POMDP Formulation for Energy Harvesting Network with Wireless Distributed Computing

Mohammed Alfaqawi, Jean Pierre Cances, Vahid Meghdadi, Oussama Habachi, Mohamed Habaebi, Mohammad U. Siddiqi, Md.Rafiqul Islam, Sheroz Khan

POMDP Formulation for Energy Harvesting Network with Wireless Distributed Computing

Mohammed I. M. Alfaqawi[1], Mohamed H. Habaebi[1], Oussama Habachi[2], Vahid Meghdadi[2], Mohammad U. Siddiqi[1], Md. Rafiqul Islam[1], Jean Pierre Cances[2] and Sheroz Khan[1]

[1] Electrical and Computer Engineering Department, International Islamic University Malaysia, Malaysia

[2] Ecole Nationale Superieure d Ingenieurs de Limoges, XLIM Institut de recherché, France

mohammedalfaqawi@gmail.com

Abstract— **Energy harvesting network (EHN) is a trending topic in wireless communications that enables wireless devices to harvest energy from ambient sources such as RF waves. However, EHN faces energy causality constraint that turns the energy harvester into sleep mode. This, in turn, delays the processing of next task until harvesting the required energy for processing. The objective of this research is to prolong the active mode of energy harvesters by employing a novel scheme of EHN that combines local computing and wireless distributed computing in one EHN. Accordingly, a novel EHN algorithm that employs partially observable Markov decision process (POMDP) is proposed in order to automate the decision of the energy harvester based on the channel variations and stored energy. Herein, the scheme of the proposed EHN is explained first. Then, the problem of the proposed EHN is formulated as POMDP. Furthermore, a policy is found in order to estimate the validity of the proposed EHN against the conventional EHN.**

Keywords—Energy harvesting network; local computing; wireless distributed computing; POMDP.

I. Introduction

Energy harvesting network (EHN) is a promising green communication technology that has various benefits such as unlimited power, does not require batteries replacement or cables connections and reducing CO_2 emissions [1], [2]. Recently, several research works on EHN have been proposed where the energy harvesting node processes it's task locally, i.e. local computing-EHN (LC-EHN), such as [1], [3]-[6] and the references therein. However, due to energy causality constraint, EHN is expected to switch into sleep mode faster when the energy harvester processes a computationally intensive task that requires high level of processing energy. This, in turn, will delay the processing of EHN's tasks until it operates in active mode again. Therefore, the main objective of this research is to prolong the active mode of energy harvesters by proposing a novel scheme that classifies EHN regions into two networks. The first one is local computing-EHN (LC-EHN) where the energy harvester processes the whole task locally. Meanwhile, the second is a novel proposed network that shares the computations of the energy harvester's tasks with other nodes wirelessly. This proposed network is referred to wireless distributed computing-EHN (WDC-EHN).

Various benefits could be gained due to employing WDC in conjunction with EHN. These benefits and challenges of applying WDC are discussed in detail in [7]-[9]. One of the key benefits of the proposed WDC-EHN is reducing the consumed energy in processing of the energy harvester particularly when the required energy for processing is high. However, even though employing WDC concept with EHN might reduce the consumed energy in processing, WDC-EHN would still lose energy when transmitting the distributed tasks through the wireless channel. Furthermore, in WDC network, master-slave cluster is typically employed where the master node transmits a pilot message to slave nodes which, in turn, estimate the channel state information (CSI). Then, the slave nodes transmit feedback CSI to the master node. Feedback CSI is crucial in case of collaborating networks in order to maintain reliable communication. However, feedback CSI will cause a communication overhead at the master node as well as a round-trip delay [10]. For these specific reasons, partially observable Markov decision process (POMDP) has been utilized to enable the energy harvesting nodes to act under uncertain channel conditions.

Herein, a system model for the proposed EHN has been described first. Then, POMDP framework has been utilized in order to automate the selection between LC-EHN and WDC-EHN as well as to nominate the best collaborating nodes in WDC-EHN that require less transmission energy. By modelling Rayleigh fading channel as finite state Markov chain (FSMC), herein, POMDP acts under imperfect CSI to obtain a policy that prolongs the duration of EHN's active mode. In order to prolong the duration of the active mode, a reward is assigned to each operational region of EHN, i.e. sleeping, LC-EHN and WDC-EHN. Due to the high level of POMDP's computational complexity, POMDP model is transferred and solved as a belief MDP.

The rest of this research paper is presented as follows, Section II. highlights the related research works. Section III. compares between the conventional EHN and the proposed EHN system models. Section IV. formulates the proposed novel EHN as POMDP problem. Section V. verifies and analyses the proposed EHN algorithm. Conclusion and future work are drawn in Section VI.

II. Related work

In [11], POMDP framework was employed in order to prolong the lifetime of EHN by deciding when the energy

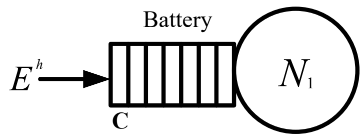

Fig. 1. The communication graph for LC-EHN.

harvester turns into active mode to harvest energy and when to turn into sleep mode to save the stored energy. Furthermore, a threshold-base optimal policy has been found to maximize the saved energy. Meanwhile, in [12], the transmission of the most valuable information was maximized by adapting the data rate and data size under delay and energy constraints. The proposed algorithm, in [12], was considered as reward maximization problem that was investigated in case of AWGN without considering any channel fading model.

In this research, Rayleigh faded channel has been considered and represented as FSMC. This model of Rayleigh fading has been described in [13]. FSMC Rayleigh fading model has been employed in [14] with POMDP framework in order to propose relay selection algorithms that mitigate the performance degradation due to imperfect CSI knowledge. On the other hand, the same channel model is utilized in [15] that proposes an online heuristic evaluation post-decision state for packet transmission in point to point communication model. Even though a FSMC Rayleigh fading channel has been employed in [15], the CSI is considered as priori knowledge. Therefore, the problem of packet transmission has been modeled as constraint Markov process decision (CMDP). Similarly, in [16], given imperfect CSI by considering Rayleigh fading channel as FSMC, the problem of maximizing the transmission throughput has been formulated as POMDP. In [16], buffer and channel adaptive transmission scheme was proposed where the system state is composed of the number of packets in the buffer and channel state. However, in [16], obtaining an optimal policy for the possible observations and belief channel states has been described as impossible due to the required infinite time and memory. Therefore, assumptions have been considered in order to approximate POMDP's optimal policy. Obtaining an optimal policy of the proposed EHN, herein, is found to require infinite time and memory because the number of states is higher than the states number in [16]. The high level of POMDP's computational complexity was highlighted in [17]-[19].

III. System Model

This section, firstly, differentiates between the conventional model of EHN and the proposed WDC-EHN scheme. Then, the channel model among the collaborating nodes in WDC-EHN is described.

A. The Conventional Energy Harvesting Network Scheme.

Typically, EHN considers only LC-EHN scheme. This scheme is composed of an energy harvesting node that harvests energy from ambient sources and then stores the harvested energy in battery. The stored energy $e^{st} \in \boldsymbol{E^{st}}$ will be utilized later, by the node, in processing the node's tasks. Fig.1 presents LC-EHN scheme that consists of one node, i.e. N_1, where N_1 is assumed that does not supplied by a fix power source and equipped by an energy harvester. At each time slot t, the harvested energy $e_t^h \in \boldsymbol{E^h}$ will be stored in battery of limited capacity C.

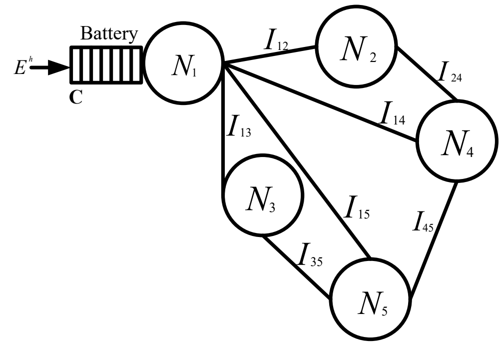

Fig. 2. The communication graph of WDC-EHN.

The harvested energy of node N_1 is assumed as Bernoulli process with probability mass function (PMF) as follows [5],

$$p(e_t^h) = \begin{cases} 1 - p_h, & e_t^h = 0. \\ p_h, & e_t^h = e_h. \end{cases} \quad \textbf{(1)}$$

On the other hand, the stored energy e_t^{st}, at slot time t, can be computed as given,

$$e_t^{st} = e_{t-1}^{st} - e_{t-1}^{c} + e_{t-1}^{h} \quad \textbf{(2)}$$

where e_t^c is the consumed energy in case of either LC-EHN or WDC-EHN.

At the beginning of each time slot t, the conventional EHN operates in two modes, i.e. sleep mode and active mode. Typically, in active mode, N_1 considers only LC-EHN where e_t^c will be consumed in processing the task locally as follows,

$$e_t^p(f) = P^p(f) \cdot T^p(f) \quad \textbf{(3)}$$

$$P^p(f) = af^3 + bf^2 + cf + d \quad \textbf{(4)}$$

where $e_t^p \in \boldsymbol{E^p}$ represents the required energy for processing the task $w_t \in \boldsymbol{W}$ and $\boldsymbol{W}$ is the set of tasks. T^p is the required processing time at clock frequency f. P^p is the consumed power in processing w_t. a, b, c and d are all positive and dependent on the processor specifications [8].

B. The Proposed Scheme for Wireless Distributed Computing-Energy Harvesting Network.

The proposed EHN operates also in two modes, i.e. sleep and active modes, however in active mode EHN might operates as LC-EHN or WDC-EHN. To the best of our knowledge, the concept of WDC-EHN is proposed for the first time in this research. In [7]-[9], master-slave cluster were employed for collaboration in WDC network. However, the master node in the proposed scheme is assumed to be equipped with energy harvester. Fig. 2 shows the communication scheme, denoted by $\boldsymbol{G}$, of WDC-EHN. The vertices of $\boldsymbol{G}$ represents the master-slave nodes where the master node is N_1, the slave or cooperating nodes are $\boldsymbol{N_g} = \{N_2, N_3, \dots, N_m\}$ and m is the number of slave nodes. I_{lj} represents the wireless links between the cooperating nodes N_l and N_j.

The wireless link I_{lj} is considered as Rayleigh fading channel where the received instantaneous SNR γ is distributed exponentially as follows [13],

$$p_\gamma(\gamma) = \frac{\gamma}{\gamma_0^2} e^{-\frac{\gamma}{\gamma_0}} \quad ,\gamma \geq 0 \tag{5}$$

where γ_0^2 is the average SNR. In [13], γ was represented as FSMC and assigned into $K+1$ non-overlapping thresholds $\{\Gamma_k\}_{k=0}^{K+1}$ where k represents the state of γ, $\Gamma_0 = 0$ and $\Gamma_{K+1} = \infty$. Consider the boundaries of $\{\Gamma_k\}_{k=0}^{K+1}$ are given, the transition probabilities from current state k to an adjacent state $k+1$ or $k-1$ can be computed as follows [13],

$$P_{k,k+1} \approx \frac{N(\Gamma_{k+1})\tau}{\pi_k}, \quad k = 0,1, \dots, K-1 \tag{6}$$

$$P_{k,k-1} \approx \frac{N(\Gamma_k)\tau}{\pi_k}, \quad k = 1, \dots, K \tag{7}$$

where τ is the packet time period, π_k is the steady state probability and can be computed as follows,

$$\pi_k = \int_{\Gamma_k}^{\Gamma_{k+1}} p_\gamma(\gamma) d\gamma = e^{-\frac{\Gamma_k}{\gamma_0}} - e^{-\frac{\Gamma_{k+1}}{\gamma_0}} \tag{8}$$

$N(\Gamma_k)$ represents the level crossing rate of the received SNR and can be estimated as follows,

$$N(\Gamma_k) = \sqrt{\frac{2\pi\Gamma_k}{\gamma_0}} f_m e^{-\frac{\Gamma_k}{\gamma_0}} \tag{9}$$

where f_m is the maximum Doppler frequency. Meanwhile, $P_{k,k} = 1 - P_{k,k-1} - P_{k,k+1}$ and $P_{K+1,K+2} = P_{0,-1} = 0$ [13].

In case of WDC-EHN, N_1 segments the task w_t, first, into subtasks w_j, where $2 \leq j \leq m$. Then, each subtask will be transmitted to a corresponding slave node and consumes e_t^c energy unit in transmission as follows [20],

$$e_t^T = w_j \cdot \frac{\rho\sigma^2}{\gamma B A_t} \tag{10}$$

where $e_t^T \in \boldsymbol{E^T}$ represents the required energy to transmit w_j bits wirelessly to the cooperating nodes. ρ is the targeted signal to interference ratio (SIR), σ^2 is the variance of thermal noise, B is the channel bandwidth and A_t is the attenuation factor.

IV. THE PROPOSED ENERGY HARVESTING NETWORK

Considering N_1 is the agent, the novel EHN can be formulated as POMDP tuples $< S, A, O, T, Z, R, \lambda, b >$ as follows,

A. POMDP States of The Novel EHN

Consider the amount of stored energy of the agent N_1 is $e_1^{st} \in \boldsymbol{E_1^{st}}$ and the required amount of energy for processing the task w_t by using LC-EHN is $e^p \in \boldsymbol{E^p}$. On the other hand, based on (10), for constant channel bandwidth, the required energy to transmit w_j, from N_1 to N_j, wirelessly is $e_{1j}^T \in \boldsymbol{E^T}$ that is dependent on γ. Accordingly, when the state of γ is located in any interval $[\Gamma_k, \Gamma_{k+1})$, a corresponding $(e_{1j}^T)_k$ will be required to transmit w_j wirelessly, i.e. $f: [\Gamma_k,\ \Gamma_{k+1}) \rightarrow (e_{1j}^T)_k$. When γ transits to any adjacent interval, e_{1j}^T will vary based on the corresponding state of γ. Therefore, for $j \in \{2,3, \dots, m\}$ and $k \in \{0,1, \dots, K+1\}$, $(e_{1j}^T)_k = f(\Gamma_k)_j$ which represents the required energy to transmit w_j from N_1 to N_j at channel gain Γ_k. For the sake of simplicity, $(e_{1j}^T)_k$ and $f(\Gamma_k)_j$ will be denoted as e_{1j}^T and $f(\Gamma_k)$. Accordingly, the set of states of N_1 is represented as follows,

$$\boldsymbol{S_1} = \{\boldsymbol{E_1^{st}}, \boldsymbol{E^p}, \boldsymbol{E^T}\} \tag{11}$$

where the battery capacity of N_1 is segmented into slots such that $\boldsymbol{E_1^{st}} = \{0,1,2, \dots, C\}$. Furthermore, the required energy for processing the task is quantized as $\boldsymbol{E^p} = \{1,2, \dots, C\}$, where the maximum required energy for processing is assumed to be equal to C. Meanwhile, the corresponding set of required energy for transmission from N_1 to N_j is $\boldsymbol{E^T} = \{f(\Gamma_0), f(\Gamma_1), \dots, f(\Gamma_{K+1})\}^m$. Accordingly, the total number of N_1's states is $\kappa = |\boldsymbol{S_1}| = (C \cdot (C+1) \cdot (K+1)^m)$. For $s_1 \in \boldsymbol{S_1}$ where $s_1 = (e_1^{st}, e^p, e_{12}^T, \dots, e_{1m}^T)$, e_1^{st} and e^p are known to N_1 while, due to channel randomness, e_{1j}^T is unknown. For this specific reason, the proposed EHN is considered as a POMDP where the current state at time t is partially known. Therefore, N_1 formulates the uncertainty of the state by a statistical distribution that is referred to as belief function $\boldsymbol{b}$. At time slot t, the belief vector over the state is defined as $\boldsymbol{b_{S_1}}(t) = \{b_{s_1}(t), b_{s_2}(t), \dots, b_{s_\kappa}(t)\} \in [0,1]^\kappa$.

B. POMDP Actions of The Novel EHN

$\boldsymbol{A}$ represents the set of actions to be taken by N_1. The set of actions $\boldsymbol{A} = \{\boldsymbol{A_1}, \boldsymbol{A_2}\}$ represents the modes of the novel proposed EHN where sleeping, LC-EHN and WDC-EHN are combined in one operational EHN as follows,

$$\boldsymbol{A_1} = \{a_{sleeping}, a_{LC-EHN}\}$$
$$\boldsymbol{A_2} = \{a_{WDC-EHN_x}\}_{x=1}^{z} \tag{12}$$

At each time t, N_1 chooses an action $a(t) \in \boldsymbol{A}$ to decide whether to turn into sleep mode or active mode. The action $a(t) = a_{sleeping}$ represents sleep mode while $a(t) = a_{LC-EHN}$ is the action taken by N_1 to process the task locally. On the other hand, $a(t) \in \boldsymbol{A_2}$ is an action of N_1 to operate as WDC network. In this case, N_1 nominates q nodes to collaborate with such that $\boldsymbol{N_q} \subset \boldsymbol{N_g}$ and $q \leq m$, segments the task into subtasks and transmits the subtasks wirelessly to the permutations of q nominated nodes. For instance, when $m = 5$ and $q = 2$, the possible combinations of two nodes is $\boldsymbol{N_2} = \{\{N_2, N_3\}, \{N_2, N_4\}, \dots, \{N_5, N_4\}\}$ where each set is equivalent to an action $a_{WDC-EHN_x} \in \boldsymbol{A_2}$. In this example, the number of WDC actions is equal to $C_5^2 = 10$ where $C_m^q = \frac{m!}{(m-q)! \cdot q!}$. Accordingly, the total number of possible WDC actions is $z = \sum_{q=1}^{m} C_m^q$. Meanwhile, the total number of N_1 actions is $|\boldsymbol{A}| = 2 + z$.

C. POMDP Transition Probability of The Novel EHN

When N_1 takes an action $a(t) \in \boldsymbol{A}$ at current state $\boldsymbol{S_1} = \boldsymbol{S_1}(t)$ and belief $\boldsymbol{b_{S_1}}(t)$, N_1 transits to a future state $\boldsymbol{S_1'} = \boldsymbol{S_1}(t+1)$, receives an observation, immediate reward and update the belief as follows,

$$\boldsymbol{b}_{\boldsymbol{S_1}}(t+1) = \frac{Z(\boldsymbol{S}'_1, a, o)\sum_{\boldsymbol{S_1}} T(\boldsymbol{S_1}, a, \boldsymbol{S}'_1)\boldsymbol{b}_{\boldsymbol{S_1}}(t)}{p(o'|a, \boldsymbol{b}_{\boldsymbol{S_1}}(t))} \quad \textbf{(13)}$$

where $\boldsymbol{b}_{\boldsymbol{S_1}}(t+1) \in [0,1]^{\kappa}$, $Z(\boldsymbol{S}'_1, a, o)$ is the observation probability, $p(o'|a, \boldsymbol{b}_{\boldsymbol{S_1}}(t))$ is a normalization factor and $T(\boldsymbol{S_1}, a, \boldsymbol{S}'_1)$ represents the transition probability where $T(\boldsymbol{S_1}, a, \boldsymbol{S}'_1) = p(\boldsymbol{S}'_1|\boldsymbol{S_1}, a) \in [0,1]^{\kappa\times\kappa}$ and can be computed as stated in **(14)**.

Due to the certain knowledge of e_1^{st} and e^p, the transition among the different states in the proposed EHN depends only on the channel state and the harvesting probability. Considering the current state of the channel is k and $p(e_t^h = e_h)$, $T(\boldsymbol{S_1}, a, \boldsymbol{S}'_1)$ has two main transitions as follows *1)* When $a(t) = a_{LC-EHN}$ or $a(t) = a_{sleeping}$, in this case N_1 will not receive an observation of the channel. Accordingly, each slave node $N_j \in \boldsymbol{N_g}$ has three possibilities for state transitions $(P_{k,k+i} \cdot p_h)_{1j}$ to move from channel state k to an adjacent state $k+i$, where $i = \{-1,0,1\}$. *2)* When $a(t) = a_{WDC-EHN_x}$, N_1 nominates the cooperating nodes $\boldsymbol{N_q} \subset \boldsymbol{N_g}$. Accordingly, N_1 receives an observation of the channel state of the nominated nodes $\boldsymbol{N_q}$. In this case, N_1 knows the exact channel state of $\boldsymbol{N_q}$ and thus predicts the channel state for the nominated nodes as the multiplication of transfer from channel state k to $k+i$. For each q, every permutation set of $\boldsymbol{N_q}$ combines different nominated nodes but has the same number of elements. Therefore, the multiplication is considered over the element index I in the set $\boldsymbol{N_q}$. For instance, when $m = 5$ and $a = a_{WDC-EHN_2}$, for any combination of two nodes, e.g. $\{N_2, N_4\} \subset \boldsymbol{N_2}$, $T(\boldsymbol{S_1}, a, \boldsymbol{S}'_1)$ is the multiplication of the probability of the first nominated node, at $I = 1$, to be in channel state k then move to $k+i$ with the probability of the second nominated node, at $I = 2$, to be in channel state k then move to $k+i$.

D. POMDP Observations of The Novel EHN

The channel observation is crucial for N_1 in order to allocate the equivalent gain threshold Γ_k, $e_{1j}^T = f(\Gamma_k)$ and the possible transitions from $\boldsymbol{S_1}$ to $\boldsymbol{S}'_1$. Given a current state $\boldsymbol{S_1}$ and an action $a(t) = a_{WDC-EHN_x}$ that leads to a future state $\boldsymbol{S}'_1$. N_1 makes an observation $o \in \boldsymbol{O}$ as follows,

1- N_1 transmits w_j to the nominated nodes $\boldsymbol{N_q}$ through the wireless channel.
2- $\boldsymbol{N_q}$ process the task and estimate γ_{1j}.
3- $\boldsymbol{N_q}$ transmit the result of the task and feedback γ_{1j} to N_1.

The received γ_{1j} represents the observation of N_1. The observation probability $Z(\boldsymbol{S}'_1, a, o) = p(o|\boldsymbol{S}'_1, a)$ is represented as follows,

$$Z(\boldsymbol{S}'_1, a, o) = p(o = \theta|\boldsymbol{S}'_1, a) = \begin{cases} 1, & b(\boldsymbol{S}'_1) > 0;\ a(t) = a_{LC-EHN} \\ & or\ a(t) = a_{sleeping};\ \theta = o_{\gamma_{1j}}. \\ 1, & a(t) = a_{WDC-EHN_x}; \theta = o_{\gamma_{1j}}. \\ 0, & otherwise. \end{cases} \quad \textbf{(15)}$$

E. The Immediate Reward of the Proposed EHN

For each action $a \in \boldsymbol{A}$, N_1 receives an immediate reward $R: \boldsymbol{S_1} \times \boldsymbol{A}$ that is given by,

$$R_t(\boldsymbol{S_1}, a) = \begin{cases} -\delta_s + p_h e_h, & a(t) = a_{sleeping}; p(e_t^h = e_h). \\ -e^p + p_h e_h, & a(t) = a_{LC-EHN}; p(e_t^h = e_h). \\ -\sum_{I=1}^{q} \delta_q e_{1I}^T + p_h e_h, & a(t) = a_{WDC-EHN_x}; p(e_t^h = e_h). \end{cases} \quad \textbf{(16)}$$

In case of $a(t) = a_{LC-EHN}$, the immediate reward of N_1 is the dissipated energy in processing the task locally while, in case of $a(t) = a_{WDC-EHN_x}$, $R_t(\boldsymbol{S_1}, a)$ is the amount of consumed energy to transmit the subtasks wirelessly from N_1 to $\boldsymbol{N_q}$. Herein, the task will be distributed uniformly. Therefore, for each $N_j \in \boldsymbol{N_q}$, the number of bits w_t will be distributed equally among the nominated nodes such that $w_j = w$. In this case, e_{1I}^T is the required amount of energy to transmit w to each nominated node $N_j \in \boldsymbol{N_q}$ and can be computed using (10). On the other hand, $a(t) = a_{sleeping}$ will not cause energy dissipation. Nevertheless, selecting $a(t) = a_{sleeping}$ repeatedly would delay the processing of next task. Therefore, a penalty δ_s is considered in case of sleeping mode in order to control the delay of processing next task. Meanwhile, δ_q is a tuning penalty for each $\boldsymbol{N_q}$, where $0 < \delta_q \leq 1$. For instance, when $q = 1$, the proposed EHN will be as similar as to the collaboration models in cloud computing technology where the node transmits the task onto a cloud which is a central unit, such as in [21]. However, herein, the slave nodes has limited resources. Tuning δ_q is a task scheduling problem which is out of this research scope.

F. The Optimal Policy of the Proposed EHN

The optimal policy of POMDP is a set of actions to be taken at each corresponding belief vector that results in

$$p(s'_1|s_1, a) = \begin{cases} 0, & e_{t+1}^{st} \neq e_t^{st} + e_t^h\ or\ e_{t+1}^{st} \neq e_t^{st};\ a(t) = a_{sleeping}. \\ 0, & e_{t+1}^{st} \neq e_t^{st} - e_t^c + e_t^h\ or\ e_{t+1}^{st} \neq e_t^{st} - e_t^c; \\ & a(t) = a_{LC-EHN}\ or\ a_{WDC-EHN_{1q}}. \\ \left[\left\{(P_{k,k+i} \cdot p_h)_{1j}\right\}_{j=2}^{m}\right]^{\kappa\times\kappa}, & a(t) = a_{sleeping}\ or\ a(t) = a_{LC-EHN};\ p(e_t^h = e_h). \\ 0, & a(t) = a_{WDC-EHN_x};\ p(e_t^h = e_h);\ \boldsymbol{N_u} = \boldsymbol{N_q} \backslash \boldsymbol{N_g}. \\ \left[\prod_{I=1}^{q} (P_{k,k+i})_I \cdot p_h\right]^{q\times q}, & a(t) = a_{WDC-EHN_x};\ p(e_t^h = e_h); \boldsymbol{N_q} \subset \boldsymbol{N_g}. \end{cases} \quad \textbf{(13)}$$

maximizing the expected long-term reward $J_\lambda^\pi(\boldsymbol{b}_{S_1}) = E[\sum_{t=0}^{\infty} \lambda^t R_{\pi_t}(\boldsymbol{b}_{S_1}(t)|\boldsymbol{b}_{S_1}(1))]$ where λ is a geometric discount factor, $\boldsymbol{b}_{S_1}(1)$ is an initial belief and π is the policy. The policy could be stationary π or non-stationary π_t. Herein, the required policy to maximize $J_\lambda^\pi(\boldsymbol{b}_{S_1})$ is a stationary policy because POMDP framework for the proposed EHN has a discrete state and action space, the proof results is presented in Theorems 8.10.9 and 8.10.7 of [18].

Herein, POMDP framework of the proposed EHN is solved as belief MDP by mapping the states to actions, i.e. $\pi: \boldsymbol{S_1} \rightarrow \boldsymbol{A}$. In order to find the optimal policy π^* that maximizes the saved energy by the energy harvester N_1, the value iteration function $V(\boldsymbol{S_1})$ of each policy has to be computed, first, as follows,

$$V_\pi(\boldsymbol{S_1}) = R(\boldsymbol{S_1}, a(\pi)) + \lambda \sum_{\boldsymbol{S}_1' \in \boldsymbol{S}} T(\boldsymbol{S_1}, a(\pi), \boldsymbol{S}_1') \sum_{o_i \in O} Z(\boldsymbol{S}_1', a(\pi), o) V_{o(\pi)}(\boldsymbol{S}_1') \quad \textbf{(17)}$$

Then, $V_\pi(\boldsymbol{b}_{S_1})$ is computed for each belief vector as follows,

$$V_\pi(\boldsymbol{b}_{S_1}) = \sum_{\boldsymbol{S_1} \in \boldsymbol{S}} \boldsymbol{b}_{S_1} V_\pi(\boldsymbol{S_1}) \quad \textbf{(18)}$$

The optimal policy π^* is the policy that achieves maximum $V(\boldsymbol{b}_{S_1})$ and given by,

$$\pi^* = \underset{\pi \subset \boldsymbol{A}}{\operatorname{argmax}} V_\pi(\boldsymbol{b}_{S_1}) \quad \textbf{(19)}$$

V. NUMERICAL RESULTS

In order to verify the proposed EHN, the channel thresholds Γ_k were modelled as in [13]. The range of SNR is assigned to 11-states of FSMC when $f_m = 8.7963\ Hz$, transmission rate $= 100\ kb/s$ and modulation scheme is $\pi/4$-DQPSK. The transmission parameters are assumed as in [20] where $\sigma^2 = 5 \times 10^{-15}$, $B = 10MHz$ and $A_t = 1.916 \times 10^{-14}$. Due to the high number of states κ, the proposed model for EHN can be solved offline [16], [22]. Herein, the complexity is reduced by assuming that, for each q, N_1 is considered to collaborate with the set of $\boldsymbol{N_q}$ that has the best channel state. Accordingly, z is reduced into the number of cooperating nodes, i.e. $z = m$.

The energy harvester's battery is segmented into 5-slots, $w_t = 300$ symbol, $\delta_s = 0.4$, $\delta_q = 1$ and $\lambda = 0.9$. The consumed energy of the proposed EHN is estimated, as shown in Fig. 3, for $m = 4$, 600-stages and at various required values of energy to process each symbol of data, i.e. e^p. The proposed EHN is compared against the conventional EHN as shown in Fig. 4 at $m = 2$, $m = 4$ and various values of $e^p = [10^{-7}, 10^{-10}]\ J/symbol$. At low e^p values, e.g. $e^p = 10^{-10}\ J/symbol$ to $10^{-9} J/symbol$, the proposed EHN is found to consume energy as same as the conventional EHN where the tasks are executed locally. Meanwhile, with increasing e^p, the conventional EHN is found to consume more energy than the proposed EHN.

In order to, further, investigate the performance of the proposed EHN, the policy is found for each belief vector at $\lambda = 0.016$, $\delta_s = 3.4733 \times 10^{-4}$ and $\delta_q = \frac{1}{2+0.1q}$ as well as at three SNR's ranges $\gamma =]-13{,}2]$, $\gamma =]2{,}7]$ and $\gamma =]10{,}15]$. Each set of γ is assigned to 4-states of channel gains. As presented in Table I, at some belief vectors, $a(t) = a_{WDC-EHN_q}$ is found as the best action to maximize the expected long-term reward. Meanwhile, with increasing the values of SNR range, the total number of $a_{WDC-EHN_q}$, i.e. z, is slightly increasing. In order to evaluate the policies in Table I the consumed energy in case of the proposed EHN is estimated. Fig. 5 presents the consumed energy in case of the

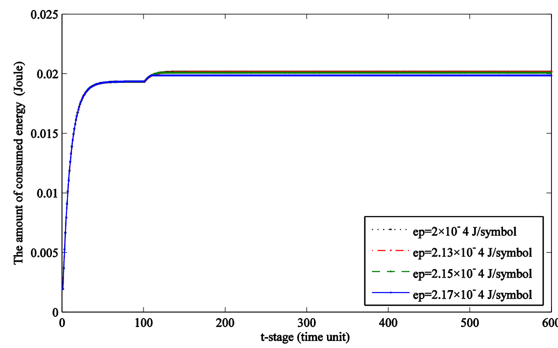

Fig.3. The consumed energy of the proposed EHN at $m = 4$.

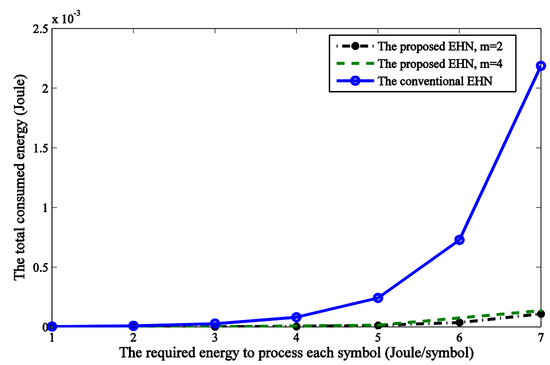

Fig.4. The consumed energy of the proposed and conventional EHN at $m = 2$ and 4.

TABLE I. THE SELECTION OF EACH ACTION

	The number of times each action is selected					
Action index	***1***	***2***	***3***	***4***	***5***	***6***
Action	$a_{sleeping}$	a_{LC}	$a_{WDC-EHN_1}$	$a_{WDC-EHN_2}$	$a_{WDC-EHN_3}$	$a_{WDC-EHN_4}$
$\gamma =]-13{,}2]$.	18	418	22	15	117	6
$\gamma =]2{,}7]$.	25	412	13	13	119	18
$\gamma =]10{,}15]$.	14	420	17	12	118	19

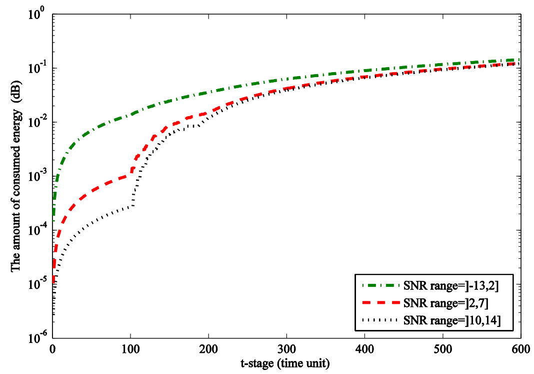

Fig.5. The consumed energy of the proposed EHN at various SNR ranges.

proposed EHN at distinct SNR ranges. The expected long-term reward is computed for 600-stages in case of the proposed EHN for each policy in Table I. The proposed EHN is found to save more energy when the wireless channel among the collaborating nodes has high SNR values.

VI. CONCLUSIONS AND FUTURE DIRECTIONS

A novel EHN that combines LC-EHN and WDC-EHN is proposed by employing POMDP framework. The proposed EHN is found to outperform the conventional EHN in terms of the consumed energy. Therefore, the proposed EHN is found to prolong the active mode of energy harvesters. This research introduces various challenges and open issues. Herein, a policy is found to prolong the active mode of the energy harvester while obtaining the optimal policy that maximizes the active mode will be part of future work. Furthermore, the workload distribution is a task scheduling problem that distributed, herein, uniformly. Therefore, proposing a novel task scheduling algorithm that considers the proposed EHN states and the amount of stored energy in each cooperating node is an open issue for future work.

REFERENCES

[1] Yin Sixing; Zhang Erqing; Qu Zhaowei; Yin Liang; Li Shufang, "Optimal Cooperation Strategy in Cognitive Radio Systems with Energy Harvesting,"*Wireless Communications, IEEE Transactions on*, vol.13, no.9, pp.4693,4707, Sept. 2014.

[2] M. I.M. Alfaqawi, M. Hadi Habaebi, M. Siddiqi, Md. Rafiqul Islam, Alhareth Zyoud and M. Al-Shibly "Survey On Energy Harvesting Cognitive Radio Network," *ARPN Journal of Engineering and Applied Sciences, vol. 10, no. 23, December 2015.*

[3] Park, Sungsoo; Kim, Hyungjong; Hong, Daesik, "Cognitive Radio Networks with Energy Harvesting," *IEEE Transactions on Wireless Communications*, vol.12, no.3, pp.1386,1397, March 2013.

[4] Zheng Gan; Ho, Z.; Jorswieck, E.A.; Ottersten, B., "Information and Energy Cooperation in Cognitive Radio Networks," *IEEE Transactions on,Signal Processing,* vol.62, no.9, pp.2290,2303, May1, 2014.

[5] Park Sungsoo, Lee Seokwon, Kim Beomju, Hong Daesik, and Lee Jemin. 2011. Energy-efficient opportunistic spectrum access in cognitive radio networks with energy harvesting. In *Proceedings of the 4th International Conference on Cognitive Radio and Advanced Spectrum Management* (CogART '11). ACM, New York, NY, USA, , Article 62 , 5 pages.

[6] Lu Xiaomei; Xu Wenjun; Li Shengyu; Lin Jiaru; He Zhiqiang, "Simultaneous information and power transfer for relay-assisted cognitive radio networks," *IEEE International Conference on Communications Workshops (ICC), 2014*, vol., no., pp.331,336, 10-14 June 2014.

[7] D. Datla, Xuetao Chen, T. Tsou, S. Raghunandan, S.M. Shajedul Hasan, J.H. Reed, C.B. Dietrich, T. Bose, B. Fette and J. Kim, "Wireless distributed computing: a survey of research challenges," *IEEE Communications Magazine,* pp. 144- 152, Jan 2012.

[8] D. Datla, H.I. Volos, S.M. Hasan, J.H. Reed and T. Bose, "Wireless distributed computing in cognitive radio networks," *Elsevier Ad Hoc Networks,* pp. 845-856, 2011.

[9] M. Alfaqawi, M. H. Habaebi, J. Chebil, "Wireless Distributed Computing For Cyclostationary Feature Detection" *Elsevier Digital Communications and Networks journal, Volume 2, Issue 1, Pages 47-56, February 2016.*

[10] A. Adhikary, H. C. Papadopoulos, S. A. Ramprashad and G. Caire, "Multi-user MIMO with outdated CSI: Training, feedback and scheduling,"*Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, Monticello, IL, 2011, pp. 886-893.

[11] 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* , vol.PP, no.99, pp.1-1, 2016.

[12] J. Gong, X. Zhong and C. Z. Xu, "Maximizing Rewards in Wireless Networks with Energy and Timing Constraints for Periodic Data Streams," *in IEEE Transactions on Mobile Computing*, vol. 9, no. 8, pp. 1187-1200, Aug. 2010.

[13] Qinqing Zhang and S. A. Kassam, "Finite-state Markov model for Rayleigh fading channels," *in IEEE Transactions on Communications*, vol. 47, no. 11, pp. 1688-1692, Nov 1999.

[14] F. Babich, A. Crismani and L. Hanzo, "Relay selection schemes relying on adaptive modulation and imperfect channel knowledge for cooperative networks," *2012 IEEE International Conference on Communications (ICC)*, Ottawa, ON, 2012, pp. 4300-4304.

[15] Juan Zhang, Hong Jiang, Zhenhua Huang, Chunmei Chen, and Hesong Jiang, "Study of Multi-Armed Bandits for Energy Conservation in Cognitive Radio Sensor Networks," *Sensors*, vol. 15, no. 4, pp. 9360–9387, 2015.

[16] A. T. Hoang and M. Motani, "Cross-layer adaptive transmission with incomplete system state information," *in IEEE Transactions on Communications*, vol. 56, no. 11, pp. 1961-1971, November 2008.

[17] Leslie Pack Kaelbling, Michael L. Littman, Anthony R. Cassandra, Planning and acting in partially observable stochastic domains, *Artificial Intelligence*, Volume 101, Issue 1, Pages 99-134, 1998.

[18] M. L. Putterman, "Markov Decision Process Discrete Stochastic Dynamic Programming," *WILEY Series in Prob. and Stat*., 2005.

[19] O. Habachi, H. P. Shiang, M. van der Schaar and Y. Hayel, "Online Learning Based Congestion Control for Adaptive Multimedia Transmission," *in IEEE Transactions on Signal Processing*, vol. 61, no. 6, pp. 1460-1469, March15, 2013.

[20] C. Pandana and K. J. R. Liu, "Near-optimal reinforcement learning framework for energy-aware sensor communications," *in IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 788-797, April 2005.

[21] F. Liu et al., "Gearing resource-poor mobile devices with powerful clouds: architectures, challenges, and applications," *in IEEE Wireless Communications*, vol. 20, no. 3, pp. 14-22, June 2013.

[22] Khodaian, M., Pérez, J., Khalaj, B.H., Pedro M Crespo,"Adaptive access and rate control of CSMA for energy, rate, and delay optimization," *EURASIP Journal on Wireless Communications and Networking*, Volume 2012, Issue 1, pages 27, 2012.