U.R.S.I. FRANCE JS 16

ENERGY AND RADIO SCIENCE

Bayesian Upper Confidence Bound Algorithm Based Decision Making Policy for RF Energy Enabled Wireless Sensor Nodes

S. J. Darak*, Christophe Moy**, Jacques Palicot**

** ECE, Indraprastha Institute of Information Technology, Delhi (IIIT-Delhi), India, sumit@iiitd.ac.in*
*** SCEE, CentraleSupélec/IETR, CS 47601, 35576 Cesson-Sévigné Cedex, France, (christophe.moy, jacques.palicot}@centralesupelec.fr*

Keywords: Multi-armed bandit algorithm, RF energy harvesting, Wireless sensor nodes

Abstract

Emerging applications such as Smart grids and Internet of Things are being considered as an integral part of 5G cellular communication standard. To bring these applications to life, lightweight wireless sensor nodes (WSNs) with ability to sustain in remote locations with low maintenance cost are desired. Radio frequency energy harvesting (RFEH) technique allows WSNs to harvest ambient RF energy transmitted by sources such as base stations, TV tower, WiFi access points, dedicated RF energy sources etc. thereby making them self-sustainable. In this paper, a new decision making policy (DMP) using Bayesian Upper Confidence Bound algorithm for efficient RFEH at WSNs is proposed. The proposed DMP offers WSNs an intelligent way to characterize frequency bands based on their RF potential and select optimum frequency band. Simulations results show that the proposed DMP leads 13% to 47% improvement in average RF energy harvested in a given time over other DMPs. Furthermore, 13% to 71% reduction in subband switching cost and hence, power consumption makes the proposed DMP suitable for the design of lightweight and self-sustainable WSNs.

Introduction

Upcoming 5G cellular communication standard is expected to support infrastructure for various applications such as Smart grids, Internet of Things, Environmental sensing etc [1-3]. In order to bring these applications to life, research efforts are mainly focused on: 1) Wireless sensor nodes (WSNs) capable of sensing the desired information and communicating it to the appropriate location, and 2) Big data algorithms to analyze the received information from all WSNs [2,3]. It is expected that there will be at least 50 billion WSNs deployed worldwide by 2020 [2,3]. Though deployment of such a huge number of WSNs, even at remote locations, is not a challenging task, attention must be paid towards making these battery-operated WSNs self-sustainable with low maintenance cost, environment friendly and light-weight [2,3]. One promising solution is to design WSNs capable of harvesting energy from the environment to improve battery life or even make them battery free and hence, environmental friendly. The design of such lightweight WSNs with energy harvesting capability is a challenging research problem.

Various energy harvesting approaches such as solar, vibrational, thermal, RF etc. have been investigated for WSNs [4-6]. Among these approaches, RFEH is the recent but promising one as it facilitates the conversion of received RF energy from ambient RF sources such as base stations, TV towers, access points, dedicated RF sources etc. into electricity that can be stored and used later for data transmission [4-6]. This is exciting, because in addition of enabling data communication at farther distances, broadcast nature of RF signals also make them valuable for extending the battery life of WSNs. From architecture perspective, RFEH circuits have the advantages of easier integration with RF front-end of WSN terminal thereby making use of same antenna for RFEH as well as data transmission tasks [5,6]. Experimentally, it has been demonstrated that energy in the order of microwatts can be harvested from RFEH circuits [5,6]. In near future, further improvement in harvested RF energy can be expected which will make RFEH feasible for WSNs with resource intensive operations in these emerging applications [1-6].

WSNs with RFEH capability need intelligence to characterize various RF sources (i.e., frequency subbands) based on their RF potential [7]. The cooperative approach for characterization of RF sources, where WSN or central controller share information about RF potential of various subbands with other WSNs, is not efficient. This is because, RF energy in a given subband varies significantly with the location and distance from the RF source [4-6]. Hence, subband which is optimal for one WSN may not be the same for other WSN unless they are located very close to each other. Furthermore, WSNs incur penalty when they switch from one frequency subband to another [7]. This penalty is in terms of dynamic power consumption and time for RFEH due to the reconfiguration overhead. The design of decision making policy (DMP) for WSNs to accurately characterize frequency subbands and choose optimum subband without compromising on switching cost is the focus of the work presented in this paper.

In this paper, a new DMP is proposed which enables WSNs to accurately estimate the RF potential of each frequency subband and choose optimal subband for RFEH. The proposed DMP is designed using Bayesian Multi-Armed Bandit (MAB) algorithm, called Bayesian Upper Confidence Bound (BUCB) algorithm [8]. Simulations results show that the proposed DMP leads 13% to 47% improvement in average RF energy harvested in a given time over other DMPs. Furthermore, 13% to 71% reduction in subband switching cost and hence, power consumption makes the proposed DMP suitable for the design of lightweight and self-sustainable WSNs.

The paper is organized as follows. In Section 1, the assumed model of the WSN network is discussed. The proposed DMP is presented in Section 2. The simulation results are given in Section 3 and Section 4 concludes the paper.

1. Network Model

Consider *N* uniform bandwidth frequency bands of the wideband electromagnetic spectrum to which RFEH circuit can be matched to harvest RF energy. The bandwidth of analog front-end of WSNs is assumed to be equal to the bandwidth of these bands and it is denoted by B_{afe}. Consider the normalized RF power of n^{th} frequency band be μ_n where $n \in \{1, 2, .., N\}$. Average RF power of any subband, $\mu_n \; \forall n$, is assumed to evolve as an i.i.d. random process with variance σ_n^2, stationary and unknown to WSNs. In each time slot, WSN chooses one of the frequency band and performs RFEH after tuning its antenna and other hardware units to the chosen frequency band. Then, k^{th} time slot consists of two sub-slots as given below

$$t_k = \Delta t = t_{1k} + t_{2k} \tag{1}$$

where Δt is duration of time slot, t_{1k} is the time required for subband selection, hardware and protocol reconfiguration, antenna adjustments etc. and t_{2k} is the time available for RFEH over the chosen subband. Let P_k^* be the total RF power harvested over the bandwidth B_{afe} in time slot k using genie-aided DMP (i.e. DMP where WSN knows which subband is optimum for RFEH and hence, chooses the optimum subband in each time slot). Similarly, let P_k be the total RF power harvested in time slot k using any other DMP. Then, total loss in terms of average harvested RF energy, U_K, up to K time slots is given by,

$$\begin{aligned} U_K = P_K^* - P_K &= \sum_{k=1}^{K} \mathbb{E}[P_k^* - P_k] \\ &= K \cdot \mu^* \cdot \eta(\mu^*) \cdot \Delta t - \sum_{k=1}^{K} \mathbb{E}\left[\mu_{n_k} \cdot \eta\left(\mu_{n_k}\right) \cdot (\Delta t - t_{1k})\right] \end{aligned} \tag{2}$$

where μ^* is the average RF power of optimum subband, n_k is the subband chosen by WSN in the k^{th} time slot, $\eta(x)$ is the efficiency of RFEH circuit for input RF power of x and $\mathbb{E}$ is an expectation operator. For any DMP, U_k should be as small as possible. The subband switching cost (SSC) of the DMP is given by Eq. 3 and it should also be as minimum as possible.

$$SSC = \sum_{k=2}^{K} \mathbb{E}\left[1_{\{n_k \neq n_{k-1}\}}\right] \tag{3}$$

where an indicator function: $1_{\{logical\ expression\}} = 1$ if logical expression=true else 0.

2. Proposed Decision Making Policy

In this section, proposed DMP is presented which enables WSNs to characterize different frequency bands based on their RF energy potential. The framework of the proposed DMP is shown in Fig. 1. The proposed DMP involves two decision making tasks: 1) To decide whether to continue RFEH in the frequency band same as that chosen in the previous time slot ($skip\_DM = 1$), 2) If not (i.e., $skip\_DM = 0$), identify another frequency band using BUCB algorithm. Based on the feedback from RFEH circuit about the amount of RF energy harvested, the parameters of BUCB algorithms are updated at the end of time slot.

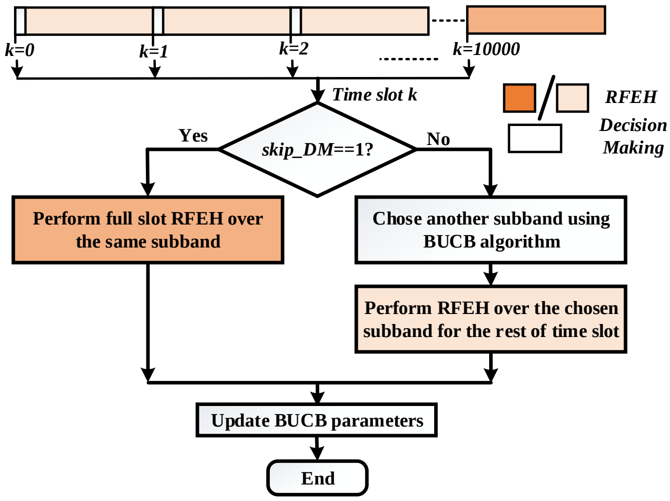

Fig. 1. Proposed decision kaking policy framework.

2.1. Task 1: To decide the value of $skip\_DM$

The benefits of skipping decision making (i.e., $skip\_DM = 1$) are increase in RFEH duration and no need of subband switching. When $skip\_DM = 1$, the RFEH duration is t_{1k} units longer than the same when $skip\_DM = 0$. Higher duration means higher total harvested RF energy. The subband switching incurs penalty in terms of dynamic power consumption which can be avoided when $skip\_DM = 1$. However, such decision (i.e., $skip\_DM = 1$) needs to be taken only when all subbands are characterized with sufficient accuracy. Otherwise, it might lead to the higher regret due to the consecutive selection of sub-optimum subbands. This is because, t_{1k} is usually much smaller than t_{2k} which means that harvested RF energy when $skip\_DM = 1$ might be lower than the harvested energy when $skip\_DM = 0$ if subbands are not chosen properly. In the proposed DMP, such decision making is based on Chebyshev inequality which is given by

$$P\{|\mu_n - \bar{\mu}_n| > \epsilon\} \leq \frac{\sigma_n^2}{T_{n,k} \cdot \epsilon^2} \tag{4}$$

In Eq. 4, $T_{n,k}$ is the number of times the n^{th} subband is chosen by WSN up to time slot, k. Let $X_{n,k}$ be the total normalized RF energy harvested over $T_{n,k}$ time slots during which n^{th} subband is chosen by WSN. Then, estimated RF potential of the subband, $\bar{\mu}_n$, is given by

$$\bar{\mu}_n = \frac{X_{n,k}}{T_{n,k}} \tag{5}$$

Eq. 4 indicates that if subband is chosen sufficient number of times, then its estimated RF potential (i.e., $\bar{\mu}_n$) and actual RF potential (i.e., μ_n) are closed to each other. Based on this observation, it can also be concluded that the change in the value of $\bar{\mu}_n$ from one time slot to another decreases as $T_{n,k}$ increases. In the proposed DMP, in any k^{th} time slot, if the difference between $\bar{\mu}_{n_{k-1}}$ and $\bar{\mu}_{n_k}$ is sufficiently small and estimated RF potential of the subband (i.e., $\bar{\mu}_{n_k}$) is highest among all subbands, then $skip\_DM = 1$. If $skip\_DM$ is changed from 0 to 1 in the k^{th} time slot, it remains equal to 1 for subsequent $\log(k)$ number of time slots. Then, in the subsequent $(\log(k) + 1)^{th}$ time slot, subband selection is governed by BUCB algorithm which is discussed in the next section.

2.2. Task 2: Subband Characterization and Selection Using BUCB Algorithm

The second task of the proposed DMP is to select the appropriate subband in each time slot so that the regret, U_K, given by Eq. 2, is as minimum as possible. To select optimum subband, DMP needs to characterize all subbands accurately. In the proposed DMP, BUCB algorithm is used to characterize and select appropriate subband. BUCB algorithm belongs to the class of multi-armed bandit family which includes other algorithms such as UCB, KL-UCB where KL stands for Kullback-Leibler etc [7-9]. All these algorithms are based on exploitation-exploration trade-off where exploitation refers to the selection of optimum subband (i.e. greedy approach) in each time slot while exploration refers to the selection of subbands which has been chosen fewer number of times in the past in order to characterize them well [8]. The above algorithms are optimal which means that they satisfy following condition [8]

$$\liminf_{k\to\infty}\frac{\mathbb{E}[T_{n,k}]}{\ln k}\geq\frac{1}{\mathrm{K}(\mu_n,\mu^*)},\forall n \tag{6}$$

where $K(p,q)$ denotes the Kullback-Leibler divergence factor. Algorithms which satisfies Eq. 6 offer good trade-off between exploration and exploitation.

The proposed DMP is designed using BUCB algorithm due to its advantages such as lower regret, lower switching cost and lower computational complexity over other algorithms [8,9]. In the proposed DMP, all subband are selected once in the beginning. Then, at subsequent time slots, BUCB algorithm calculates quality index, $G(n,k)$, for each subband which is given by [8],

$$G(n,k)=Q\left\{1-\frac{1}{k};Beta\left[X_{n,k}+1,T_{n,k}-X_{n,k}+1\right]\right\} \tag{7}$$

Since higher the quality index, $G(n,k)$, higher is the RF potential of the subband, the subband having highest quality index is selected in each time slot. Based on the feedback from RFEH circuit, $X_{n,k}$ is updated at the end of time slot. For example, if $\bar{P}_k$ is the normalized RF power harvested from the chosen subband in the k^{th} time slot, then $X_{n,k}=\bar{P}_k$. When $skip\_DM=1$, BUCB algorithm remains idle but $X_{n,k}$ is still updated at the end of time slot. In the next section, simulation results are presented.

3. Simulation Results

In this section, simulation results are presented to evaluate and compare the performance of the proposed DMP with other DMPs. Other DMPs include DMP employing randomization approach, UCB algorithm and BUCB algorithm for subband selection. WSN consists of RFEH circuit which can harvest energy from any one of the 9 subbands. The RF potential of these subbands is governed by distribution, μ. Here, we consider four different distributions of μ which are given by

Case 1: Random distribution
Case 2: {0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90}
Case 3: {0.05, 0.10, 0.35, 0.55, 0.70, 0.75, 0.80, 0.85, 0.90}
Case 4: {0.05, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.50, 0.55}

The Case 1 considers random distribution where RF energy is randomly allocated to subbands in each time slots and hence, it is difficult to accurately characterize subbands. The efficiency of RFEH circuit for a given incident RF power are obtained from powercast RFEH circuit specifications [6]. Note that efficiency of RFEH increases with the increase in incident RF power but the relationship is not linear. Also, minimum normalized RF energy is 0.2. This means that if the incident normalized RF energy is below 0.2, RFEH circuit fails to harvest any RF energy in that time slot. Each numerical result reported hereafter is the average of values obtained over 25 independent experiments and each experiment consider a time horizon of 10000 iterations.

In Fig. 2, average RF energy harvested in percentage with respect to genie-aided DMP are shown for Case 1, Case 2, Case 3 and Case 4. It can be observed that all DMPs harvests almost equal energy in Case 1 due to difficulty in subband characterization because of random distribution of RF energy. The harvested RF energy is highest in Case 3 due to better RF potentials of subbands. Similarly, harvested RF energy is lowest in Case 4 when compared to Case 2 and Case 3. Among all DMPs, random selection based DMP leads to poor performance in all the cases. DMP using BUCB is superior to DMP using UCB while the proposed DMP with intelligent choice of skipping decision making in specific time slots leads to further improvement in harvested RF energy. Numerically, the proposed DMP offers 47%, 15% and 13% improvement in average harvested RF energy over random selection based DMP, DMP using UCB and DMP using BUCB, respectively.

Next, various DMPs are compared in terms of the average SSC. As shown in Fig. 3, SSC of the proposed DMP is lowest while that of random selection based DMP is highest. Even in Case 1, the proposed DMP offers lower SSC over other DMPs. Numerically, proposed DMP leads to 71%, 49%, 13% reduction in average SSC over random selection based DMP, DMP using UCB and DMP using BUCB, respectively.

To summarize, simulation results show that the random selection based DMP is not suitable for RFEH enabled WSNs even when RF potential of subbands are randomly distributed. Among other DMPs, BUCB based DMPs offer superior performance over UCB based DMP in terms of average harvested RF energy and average SSC. Note that SSC of UCB is highest in Case 1. The proposed DMP with smart decision making approach leads to significant improvement in harvested RF energy as well as SSC when compared to BUCB based DMP. Thus, proposed DMP is superior and energy-efficient.

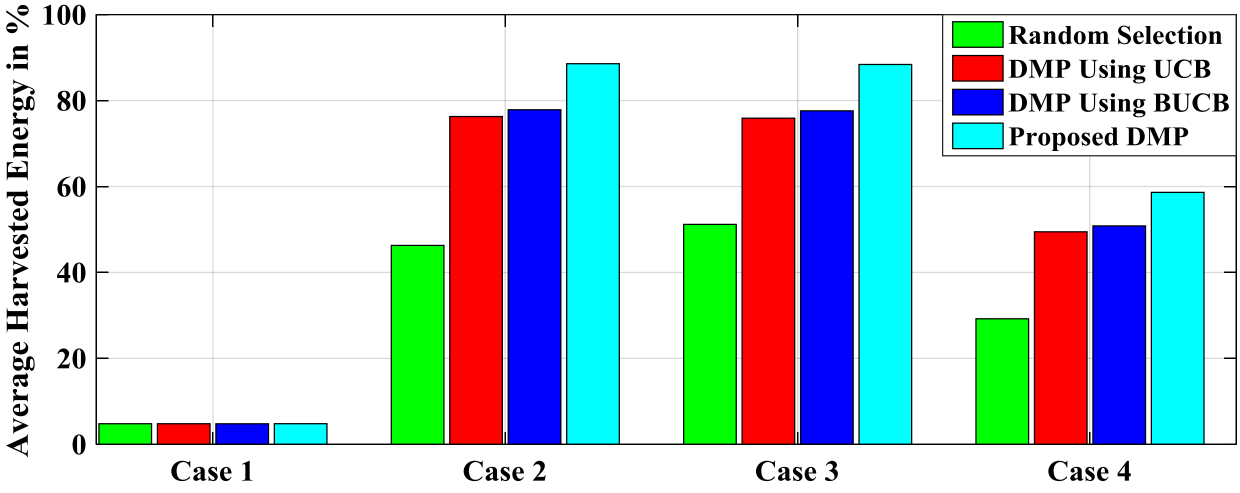

Fig. 2. Average harvested RF power using different DMPs in % with respect to the power harvested using Genie-aided DMP for four different cases of distributions, μ.

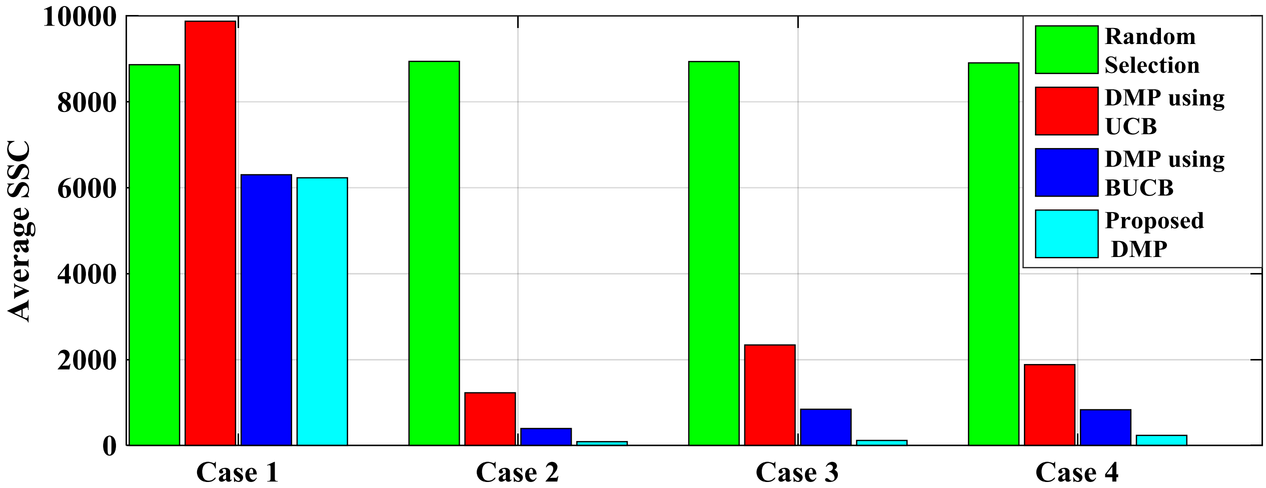

Fig. 3. Average SSC of various DMPs for four different cases of distributions, μ.

4. Conclusions

In this paper, a new decision making policy (DMP) using Bayesian Upper Confidence Bound algorithm for efficient RFEH at WSNs is proposed. The proposed DMP offers WSNs an intelligent way to characterize frequency bands based on their RF potential and select optimum frequency band. Simulations results show that the proposed DMP leads 13% to 47% improvement in average RF energy harvested in a given time over other DMPs. Furthermore, 13% to 71% reduction in subband switching cost and hence, power consumption makes the proposed DMP suitable for the design of lightweight and self-sustainable WSNs. Future works include validation of functionality of various DMPs using real radio signals and RFEH circuits.

Acknowledgements

The authors would like to thank Department of Science and Technology (DST), Government of India for INSPIRE fellowship in support of this work.

Citations

[1] J. Thompson, X. Ge, H. Wu, R. Irmer, H. Jiang, G. Fettweis and S. Alamouti, “5G Wireless Communication Systems: Prospects and Challenges”, *IEEE Communication Magazine*, vol. 52, no. 2, pp. 62-64, Feb. 2014.

[2] M. Alsheikh, D. Hong, D. Niyato, H. Tan and S. Lin, “Markov Decision Processes with Applications in Wireless Sensor Networks: A Survey”, *IEEE Communications Surveys & Tutorials*, vol. 3, no. 4, pp. 1239-1267, Aug. 2015.

[3] M. Alsheikh, S. Lin, D. Niyato and H. Tan, “Machine Learning in Wireless Sensor Networks: Algorithms, Strategies, and Applications”, *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 1996-2018, Nov. 2014.

[4] 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.
[5] Drayson Technologies, "RF Energy Harvesting for the Low Energy Internet of Things, " *white paper*, 2015.
[6] http://www.powercastco.com/documentation/
[7] 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 European Signal Processing Conference* (EUSIPCO), pp. 1236–1240, Nice, France, Aug. 2015.
[8] E. Kaufmann, O. Cappe and A. Garivier, "On Bayesian Upper Confidence Bounds for Bandit Problems," in 1*5th Int. Conf. on Artificial Intelligence and Statistics*, pp. 592–600, Canary Islands, Apr. 2012.
[9] S. J. Darak, A. Nafkha, C. Moy and J. Palicot, "Is Bayesian Multi-armed Bandit Algorithm Superior?: Proof-of-Concept for Opportunistic Spectrum Access in Decentralized Networks, accepted in *11th International Conference on Cognitive Radio Oriented Wireless Networks* (CROWNCOM), Grenoble, France, May 2016.