. Kim, "Opportunistic Channel Access and RF Energy Harvesting in Cognitive Radio Networks, " *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 11, pp. 2039-2052, Dec. 2014.

21. D. Niyato, P. Wing and D. I. Kim, "Performance Optimization for Cooperative Multiuser Cognitive Radio Networks with RF Energy Harvesting Capability, " *IEEE Transactions on Wireless Communication*, vol. 14, no. 7, pp. 3614-3629, Jul. 2015.

22. S. J. Darak, H. Zhang, J. Palicot and C. Moy, "An Efficient Policy for D2D Communications and Energy Harvesting in Cognitive Radios: Go Bayesian!," in 23th *Euro-*

pean Signal Processing Conference (EUSIPCO), pp. 1236–1240, Nice, France, Aug. 2015.

23. S. J. Darak, H. Zhang, J. Palicot and C. Moy, "Decision Making Policy for RF Energy Harvesting Enabled Cognitive Radios in Decentralized Wireless Networks," *Digital Signal Processing*, vol. 60, pp. 33-45, Jan. 2017.

24. P. Auer, N. Cesa–Bianchi and P. Fischer, "Finite–time Analysis of the Multiarmed Bandit Problem," *Machine Learning*, vol. 47, no. 2, pp. 235–256, 2002.

25. A. Garivier and O. Cappé, "The KL-ucb Algorithm for Bounded Stochastic Bandits and Beyond, " *Conference On Learning Theory (COLT)*, pp. 359-376, Budapest, Hungary, July 2011.

26. J. Komiyama, J. Honda and H. Nakagawa, "Optimal Regret Analysis of Thompson Sampling in Stochastic Multi-armed Bandit Problem with Multiple Plays," in 32nd *International Conference on Machine Learning (ICML)*, pp. 1152–1161, Lille, France, July 2015.

27. E. Kaufmann, O. Cappé and A. Garivier, "On Bayesian Upper Confidence Bounds for Bandit Problems," in 15th *Int. Conf. on Artificial Intelligence and Statistics*, pp. 592–600, Canary Islands, Apr. 2012.

28. http://www.powercastco.com/products/development-kits/

29. A. Anandkumar, N. Michael, A. Tang and A. Swami, "Distributed Algorithms for Learning and Cognitive Medium Access With Logarithmic Regret," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, pp. 731–745, April 2011.

30. S. J. Darak, S. Dhabu, C. Moy, H. Zhang, J. Palicot and A. P. Vinod , "Decentralized Spectrum Learning and Access for Heterogeneous Cognitive Radio Networks," *Elsevier Digital Signal Processing*, vol. 37, pp. 13–23, Feb. 2015.

31. N. Modi, C. Moy, P. Mary and S. J. Darak, "Proof-Of-Concept: Spectrum And Energy Efficient Multi-User Cr Network Via Vacancy And Quality Based Channel Selection," accepted in *32nd General Assembly and Scientific Symposium of the URSI (URSI-GASS)*, Canada, Aug. 2017.

32. S. Bukhari, M. H. Rehmani, and S. Siraj, "A Survey of Channel Bonding for Wireless Networks and Guidelines of Channel Bonding for Futuristic Cognitive Radio Sensor Networks," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 924-948, Nov. 2016.

33. Z. Chang, J. Gong, Y. Li, Z. Zhou, T. Ristaniemi, G. Shi, Z. Han and Z. Niu, "Energy Efficient Resource Allocation for Wireless Power Transfer Enabled Collaborative Mobile Clouds," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 12, pp. 3438-3450, Dec. 2016.

34. M. Hanawal, D. Nguyen and M. Krunz, "Jamming Attack on In-Band Full-duplex Communications: Detection and Countermeasures," in 35th *IEEE Annual International Conference on Computer Communications*, pp. 1–9, San Francisco, April. 2015.

35. B. Nikfar, S. Maghsudi and A. J. Vinck, "Multi-armed Bandit Channel Selection for Power Line Communication," in *IEEE International Conference on Smart Grid Communications*, pp. 19–24, Florida, Nov. 2015.

36. W. Tang, S. Bi, and Y. J. Zhang, "Online Charging Scheduling Algorithms of Electric Vehicles in Smart Grid: An Overview," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 76-83, Dec. 2016.