# Robust Bayesian Learning for Wireless RF Energy Harvesting Networks

Nof Abuzainab[1], Walid Saad[1], and Behrouz Maham [2]
[1]Wireless@VT, Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA
Emails:{nof, walids}@vt.edu
[2] School of Engineering, Nazarbayev University, Astana, Kazakhstan, Email: behrouz.maham@nu.edu.kz

***Abstract*—In this paper, the problem of adversarial learning is studied for a wireless powered communication network (WPCN) in which a hybrid access point (HAP) seeks to learn the transmission power consumption profile of an associated wireless transmitter. The objective of the HAP is to use the learned estimate in order to determine the transmission power of the energy signal to be supplied to its associated device. However, such a learning scheme is subject to attacks by an adversary who tries to alter the HAP's learned estimate of the transmission power distribution in order to minimize the HAP's supplied energy. To build a robust estimate against such attacks, an unsupervised Bayesian learning method is proposed allowing the HAP to perform its estimation based only on the advertised transmisson power computed in each time slot. The proposed robust learning method relies on the assumption that the device's true transmission power is greater than or equal to advertised value. Then, based on the robust estimate, the problem of power selection of the energy signal by the HAP is formulated. The HAP optimal power selection problem is shown to be a discrete convex optimization problem, and a closed-form solution of the HAP's optimal transmission power is obtained. The results show that the proposed robust Bayesian learning scheme yields significant performance gains, by reducing the percentage of dropped transmitter's packets of about $85\%$ compared to a conventional Bayesian learning approach. The results also show that these performance gains are achieved without jeopardizing the energy consumption of the HAP.**

## I. Introduction

Radio frequency (RF) energy harvesting is one of the most promising technologies to operate massive self-powered networks, such as the Internet of Things (IoT) [1]. The reliance on RF signals for energy supply makes RF energy harvesting favourable since it will be easy to be implemented and integrated into current wireless systems. Moreover, RF energy harvesting offers a reliable method to supply energy, as opposed to traditional energy harvesting techniques that rely on ambient sources such as solar or wind in which the amount of energy harvested strongly depends on environmental factors. The RF energy source can be a wireless access point, known as a hybrid access point (HAP), that can be configured to provide simultaneous communication and energy supply. Another example of such RF energy sources could be a power beacon that operates independently from the HAP. Thus, the reliable energy supply provided by these dedicated RF energy sources allows the network to better serve devices with stringent quality-of-service (QoS) requirements. However, RF energy supply incurs extra energy expenditure by the HAP since the HAP must send dedicated signals for energy transfer to its associated devices. Hence, one of the main technical challenges in a wireless powered communication network (WPCN) is to find energy efficient resource allocation mechanisms that determine the optimal energy that should be supplied by the HAP to its associated devices in order to meet their QoS requirements, as pointed out in [2].

There has been considerable interest in designing energy efficient wireless resource allocation schemes suitable for WPCNs [3]–[5]. In [3] and [4], a WPCN composed of a HAP serving multiple mobile users in a time division manner is considered. A centralized approach for maximizing the total network throughput in a WPCN is adopted in [3] by finding the optimal time fraction allocated to each user. The authors in [4] also propose a centralized approach for maximizing the proportional fairness sum of users' rates by finding the optimal transmission power and the harvesting duration for each user. A distributed noncooperative game theoretic approach is proposed in [5] which considers a WPCN composed of several source destination pairs operating in the same frequency band. Thus, each source finds the minimum transmit power that meets the QoS and harvesting constraints of its associated destination.

However, some of these works assume that the HAP transmits an energy signal with fixed power, and that the device consumes all of the harvested energy for transmission in the same slot. Moreover, the existing literature typically assumes that the HAP knows apriori the QoS and harvesting requirements of all its associated devices. These assumptions are not very realistic especially in emerging IoT systems in which devices have very diverse characteristics and requirements. Further, the transmission power of each wireless device depends on its adopted power control policy which is often a function of the device's QoS requirements and its traffic characteristics. One promising approach is to use machine learning techniques [6], [7] in order to form a more realistic estimate of the distribution of the transmission power consumed by the wireless device. This enables the HAP to predict the transmission energy consumed by each associated wireless device and determine the required energy to be supplied for the device.

There is still little prior work that considers learning for

This research was supported by the US National Science Foundation under Grant CNS-1524634.

978-3-901882-90-6/17 ©2017 IFIP

RF energy harvesting [8]–[10], and most of this prior art has considered learning at the device's end. In [8], two algorithms based on supervised machine learning techniques: the linear regression (LR) and the decision trees (DT) are proposed in order to predict the RF energy that can be harvested in a certain frequency band and a given time slot. In [9], the problem of energy efficient RF energy harvesting for a wireless device is considered. To this end, an unsupervised Bayesian learning approach is proposed that allows the energy harvesting wireless device to predict the ambient RF energy availability in each time slot. Then, based on the predicted RF energy, the optimal sleep and harvesting policy are determined to minimize the consumed energy. An online convex optimization method is proposed in [10] to allow an energy harvesting wireless transmitter to predict the energy available in a current time slot based on measurements from previous time slots.

Despite its benefits, learning can be vulnerable to a man-in-the-middle (MITM) attack by a malicious user that can alter the data used by the learning algorithm and, consequently, degrade the system performance. MITM attacks constitute serious threat in the emerging IoT systems [11] due to the fact that many IoT machine type devices have limited computational capabilities and can not implement strong security mechanisms. Thus, their security can be easily compromised. In WPCNs, an adversary, through an MITM attack, can modify the transmission power consumption profile of the wireless device. Thus, the HAP will be misled into supplying less energy to the associated wireless device, which will eventually exhaust the device's battery. Hence, learning algorithms for WPCNs must be designed to be robust against such attacks. Existing works that study security for RF energy harvesting considered either jamming attacks [12], [13] or eavesdropping [14], [15]. To the best of our knowledge, there is still no work that considers attacks on learning within RF energy harvesting networks.

The main contribution of this paper is to introduce a novel learning scheme for RF energy harvesting that allows the HAP to form a reliable estimate of the power consumption profile of each associated wireless device. The proposed learning scheme is based on unsupervised Bayesian learning, and it relies only on the received power from the wireless device in each time slot and on channel state information (CSI). Thus, it does not result in extra comunications and energy costs. However, the dependence of the proposed learning scheme on the device's received power makes it subject to attacks by an adversary that is interested in depleting the battery of the wireless device. The adversary can achieve this end by altering the formed estimate of the power consumption profile through performing MITM attack. To counter such attacks, the estimate is built by the HAP based on the assumption that the true value of the transmission power of the wireless transmitter is censored by a potential malicious user, and that true transmission power value is greater than or equal to the advertised value. Then, based on this robust estimate, the HAP determines the transmission power of the energy signal to be delivered to its associated device such that the HAP's payoff is maximized. We formulate the problem of optimal power selection by the HAP as discrete convex optimization problem, and we obtain a closed-form expression for the optimal transmission power. The results show that the proposed robust Bayesian learning scheme yields significant performance gains, by reducing the percentage of dropped transmitter's packets of about $85\%$ compared to the conventional Bayesian learning approaches. The results also show that these performance gains are achieved without jeopardizing the energy consumption of the HAP.

The paper is organized as follows. Section I presents the system model. Section II presents the attacker model. Section III presents the defensive learning strategy of the HAP and the power selection mechanism of the energy signal. Section IV presents the simulation results. Finally, conclusions are drawn in section V.

## II. System Model

Consider a WPCN composed of a HAP [2] serving a set of wireless devices over orthogonal frequency channels. For each device $i$, the HAP can act as both an energy supplying device that performs wireless power transfer to the device and as an access point that collects the information from the device. The HAP is connected to a constant power supply such as a smart grid whereas the device is not connected to any additional energy supply, and, hence, it relies on the energy harvested from the HAP.

Simultaneous uplink and downlink transmissions are assumed [3] where the HAP transmits the energy signal and receives the uplink transmission from the device over two separate frequency bands. In each time slot $t$ of duration $T$ seconds, the HAP transmits an energy signal with power $P_{at}$ from a discrete set $\mathcal{P}_a$ to device $i$. In the set $\mathcal{P}_a$, the power values are multiples of $h$ where $0 \leq h \leq 1$. During the uplink, device $i$ uses the harvested energy from the previous time slots to transmit its data to the HAP with power $P_{it}$ which takes value from a discrete set $\mathcal{P}_i$. In the uplink phase, device $i$ determines the value of $P_{it}$ based on its QoS requirements, its traffic characteristics as well as the energy available in its battery. The uplink and downlink channels between device $i$ and the HAP are modeled as block Rayleigh fading channels with coefficients $h_{D,it}$ and $h_{U,it}$ for downlink and uplink, respectively. These channel gains do not change within time slot $t$. Thus, the amount of energy harvested by device $i$ at time slot $t$ is $E_{it} = \eta|h_{D,it}|^2 P_{at} T$ where $0 < \eta < 1$ is the energy harvesting efficiency.

In our model, the devices served by the HAP have heterogeneous traffic characteristics and QoS requirements. Hence, it is not practical to assume that the HAP has prior knowledge of the traffic characteristic and the QoS requirement of each device and, consequently, its power consumption profile. Instead, the HAP uses unsupervised Bayesian learning in order to estimate the power/energy consumption distribution of each device. The HAP relies only on the received signal power $P_{r,it}$ in the uplink in order to update its estimate in each time period $t$. By using such a method, there is no need for the device

to explicitly send energy requests to the HAP. Such energy requests would waste the energy stored in the device. Thus, assuming that the HAP has full channel state information, it computes the transmission power consumed by the device in time slot $t$ as: $P_{it} = \frac{P_{r,it}}{|h_{U,it}|^2}$.

In nonparametric Bayesian learning, the Dirichlet distribution [16] is often used to model a parameter with unknown distribution since it is a conjugate prior of the multinomial distributon. Given that $N_1, N_2, ..., N_K$ independent observations of events $E_1$, $E_2$,...,$E_K$ are made, and under the assumption that the prior distribution of the probability vector $\boldsymbol{p} = (p_1, p_2, ..., p_K)$ of the events $E_1, E_2, ..., E_K$ is Dirichlet distributed with parameter $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_K)$ ($\alpha_1$, $\alpha_2$,...,$\alpha_K > 0$), the posterior probability $\boldsymbol{p}$ given the observations $\boldsymbol{N} = (N_1, N_2, ..., N_K)$ will follow a Dirichlet distribution of order $K$ with parameter $\boldsymbol{\alpha} + \boldsymbol{N}$ as follows

$$f(\boldsymbol{p}|\boldsymbol{N}) = \prod_{i=1}^{K} {p_i}^{\alpha_i + N_i - 1} \cdot \frac{1}{B(\alpha_i + N_i)}, \tag{1}$$

where $\Gamma(.)$ is the gamma function and $B(.)$ is a normalizing factor given by

$$B(\boldsymbol{x}) = \frac{\prod_{i=1}^{K} \Gamma(x_i)}{\Gamma(\sum_{i=1}^{K} x_i)}. \tag{2}$$

The posterior expected probability $\mathbb{E}[p_i|\boldsymbol{N}]$ of observing event $E_i$ given the observations will then be

$$\mathbb{E}[p_i|\boldsymbol{N}] = \frac{\alpha_i + N_i}{\sum_{j=1}^{K} \alpha_j + N_j}. \tag{3}$$

In our system, during each slot $t$, the HAP seeks to estimate the probability distribution of the power consumption profile of device $i$ based on the observations in the previous time slots to determine the suitable energy signal transmission power $P_{at}$. The observation made in each time period $t$ is the transmission power value $P_{it}$ that is computed from the received uplink signal power $P_{r,it}$ of the device. The HAP only considers positive transmission power values to form its estimate. This is because many wireless devices, especially machine type devices, have very bursty traffic characteristic, and thus, the number of time slots when the device transmits will be negligble compared to when the device does not transmit.

In the considered model, we assume that a malicious adversary seeks to launch a MITM attack on the HAP's learning mechanism so as to deplete the battery of the device by altering the learned power consumption distribution by the HAP. In the context of learning, this is commonly known as data poisoning. Next, we explain how the malicious user will interact with the studied HAP's learning scheme.

## III. Attacker Model

The malicious user attempts to alter the power consumption distribution learned by the HAP by performing a wireless MITM attack [17]. In a wireless MITM, a malicious user is assumed to be capable of impersonating the HAP to lure the device to connect to it. After the device connects to the malicious user, the adversary adverstises a different uplink frequency band in order for the HAP not to receive the transmissions directly from the device. Then, the advsersary obtains the device's information (such as device ID and security parameters) and subsequently impersonates the device. In this attack, the adversary intercepts the uplink packet transmitted by the device in time slot $t$ and retransmits the packet to the HAP with the minimum possible power value $P_{mt} \in \{P_j \in \mathcal{O}_t \text{ s.t. } P_j \leq P_{it}\}$ that maintains a low risk of being detected by the HAP. The set $\mathcal{O}_t \subset \mathcal{P}_i$ is the set of device's transmission power values oberved by the malicious user up to time slot $t$. Full channel state information is assumed to be available at the malicious user, and, thus, the malicious user can perfectly recover the transmission power value $P_{it}$ of the device from the received power value $P_{r,it}$.

In our model, the adversary has no knowledge of the HAP's exact attack detection method or its defensive strategy. Thus, in order to limit the risk of being detected by the HAP, the malicious user chooses the minimum transmission power $P_{mt}$ at each time slot $t$ such that the Kullback–Leibler (KL) distance between the estimates that are based on the real and modified power values respectively do not exceed a predefined value $r$. Let $R_t$ and $M_t$ the estimates of the probability distribution based on the real and modified power values respectively at time slot $t$. The malicious user uses the conventional Bayesian learning method based on Dirichlet distribution described in Section I to determine the estimates $R_t$ and $M_t$. For the attacker, the value of $r$ captures the risk of being detected by the HAP. Here, the higher the value of $r$, the higher the probability that the attacker will be detected by the HAP. Thus, the attacker selects the transmission power according to the following optimization problem

$$\min_{P_{mt}} P_{mt}$$
$$\text{s.t. } D_{KL}(R_t||M_t) \leq r, \;\; 0 \leq P_{mt} \leq P_{it}, \; P_{mt} \in \mathcal{O}_t. \tag{4}$$

In the studied system, the attacker has no prior information on the power consumption distribution, and, hence, the prior distribution of the probabilities of transmissions with powers in $\mathcal{O}_t$ is assumed to be uniform, i.e., Dirichlet with parameter $\mathbf{1}$. For a set of observed transmission power values $\mathcal{O}_t$, let $\phi_{i,t}$ be the number of occurrences of transmission power value $P_i \in \mathcal{O}_t$ up to time slot $t$ and $\omega_{i,t}$ be the number of times the malicious user transmits with power value $P_i$ up to time slot $t$. Define the vectors $\boldsymbol{\phi}_t = (\phi_{i,t})_{P_i \in \mathcal{O}_t}$ and $\boldsymbol{\omega}_t = (\omega_{i,t})_{P_i \in \mathcal{O}_t}$. Thus, the posterior distributions $R_t$ and $M_t$ follow the Dirichlet distribution with parameters $\mathbf{1} + \boldsymbol{\phi}_t$ and $\mathbf{1} + \boldsymbol{\omega}_t$ respectively.

Thus, the expected probabilities $\bar{p}_{i,t}$ and $\bar{q}_{i,t}$ of observing power value $P_i$ based on the estimates $R_t$ and $M_t$ are, respectively, $\bar{p}_{i,t} = \frac{\phi_{i,t}+1}{\sum_{j \in \mathcal{O}_t} \phi_{j,t} + |\mathcal{O}_t|} = \frac{\phi_{i,t}+1}{t+|\mathcal{O}_t|}$ and $\bar{q}_{i,t} = \frac{\omega_{i,t}+1}{\sum_{j \in \mathcal{O}_t} \omega_{j,t} + |\mathcal{O}_t|} = \frac{\omega_{i,t}+1}{t+|\mathcal{O}_t|}$. The KL distance of $R_t$ and $M_t$ is then given by

$$D_{KL}(R_t||M_t) = \sum_{i \in \mathcal{O}_t} \bar{p}_{i,t} \log \frac{\bar{p}_{i,t}}{\bar{q}_{i,t}} = \sum_{i \in \mathcal{O}_t} \frac{\phi_{i,t}+1}{t+|\mathcal{O}_t|} \log \frac{\phi_{i,t}+1}{\omega_{i,t}+1}. \tag{5}$$

Thus, the KL distance $D_{KL}(R_t||M_t)$ depends on the power value $P_{mt}$ chosen by the malicious user at time slot $t$. The following proposition provides a simplified version of the constraint on the KL distance in (4) in order to avoid computing the KL distance for each power value $P_{mt}$ to find the miminum transmission power $P^*_{mt}$.

**Proposition 1.** Let $P_l$ be the observed device power value at time slot $t$. The attacker selects the minimum transmission power $P^*_{mt} = P_k < P_l$ at time slot $t$ such that

$$\frac{\omega_{k,t-1}+1}{\omega_{k,t-1}+2} \leq e^{\frac{(t+|\mathcal{O}_t|)\cdot r-(t-1+|\mathcal{O}_{t-1}|)\cdot r'+\kappa_{l,t-1}}{\phi_{k,t-1}+1}}, \quad (6)$$

where

$\kappa_{l,t-1} =$
$(\phi_{l,t-1}+1)\log(\frac{\phi_{l,t-1}+1}{\omega_{l,t-1}+1}) - (\phi_{l,t-1}+2)\log(\frac{\phi_{l,t-1}+2}{\omega_{l,t-1}+1})$.
Otherwise, The attacker chooses $P^*_{mt} = P_l$.

*Proof.* First, let $\omega'_{k,t-1} = \omega_{k,t-1}+1$ and $\phi'_{k,t-1} = \phi_{k,t-1}+1$. The KL distance at time slot $t$ is given by

$$D_{KL}(R_t||M_t) = \sum_{i\in\mathcal{O}_t} \bar{p}_{i,t}\log\frac{\bar{p}_{i,t}}{\bar{q}_{i,t}}$$
$$= \frac{(\phi'_{l,t-1}+1)}{t+|\mathcal{O}_t|}\log\big(\frac{\phi'_{l,t-1}+1}{\omega'_{l,t-1}}\big) + \frac{\phi'_{k,t-1}}{t+|\mathcal{O}_t|}\log(\frac{\phi'_{k,t-1}}{\omega'_{k,t-1}+1})$$
$$+ \sum_{i\in\mathcal{O}_t, i\neq l,k}\frac{\phi'_{i,t-1}}{t+|\mathcal{O}_t|}\log\frac{\phi'_{i,t-1}}{\omega'_{i,t-1}}. \quad (7)$$

Let $F(R_t||M_t) = (t+|\mathcal{O}_t|)\cdot D_{KL}(R_t||M_t)$. Then,

$$F(R_t||M_t) - F(R_{t-1}||M_{t-1})$$
$$= (\phi'_{l,t-1}+1)\log(\frac{\phi'_{l,t-1}+1}{\omega'_{l,t-1}}) - \phi'_{l,t-1}\log(\frac{\phi'_{l,t-1}}{\omega'_{l,t-1}})$$
$$+\phi'_{k,t-1}\log(\frac{\omega'_{k,t-1}}{\omega'_{k,t-1}+1}). \quad (8)$$

Given that the value of the divergence at time slot $t-1$ is $D_{KL}(R_{t-1}||M_{t-1}) = r' \leq r$. The constraint on the divergence at time slot $t$ translates to
$F(R_t||M_t) - F(R_{t-1}||M_{t-1}) \leq (t+|\mathcal{O}_t|)r - (t-1+|\mathcal{O}_{t-1}|)r'$.

From (8), we get the constraint

$\log(\frac{\omega'_{k,t-1}}{\omega'_{k,t-1}+1}) \leq \frac{(t+|\mathcal{O}_t|)\cdot r-(t-1+|\mathcal{O}_{t-1})\cdot r'+\kappa_{l,t}}{\phi'_{k,t-1}}$
where $\kappa_{l,t} = \phi'_{l,t-1}\log(\frac{\phi'_{l,t-1}}{\omega'_{l,t-1}}) - (\phi'_{l,t-1}+1)\log(\frac{\phi'_{l,t-1}+1}{\omega'_{l,t}})$.

Thus, we get
$\frac{\omega'_{k,t-1}}{\omega'_{k,t-1}+1} \leq e^{\frac{(t+|\mathcal{O}_t|)\cdot r-(t-1+|\mathcal{O}_{t-1}|)\cdot r'+\kappa_{l,t}}{\phi'_{k,t-1}}}$,
and
$\frac{\omega_{k,t}+1}{\omega_{k,t}+2} \leq e^{\frac{(t+1+|\mathcal{O}_t|)\cdot r-(t-1+|\mathcal{O}_{t-1}|)\cdot r'+\kappa_{l,t-1}}{\phi_{k,t-1}+1}}$.

□

Proposition 1 transforms the constraint on the KL distance given by (5) to a constraint on $\omega_{k,t-1}$ – the number of times the malicious user transmits with a power value $P_k$ up to time slot $t-1$. Thus, finding the optimal power value for the optimization problem does not require computing the KL distance to check the constraint for each power value $P_k$. It suffices to check the constraint on $\omega_{k,t-1}$ given by (6).

Thus, by transmitting with a power value less than $P_{it}$, the attacker misleads the HAP into believing that the device is consuming a lower transmission power. To thwart such attacks, the HAP, on the other hand, utilizes a defensive/robust learning mechanism. The details of the learning mechanism are explained in the following section.

## IV. HAP Defensive Strategy

### A. HAP Information Censoring Based Learning Mechanism

In order to reduce the effect of a potential MITM on the updated estimate of probability distribution at each time slot $t$, the HAP assumes that the true transmission power of the device is higher than the transmission power computed from the received signal at time period $t$ i.e. the true transmission power belongs to the set $\{P_j \in \Omega_t \text{ s.t. } P_j \geq P_{mt}\}$ where $\Omega_t$ is the set of power values observed by the HAP up to time slot $t$. Thus, the HAP constructs an estimate of the power consumption distribution based on this belief. In this case, the observation of the true transmission power of the device is considered to be censored. The general definition of a censored observation [18] is given next.

**Definition 1.** An observation is said to be *censored* when it is not fully observable but rather it is reported that it belongs to a subset $\mathcal{C}$ of the set of events $\{E_1, E_2, ..., E_K\}$.

Thus, in the case of censored observations, the estimate of the probability distribution of the events $\{E_1, E_2, ..., E_K\}$ will depend on the received reports about the censored observations [18]. In our problem, the report at time slot $t$ is that the true transmission power of the device belongs to the set $\mathcal{C}_t = \{P_j \in \Omega_t \text{ s.t. } P_j \geq P_{mt}\}$. In this case, the joint distribution of the probabilities $\boldsymbol{p}$ of the transmission powers in $\Omega_t$ depends on $\lambda_{\mathcal{C}|k}$, the conditional probability of getting a report $\mathcal{C}$ given that the actual transmission power is $P_k$. Denote by $\Lambda$ the matrix of $\lambda_{\mathcal{C}|k}\ \forall \mathcal{C}, k$ and $\mathbb{C}_t$ the set of reports up to time slot $t$. Then, the likelihood of the reports given $\boldsymbol{p}$ and $\Lambda$ will be

$$f(\{\mathcal{C}_k\}_{k=1}^t|\boldsymbol{p},\Lambda) = \prod_{\mathcal{C}\in\mathbb{C}_t}\Big(\sum_{i \text{ s.t.} P_i\in\mathcal{C}} p_i\lambda_{\mathcal{C}|i}\Big)^{N_{\mathcal{C},t}}. \quad (9)$$

where $\boldsymbol{N}_{\mathbb{C}_t} = (N_{\mathcal{C},t})_{\mathcal{C}\in\mathbb{C}_t}$ is the vector of counts of observed reports up to time slot $t$ and $N_{\mathcal{C},t}$ is the number of times the set $\mathcal{C}$ is reported up to time slot $t$. As seen in (9), the likelihood $f(\{\mathcal{C}_k\}_{k=1}^t|\boldsymbol{p},\Lambda)$ depends on $\mu_{\mathcal{C},i} = p_i\lambda_{\mathcal{C}|i}$ the joint probability of receiving report $\mathcal{C}$ when the true transmission power is $P_i$. Let $\boldsymbol{\mu}$ be the matrix of $\mu_{\mathcal{C},i}\ \forall \mathcal{C}, i$. The joint outcomes $(P_{it}, \mathcal{C}_t)$ at time slot $t$ are then distributed with parameter $\boldsymbol{\mu}$. Hence, as shown in [18], we can assume that the prior distribution of $\boldsymbol{\mu}$ follows a Dirichlet distribution with parameter $\boldsymbol{a}$ where each entry $a_{\mathcal{C},i}$ is the parameter corresponding to $\mu_{\mathcal{C},i}$. Consequently, the prior distribution of the probability vector $\boldsymbol{p}$ at time slot $t$ follows a Dirichlet distribution with parameter $\boldsymbol{\beta}_t = (\beta_{i,t})_{i=1}^K$ where $\beta_{i,t} = \sum_{\mathcal{C}\in\mathbb{C}_{i,t}} a_{\mathcal{C},i}$ and $\mathbb{C}_{i,t}$ is the set of all reported sets that include $P_i$ up to time slot $t$.

Under these assumptions, the posterior distribution of the probability vector $\boldsymbol{p}$ of transmission powers in $\Omega_t$ at each time slot $t$ is shown [18], [19] to belong to a class of generalized Dirichlet distributions and is thus given by $f(\boldsymbol{p}|\boldsymbol{N}_{\mathbb{C}_t}, \Lambda) = D(\boldsymbol{\beta_t}, \Lambda, \boldsymbol{N}_{\mathbb{C}_t})$. In general, the distribution $D(\boldsymbol{b}, Z, \boldsymbol{d})$ has a probability mass function

$$g(\boldsymbol{p}, \boldsymbol{b}, Z, \boldsymbol{d}) = \frac{f(\boldsymbol{p}, \boldsymbol{b}) \prod_k (\sum_i z_{ki} p_k)^{d_i}}{R(\boldsymbol{b}, Z, -\boldsymbol{d})}, \tag{10}$$

where $f(\boldsymbol{p}, \boldsymbol{b})$ is the pdf of a Dirichlet distribution with parameter $\boldsymbol{b}$ and $R(\boldsymbol{b}, Z, -\boldsymbol{d})$ is a Carlson's bidimensional hypergeometric function which can be expressed as $R(\boldsymbol{b}, Z, -\boldsymbol{d}) = \frac{B(Z'\boldsymbol{b}+\boldsymbol{d})}{B(Z'\boldsymbol{b})}$ where $B(.)$ is the normalizing factor of the Dirichlet distribution given by (2). The posterior mean of the probability $p_i$ of transmitting with power value $P_i$ given the reports counts $\boldsymbol{N}_{\mathbb{C}_t}$ is [18]

$$\mathbb{E}[p_i|\boldsymbol{N}_{\mathbb{C}_t}] = \frac{a_{\sum}}{a_{\sum}+t}\mathbb{E}[p_i] + \frac{t}{a_{\sum}+t}\Big(\frac{N_{\{i\},t}}{t} + \sum_{\mathcal{C}\in\mathbb{C}_{i,t}\setminus\{i\}} \frac{N_{\mathcal{C},t}}{t}\frac{a_{\mathcal{C},i}}{\sum_{i\in\mathcal{C}} a_{\mathcal{C},i}}\Big), \tag{11}$$

where $a_{\sum}$ is the sum of elements of the Dirichlet hyperparameter $\boldsymbol{a}$, and $\mathbb{E}[p_i]$ is the expectation of the prior distribution. Since the prior distribution is Dirichlet with paramter $\boldsymbol{\beta}_t$, the expectation is $\mathbb{E}[p_i] = \frac{\beta_{i,t}}{\sum_j \beta_{j,t}}$ .

Let $I_t$ be the updated estimate probability distribution by the HAP by the end of time slot $t$. Based on the formed estimate $I_t$ at the end of time slot $t$, the HAP selects in the subsequent time slot $t+1$, the transmission power of the energy signal $P_{a,t+1}$ that maximizes its utility while ensuring that the battery of the device is not depleted, as explained next.

### *B. Energy Signal Power Selection*

During the downlink at slot $t$, the HAP uses the last updated estimate $I_{t-1}$ of the power consumption distribution to decide on the power value $P_{at}$ of the energy signal. Since in the first time slot the HAP has not received any observations, it transmits with the maximum power $P_{a,\max}$. In the subsequent time slots, the objective of the HAP is to find the optimal transmission power value $P^*_{at}$ that maximizes its utility while not depleting the device's battery. To achieve this end, the HAP selects the transmission power such that the energy supplied is greater than the expected transmission energy consumed by the device. In time slot $t$, $I_{t-1}$ is the most updated estimate of the power consumption probability distribution at the HAP. Hence, the HAP assumes that each $P_{ik}$ is distributed according to $I_{t-1}$. Thus, the constaint is given by

$$\eta T\Big(\sum_{k=1}^{t-1} |h_{D,ik}|^2 P^*_{ak} + |h_{D,it}|^2 P_{at}\Big) \geq \sum_{k=1}^{t} \mathbb{E}[P_{ik}] \cdot T, \tag{12}$$

where $P^*_{ak}$ is the chosen transmission power value of the energy signal transmitted by the HAP at time slot $k$ ($1 \leq k \leq t-1$) and the expectation is with respect to the distribution $I_{t-1}$. Thus, the expected transmission power value of the device $\mathbb{E}[P_{ik}] (1 \leq k \leq t-1)$ is given by $\mathbb{E}[P_{ik}] = \sum_{i\in\Omega_t} \bar{p}_{i,t} P_i$ where $\bar{p}_{i,t}$ is the posterior expectation $\mathbb{E}[p_i|\boldsymbol{N}_{\mathbb{C}_t}, \Lambda]$ given by (11).

Since the transmission power values are positive, the constraint (12) becomes

$$P_{at} \geq \left[\frac{\sum_{k=1}^{t} \mathbb{E}[P_{ik}] - \eta \sum_{k=1}^{t-1} |h_{D,ik}|^2 P^*_{ak}}{\eta |h_{D,it}|^2}\right]^+. \tag{13}$$

Further, since $P_{at} \in \mathcal{P}_a$, the lower bound on $P_{at}$ is redefined as

$$P_{at,LB} = h \cdot \left\lceil \frac{\left[\frac{\sum_{k=1}^{t} \mathbb{E}[P_{ik}] - \eta \sum_{k=1}^{t-1} |h_{D,ik}|^2 P^*_{ak}}{\eta |h_{D,it}|^2}\right]^+}{h} \right\rceil. \tag{14}$$

The payoff of the HAP is expressed in terms ot its utility which is the energy harvested by the device minus the cost $C(P_{at})$ of transmitting the energy signal. The cost $C(P_{at})$ is typically defined as [20] $C(P_{at}) = aP^2_{at} + bP_{at}$ where the values of $a$ and $b$ ($a, b > 0$) depends on the characteristics of the HAP. Hence, the payoff of the HAP is

$$U_{at}(P_{at}) = \eta |h_{D,it}|^2 P_{at} - C(P_{at}). \tag{15}$$

Let $\xi_t = \eta |h_{D,it}|^2$, the payoff becomes

$$U_{at}(P_{at}) = (\xi_t - b)P_{at} - aP^2_{at}. \tag{16}$$

Hence to find the optimal power value $P^*_{at}$, the HAP solves the following optimization problem

$$\max_{P_{at}} U_{at}(P_{at}) \text{ s.t. } P_{at} \geq P_{at,LB}, P_{at} \in \mathcal{P}_a. \tag{17}$$

The optimal solution $P^*_{at}$ is found by first showing that the payoff function $U_{at}(P_{at})$ is discrete concave in $P_{at}$. Then, the relaxed continuous version of the optimization problem in (17) is considered and its closed form solution $P^c_{at}$ is obtained. Based on the solution of the continous version of the problem, the optimal solution of the original problem is obtained.

**Proposition 2.** The payoff $U_{at}$ is discrete concave in $P_{at}$.

*Proof.* A univariate discrete function $f : \mathbb{Z} \to \mathbb{R}$ is discrete concave if $f(x-1) + f(x+1) \leq 2f(x)$. Thus, the standard definition of discrete convexity/concavity assumes that a discrete function $f$ is defined over the set $\mathbb{Z}$ while the set $\mathcal{P}_a$ is not necessarily $\mathbb{Z}$ but it is assumed that in $\mathcal{P}_a$, the power values are multiples of $h$ where $0 \leq h \leq 1$. In order to show that $U_a$ is discrete concave, the variable $P_a$ is transformed into a variable $P'_a$ defined in a subset in $\mathbb{Z}$ by defining $P'_a = \frac{P_a}{h}$. By substituting the $P_a$ in terms of $P'_a$ in terms of the utility function $U_{at}$, we get $U_{at}(P'_a) = (\xi_t - b)hP'_a - ah^2 P'^2_a$. The payoff $U_{at}(P'_a)$ is discrete concave in $P'_a$ since

$$U_{at}(P'_a - 1) + U_{at}(P'_a + 1) = 2(\xi_t - b)hP'_a - 2ah^2(P'^2_a + 1) \leq 2(\xi_t - b)hP'_a - 2ah^2 P'^2_a = 2U_{at}(P'_a). \quad \square$$

A consequence of this proposition is that any local maximum is a global maximum of the optimization problem in (17). In order to characterize the optimal solution, we consider the relaxed continuous version of the problem in (17).

**Remark 1.** The optimal solution for the relaxed optimization problem is

$$P_{at}^{c} = \begin{cases} \frac{\xi_t - b}{2a}, & \text{if } P_{at,LB} \leq \frac{\xi_t - b}{2a} \leq P_{a,\max}, \\ P_{at,LB}, & \text{if } \frac{\xi_t - b}{2a} < P_{at,LB}, \\ P_{a,\max}, & \text{otherwise}, \end{cases} \tag{18}$$

where $P_{a,\max}$ is the maximum power value in $\mathcal{P}_a$.

*Proof.* It can be easily shown that the utility function is continuous strictly concave in $P'_a$ since the second order partial derivative is $-h^2 a$. Also, the value of $P'_a$ at which the derivative of the utility function is zero is $P'_a = \frac{\xi_t - b}{2ah}$. Also, the only constraints of the optimizations are bound constraints on $P'_a$. The results then follows from the concavity of $U_{at}$ and the bound constraints. □

**Proposition 3.** The optimal power $P^*_{at}$ of the HAP is

$$P^*_{at} = \begin{cases} h \cdot \max(\lceil \frac{(\xi_t - b)}{2ah} \rceil, \lfloor \frac{(\xi_t - b)}{2ah} \rfloor), & \text{if } \frac{P_{at,LB}}{h} \leq \frac{(\xi_t - b)}{2ah} \leq \frac{P_{a,\max}}{h}, \\ P_{at,LB}, & \text{if } \lceil \frac{(\xi_t - b)}{2ah} \rceil < \frac{P_{at,LB}}{h}, \\ P_{a,\max}, & \text{otherwise.} \end{cases} \tag{19}$$

*Proof.* When $\frac{P_{at,LB}}{h} \leq \frac{(\xi_t - b)}{2ah} \leq \frac{P_{a,\max}}{h}$, we have $\frac{P_{at,LB}}{h} \leq \lceil \frac{(\xi_t - b)}{2ah} \rceil, \lfloor \frac{(\xi_t - b)}{2ah} \rfloor \leq \frac{P_{a,max}}{h}$ since $\frac{P_{at,LB}}{h}$ and $\frac{P_{a,max}}{h}$ are integers. Since $U_{at}()$ is strictly concave for continuous values of $P'_a$, the payoff for any integer power value $d$ will be less than the payoff of using the power value $m = \max(\lceil \frac{(\xi_t - b)}{2ah} \rceil, \lfloor \frac{(\xi_t - b)}{2ah} \rfloor)$ i.e. $U_a(m) \leq U_a(d)$. Hence in this case, the optimal power value $P'^*_a$ is $\max(\lceil \frac{(\xi_t - b)}{2ah} \rceil, \lfloor \frac{(\xi_t - b)}{2ah} \rfloor)$ and the corresponding optimal value in $\mathcal{P}_a$ is $P^*_a = h \cdot \max(\lceil \frac{(\xi_t - b)}{2ah} \rceil, \lfloor \frac{(\xi_t - b)}{2ah} \rfloor)$. For the case when $\lceil \frac{(\xi_t - b)}{2ah} \rceil$ is less than the $P_{at,LB}$, the optimal power value is $P'^*_a = \frac{P_{at,LB}}{h}$ since the payoff of any other power value greater than $P_{at,LB}$ is less than $U_{at}(P_{at,LB})$ due to the discrete concavity of $U_{at}$ and the corresponding optimal value in $\mathcal{P}_a$ is $P^*_a = P_{at,LB}$. Using the same concavity argument for the last case i.e. when $\lceil \frac{(\xi_t - b)}{2ah} \rceil \geq \frac{P_{a\,max}}{h}$, the optimal value is $P'_a = \frac{P_{a,\max}}{h}$ and the corresponding power value in $\mathcal{P}_a$ is $P_a = P_{a,\max}$. □

Proposition 3 shows that when $\xi_t$, the product of the device battery efficiency and the channel gain, is considerably greater than the energy cost parameters $a$ and $b$, the utility of the HAP becomes higher than the cost and thus the HAP transmits with maximum power. Also, if $\xi_t$ is considerable smaller than the energy cost parameters $a$ and $b$, the HAP's cost becomes higher than its utility and the HAP transmits with the lowest feasible power. Otherwise, the HAP transmits with the optimal power that maximizes its payoff.

## V. Simulation Results

For our simulations, we set $W = 10$ kHz, $T = 2$ msec, $N_0 = -137$ dBm , $\eta = 0.8$, $P_{a,\max} = 2W$, $h = 0.1$, and $\mathcal{P}_i = \{0.1, 0.2, 0.3, 0.4\}$ W. The wireless transmitter is considered to be a video surveillance device [21] which generates UDP packets of size $M = 1000$ bits. The packets are generated according to a Poisson distribution of rate 30 packets/sec. The video surveillance device chooses its transmission power in each time slot such that the received SNR is greater than or equal to the required threshold $\gamma_R$ to decode the packet at the HAP. Assuming that the achieved rate and SNR are related by Shannon's capacity formula, the threshold is thus chosen such that $\frac{M}{T} = W \log(1 + \gamma_R)$. The values considered for the energy cost parameters $(a, b)$ of the HAP are $(1, 1)$ and $(1.5, 1.5)$ respectively [20]. In each time slot, the device drops the packet if it does not have enough energy to transmit it.

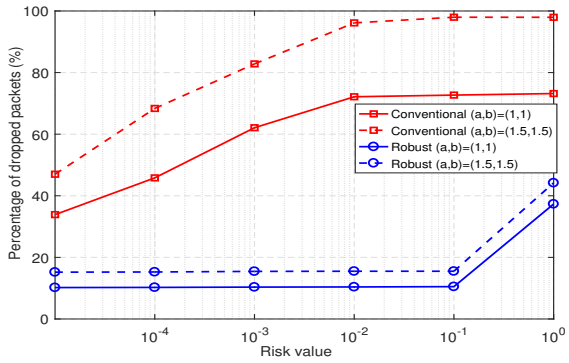


Fig. 1: Percentage of packets lost vs. risk value

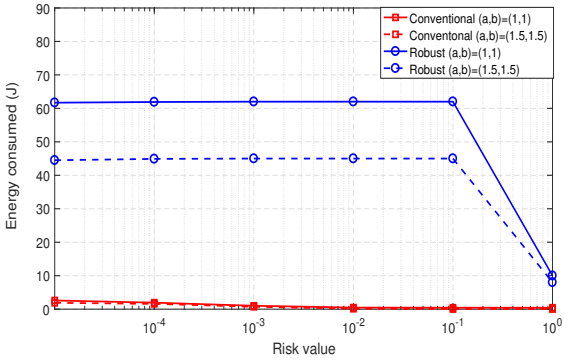


Fig. 2: Energy consumed vs. risk value

Each simulation run simulates the network for 100000 time slots, i.e., 100 seconds. In each run, the percentage of packets dropped by the surveillance device and the energy consumed by the HAP are computed when the HAP uses the robust and conventional learning strategies respectively for the considered values of the energy cost parameters. Then, the average percentage of packets dropped and the average energy consumed by the HAP is computed from 1000 simulation runs. The simulation is performed for two scenarios. The first is when the malicious user's risk value takes values $0, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1$ respectively while the fading variance of the channel between the HAP and the device for both uplink and downlink is set to 0.3. The second scenario is when the fading variance is varied between 0.3 and 0.9 in steps of 0.1 while the risk value is set to be $r = 0.01$.

Fig. 1 shows the percentage of dropped packets for both the conventional and robust learning approaches versus the risk

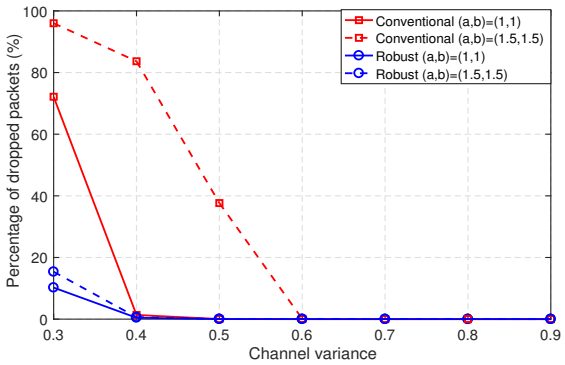


Fig. 3: Percentage of packets lost vs. channel variance

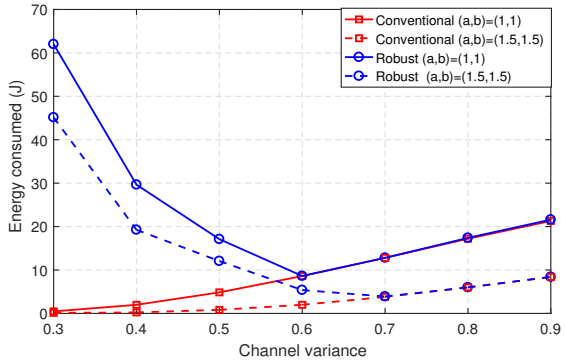


Fig. 4: Energy consumed vs. channel variance

value when the values of the HAP's energy cost paramters $(a, b)$ are $(1, 1)$ and $(1.5, 1.5)$ respectively. First, for the conventional learning approach, and when the energy cost parameters values $(a, b)$ are $(1, 1)$, the percentage of dropped packets with no attack (r = 0) is $33\%$. This percentage increases with the risk and reaches $72\%$ for risk values greater than or equal to $0.01$. When the cost parameters $(a, b)$ increases to $(1.5, 1.5)$, the percentage of dropped packets increases for all considered risk values reaching up to $97\%$ when the risk value greater than or equal to $0.01$. In contrast, for the proposed approach, when $(a, b)$ are set to $(1, 1)$, the percentage of dropped packets $10.25\%$ when no attack occurs, and it remains around $10\%$ when the attacker's risk increases to $0.1$. A more pronounced increase occurs when the risk value is $1$ as the percentage of dropped packets attains $37\%$. For such a high risk value, the optimal strategy of the attacker is to transmit with minimum power, which affects the effectiveness of the robust approach. However, in practice, the attacker will only choose low risk values in order not to be detected, and hence, the percentage of dropped packets when $r = 1$ will not be attained. When the energy cost parameters increase to $(1.5, 1.5)$, the percentage of dropped packets is around $15\%$ for a risk value less than or equal to $0.1$ and increases to $44\%$ when the risk value is one. Thus, Fig. 1 shows that the proposed robust learning strategy constitutes a better learning approach than the conventional learning approach even when no attack happens. Further, the proposed robust learning approach is more robust to changes in the risk values unlike the conventional learning approach that is sensitive to slight variations in the risk value. From Fig. 1, we can also see that, the proposed aproach can achieve a performance gain, in terms of the percentage of dropped packets, which can reach up to $85\%$ at $r = 0.1$ compared to the conventional learning approach.

Fig. 2 shows the energy consumed for the conventional and robust learning approaches versus the risk value for different energy cost parameters. As shown in Fig. 2, the conventional learning approach maintains low energy consumption. When the energy cost paramteres are $(1, 1)$, the energy consumed is $2.59$ J when no attack happens and drops to $0.467$ J for a risk value greater than or equal to $0.01$. When $(a, b)$ increases to $(1.5, 1.5)$, the energy consumed decreases to $1.9$ J when no attack happens and drop to $0.03$ for a risk value higher than $0.3$. On the other hand, the proposed robust strategy exhibits higher energy consumption. This is because the robust approach overestimates the transmission power consumed by device, which results in increasing the energy delivered to the device in each time slot. When the value of the energy cost paramteres $(a, b)$ is $(1, 1)$, the energy consumed is around $62$ J for a risk value lower than or equal to $0.1$ and drops to $10$ J for a risk value equal to one. When $(a, b)$ increases to $(1.5, 1.5)$, the consumed energy decreases to $45$ J for a risk value less than $0.1$ and drops to $8$ J for a risk value equal to one. Thus, the results in Fig. 2 show the tradeoff between maintaining a good performance in terms of the percentage of dropped packets and the energy consumed. Yet, the energy consumed by the robust learning is lower than the energy consumed when the HAP transmits with fixed maximum power in each time time slot. For the considered simulation values of the system parameters, the energy consumed by the fixed power policy is $200$ J. Hence, the robust learning strategy can reach a gain in terms of energy efficiency up to $77\%$ while maintaining a low percentage of dropped packets.

Fig. 3 shows the percentage of dropped packets for both the conventional and robust learning approaches versus the channel variance value for the considered values of the HAP's energy cost paramters. First, for the conventional learning appoach and when the values of the energy cost parameters $(a, b)$ are $(1, 1)$, the percentage of dropped packets decreases significantly from $72\%$ to $1.35\%$ when the value of the channel variance increases from $0.3$ to $0.4$. Then, the percentage of dropped packets tend to zero as the variance increases further. This is due to the fact that, when the channel quality improves, the received energy by the device increases, and the transmit power required by the device to deliver the packet successfully decreases which allows for more successfull transmissions. Moreover, for energy cost parameters $(1, 1)$, the probability that the HAP's energy cost becomes lower than its utility increases. Thus in this case, the HAP is more likely to transmit with maximum power $P_{a,\max}$. Next when the HAP's energy cost parameters $(a, b)$ are increased from $(1, 1)$ to $(1.5, 1.5)$, the percentage of dropped packets using the conventional learning approach increases considerably when the channel variance value is less than or equal to $0.5$. This is due to the fact that, for higher values of the cost parameters, the energy cost increases, and the HAP is more likely to

supply less energy to the device. Thus, the HAP's conservative strategy combined with the altered estimate by the attacker will yield a significant increase in packet loss. On the other hand, the robust learning strategy maintains a low to negligible percentage of packet loss for both considered values of energy cost parameters. The increase in the percentage of packet loss due to the increase in the cost parameters is slightly obervable when the value of the channel variance is $0.3$ where the increase is from $10\%$ to $15\%$. However, the percentage of dropped packets is negligible for higher values of the channel variance. Clearly, from Fig. 3, we can see that the proposed approach is more robust to more conservative energy policies by the HAP under different channel conditions.

Fig. 4 shows the energy consumed for both the conventional and robust learning approaches as a function of the channel variance. For the conventional learning approach and for cost parameters $(1, 1)$, the energy consumed is only $0.46$ J when the value of channel variance is 0.3 due to the high packet loss as shown in Fig. 3. Then, the energy consumed increases with the channel variance. This is because the percentage of packets lost decreases with the channel variance, as shown in Fig. 3, which implies that the device is transmitting successfully more packets and requires the HAP to transmit more energy. When the cost parameters increases to $(1.5, 1.5)$, the energy consumed by the HAP decreases since the HAP adopts a more conservative energy policy. For the robust learning approach, for cost parameters $(1, 1)$, the energy consumed decreases first with the channel variance when the value of the channel variance is less than or equal to 0.6. This is because when the channel quality is low, the channel gain takes low values with high probability. Thus, the HAP must spend more energy to meet each device's energy requirements. However, for values of channel variance higher or equal to $0.6$, the energy consumed starts to increase due to the increase in the number of successfully transmitted packets by the device. Also, the energy consumed using the robust strategy becomes almost equal to the energy consumed using the conventional learning approach. This is because, when the channel quality becomes high, the HAP can meet the energy requirements of the device with minimal transmission power, which makes the attacks by the malicious user ineffective. The energy consumed by the robust learning strategy exhibits a similar pattern with the channel variance value when the values of the cost paramters are $(1.5, 1.5)$ yet it is lower than the energy consumed when the value of the costs parameters are $(1, 1)$.

## VI. Conclusion

In this paper, we have introduced a robust Bayesian learning scheme for RF energy harvesting which allows the HAP to form an estimate of the transmission power consumption profile of each associated device based on the device's received power at each time slot. The proposed scheme takes into account potential man-in-the-middle-attacks by a malicious user that tries to alter the learned estimate of the HAP in order to deplete the battery of the device. Based on the learned estimate, we have considered the problem of optimal power selection by the HAP in each time slot that maximizes the HAP's payoff while meeting device's energy requirements are met. Further, we have shown that the payoff function is discrete concave and obtained a closed-form expression of the optimal power of the supplied energy signal. Our results have shown that our proposed robust Bayesian learning scheme can achieve performance gains in terms of the percentage of dropped packets by the HAP compared to the conventional Bayesian learning approaches. Also, the proposed learning scheme exhibits gains in terms of energy efficiency compared to the fixed power transmission policy .

## References


[1] L. Atzoria, A. Ierab, and G. Morabitoc, "The Internet of Things: A survey", *Computer Networks*, vol. 54, no. 15, pp. 2787-2805, 2010.

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

[3] X. Kang, C. K. Ho and S. Sun, "Full-Duplex Wireless-Powered Communication Network With Energy Causality," in *IEEE Transactions on Wireless Communications*, vol. 14, no. 10, pp. 5539-5551, Oct. 2015.

[4] Z. Hadzi-Velkov, I. Nikoloska, H. Chingoska and N. Zlatanov, "Proportional Fair Scheduling in Wireless Networks With RF Energy Harvesting and Processing Cost," in *IEEE Communications Letters*, vol. 20, no. 10, pp. 2107-2110, Oct. 2016.

[5] H. Chen, Y. Ma, Z. Lin, Y. Li and B. Vucetic, "Distributed Power Control in Interference Channels With QoS Constraints and RF Energy Harvesting: A Game-Theoretic Approach," in *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 10063-10069, Dec. 2016.

[6] F. Hu, Q. Hao, *Intelligent Sensor Networks: The Integration of Sensor Networks, Signal Processing and Machine Learning*, CRC Press, 2012.

[7] T. Park, N. Abuzainab and W. Saad, "Learning How to Communicate in the Internet of Things: Finite Resources and Heterogeneity," in *IEEE Access*, vol. 4, pp. 7063-7073, Nov. 2016.

[8] F. Azmat, Y. Chen and N. Stocks, "Predictive Modelling of RF Energy for Wireless Powered Communications," in *IEEE Communications Letters*, vol. 20, no. 1, pp. 173-176, Jan. 2016.

[9] Z. Zou, A. Gidmark, T. Charalambous, M. Johansson, "Optimal Radio Frequency Energy Harvesting with Limited Energy Arrival Knowledge," in *IEEE Journal on Selected Areas in Communications*, to appear, Aug. 2016.

[10] M. Gregori and J. Gómez-Vilardebò, "Online learning algorithms for wireless energy harvesting nodes," in Proc. of *2016 IEEE International Conference on Communications (ICC)*, Kuala Lumpur, Malaysia, 2016, pp. 1-6.

[11] Insecurity in the Internet of Things (White Paper), Symantic, Mar. 2015, retrieved Jan, 6, 2017.

[12] D. Niyato, P. Wang, D. I. Kim, Z. Han and L. Xiao, "Game theoretic modeling of jamming attack in wireless powered communication networks," in Proc. of *IEEE International Conference on Communications (ICC)*, London, Jun. 2015, pp. 6018-6023.

[13] D. Niyato, P. Wang, D. I. Kim, Z. Han and L. Xiao, "Performance analysis of delay-constrained wireless energy harvesting communication networks under jamming attacks," in Proc. of *IEEE Wireless Communications and Networking Conference (WCNC)*, New Orleans, LA, USA, Jun. 2015, pp. 1823-1828.

[14] A. El Shafie, D. Niyato and N. Al-Dhahir, "Security of Rechargeable Energy-Harvesting Transmitters in Wireless Networks," in *IEEE Wireless Communications Letters*, vol. 5, no. 4, pp. 384-387, Aug. 2016.

[15] A. Salem, K. A. Hamdi and K. M. Rabie, "Physical Layer Security With RF Energy Harvesting in AF Multi-Antenna Relaying Networks," in *IEEE Transactions on Communications*, vol. 64, no. 7, pp. 3025-3038, Jul. 2016.

[16] B. K. Wang Ng, G. L. Tian, M. L. Tang, *Dirichlet and Related Distributions: Theory, Methods and Applications*, John Wiley & Sons, 2011.

[17] Z. Chen, S. Guo, K. Zheng and Y. Yang, "Modeling of Man-in-the-Middle Attack in the Wireless Networks," in Proc. of *International Conference on Wireless Communications, Networking and Mobile Computing*, Shanghai, China, Sep. 2007, pp. 2255-2258.

[18] C. Paulino and C. Pereira, "Bayesian Methods for Categorical Data Under Informative General Censoring," in *Biometrika*, vol. 82, no. 2, pp. 439-446, Jun. 1995.

[19] J. M. Dickey, J. Jiang and J. B. Kadane, "Bayesian Methods for Censored Categorical Data," in *Journal of the American Statistical Association*, vol. 82, no. 399, pp. 773-781, Sep. 1987.

[20] Chen, Y. Li, Z. Han and B. Vucetic, "A stackelberg game-based energy trading scheme for power beacon-assisted wireless-powered communication," in Proc. of *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, South Brisbane, QLD, Australia, Apr. 2015, pp. 3177-3181.

[21] F. Y. Lin, C. Hsiao, H. Yen, and Y. Hsieh, "A Near-Optimal Distributed QoS Constrained Routing Algorithm for Multichannel Wireless Sensor Networks, " in *Sensors*, vol. 13, no. 12, Dec. 2013