An Analytical Approach towards Cooperative Relay Scheduling under Partial State Information

Huijiang Li†, Neeraj Jaggi‡ and Biplab Sikdar†

†Department of ECSE, Rensselaer Polytechnic Institute, Troy, NY 12180

‡ Department of EECS, Wichita State University, Wichita, KS 67226

***Abstract*—Energy harvesting and cooperative communication are promising solutions to overcome the power limitations of Wireless Sensor Networks (WSNs) comprising of battery-powered nodes. In order to maximize the efficiency of such systems, measured in terms of packet delivery ratio achieved over time, efficient scheduling algorithms need to be designed. In particular, relay usage scheduling is critical for addressing the trade-off between energy consumption and efficiency in the network. However, the stochastic nature of the recharge and traffic generation processes at the sensor nodes, along with partial state information availability about neighboring nodes, makes the transmission and relay scheduling problem quite challenging. To address this problem, we model the system using a stochastic framework, and formulate the scheduling problem at source sensor node, when only partial state information about the relay is available at the source, as a Partially Observable Markov Decision Process (POMDP). We characterize an approximate solution to the optimality equations, which provides us with useful insights into the system dynamics. We observe that the structure of optimal policy is quite sensitive to system parameters, which makes it unsuitable for practical deployment. Therefore, we design a simple and practical threshold based relay scheduling policy, and show using simulations that it achieves close to optimal performance.**

***Index Terms*—Cooperative communications, Energy harvesting, Relay scheduling**

I. Introduction

Wireless Sensor Networks have a wide range of applications in various fields but are typically limited by the size of the battery and the power it can store. Existing research has shown that cooperative diversity gains can be achieved in distributed networks using relaying transmissions [1], [2], resulting in either higher network capacity or lower energy consumption with the same capacity. This paper addresses the problem of scheduling transmissions in sensor networks where nodes may use cooperative communications strategies.

The cooperative communication considered in this paper is the simple discrete memoryless three-terminal relay network [3], [4]. For such networks, the capacity, the strategies on the relay, energy efficiency and distribution among the network, have been the focus of intensive research [5]. However, exploiting the cooperative gain itself is not enough to address the fundamental problem of the limited battery capacity in sensor nodes. The limited battery capability of sensor nodes is a key challenge in the widespread deployment of WSNs. Although current battery technology is incapable of facilitating sensor networks with a sufficiently long life for many applications, energy harvesting or energy scavenging has become a promising and feasible approach to address the energy supply problem [6]. However, to improve the performance of energy harvesting WSNs to acceptable levels, harvesting-aware communication policies and protocols need to be developed.

This paper considers sensor networks with energy harvesting capability and addresses the problem of scheduling cooperative, relay based communications. We consider a time-slotted source-relay-destination system, where a sensor (the source) has the option to have another sensor (the relay) help to transmit its data to the destination. All sensor nodes under consideration are equipped with energy harvesting capability. From an energy efficiency perspective, the source may achieve the same bit error rate (BER) for a lower transmission power if it uses a relay, as compared to a direct transmission.[1] However, this increases the power consumed by the relay and as a result, the relay sensor may not have energy to report its own data in the future.

In order to optimally determine if the relay should be used or not at a given time to maximize the long term ratio of the data that is successfully delivered to the total data that is generated, in addition to its own state information (e.g. current battery level), the source also needs to know the state information at the relay. It is reasonable to assume that when a relay transmits or relays data, the header of the packets includes the relay's state information. However, in periods without data, broadcasting the state information of the relay in real time represents a significant overhead. Thus the source may have to base its decision on stale relay state information. The focus of this paper is to determine optimal decision at the source despite the *partial information* availability in the system.

The rest of the paper is organized as follows. Section II describes the system model. The transmission scheduling problem is formulated as a *Partially Observable Markov Decision Process* (POMDP) in Section III and the structure of optimal reward is derived as a function of system parameters. A near-optimal, practical and simple relay usage scheduling scheme is developed in Section IV, its performance is compared with the value iteration results representing the optimal performance, and the application is extended to multi-node networks. Section V concludes the paper.

[1]The actual gain depends on the relative positions of the nodes and it is possible to have scenarios where a direct transmission may be more energy efficient. Our work is applicable for cases where relay based transmissions are preferable.

TABLE I
NOTATION: ENERGY GENERATION PROCESS AND EVENT GENERATION PROCESS

Symbol	Definition
q_{on}	probability that the sensor harvests energy in the next slot given that it harvested energy in the current slot
q_{off}	probability that no energy is harvested in the next slot given that no energy is harvested in the current slot
c	units of energy harvested per time slot
μ_{on}	the steady-state probability of energy harvesting in a slot
p_{on}	probability that an event is generated in the next slot given that an event is generated in the current slot
p_{off}	probability that no event is generated in the next slot given that an event is generated in the current slot
π_{on}	the steady-state probability of event occurrence

II. SYSTEM MODEL

We consider a WSN where each sensor can be categorized as either a source or a relay sensor. Every source sensor has a designated relay counterpart. A source sensor has two transmission modes: the *direct mode* in which the sensor transmits the packet directly to the destination and consumes δ_1^s units of energy and the *relay mode* (which consumes δ_2^s units of energy at the source) in which the packet is transmitted by the source and relayed by the relay sensor. A relay sensor also has two transmission modes: *own-traffic mode* and *relay mode*. In own-traffic mode, the relay sensor transmits its own packet to the destination consuming δ_1^r units of energy while in relay mode the relay sensor's own traffic is discarded (if any) and δ_2^r units of energy is consumed to relay another sensor node's packet. For energy consumption, we assume that direct transmission mode requires more than relay mode, that is, $\delta_1 > \delta_2$, where we have dropped the superscript (s or r to indicate source and relay sensors, respectively) to indicate that the relation holds for both source and relay sensors. We assume that the sensors are deployed in real-time monitoring scenarios. Thus no retransmissions are attempted for packets with errors and no packets are buffered.

A discrete time model is assumed where time is slotted in intervals of unit length. Each slot is long enough so that a source node and a relay node can either cooperatively transmit one data packet for the source, or both can transmit one of their own packets. At most one data packet is generated at a node in a slot. Each sensor has a rechargeable battery and an energy harvesting device. The energy generation process at each sensor is modeled by a temporally correlated, two-state process with parameters (q_{on}, q_{off}), assuming $0.5 < q_{on}, q_{off} < 1$. In the *on* state (i.e. when ambient conditions are conducive to energy harvesting), the sensor generates energy at a constant rate of c units in a time slot. In the *off* state, no energy is generated. The data packets that the sensors report to a sink are also generated according to a temporally correlated, two-state process with parameters (p_{on}, p_{off}) with $0.5 < p_{on}, p_{off} < 1$, where in the *on* state an event (i.e. data packet) is generated in each slot, and no events are generated in the *off* state. Table I summarizes the parameters for both energy harvesting and event generating processes. Note that for analytical tractability we assume the event generation process is independent across different sensors, but the model can be extended to the scenario where the event process is correlated among nodes. The parameters corresponding to the source and relay nodes are denoted with a superscript of s and r respectively (e.g. p_{on}^s).

The communication strategy of a sensor pair {source, designated relay} is governed by a policy Π that decides on the transmission mode to be used for reporting events. Denote the set of actions as $\mathcal{A} = \{0, 1, 2, 3, 4\}$. The action taken by the sensor pair in time slot t is denoted by a_t with $a_t \in \mathcal{A}$ denoting {no transmission, no transmission}, {direct, no transmission}, {relay, relay}, {direct, own-traffic}, and {no transmission, own-traffic}. A transmission action can be taken only if the corresponding sensor has enough power and an event occurs at the beginning of the slot. A node is said to be *active* in a time slot if the action is taken such that it has a packet transmission in the time slot (either its own traffic or relaying), and *inactive* if there is no transmission.

The state information of a sensor includes: (a) its current energy level, (b) whether there is an event generated at the node in current slot, and (c) whether its battery is currently recharging or not. We assume that when a sensor transmits a packet, its current state information is included in the packet's header. However, if a relay is inactive in a slot, the source will not have the updated state information of the relay. We refer to such system as a *partially observable* system.

We assume that the communication strategy is decided at the source sensor. The decision may be based on: (a) the state information of the source; (b) the partial state information of the relay obtained at the source when the relay was last active. The objective of the decision policy Π is to maximize the *packet delivery ratio*, which is defined as the long term ratio of the total number of events reported, to the total number of events generated in the network.

III. OPTIMAL SCHEDULING POLICY USING POMDP FORMULATION

For the partially observable system, we first outline the formulation of the decision problem as a POMDP, and then focus on approximating a solution to the optimality equations and the optimal reward.

A. POMDP Formulation

Denote the system state at time t by $X_t = (L_t^s, E_t^s, Y_t^s, L_t^r, E_t^r, Y_t^r)$ where $L_t^s, L_t^r \in \{0, 1, 2, \cdots, K\}$ represents the energy available at the sensors at time t. $Y_t^s \in \{0, 1\}$ equals one if the source is being charged during time interval $[t, t+1)$ and zero otherwise. Also, $E_t^s \in \{0, 1\}$ equals one if an event to be reported during time interval $[t, t+1)$ is generated at time t at the source and zero otherwise. The variables $Y_t^r, E_t^r \in \{0, 1\}$ are defined similarly for the relay, but equal one if the recharge and event processes, respectively, are *on* during time interval $[t-1, t)$. The state of the relay at time t is defined in terms of the previous slot since that is the latest information the source may have about the relay. We assume that the battery at a sensor has a finite capacity K. Denote the state space as $\mathcal{X}$.

The *system observation* at time t at the source sensor is denoted by Y_t. The source is assumed to always have complete information about itself. Thus the observation Y_t is characterized by the action taken at time $t-1$,

$$Y_t = \begin{cases} X_t & \text{if } a_{t-1} \in \{2, 3, 4\} \\ (L_t^s, E_t^s, Y_t^s, \phi_L, \phi_E, \phi_Y) & \text{if } a_{t-1} \in \{0, 1\} \end{cases}$$

where ϕ_ω denotes that a variable ω is unknown. The observation space is denoted by $\mathcal{Y}$.

In the presence of only partial observations, the optimal action depends on the current and past observations, and on past actions. Existing work has shown that a POMDP may be formulated as a completely observable MDP with the same finite action set [7], [8]. The state space for the equivalent MDP comprises of probability distributions on the original state space.

It can be shown that the structure of the POMDP in this case leads to a countable state space for equivalent MDP, which guarantees the existence of an optimal solution to the average cost (reward) optimality equation [7].[2] Further, the state of the equivalent MDP at time t, Z_t, can be represented as

$$Z_t = (L_t^s, E_t^s, Y_t^s, L^r, E^r, Y^r, i), \tag{1}$$

representing the following: (a) the relay had no transmissions in the past i slots; (b) the state of the relay when it last transmitted was (L^r, E^r, Y^r); (c) the current state at the source is (L_t^s, E_t^s, Y_t^s).

The POMDP can be transformed to an equivalent MDP with state space Δ and the optimality equations for this MDP are given by [8]:

$$\Gamma^* + h^*(Z) = \max_{a \in \mathcal{A}} \left[\bar{R}(Z, a) + \sum_{y \in \mathcal{Y}} V(y, Z, a) h^*(W(y, Z, a)) \right], \quad \forall Z \in \Delta. \tag{2}$$

where $h^*(Z)$ is the optimal reward when starting at state Z, Γ^* is the optimal average reward, $\bar{R}(Z, a)$ is the reward associated with the states Z of the equivalent MDP, $V(y, Z, a)$ is interpreted as the probability of system observation $Y_{t+1} = y$ given the past actions and observations, and $W(y, Z, a)$ represents the probability of $X_{t+1} = i$ for all $i \in \mathcal{X}$, given the system observation Y_{t+1}, past system observations and actions. The optimal reward (in the optimality equation) of equivalent MDP is same as that in the original POMDP [7].

[2] For more detail of formulations and evaluations of MDP/POMDP for wireless sensor networks with cooperative relays and recharging capability, the reader is referred to [9]. The focus of this paper is towards analyzing the optimal reward and designing practical near-optimal and practical policies.

B. Optimal Performance Characterization

In this subsection, we characterize the structure of the optimal reward Γ^* as a function of system parameters by analyzing some simple cases. The structure reveals how the optimal reward depends on the system parameters and is used to design practical and efficient relay usage scheduling policies.

Consider a state $Z = (L^s, E^s, C^s, L^r, E^r, C^r, SI)$, $Z \in \Delta$, where Δ is the state space of the equivalent MDP. The $h^*(Z)$ can be expressed approximately as

$$h^*(Z) \cong \zeta_1 L^s + \zeta_2 L^r + \eta_1 C^s + \eta_2 C^r + \Psi_1 E^s + \Psi_2 E^r + \gamma SI. \tag{3}$$

Note that we are ignoring the cross terms here (e.g. $E^s E^r$), and the unknowns are given by ζ_1, ζ_2, η_1, η_2, Ψ_1, Ψ_2 and γ. We numerically solve (2) using value iteration, and the numerical solution matches closely with this approximate solution, thus justifying the linearity assumption in (3). Using some example scenarios, we find the relationship between the above mentioned unknowns and express them as functions of system parameters (such as p_{off}^s).

For example, in state $Z = (0, 0, 0, 0, 0, 0, 0)$ the only feasible action is 0, and it is therefore the optimal action. With the combination of different state transitions in the event and energy generation processes at both the source and the relay sensor, the set of possible next states is,

$$\begin{aligned} \{Z_1' &= (0, 0, 0, 0, 0, 0, 1), Z_2' = (c, 0, 1, 0, 0, 0, 1), \\ Z_3' &= (0, 1, 0, 0, 0, 0, 1), Z_4' = (c, 1, 1, 0, 0, 0, 1)\}. \end{aligned}$$

Substituting (3) into the optimality equation and simplifying, we get

$$\Gamma^* = \gamma + (1 - q_{off}^s)(\zeta_1 c^s + \eta_1) + (1 - p_{off}^s)\Psi_1. \tag{4}$$

Similarly, the only feasible and thus optimal action in state $Z = (0, 0, 1, 0, 0, 0, 0)$ and state $Z = (0, 1, 0, 0, 0, 0, 0)$ is 0, and we have

$$\Gamma^* + \eta_1 = \gamma + q_{on}^s(\zeta_1 c^s + \eta_1) + (1 - p_{off}^s)\Psi_1. \tag{5}$$

$$\Gamma^* + \Psi_1 = \gamma + (1 - q_{off}^s)(\zeta_1 c^s + \eta_1) + p_{on}^s \Psi_1. \tag{6}$$

Substituting (4) into (5) and (6) respectively, we get

$$\eta_1 = \frac{\zeta_1 c^s (q_{on}^s + q_{off}^s - 1)}{2 - q_{on}^s - q_{off}^s}, \quad \Psi_1 = 0. \tag{7}$$

Proceeding in the same way, we consider other states where the optimal actions are obvious and unique, and use the optimality equation (2) to develop the relationships between above unknowns and system parameters. Due to space constraints, we only list the results we have as following:

$$\zeta_1 = \frac{\theta^s}{\delta_1^s}, \quad \Psi_2 = 0, \quad \zeta_2 = \frac{\theta^r}{\delta_1^r} \tag{8}$$

$$\eta_2 = \frac{\zeta_2 c^r (q_{on}^r + q_{off}^r - 1)}{2 - q_{on}^r - q_{off}^r}, \quad \gamma = \frac{\theta^r c^r}{\delta_1^r} \mu_{on}^r \tag{9}$$

$$\Gamma^* = \frac{\theta^r c^r}{\delta_1^r}\mu_{on}^r + \frac{\theta^s c^s}{\delta_1^s}\mu_{on}^s. \quad (10)$$

Γ^* is the optimal average reward, and thus corresponds to the maximum achievable performance. Specifically in this model, Γ^* equals the number of events reported per unit time. Since in the long run, the number of events occurred per unit time equals $\pi_{on}^s + \pi_{on}^r$, denoting the optimal policy as Π_{opt} and optimal performance as $U(\Pi_{opt})$, we have,

$$U(\Pi_{opt}) = \frac{\Gamma^*}{\pi_{on}^s + \pi_{on}^r} = \frac{\frac{\theta^s c^s}{\delta_1^s}\mu_{on}^s + \frac{\theta^r c^r}{\delta_1^r}\mu_{on}^r}{\pi_{on}^s + \pi_{on}^r}. \quad (11)$$

The analytical solution presented above can be used to derive optimal action in different states. For every state, using the structure of $h^*(Z)$ function given in (3), we can evaluate the optimality equation (2) for all feasible actions in order to determine the optimal action in this state. For example, the optimal action for state $Z = (\delta_2^s, 1, 0, L, 0, 0, 0, 0)$ is action 2 while action 0 and 4 are optimal in state $Z = (L^s, 0, 0, L^r, 0, 0, 0, 0)$. The evaluation process is omitted here due to the space constraints. It is clear that the choice of optimal action (and hence the optimal policy) is very sensitive to system parameters. The same observation is obtained from the results of value iteration algorithm as well. Thus the optimal policy is not only difficult to characterize but also not suitable for practical deployment. Therefore, we next focus on outlining a *near-optimal and practical* scheme that could be used by a source node in practice and is not very sensitive to system parameters.

IV. A Practical (near-optimal) Relay Scheduling Scheme

In this Section, we design a practical relay scheduling policy and show (using analysis and simulations) that the proposed policy performs close to optimal in most scenarios. To simplify our discussion, we assume that the reward for a successful packet transmission at the source and the relay are the same, and the energy consumption parameters are identical at the source and the relay sensors.

If the source has no event at the beginning of a slot (i.e. $E^s = 0$), then the only feasible actions are {0,4}. We know from Section III-B that action 4 could be optimal. Since the relay's status is not completely observable at the source, even if action 4 is chosen at the source but not taken at the relay (insufficient energy level or medium is busy), the system does not have much to lose by choosing action 4. Thus we choose action 4 instead of action 0. The same action happens when the source does not have enough energy to transmit a packet even with the help of relay.

When the source has data to report, and $\delta_2^s \le L^s < \delta_1^s$, the feasible actions are {0,2,4} and action 2 is a good choice from the Section III-B. Note that with the equal energy consumption and equal reward assumption, and if $\delta_1 > 2\delta_2$, the relay usage does not cost more energy (in total) than direct transmission. The optimal action could be different from action 2 if the reward of transmitting a source packet is much less than that of transmitting a relay packet or if the total energy consumption of relay usage for a packet is much higher than that of a direct transmission.

Algorithm 1 A Practical Relay Usage Scheduling Algorithm

```
if E^s = 0 OR L^s < δ_2^s then
    â = 4;
else if E^s = 1 AND L^s ≤ T_H then
    â = 2;
else
    â = 3;
end if
```

When the source has data to report, and $L^s \ge \delta_1^s$, the feasible actions are {0,1,2,3,4}. When both action 1 and 3 are feasible, it is preferable to choose action 3 even though the relay may not transmit (as discussed earlier in the section). From the results of value iteration, we see that when the source has enough energy for transmission, action 4 is not optimal. Thus, either action 2 or 3 is the best choice and the question of interest is *whether to use relay or not?* We propose a practical decision choice here. Consider a threshold energy level T_H, when the source's energy level is less then T_H, relay usage is scheduled; otherwise source and relay are scheduled to transmit their own data. The algorithm is summarized in Algorithm 1.

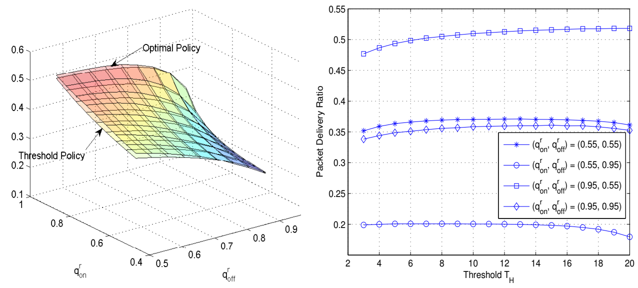
For any state $Z = (L^s, E^s, C^s, L^r, E^r, C^r, SI)$, let $\hat{a}$ denote the chosen action and let T_H denote the energy level threshold such that $\delta_1^s \le T_H \le K$.

Note that this simple threshold based scheme is not only easier to deploy, but its performance is also close to that of the optimal policy, as shown in Section IV-A. In addition, unlike the optimal policy which is very sensitive to various system parameters, the threshold scheme has only one tunable parameter T_H. The simplicity in fact provides better analytical tractability.

A. Simulation Results

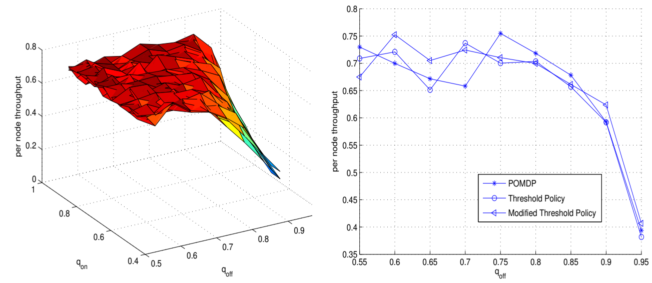
We compare the performance of the optimal policy obtained using value iteration algorithm with that of the proposed threshold based scheme, and discuss the choice of a good threshold. We consider a network where a three-node-group model (i.e. source, relay and destination) is applied and each group is independent of others. The simulations are based on a single group and the results can be generalized. The simulations were done using a simulator developed by us, primarily because energy harvesting is not well supported in existing simulators. All simulations were run for a duration of 5×10^6 time units.

Figure 1 (a) compares the *packet delivery ratio* between the optimal policy and the threshold policy. The threshold used is half of the battery capacity, that is, $T_H = K/2$. In the simulation scenario, the event occurrence parameters are fixed for both the source and the relay. The steady-state probability of event occurrence at the source $\pi_{on}^s = 3/4$ and at the relay $\pi_{on}^r = 2/3$. The recharge process parameters (q_{on}^s, q_{off}^s) of the source are fixed while those of the relay (q_{on}^r, q_{off}^r)



(a) Packet delivery ratio (b) Effect of the T_H

Fig. 1. (a) The packet delivery ratio comparison between the optimal policy and the threshold based policy. (b) Effect of T_H on the packet delivery ratio. Parameters used: $p^s_{on} = 0.9$, $p^s_{off} = 0.7$, $q^s_{on} = p^r_{on} = 0.85$, $q^s_{off} = p^r_{off} = 0.7$, $c^s = c^r = 1$, $\delta^s_1 = \delta^r_1 = 3$, $\delta^s_2 = \delta^r_2 = 1$, $K = 20$, $T_H = 10$.



(a) all q_{on} and q_{off} combinations (b) fixed $q_{on} = 0.8$

Fig. 2. The average per-node throughput comparison between the POMDP policy, the threshold-based policy and the modified threshold-based policy in a multi-node network with 50 nodes. Parameters used: $q^s_{on} = q^r_{on} = q_{on}$, $q^s_{off} = q^r_{off} = q_{off}$, $p^s_{on} = p^r_{on} = 0.6$, $p^s_{off} = p^r_{off} = 0.9$, $\delta^s_1 = \delta^r_1 = 3$, $\delta^s_2 = \delta^r_2 = 1$, $c^s = c^r = 1$, $K = 20$, $T_H = 10$.

are changing. The result shows that the performance of the threshold policy is very close to that of the optimal policy.

Figure 1 (b) depicts the performance for various choices of the threshold. The performance trend observed while varying the threshold depends upon the choice of system parameters. Particularly, if relay has surplus energy a larger threshold performs better (since relay usage is more often scheduled) and vice-versa. However, $T_H = K/2$ seems to be a good choice under various parameter settings.

B. Multi-node Networks and Simulation Results

The practical scheduler can be applied to a more realistic multi-node network. Effectively, each node tries to maximize its own quality of coverage in a distributed way. The results show that the threshold scheduling scheme performs very well in multi-node network scenarios as well.

In the network, a node may act as a source node for its packets, a destination for receiving other node's packets, or a relay as well for its neighbor's packet. Considering the dynamic and complex characteristics such networks, we modify the practical relay scheduling scheme proposed in Section IV as: **line 2:** $\Rightarrow \hat{\mathbf{a}} = \mathbf{0}$**; line 6:** $\Rightarrow \hat{\mathbf{a}} = \mathbf{1}$. In other words, a node only decides its own transmission scheme. This is more suitable in multi-node networks where data exchange happens distributively and possibly simultaneously.

We consider a network with 50 nodes, spread randomly over a 1000×1000 square meter region. The transmission range of each sensor is 100 meters. We consider one-hop traffic and the source picks the relay at random from available candidates. Random backoff based MAC protocol with request-to-send (RTS)/clear-to-send (CTS) is used by each node.

Figure 2 shows the per-node throughput averaged over all nodes in the network. The throughput is defined as the ratio of the average number of packets successfully transmitted by a node to the total number of packet generated by the node. Parameters are chosen such that traffic rate is relatively low compared to energy harvesting rate, which makes the usage of energy harvesting applicable. The figure compares the performance of the optimal policy obtained by solving the POMDP with value iteration for three-node network model (referred as POMDP policy), the threshold policy and the modified threshold policy. It shows that the performances are very close to each other and the modified threshold policy in general works slightly better than the threshold policy in a multi-node network. Figure 2 (b) shows that POMDP is slightly worse than threshold polices at some parameter settings. One of the possibility is that our models do not consider the scenario where a source node may in turn act as the relay for another node. However it has been shown that POMDP still achieves better performance as compared to direct transmissions.

V. Conclusions

This paper addresses the problem of developing transmission and relay usage strategies for WSNs in the presence of energy harvesting sensor nodes. The problem is formulated as a POMDP. An approximate solution to the optimality equations and the optimal reward is outlined. Due to the sensitivity of the optimal scheduler to system parameters, a practical threshold based relay usage scheduler is proposed, analyzed and evaluated.

References

[1] J. Laneman, D. Tse and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory,* vol. 50, no. 12, pp. 3062-3080, December 2004.

[2] Y. Yao, X. Cai and G. Giannakis, "On Energy Efficiency and Optimum Resource Allocation of Relay Transmissions in the Low-Power Regime," *IEEE Trans. Wireless Comm.,* vol. 4, no. 6, November 2005.

[3] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.,* vol. 3, pp. 120-154, 1971.

[4] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory,* vol. 25, no. 5, pp. 572-584, September 1979.

[5] L. Guo, X. Ding, H. Wang, Q. Li, S. Chen and X. Zhang , "Cooperative Relay Service in a Wireless LAN," *IEEE Jour. Sel. Areas Comm.*, vol. 25, no. 2, pp. 355-368, February 2007.

[6] J. Paradiso and T. Starner, "Energy scavenging for mobile and wireless electronics," *IEEE Perv. Comp.,* vol. 4, no. 1, pp. 18-27, Jan-Mar 2005.

[7] A. Cassandra, L. Kaelbling and M. Littman, "Acting optimally in partially observable stochastic domains," *Proc. National Conference on Artificial Intelligence,* vol. 2. pp. 1023-1028, Seattle, WA, 1994.

[8] E. Fernandez-Gaucherand, A. Arapostathis and S. Marcus, "On the average cost optimality equation and the structure of optimal policies for partially observable Markov decision processes," *Annals of Operations Research,* vol. 29, no. 1-4, pp. 439-470, April 1991.

[9] H. Li, N. Jaggi and B. Sikdar, "Relay Scheduling for Cooperative Communications in Sensor Networks with Energy Harvesting," *IEEE Trans. Wireless Comm.,* vol. 10, no. 9, pp. 2918-2928, September 2010.