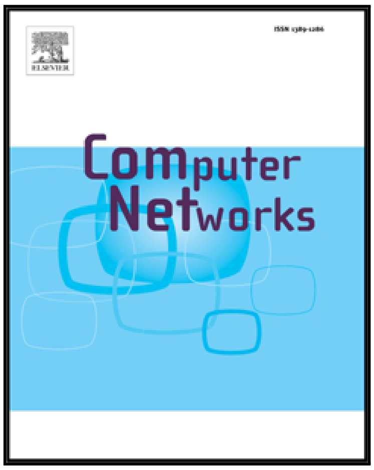
Accepted Manuscript



Adaptive Learning-Enforced Broadcast Policy for Solar Energy Harvesting Wireless Sensor Networks

Khiati Mustapha, Djamel Djenouri

PII: S1389-1286(18)30521-8
DOI: 10.1016/j.comnet.2018.07.016
Reference: COMPNW 6546

To appear in: *Computer Networks*

Received date: 29 January 2018
Revised date: 24 May 2018
Accepted date: 9 July 2018

Please cite this article as: Khiati Mustapha, Djamel Djenouri, Adaptive Learning-Enforced Broadcast Policy for Solar Energy Harvesting Wireless Sensor Networks, *Computer Networks* (2018), doi: 10.1016/j.comnet.2018.07.016

Adaptive Learning-Enforced Broadcast Policy for Solar Energy Harvesting Wireless Sensor Networks

Khiati Mustapha[1], Djamel Djenouri[2,*]
[1] USTHB,Algiers, Algeria
[2] CERIST Research Center, Algiers, Algeria,
mkhiati@usthb.dz, ddjenouri@acm.org
Correspandance authors address: CERIST Research Center, Rue des Freres Aissou, Ben-aknoun, BP 143 Algiers, Algeria.

Abstract

The problem of message broadcast from the base station (BS) to sensor nodes (SNs) in solar energy harvesting enabled wireless sensor networks is considered in this paper. The aim is to ensure fast and reliable broadcast without disturbing upstream communications (from SNs to BS), while taking into account constraints related to the energy harvesting (EH) environment. A new policy is proposed where from the one hand, the BS first selects the broadcast time-slots adaptively with the SNs schedules (to meet active periods that are constrained by EH conditions), and from the other hand, SNs adapt their schedules to enable optimal selection of the broadcast time-slots that minimizes the number of broadcasts per message and the latency. Compared to the existing solutions, this enables fast broadcast and eliminates the need of adding message overhead to the broadcast message. For this purpose, an analytical energy model, a hidden Markov model(HMM), Baum-Welch learning algorithm, and a heuristic algorithm of the minimum covering set problem (MCS) are proposed and combined in a unique solution. The proposed solution is analyzed and compared with a state-of-the-art approach. The results confirm that the former has the advantage of performing the broadcast operation more reliably and in lower delay.

Keywords: Wireless Sensor Networks, Energy Harvesting, Green Communication, Broadcast, HMM

1. Introduction

The interconnection of heterogeneous wireless networks will enable large scale deployment in the near future and provide unique applications and services, such as health care, smart buildings and grids, environment monitoring, smart cities, smart object tracking, etc. One of the challenges that prevent large scale deployment is energy limitation. Most existing network protocols assume the use of portable and limited batteries, while many applications are expected to be deployed in unintended, sometimes hostile environments, which makes batteries replacing unfeasible. Environment energy harvesting emerges as the appropriate alternative for battery replacement. The ambient energy harvesting ability of wireless devices is expected to enable self-sustaining systems. However, satisfying communication requirements while tuning activities to the available energy and energy profiles is a nontrivial problem. Existing protocols and architectures should be revisited and rebuilt upon an energy model that reflects the real world constraints for harvesting [19, 20, 10]. To our knowledge, this paper is the first that considers broadcast under slotted EHWSN and proposes ADAPCAST(Adaptive Learning-Enforced Broadcast Policy for Solar Energy Harvesting Wireless Sensor Networks), a policy that minimizes the number of time-slots allocated to the broadcast while considering the EH constraints. The main contributions of ADAPCAST is the use of a new energy model suitable for solar EH environment and of a Hidden Markov Model (HMM)[28] that faithfully reflects the nodes behavior activities under this environment. The proposed policy considers local broadcast (one-hope) and targets the selection of the optimal time-slots within the frame for broadcasting the message from the BS to SNs (on the downstream links) while avoiding interfering with the upstream communications to the best effort. In particular, the policy attempts to match the broadcast time-slot with all active periods of SNs (to the best effort), i.e., to ensure SNs are ready for reception. The policy also allows nodes to adapt their activity in a way that optimizes the broadcast time-slot selection. Contrary to the proposed solution in [20] that uses the concept of erasure coding and relays on adding packets to increase reliability, the solution proposed herein minimizes the number of time-slots dedicated to the broadcast operation without augmenting the packet payload with overhead. This has a positive impact on the minimization of the number of transmissions required for broadcast at the BS, and it allows SNs to preserve their energy and to allocate up-link communication time-slots. A Hidden Markov Model (HMM)[28], Baum-Welch learning algorithm [12], and an heuristic algorithm of the minimum covering set(MCS)[30] are applied for this purpose. Rather than using Bernoulli distribution (used in [20]), the

[1]Correspandance to ; Djamel Djenouri, CERIST Research Center, Rue des Freres Aissou, Ben-aknoun, BP 143 Algiers, Algeria.
E-Mail: ddjenouri@acm.org, djenouri@gmail.com

proposed policy uses an accurate solar energy model that faithfully reflects the realistic behavior of energy and tracks the energy consumption at the SN level. The energy model is composed of two sub-models. The first one is devoted to the incoming energy and uses a Discrete Markov Chain (DMC), while the second one is a sub-analytic model that introduces a gain function to estimate the energy consumption per time unit. We analyzed the proposed solution and compared it with RMBMEH (Reliable Multicast and Broadcast Mechanisms for Energy Harvesting Devices)[20], which is to the best of our knowledge the only work from the literature that considers broadcasting in EH environments. The results confirm that the proposed solution outperforms the one in [20] in reducing the time and cost required for the broadcast. This is by minimizing the number of necessary time-slots, which– in addition to reducing the broadcast latency– improves energy efficiency.

The reminder of this paper is organized as follows. Section 2 lists the related work. Section 3 presents the assumptions, the network, the energy models and the hidden Markov model. The proposed solution is described in details in Section 4, whereas Section 5 evaluates the proposed solution by simulations and mathematical analysis and compares it to [20]. Finally, Section 6 draws the conclusions and highlights future directions.

2. Related Work

Developing models and protocols that are appropriate for wireless sensor networks in energy harvesting environments has been attracting research works in recent years. Yang and Ulukus. [32] proposed a policy that adaptively changes the transmission rate according to the traffic load and available energy subject to causality constraints. This is to minimize the transmission completion time in a single-user additive white gaussian noise channel. Ozel et al. [27] proposed a policy for M-user AWGN (additive white gaussian noise) broadcast channel, which is subject to the causality of energy arrivals as well as the finite rechargeable battery capacity constrained by energy overflow. Beside the finite energy battery, QoS of wireless links in EH environments has been considered in [11], where finite state Markov chain has been used to model the communication channel. The effective capacity for power allocation policies is analyzed to characterize the QoS performance of energy harvesting wireless links using finite state Markov chain. The solution derives the effective capacity for the power allocation policies and relies on the fact that the more stringent is the effective capacity, the faster is the decay of the transmission rate. The major disadvantage of these solutions is that they do not consider the neighborhood of nodes and that each node tries to adapt to the EH environment in order to be able to send or receive, and this without any information on the status of nodes by which it is surrounded. Zhang et al. [34] presented an analytical approach that addresses the problem of the system design of energy harvesting capable wireless devices. This is by fitting the sizes of buffers for energy and data, as well as the size of the harvester, to the specified delay and loss requirements. A stochastic model that takes into account energy harvesting and event arrival processes is proposed for this purpose. The proposed policy in[21] is based on the available information on channel conditions and energy. It considers EH cooperative communication systems where the relay selection rule depends on the relative throughput gain for better use of the available information and improves the performance of the EH cooperative communication system. Several other problems in energy harvesting WSN have been revisited such as duty cycle management [2], clustering [33], relay node deployment [24], [6], etc.

Many solutions have been proposed for effective message broadcast in wireless sensor networks, e.g, [5], [17], etc., but few has been done in energy harvesting environments. Kuan et al. To our knowledge, [20] and [17] are the only solutions the only work that considers broadcasting in EH environments. The authors in [20] addressed the trade-off between reliability and throughput and proposed a down-link broadcast policy for energy harvesting that uses erasure coding to guarantee the transmission reliability for EH wireless devices. The policy maximizes both the throughput and the probability of receiving broadcasted packets by configuring the broadcast period and erasure coding parameters. In order to cope with the problem of energy deficiency, two mechanisms (energy-aware receiving and early termination) have been also proposed. The solution proposed in this work has many shortcomings. If the number of the received packets does not reach the required value then the message is dropped, which means the energy devoted to the reception of those packets is wasted. Another problem is related to the use of the erasure coding that causes important overhead to every broadcast message, which affects the broadcast latency. Further, the number of packets that must be received is determined beforehand, which is not always possible in EH environment. In this paper, we provide a comprehensive extension of our preliminary solution of [17].

3. Assumptions and Models

3.1. Problem Statement and Network Model

The proposed solution addresses the problem of broadcasting in asynchronous and homogeneous energy harvesting wireless sensor networks (EHWSNs). We assume that only one message has to be broadcasted from the BS to the SNs (over downstream links) for each frame, and that one time-slot is sufficient to receive the message. We assume that the BS is energy unconstrained and plugged into a continuous power source, while SNs are equipped with solar EH capability (that are attached to rechargeable batteries) and located within the communication range of the BS, i.e start topology. The SNs use a slotted CSMA(Carrier Sens Multiple Access) based MAC protocol such as [14, 22] to access the shared communication medium. At the beginning, the BS is supposed to have the knowledge about the radio duty cycle and energy arrivals rate of each SN. This knowledge is periodically updated. SNs are supposed synchronized by applying some underlying synchronization protocol, such as [31, 7, 23]. Operations under low data rate applications and events triggered mode are considered in this paper, which are more appropriate in energy limited environments. The communication between the BS and SNs (upstream/downstream link) is assumed to follow a single-hop fashion.

3.2. Energy Model

3.2.1. Energy Consumption Model

The energy consumption model at the node level is presented in this section. The energy consumption is strongly related to the activities of the node and varies depending on the type of their activities. In particular, the radio is a major source of energy consumption, which makes it largely dependent on parameters related to the MAC protocol (duty cycle, length of the frame, etc.). We introduce two models for deriving the energy consumption at node level. The first one concerns the radio activity and relies on a discrete Markov chain(DMC), whereas the second one deals with micro-controller activities that do not require any radio communication such as sensing, and data processing.

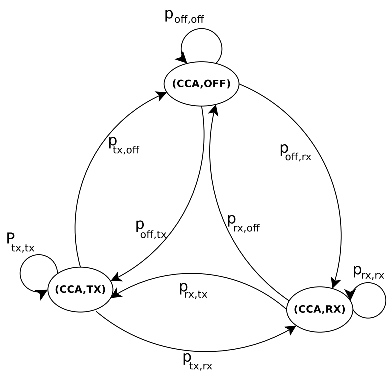


Figure 1: Radio Energy Consumption Model

Radio Activity Model:
The energy consumed by the radio (transceiver) is due to the reception (RX), transmission (TX), or carrier-sensing. Since this largely depends on the MAC protocol to be used, we consider (without loss of generality) B-MAC [14] as the underlaying MAC protocol. The proposed model can be amended to fit any protocol slotted MAC protocol. B-MAC, and all preamble-sampling MAC protocols, uses periodic sampling of the channel at the beginning of every time-slot, which is known as clear channel assessment (CCA). We propose an off-line Discrete Markov Chain (DMC), which is denoted by, Φ, and presented in Fig.1. Every sate represents the activity of the node during a single time-slot, which can be transmission (TX), reception (RX), or turning off the radio (off). A single activity is assumed to take place during a time-slot. The CCA that the node performs before switching to any state at the beginning of the time-slot is considered as part of the state in Φ (which explains the use of the prefix *CCA* in the states' labels). The states represented by the DMC are: 1) the transmission state, (CCA,TX), where the radio is kept on for transmission After CCA, 2) the reception state (CCA, RX) where the radio is kept on for reception, and, 3) (CCA, OFF) where the radio is turned off immediately after CCA for the whole time-slot. The total power consumed at each state is given by:

$$P_{state} = \alpha P_{act} + \beta P_{cca}, \tag{1}$$

where P_{cca} and P_{act} stands for the power consumed for performing CCA and the activity of the state TX, RX, OFF, respectively, α stands for the ratio of the period of time within the time-slot duration devoted to the corresponding activity, and β designates the one spent for performing CCA, ($\alpha + \beta = 1$).

The transition probabilities are determined on the basis of the following assumptions: 1) the arrival of the messages is assumed to be Poisson distributed random variables (r.v) for both incoming and out going messages with rates λ_1, and λ_2, respectively, 2) RTS/CTS and ACK[14] mechanisms are considered disabled, 3) The B-MAC protocol used in our solution is amended to use the time unit of the proposed DMC, i.e., one time-slot.

The matrix Q_Φ determines the transition probabilities between the different DMC states, which are calculated using the Poison distribution. Two Poisson processes with parameters, λ_1 and λ_2, are considered. Giving that the r.v. of the two processes are independent, the transition probabilities to any state are equal (independently from the origin). For the sake of simplifying the notations, we denote (CCA, OFF), (CCA, RX), (CCA, TX) respectively by e_1, e_2, e_3. Without loss of generality, we prioritize handling the reception vs. the transmission, i.e., when having a reception and a transmission events simultaneously, the reception is handled first.

The transition probability to the state e_2 is the probability to have an incoming message, which is $1 - e^{-\lambda_1}$. The transition probability to e_3 is the probability to have an outgoing message and not having an incoming message, which is $(1 - e^{-\lambda_2}) * e^{-\lambda_1}$. Final, the probability to transit to e_1 is the probability to have no message i.e., 0 incoming message and 0 outgoing message, which is $e^{-\lambda_1} * e^{-\lambda_2} = e^{-(\lambda_1+\lambda_2)}$ (the product as the two events are independent). This is also the complement to 1 of the sum of the first two transitions. Q_Φ is then given by,

$$Q_\Phi = \begin{bmatrix} e^{-(\lambda_1+\lambda_2)} & 1 - e^{-\lambda_1} & (1 - e^{-\lambda_2}) \,.\, e^{-\lambda_1} \\ e^{-(\lambda_1+\lambda_2)} & 1 - e^{-\lambda_1} & (1 - e^{-\lambda_2}) \,.\, e^{-\lambda 1} \\ e^{-(\lambda_1+\lambda_2)} & 1 - e^{-\lambda_1} & (1 - e^{-\lambda_2}) \,.\, e^{-\lambda_1} \end{bmatrix}$$

The equilibrium state probability vector, Π,is calculated by solving ΠQ_Φ = Π. Using the equilibrium state, the radio power consumption is approximated as follows,

$$P_{radio} = \sum_{i \in S} P_i \pi(i) \ , \tag{2}$$

where S is the set of the DMC states.

The average energy consumption for each time-slot is then,

$$E_{Radio} = P_{radio} T^u, \tag{3}$$

where T^u is the time unit (time-slot duration).

Microcontroller Model:
Remember that we consider event-triggered low-data

rate applications. The microcontroller (CPU) is then put in the sleep state until an event is triggered, i.e., a request for sensing or a data processing. It returns again to the sleep as soon as the event handling ends. Therefore, three states for the CPU operation can be distinguished, 1) "Sensing" (sen), 2) "Sleeping" (sl), and 3) "Data-processing" (dp). The energy consumption may be expressed as follows:

$$P_{cpu} = P_{sen}R_{sen} + P_{dp}R_{dp} + P_{sl}(1 - (R_{sen} + R_{dp})), \quad (4)$$

where P and R stand for the power consumption and ratio of the respective mode, respectively. Similarly to the energy consumed by the radio, the energy drained by the CPU during a time-slot is given by,

$$E_{Cpu} = P_{cpu}T^u. \quad (5)$$

The energy consumed by the node during every time-slot of the frame is simply: $C = E_{Cpu} + E_{Radio}$.

3.2.2. Energy Harvesting Model

Solar energy is quite stable source when available, i.e., it smoothly changes over time. To reflect this, we introduce a sensitive analytical gain model for the incoming solar energy. Let us consider a generic set, S_N, of N nodes, where we refer by Ψ_k as the gain associated with node $k \in S_N$, which designates the average amount of harvested energy for each time unit(T^u). The geographical locations and the weather condition make the gain Ψ slightly different from node to another. We assume that Ψ_k is given by $\Psi_k = \bar{\Psi} + \eta_k$ [29], where $\bar{\Psi}$ is a gain component common to all nodes, and η_k is an additive random displacement uniformly distributed in : [-$\theta\bar{\Psi}$,$\theta(1-\bar{\Psi})$], where $\theta \in$ [0,1] models the influence of the aforementioned constraints on the incoming energy. In this model, $\theta = 0$ corresponds to the fully correlated case as all node gains collapse to $\bar{\Psi}$. Conversely, $\theta = 1$ gives the i.i.d. case. For a given (θ, $\bar{\Psi}$) pair, all gains fall in the interval $[\Psi_{min}, \Psi_{max}]$, where, Ψ_{min} =$\bar{\Psi} - \theta\bar{\Psi}$, and Ψ_{max} =$\bar{\Psi} + \theta(1-\bar{\Psi})$. The gain set can thus be specified in terms of either (θ,Ψ) or ($[\Psi_{min}, \Psi_{max}]$).

Note also that by specifying θ and Ψ or, equivalently, Ψ_{min} and Ψ_{max}, we only know that all gains are uniformly and independently distributed in the subset $[\Psi_{min}, \Psi_{max}] \subseteq$ [0,1], where each node at the beginning will be independently picking a value (gain) uniformly distributed from this interval as incoming energy. The obtained gain is then multiplied by the unit rate of the incoming energy.

3.2.3. Effective Energy

The effective energy is cumulated from the difference between the incoming energy (uniform random variable) and the consumed energy, which is a constant value derived from the DMC presented above. The effective energy is formulated by the following equation:

$$E_{eff}(t_W) = E_{eff}(t_{W-1}) + (E_{in}(t_W) - E_{out}(t_W)), \quad (6)$$

where t designates the index of the time-slot, E_{in} and E_{out} the harvested and consumed energy, respectively, during the time-slot t. We suppose that $E_{eff}(t_0) = E_{min}$, which represents the initial amount of the effective energy in the battery. Let us denote, $X_i = E_{in}(t_i) - E_{out}(t_i)$, $i \in \{1, \ldots, W\}$ where W is the number of time-slots, $X_0 = E_{eff}(t_0)$. Eq 6 yields the following recurrent equation,

$$E_{eff}(t_i) = E_{eff}(t_{i-1}) + X_i. \quad (7)$$

By successive substitutions of the terms $E_{eff}(t_{i-1})$, $i = \{1, \ldots, W\}$, we get,

$$E_{eff}(t_W) = \sum_{i=0}^{W} X_i. \quad (8)$$

This is the sum of W continuous uniform r.v. $(X_0 + ... + X_W)$ in the interval [E_{min}, E_{max}] where E_{min} is as aforementioned the initial effective energy, and E_{max} is the maximum amount of energy that can be stored in the battery. Nevertheless, the distribution of this sum is unknown. In the following we will normalize these r.v. (X_i) to fall into a known distribution. As the coming energy in every time slot is bounded by $[\Psi_{min}, \Psi_{max}]$, the effective energy is then bounded in the interval $E_{eff}(t_i) \in [X_0 + i(\Psi_{min} - C), X_0 + i\Psi_{max}]$, $i \in \{1, \ldots, W\}$. From this interval, we define a normalized r.v., $\delta(t_i)$, as follows (proof of this is given in the appendix),

$$\delta_i = \frac{E_{eff}(t_i) - X_0 - i\,(\Psi_{min} - C)}{i\,(\Psi_{max} - \Psi_{min} + C)}, \quad for\ i \in \{1, ..., W\}, \quad (9)$$

which follows a continuous standard uniform distribution (U~[0,1]). The sum $(\sum \delta_i)$ follows the Irwin-Hall probability distribution [15]. Their probability density function (PDF) and cumulative density function (CDF) are given by E.q. 10 and E.q. 11, res,

$$f_{pdf}(x) = \frac{1}{(W-1)!}\sum_{k=0}^{\lfloor x \rfloor}(-1)^k \binom{W}{k}(x-k)^{W-1} \quad (10)$$

$$f_{cdf}(x) = \frac{1}{(W)!}\sum_{k=0}^{\lfloor x \rfloor}(-1)^k \binom{W}{k}(x-k)^{W} \quad (11)$$

3.3. Proposed Hidden Markov Model (HMM)

The proposed HMM (presented in Fig.2) is a left-to-right HMM [4]. It models the radio activity of the SNs under solar EH constraints using the proposed energy model (sec. 3.2). This HMM is scalable since the number of its states is independent from the number of SNs.

The HMM model (γ) for each node in the network is defined by the following parameters $< \zeta, A, \Pi, T, E >$ where :

- ζ : is the set of the hidden states, where at any time-slot, i, ζ_i may be in one of the following states: 1) $e_i(0,0)$ if the effective energy is less than the

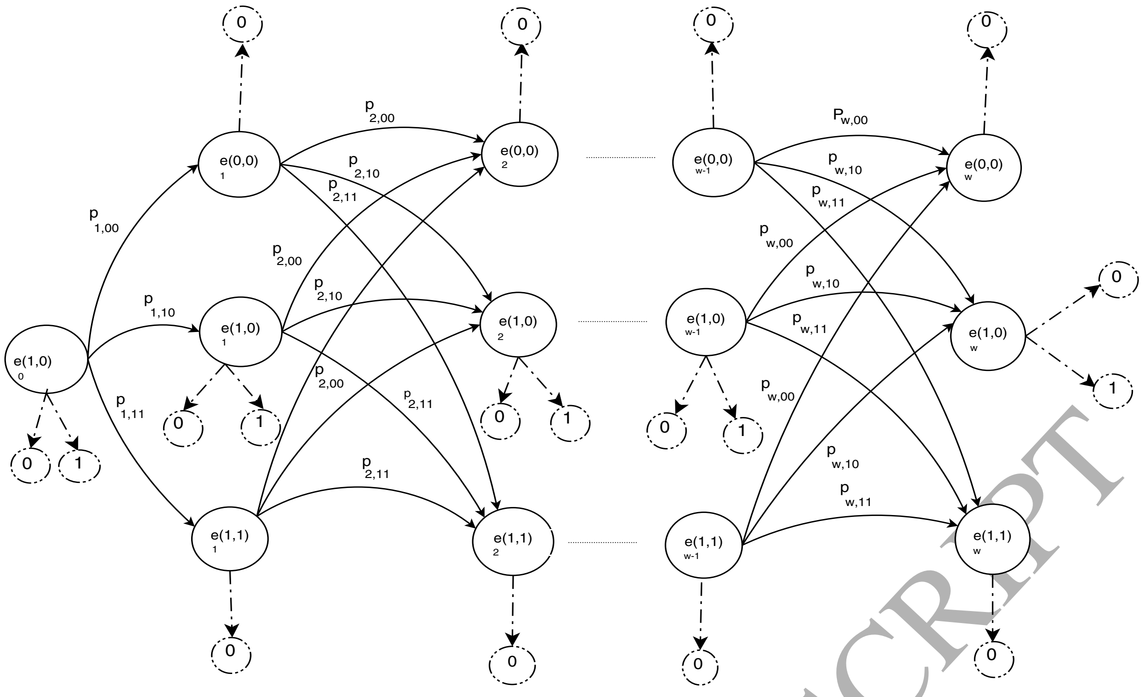


Figure 2: HMM graph

threshold and there is no activity, 2) $e_i(1,0)$ if the effective energy is no less than the threshold and there is no activity, 3) $e_i(1,1)$ if the effective energy is no less than the threshold with presence of activity.;

- A : stands for the alphabet symbols or observations $A\{0,1\}$, issued by hidden states, where $"0"$ stands for the absence of any radio activity (denoted by "$Act = 0$") and $"1"$ for the presence of activity ($Act = 1$).

- T : the matrix of the transition probabilities between the HMM hidden states, i.e., $T\,(i,j) = p\,(X_t = j\,|\,X_{t-1} = i)$, where $X(t)$ is a r.v that designates the hidden states at time t. It takes its values from ζ.

- Π: Probability distribution of the initial state (the state from where the HMM starts), Π (i) = $p\,(X_1)$, given that :

$$\sum_{i=1}^{H_s} \pi(i) = 1, \qquad (12)$$

where H_s stands for the number of the possible hidden states of the HMM.

- E : the matrix of the emission probabilities of the symbols in A by the hidden states of the HMM. Let us denote by M the number of symbols. The probability that the state i generates the observation j, say, $p\,(O_t = j\,|\,X_t = i)$, is denoted $E(i,j)$ where the matrix E has to satisfy:

$$\sum_{j=1}^{M} E\,(i, O_j) = 1 \;\; \forall i \in S. \qquad (13)$$

In the proposed HMM model, the transition probabilities relay on two dependents r.v. The first one designates the effective energy of the node that is modeled by the proposed energy model (Sec. 3.2), whereas the second one represents the radio activity during each timeslot within the frame duration and is represented with a logistic distribution [16, 8, 25, 26] (Fig. 3).

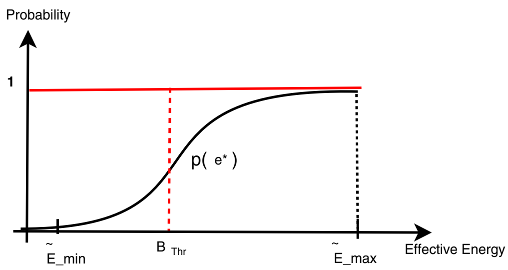


Figure 3: Logistic Distribution (Probability of presence of activity) vs. E_{eff}

As the two r.v. are dependent, conditional probabilities [1] are used to model the transition probabilities between the different HMM hidden states.

Given that, 1) the transition probability to the state $e_i(0,0)$ is, $p\,(E_{eff} < E_{thr} \cap Act = 0)$, 2) the transition probability to the state $e_i(1,0)$ is $p\,(E_{eff} \geq E_{thr} \cap Act = 0)$, 3) the transition probability to the state $e_i(1,1)$ is : $p\,(E_{eff} \geq E_{thr} \cap Act = 1)$; then transitions probabilities are:

$$\begin{cases} T\,(.,e(0,0)) = p\,(E_{eff} < E_{thr}). \\ p\,(Act = 0\,|\,E_{eff} < E_{thr}) \\ T\,(.,e(1,0)) = p\,(E_{eff} \geq E_{thr}). \\ p\,(Act = 0\,|\,E_{eff} \geq E_{thr}) \\ T\,(.,e(1,1)) = p\,(E_{eff} \geq E_{thr}). \\ p\,(Act = 1\,|\,E_{eff} \geq E_{thr}) \end{cases} \qquad (14)$$

To use the normalized variable, $\widetilde{E}_{eff}$, we rewrite E.q. 14 as follows:

$$\begin{cases} T\left(., e(0,0)\right) = \\ p\left(\widetilde{E}_{eff} < B_{thr}\right).p\left(Act = 0 \,|\, \widetilde{E}_{eff} < B_{thr}\right) \\ T(., e(1,0)) = \\ p\left(\widetilde{E}_{eff} \geq B_{thr}\right).p\left(Act = 0 \,|\, \widetilde{E}_{eff} \geq B_{thr}\right)) \\ T\left(., e(1,1)\right) = \\ p\left(\widetilde{E}_{eff} \geq B_{thr}\right).p\left(Act = 1 \,|\, \widetilde{E}_{eff} \geq B_{thr}\right) \end{cases} \tag{15}$$

Note that this is possible since, $E_{eff} < E_{thr}$ is equivalent to, $\widetilde{E}_{eff} < B_{thr}$. B_{Thr} **denotes the threshold of the effective energy corresponding to the normalized variable,δ. It is common to all SNs and given by,**

$$B_{thr} = \frac{E_{Thr}}{\max\limits_{i \in \{1,\ldots,N\}} \Psi_{max}(i) - C} \tag{16}$$

Where N is the number of SNs. The probability of presence of activity is given using the logistic distribution as shown in Fig.3, whose PDF and CDF (that are defined in the interval $[\widetilde{E}_{min}, \widetilde{E}_{max}]$) are given by E.q. (17), and E.q. (18), respectively,

$$g_z(z) = \frac{e^{-\left(\frac{z-\mu}{s}\right)}}{s(1+e^{-\left(\frac{z-\mu}{s}\right)})^2} \tag{17}$$

$$G_z(z) = \frac{1}{1+e^{-\frac{z-\mu}{s}}} \tag{18}$$

The parameter, μ, designates the mean, and s the standard deviation. For a standard logistic distribution, these parameters take the values, 0, and, 1, respectively.

The PDF is used to calculate the transition probabilities given in Eq.(15). From the equations (11,15,16,17,18) the transition probabilities are determined as follows,

$$\begin{cases} T\left(., e(0,0)\right) = f_{cdf}(z)\Big|_{\widetilde{E}_{min}}^{B_{thr}}, \\ T\left(., e(1,0)\right) = f_{cdf}(z)\Big|_{B_{thr}}^{\widetilde{E}_{eff}} . \left(1 - \left(\frac{1}{1+e^{-z}}\right)\Big|_{B_{thr}}^{\widetilde{E}_{eff}}\right), \\ T\left(., e(1,1)\right) = f_{cdf}(z)\Big|_{B_{thr}}^{\widetilde{E}_{eff}} . \left(\frac{1}{1+e^{-z}}\Big|_{B_{thr}}^{\widetilde{E}_{eff}}\right) \end{cases} \tag{19}$$

4. Solution Description

4.1. Overview

The solution uses for every node in the network a Hidden Markov Model (HMM). The BS broadcasts incoming messages based on the duty cycle of each node and the presence of activity. Based on transitions between the different states of the HMM model (detailed in sec.3.3), each node dynamically adapts its radio duty cycle by alternating between ON("1")/OFF("0") for each time-slot within the frame duration. Consequently, the HMM generates for each time-slot an observation representing the readiness of receiving the broadcast message by the node. A heuristic algorithm of the minimum covering set problem is applied to the sequences of observations of all SNs to retrieve the minimum set of time-slots that provides maximum coverage of SNs among active periods, i.e., ensure maximum coverage of SNs with the minimum number of intersections between SNs observations where the radios are "ON".

The optimal set defines the most preferred sequence emitted by each HMM in order to efficiently broadcast a message. Once the most preferred sequences emitted by all HMMs in hand, the BS executes the Baum-Welch algorithm with theses sequences as input. The algorithm adjusts the initial HMM model parameters and derives accordingly for every node a new HMM model that is most likely to produce the sequence that meets the obtained minimum set of time-slots required for the broadcast. By considering the EH model of every node, the solution has the advantage of increasing chances to reach all nodes with the minimum number of broadcast count, and thus minimum latency. The broadcast count is defined as the number of transmissions (time-slots) required to make a broadcast in a frame (to reach all the nodes if possible, or at least the maximum number of nodes). Another advantage of the proposed solution is that the selection of the optimal set allows nodes to release the remaining time-slot for other scheduled activities and services, such as data gathering and the communication of the sensed data to the BS.

4.2. BS Level

The proposed solution is centralized as the BS maintains and updates the required information to run the HMM for every node. The BS level runs in three phases. i) The first phase generates observation sequences by running the HMMs. iii) The second phase consists in applying an optimal set selection algorithm as a heuristic of the covering set problem to determine the optimal broadcast time-slots. Finally, iii) The third phase uses the Baum-Welch learning algorithm to adjust the initial HMMs of SN's to increase the likelihood of meeting the most appropriate sequences that produce the broadcast time-slots obtained in phase (ii). The BS communicates the derived HMM models to theirs corresponding nodes. This process is executed at the initialization of the network, and the obtained scheme is used for a long period until a significant change on the HMMs parameters takes place. Another case that requires the re-execution of the three phases is when the average number of missed nodes becomes important. This information becomes available at the BS level at the last frame (i.e., before the end of each period), where nodes send the number of times they missed the broadcast messages.

4.2.1. Phase I

During the first phase, the BS runs the HMMs of the SNs. The observation sequences are generated by each HMM, which follow a discrete distribution where the emission probabilities change depending on the current hidden state. These sequences are composed of "1" and "0" values, where "1" means the SN's radio is "ON" and ready to receive the broadcast message, and "0" that the SN's radio is "OFF" or busy (presence of activity) during the corresponding time-slot. The radio behavior of each SN is determined by the emitted sequences for each time-slot within the frame duration. The decision on whether the node will turn ON or OFF the radio during the next time-slot depends on the probability of generating "0" or "1", i.e., based on the proposed energy model (Sec.3.2) and the radio activity during each time-slot. The effective energy is formulated as a ratio between the harvested energy and the consumed one. Having effective energy beyond a given threshold, say E_{thr}, along with the absence of activity allows the generation of "1" observation with high probability, and "0" otherwise. Recall that in the proposed energy model, the r.v. corresponding to the effective energy follows an Irwin-Hall distribution. The radio activity during each time-slot within the frame is then modeled with binary values $(0, 1)$ that depends on two parameters, i) the effective energy (which should be no less than E_{thr} to generate 1), and ii) the probability of presence of activity which is bounded by an upper-bond, ξ. The steps through which the HMMs generate the observation sequences during the first phase is summarized in Algorithm 1. Detailed description of the HMM model is given in Sec.3.3.

Algorithm 1 Generation of observations by each HMM at BS level

Input : HMMs of all SNs,
N: Number of sensor nodes,
δ: Ratio of effective energy at each time-slot,
f: Logistic function (Sec.3.3),
p_{act} : Probability of having an activity,
ξ : Upper-Bond for p_{act},
B_e : Bernoulli distribution ,
I: Boolean indicator on the presence of activity,
ρ : Probability of emitting 1 when there is enough energy and no radio activity (constant common to all nodes),
$FrameSize$: Size of the frame (Number of slots)
Output: Emitted sequences by each HMM.

```
for i ∈ {1, ..., N} do
    slt ← 0

    while slt ≤ FrameSize do
        Calculate δ(i)
        p_act ← f(δ(i), ξ)
        I ← B_e(p_act)
        if (δ(i) < B_Thr) then
            seq[slt] =0
        else
            if (I == 1) then
                seq[slt]=0
            end if
            if (I == 0) then
                seq[slt] =B_e(ρ)
            end if
        end if
        slt ← slt + 1
    end while
end for
```

4.2.2. Phase II

After running the HMMs of all SNs for several frames, the BS applies the set selection algorithm. We transformed the problem to the set covering problem (a well known problem in the literature) to use existing heuristics. Intersections between nodes schedules in every time-frame are explored, i.e., of active time-slots, to retrieve the minimum number of time-slots that ensures maximum nodes coverage. This leads to a minimum number of time-slots for broadcast and releases the remaining unselected time-slots for other activities. The following explains the transformation into the set covering problem. Let us consider a set of elements {1 , 2 , . . . , n} (called the universe), and a collection S of m sets. The set cover problem is to identify the smallest subset of S that forms the universe. In our case, the problem is to find the minimum number of time-slots that cover all SNs. The universe is all the SNs, m the number of slots in a frame, and S designates the collection of sets in the frame, where each set, Slt_i, contains the nodes covered by the corresponding time-slot,i, of the frame (nodes that emit "1" in the time-slot i). Note that the SN is covered by that time-slot if it emits the observation "1" during that time-slot.

Example: This example illustrates the transformation of our problem into the minimum covering set. Let us consider a network of 10 SNs and 10 time-slots. Fig.4 depicts a possible scenario where the scratched slots indicate that the corresponding HMM emits 1 observation, and white ones that it emits 0.

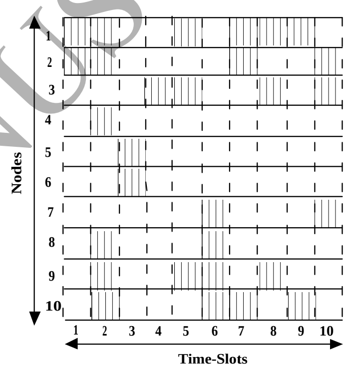


Figure 4: Illustration of problem transformaion into minimum covering set

The sets of covered nodes by each time-slot are: $Slt_1 = \{1, 2\}$, $Slt_2 = \{1, 2, 4, 8, 9, 10\}$, $Slt_3 = \{5, 6\}$, $Slt_4 = \{3\}$, $Slt_5 = \{1, 3, 9\}$, $Slt_6 = \{7, 8, 9, 10\}$, $Slt_7 = \{1, 2, 10\}$, $Slt_8 = \{1, 3, 9\}$, $Slt_9 = \{1, 10\}$, $Slt_{10} = \{2, 3, 7\}$. The smallest sub-collection of sets that form the universe in this case is $\{Slt_2, Slt_3, Slt_{10}\}$. This is known as the minimum covering set problem. Many heuristics are proposed in the literature for this problem and its variants. We are interested in the variant of retrieving the minimum set that ensures maximum coverage which is an heuristic known as greedy algorithm, i.e., equivalently set of minimum time-slots within the frame duration that ensures maximum nodes coverage (not necessary all)[3].

4.2.3. Phase III

In this phase the Baum-Welch algorithm is run for each HMM of each node, with the desired generated sequences plugged in as inputs. The Baum-Welch algorithm belongs to the class of Expectation Maximization(EM) algorithms [18, 9]. It allows to adjust the initial HMM parameters in order to obtain the one that maximizes the probabilities of reproducing the most preferred

ACCEPTED MANUSCRIPT

sequence. This is expressed by maximizing the product of the probabilities of generating the desired sequence by each HMM. If we denote by O_i the desired sequence generated by the HMM, γ_i, of node, i,the problem is to find HMMs: $\hat{\gamma} = <\hat{\gamma_1}, \hat{\gamma_2}, ..., \hat{\gamma_n}>$ that ensures:

$$\textbf{Max}(\prod_{i=1}^{N} p(O_i|\gamma_i)) \tag{20}$$

The result is a set of new HMM models obtained after adjustment of the parameters (transition and emission probabilities) of the initial HMM model. As a final step, the BS communicates the parameters of the derived HMM models to their corresponding nodes. The three phases are summarized in Algorithm 2.

Algorithm 2 : General algorithm of the three phases

Input (all phases):N: Number of Nodes, $Period\_Length$: Number of frames, $FrameSize$: Number of time-slots

.... **Phase I** ....

Input : HMMs of all Sensor nodes (SNs), Energy model, Activity model
Output : The emitted sequences by all HMMs
φ = [] []
for $F \in \{1, ..., Period\_Length\}$ **do**
  V = []
  for $i \in \{1, ..., N\}$ **do**
    $slt \leftarrow 0$
    while $slt \leq FrameSize$ **do**
      execute the HMMs ($\gamma_i(slt)$)
      $slt \leftarrow slt + 1$
    end while
    $V[i] \leftarrow The\ emitted\ sequence\ by\ node(i)\ during\ frame\ F$
  end for
  $\varphi[F] \leftarrow V$
end for
return φ %The emitted sequences by all SNs during each frame)

.... **Phase II** ....

Input : Schedules of all SNs in every time frame (φ)
Output : The subset of a minimum number of time-slots for a maximum nodes coverage
Opt =[]
for $F \in \{1, ..., Period\_Length\}$ **do**
  Runs the minimum covering set algorithm to schedule all SNs during frame (F)
  $Opt[F] \leftarrow The\ optimal\ set\ ensuring\ maximum\ nodes\ coverage$
end for
return The subset of the minimum number of time-slots among all optimal subsets in Opt

.... **Phase III** ....

Input :The optimal sequence generated by every HMM
Output: New derived HMM $\hat{\gamma}$ for each SN
for $i \in \{1, ..., N\}$ **do**
  $\hat{\gamma}(i)$ = BAUM_WELCH($\gamma(i)$)
end for
return $\hat{\gamma}$

4.3. Node Level

At the end of the third phase, every node receives its derived HMM parameters, i.e., the transition probabilities between HMM hidden states and the probability of generating output observations for each hidden state of the HMM. The parameters related to energy harvesting cannot be adjusted, as they are strongly dependent on the the weather conditions. However, parameters related to the power consumption of the nodes (powers of different components) are adjustable. The power consumption units that may be adapted are the receiving unit (e.g. the duty cycle and radio wake up periods), the microcontroller and the sensing unit (e.g., sensing sampling periods). To allow the BS updating the HMM models at the last frame of the period, every node sends to the BS the number of times it misses the broadcast message, as well as the average amount of incoming energy.

5. Performance Evaluation

In the present section, the proposed solution (ADAPCAST) is compared to RMBMEH [20]. For the latter, the source (K) and redundant data packets (R) are set respectively to 2, 3.

First, we evaluate the DMC that imitates the B-MAC behavior, while considering the CC2420 RF transceiver model [13]. The average of energy drained by the radio vs. λ_2 for different values of λ_1 is plotted in Fig.5. The results prove the steadiness and convergence of the proposed model with the increase of the two parameters.

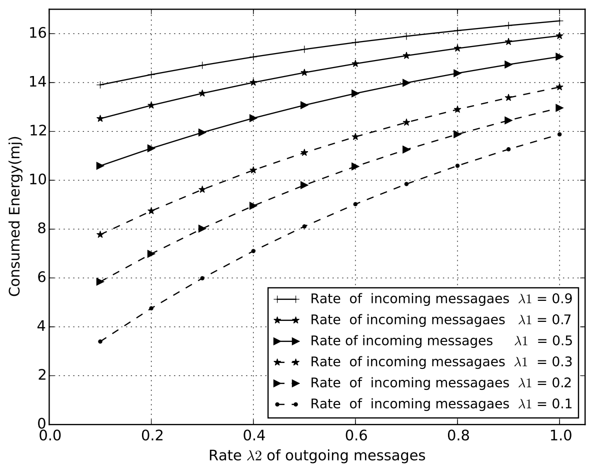


Figure 5: Radio average consumed energy per time-slot

Since the reception is more prioritized than transmission in the proposed energy model (Sec.3.2), we set the rate λ_1 of the incoming messages (receptions) to 0.7 and λ_1 of the outgoing messages to 0.5. Table 1 summaries the simulation parameters.

Parameter	Value
Number of Nodes	[25 , 200]
Poisson rates λ_1, λ_2	0.7, 0.5
Upper-bond ξ	1.0
Unit of incoming energy E_r(mj)	100
ρ (Probability of Bernouli Dist.)	0.5
θ	0.1,0.5,0.9
Threshold of effective energy E_{thr}(mj)	250
Initial effective energy(mj)	150
DATA Packet Size (Byte)	32
Time_Slot length (ms)	5
Frame length (Time_Slot)	20
B_{thr}	4

Table 1: Simulation parameters.

The incoming energy has been determined as a function of the random displacement expressed by the influence factor, θ, which has been set to 0.1, 0.5 and 0.9, respectively. The performance of the proposed solution is evaluated with respect to the following metrics, 1) broadcast count, 2) effective energy, 3) number of missed node, and 4) the broadcast delay. Three scenarios are

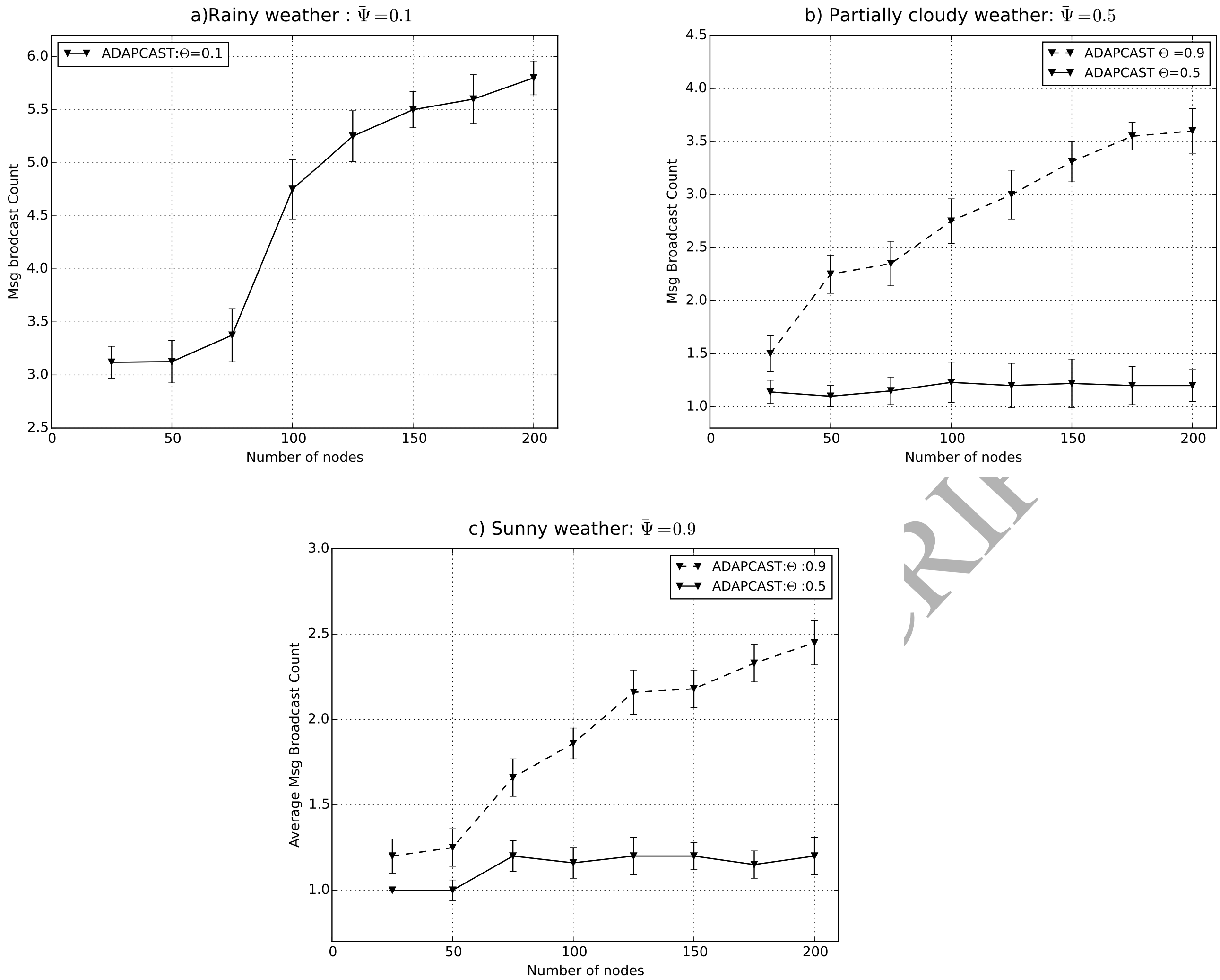


Figure 6: Broadcast count vs. number of nodes

considered, 1) rainy weather, where the available energy is very low, i.e., $\bar{\Psi} = 0.1$, and nodes are highly correlated $\Theta = 0.1$. 2) partially cloudy, where $\bar{\Psi} = 0.5$, and $\Theta \in \{0.5, 0.9\}$. 3) and sunny where there is enough energy, i.e., we set $\bar{\Psi} = 0.9$ and $\Theta \in \{0.5, 0.9\}$. The proposed energy model has been adapted to the policy in [20] for the three aforementioned scenarios. The results given in the following are averages of several runs, and plots are presented with errors bars of 95% confidence interval. Before comparing ADAPCAST with RMBMEH, we first evaluate the broadcast count, i.e., the average number of transmissions (time-slot scheduled for broadcast) of the same broadcast message per frame. Note that RMBMEH is not considered by this metric since it uses a completely different concepts and is not based on time-slot scheduling. The results are depicted in Fig 6. Fig 6 (a) shows smooth increase with the number of nodes. This increase is inevitable due to low energy at receiving nodes that are available for limited time-slots, and the increase in the number of nodes requires more time-slots to reach them. This dispersal in the availability of periods (readiness for broadcast reception) is also present for $\theta = 0.9$ in scenarios of medium and high energy, which explains the increase (Fig 6 (b) and (c)). However, for $\theta = 0.5$ where the dispersal in incoming energy is lower, the average broadcast count remains close to the optimum. In all cases, this metric remains less than 4 for medium energy scenario (resp. 3 for high energy scenario) even for a network of 200 nodes, which is quite reasonable. The impact of this performance in reducing the broadcast count is investigated in comparison with RMBMEH for the performance metrics starting the average effective energy. The results for the first scenario (Fig.7(a)) show that when SNs are fully correlated $\Theta = 0.1$ and $\bar{\Psi} = 0.1$, the average amount of effective energy in the EHWSN at the end of the frame is close to the energy threshold ($240mj$), while it is less than $210mj$ for RMBMEH (both for the original RMBMEH, and RMBMEH with reinforced energy model). This is due to the fact that in the proposed solution, nodes tend to be off in most of the time during the frame which allows to store the harvested energy until the effective energy of the SN reaches the threshold required for performing a broadcast operation. In the proposed solution, the BS schedules the optimal time-slots as a function of the available effective energy for each SN while the solution in [20]

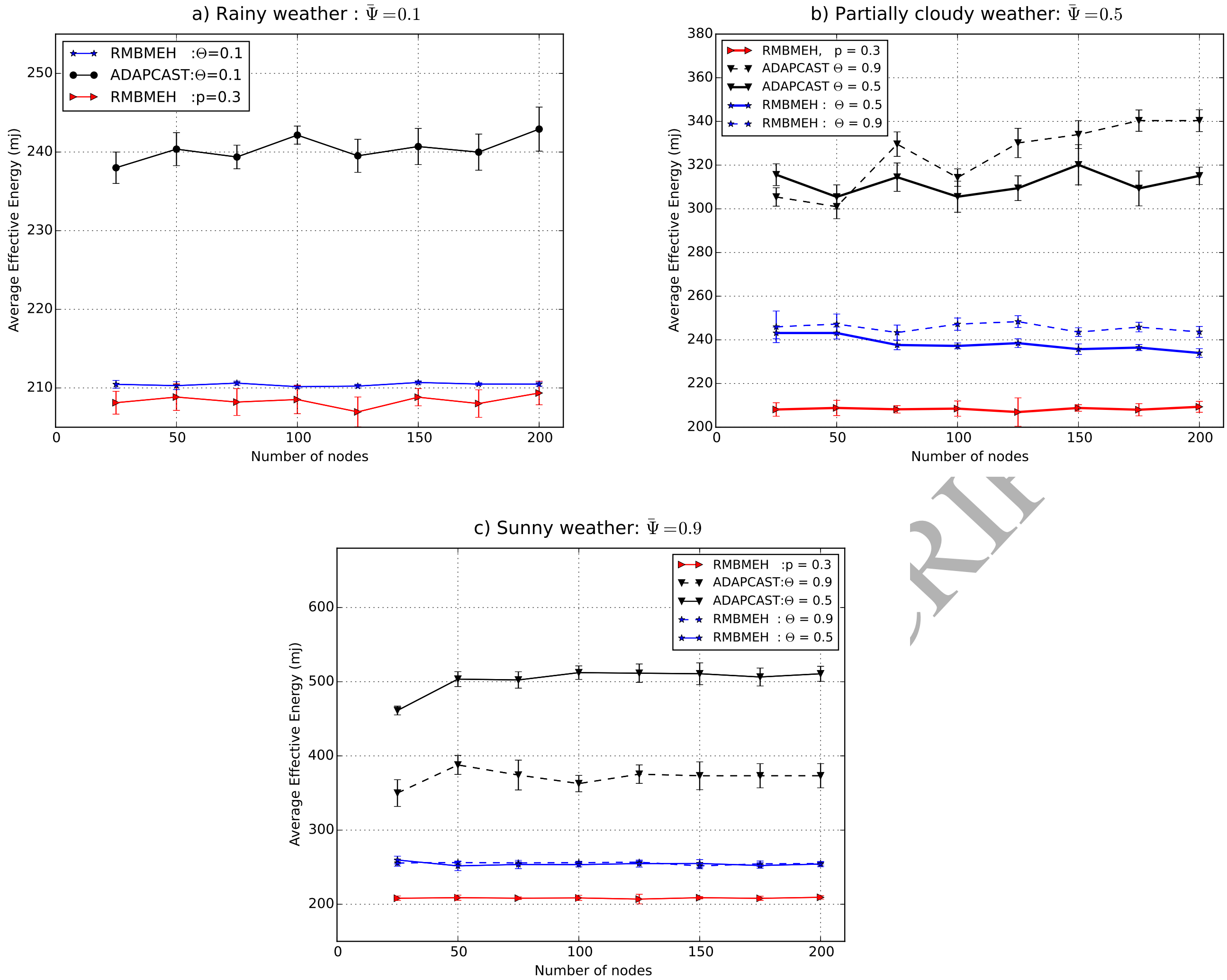


Figure 7: Average effective energy vs. number of nodes

applies the erasure coding mechanism for broadcasting that relies on the addition of control messages which requires more energy and does not take into account the activity of SNs.Fig.7(b) presents the second scenario. It shows that the proposed solution ranges above $320mj$ and clearly out performs RMBMEH that has value at the order of $210mj$, and $240mj$ when using adaptive energy model. The increase of Θ normally leads to a slight improvement for both solutions. Results of the third scenario are presented in Fig. 7(c). They are similar to the previous scenario with higher amplitudes, and with difference between RMBMEH and ADAPCAST. The superiority of the proposed solution is due to the accurate models applied and the learning approach that considers nodes activities vs. the available energy. Results also show that when applying the proposed energy model to RMBMEH enhances its performance.

Results also confirm the superiority of ADAPCAST for the number of missed nodes (Fig. 8), i.e. lesser missed nodes in all scenarios, and almost all nodes covered in the second and third scenario. More importantly, the figures show that the gap between the two solutions generally increases with the number of nodes. This is due to the use of coordinated slot in ADAPCAST.

In the first scenario (Fig.8(a)), the effective energy of most nodes is less than the threshold during all the cycles, which explains the inevitable increase of missed nodes with the number of nodes for both solutions. This problem does not exist for the second and third scenario for our solution that adapts to the nodes' energy fluctuation and increases the number of times slots to cover missing nodes, which explains the stability of ADAPCAST results in these scenarios. The results also confirm improvement of RMBMEH when using the proposed energy model.

The last metric concerns the average broadcast latency over all weather conditions. Each point of Fig.9 is averaged over several runs for the each value of Ψ. The results confirm superiority of the proposed solution with respect to performing fast broadcast. Overall, the difference between the basic RMBMEH and RMBMEH with enhanced energy model confirms the effectiveness of this model, while the difference between RMBMEH (both versions) and ADAPCAST confirms the effectiveness of the learning approach (HMM and Baum-Welch models) and the time-slot selection algorithm.

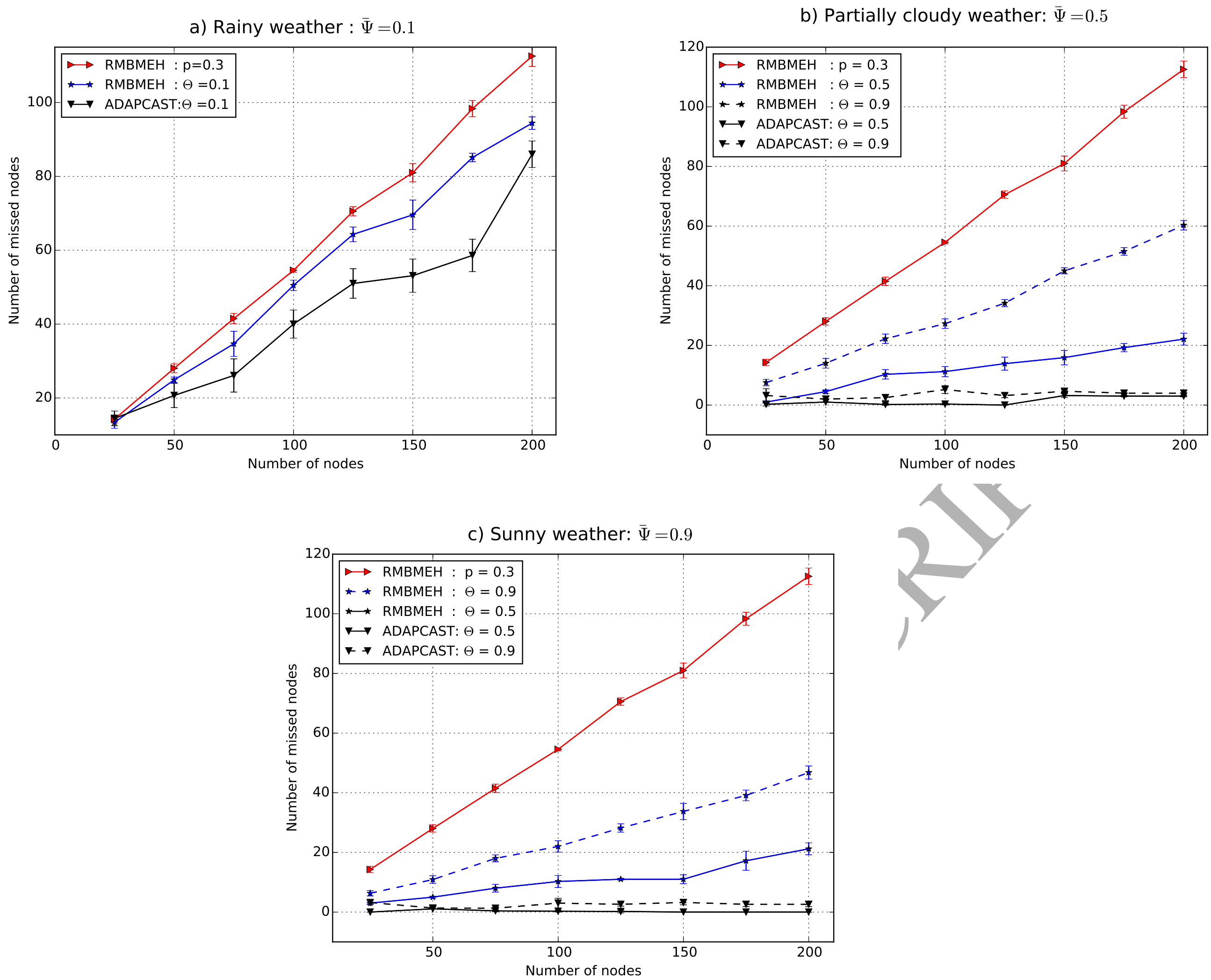


Figure 8: Number of missed nodes vs. number of nodes

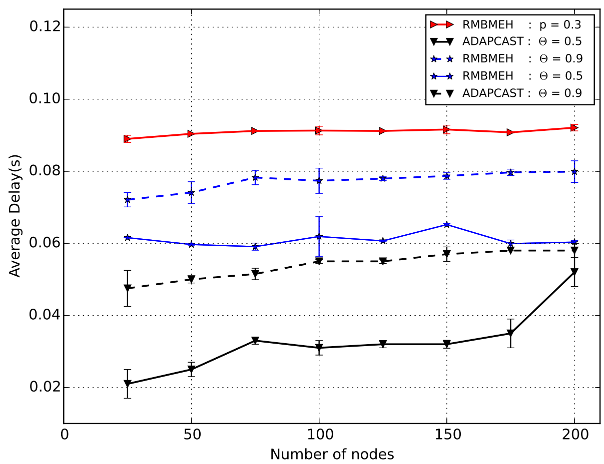


Figure 9: Average delay

6. Conclusion and Perspectives

In this paper, the problem of message broadcasting from the base station (BS) to sensor nodes (SNs) has been considered in solar energy harvesting enabled wireless sensor networks. We proposed a solution based on analytical models for both energy harvesting arrivals and energy consumption for each SN during regular time intervals within a frame. A greedy policy for calculating the optimal set of time-slots for performing broadcast operation has been proposed which relies on reducing the problem to the minimum covering set, which is a fundamental problem that has been largely treated in literature. In order to perform broadcasting during the coordinated time-slots, the policy uses a Hidden Markov Model with the Baum-Welch Estimation Maximization Algorithm to produce the most likely sequences. The proposed solution has been compared by simulation to [20], the only solution in the literature (to the best of our knowledge) that treats the same problem in energy harvesting environment.

The results demonstrated a significant performance improvement in terms of the average effective energy in the network, the number of missed nodes, and latency. A simple scenario where a single broadcast message per time frame has been considered. As a future direction, we plan to generalize the model for handling any number of messages. This will require the update of the policy and the use of some queuing theory techniques to manage the messages. Besides, tracking the energy neutral operations (ENO) to assure of SNs perpetual operations while performing the broadcast of messages is an issue to consider in this area of work.

7. Achnowledgments

This work is part of the CERIST's WSN Energy harvesting project, which is partly supported by the research grant AL16MO03 from Algerian Ministry of Higher Education and Scientific Research through the DGRSDT, for collaboration with the university of Padova. The authors would like to thank Dr Michele Rossi, from university of Padova, Italy, for hosting Mustapha in his lab and for his helpful recommendations in developing the stochastic models used in this work.

References

References

[1] Ackerman, N. L., Freer, C. E., M.ROY, D., 2010. On the computability of conditional probability. Computer Science and Artificial Intelligence Laboratory,MIT, MA, USA.

[2] Castagnetti, A., Pegatoquet, A., Le, T. N., Auguin, M., 2014. A joint duty-cycle and transmission power management for energy harvesting wsn. IEEE Trans. Industrial Informatics 10 (2), 928–936.

[3] Chen, P., G.Ding, 2008. A greedy heuristic for a generalized set covering problem. NSF grants DMS-0556091 and ITR-0326387 RRR-16.

[4] Couvreur, C., 1996. Hidden markov models and their mixtures. Technicall report, catholic university of Louvain, Faculty of Sciences.

[5] Di Pietro, R., Oligeri, G., Apr 2017. Enabling broadcast communications in presence of jamming via probabilistic pairing. Computer Networks 116 (C), 33–46.

[6] Djenouri, D., Bagaa, M., 2017. Energy-aware constrained relay node deployment for sustainable wireless sensor networks. IEEE Trans. on Sustainable Computing 2 (1), 30–42.

[7] Djenouri, D., Merabtine, N., Mekahlia, F. Z., Doudou, M., 2013. Fast distributed multi-hop relative time synchronization protocol and estimators for wireless sensor networks. Ad Hoc Networks 11 (8), 2329–2344.

[8] D.Juki, R.Scitovski, 2003. Solution of the least-squares problem for logistic function. Journal of Computational and Applied Mathematics (156), 159–177.

[9] Do, C. B., Batzoglou, S., Augst 2008. What is the expectation maximization algorithm? Nature Publishing Group, Nature biotechnology 26 (08).

[10] Fafoutis, X., Dragoni, N., 2012. Analytical comparison of mac schemes for energy harvesting - wireless sensor networks. IEEE 978-1-4673-1786-3.

[11] Gong, J., Zhang, S., Wang, X., Zhou, S., Niu, Z., 2014. Supporting quality of service in energy harvesting wireless links: The effective capacity analysis. ICC Workshop on Energy Efficiency in Wireless Networks & Wireless Networks for Energy Efficiency 14 (978).

[12] Hu, P., Zhou, Z., Liu, Q., Li, F., 2007. The hmm-based modeling for the energy level prediction in wireless sensor networks. In: Second IEEE Conference on Industrial Electronics and Applications. pp. 2253–2258.

[13] Instruments, T., 2004. 2.4 GHz IEEE 802.15.4 / Zigbee-ready RF Transciever.
URL `http://www.ti;com`

[14] J., P., J., H., D., C., November 2004. Versatile low power media access for wireless sensor networks. In: Second ACM Conference on Embedded Networked Sensor Systems (SenSys). Baltimore, Maryland, USA, pp. 95–107.

[15] J.Marengo, L., Farnsworth, Stefanic, L., 2017. A geometric derivation of the irwin-hall distribution. International Journal of Mathematics and Mathematical Sciences :IJMMS, 6 pages.

[16] J.Navarro, D., 2006. Similarity, distance and categorization: A discussion of smith's (2006) warning about colliding parameters. Technicall report, School of Psychology University of Adelaide.

[17] Khiati, M., Djenouri, D., 2017. ADABCAST: adaptive broadcast approach for solar energy harvesting wireless sensor networks. In: 26th Wireless and Optical Communication Conference, WOCC 2017, Newark, NJ, USA, April 7-8, 2017.

[18] Krishnan, G. J. M. T., 2008. The EM Algorithm and Extensions. Vol. Second Edition. JOHN WILEY SONS, INC.

[19] Ku, M. L., Li, W., K, Y. C., Liu, J. R., 2016. Advances in energy harvesting communications:past, present, and future challenges. In: IEEE (Ed.), IEEE Communication Surveys And Tutorials, Second Quarter. Vol. 18. pp. 1384–1412.

[20] Kuan, C.-C., Lin, G.-Y., Wei, H.-Y., 2014. Reliable multicast and broadcast mechanisms for energy-harvesting devices. IEEE Transactions on Vehicular Technology 63 (4).

[21] Luo, Y., Zhang, J., Letaief, K., 2012. Relay selection for energy harvesting cooperative communication systems. Hong Kong Research Grant Council (610212).

[22] M., B., V., Y., E., E., R., H., 2006. X-mac: A short preamble mac protocol for duty-cycled wireless sensor networks. In: ACM SenSys. Boulder, Colorado, pp. 307–320.

[23] M, L., Y-C, W., 2010. On clock synchronization algorithms for wireless sensor networks under unknown delay. IEEE Transactions on Vehicular Technology 2010 59(1), 182–190.

[24] Mehajabin, N., Razzaque, M. A., Hassan, M. M., Almogren, A., Alamri, A., 2016. Energy-sustainable relay node deployment in wireless sensor networks. Computer Networks 104, 108–121.

[25] Nassar, M., Elmasry, A., November 2012. A study of generalized logistic distributions. Journal of the Egyptian Mathematical Society,Production and hosting by Elsevier B.V, 126–133.

[26] Olapade, A., 2004. An extended type 1 generlized logistic distribution. International Journal of Mathematics and Mathematical Sciences :IJMMS 57, 3069–3074.

[27] Ozel, O., Yang, J., Ulukus, S., 2012. Optimal broadcast scheduling for an energy harvesting rechargeable transmitter with a finite capacity battery. IEEE Transactions On Wireless Communications VOL. 11 (06).

[28] Porits, A. B., 1988. Hidden markov models : A guided tour. IEEE CH2561-9 (0000-0007).

[29] Rossi, M., Bui, N., Zorzi, M., March 2009. Cost and collision-minimizing forwarding schemes for wireless sensor networks: Design,analysis,and experimental validation. IEEE Transactions On Mobile Computing 08 (03).

[30] Spasovski, S., Bogdanova, A. M., Ss. Cyril, M., 2013. Optimization of the polynomial greedy solution for the set covering problem. In: of Computer Science, F., Engineering (Eds.), The 10th Conference for Informatics and Information Technology (CIIT 2013).

[31] Wang, Y., Mosalakanti, K., Núñez, F., Deligeorges, S., III, F. J. D., 2017. A kernel module for pulse-coupled time synchronization of sensor networks. Computer Networks 127, 161–172.
URL https://doi.org/10.1016/j.comnet.2017.08.009

[32] Yang, J., Ulukus, S., 2010. Transmission completion time minimization in an energy harvesting system. conf .inf Sciences Syst.

[33] Zhang, P., Xiao, G., Tan, H.-P., 2013. Clustering algorithms for maximizing the lifetime of wireless sensor networks with energy-harvesting sensors. Computer Networks 57 (14), 2689–2704.

[34] Zhang, S., Seyedi, A., Sikdar, B., August 2013. An analytical approach to the design ofenergy harvesting wireless sensor nodes. IEEE Transations on Wireless Communications 12 (8).

Appendix A :Derivation of the equation 9 in page 4

The r.v. representing the effective energy, E_{eff}, for each time-slot, t_i, is cumulated from the difference between the incoming energy (which is in the interval $[\Psi_{min}, \Psi_{max}]$) and the consumed energy (C). Initially, if we assume the residual energy is X_0, then the energy in the first time slot ($E_{eff}(t_1)$) will be bounded in the interval, $[X_0 + \Psi_{min} - C, X_0 + \Psi_{max}]$,. The left term represents the lower bound which is reached from X_0 when the incoming energy is minimal (Ψ_{min}) and the node is active (consumes energy C). The right term represents the upper bound which is reached when the incoming energy is maximal (Ψ_{max}) and the node is inactive (no consumption). We may get by recurrence, $E_{eff}(t_2) \in [2\,\Psi_{min} + X_0 - 2\,C, 2\,\Psi_{max} + X_0]$, $E_{eff}(t_3) \in [3\,\Psi_{min} + X_0 - 3\,C, 3\,\Psi_{max} + X_0]$, etc. In general, $\forall i \in \{1, \ldots, W\}, E_{eff}(t_i) \in [i\,\Psi_{min} + X_0 - i\,C, i\,\Psi_{max} + X_0]$. We normalize these r.v., $E_{eff}(t_i)$. Let us denote the lower and upper bounds of its interval by, a, and, b, respectively. Since $E_{eff}(t_i)$ is ($U\tilde{}[a, b]$), then $\delta_i = \frac{E_{eff}(t_i)-a}{b-a}$ is $U\tilde{}[0, 1]$. That is,

$$\forall i \in \{1, \ldots, W\}, \delta_i = \frac{E_{eff}(t_i) - X_0 - i\,(\Psi_{min} - C)}{i\,(\Psi_{max} - \Psi_{min} + C)},$$

is $U\tilde{}[0, 1]$ □

Mustapha KHIATI received the Engineer and master degrees in Computer Engineering, from the University of Science and Technology, Oran, Algeria in 1999, and USTHB, Algiers in 2012, respectively. He worked on simulation of circuits during his engineering project and on broadcasting over duty-cycled wireless sensor networks during the master degree. From 2003 to 2009, has have been with a research and development center (CRD) in Algiers and served as a software developer using C language and UNIX-based systems. He participated in developing applications for embedded systems using FPGA cards, as well as implementing wireless sensor MAC protocols for intelmote2 real platform. Currently, Khiati Mustapha leads the data communication protocols laboratory and prepares a PhD thesis.

Djamel Djenouri obtained the PhD in Computer Science from the University of Science and Technology USTHB Algiers, Algeria, in 2007. From 2008 to 2009, he was granted a post-doctoral fellowship from the European Research Consortium on Informatics and Mathematics (ERCIM), and he worked at the Norwegian University of Science and Technology (NTNU), Norway. He is currently a senior research scientist (Director of Research) at the CERIST research center in Algiers, where he is leading the wireless sensor networking group, and adjunct full professor at the EMP, Algiers. He is working on topics related Internet of things, wireless and mobile networks, machine learning and application for smart environments, smart building and green applications. He has been conducting several research projects with international collaborations as the principal investor for many of them. He participated in many international conferences worldwide and gave many keynotes and plenary-session talks. He has been granted mobility internships and visited many renowned universities including NTNU, SICS (Stockholm), university of Cape town, UPC Barcelona, JMU Liverpool, University of Padova, and the university of Oxford. He published more than 90 papers in international peer-reviewed journals and conference proceedings, two books, and he is holding two national patents. He organized workshops held in conjunction with DCOSS and GlobeCom. He also served as TPC member of many international conferences (e.g., GlobeCom, LCN, WiMob, etc.), guest editor of a special issue with Int. J. Communication Networks and Distributed Systems, and reviewer for many Journals. He is a senior member of Association of Commuting Machinery (ACM), member of the Arab/German Young Academy of Science and Humanities (AGYA), and the Algerian NPC on ICT for the H2020 program.