TRANSMISSION SCHEDULING FOR SENSOR NETWORK LIFETIME MAXIMIZATION: A SHORTEST PATH BANDIT FORMULATION

Yunxia Chen, Qing Zhao

University of California
Davis, CA 95616, USA
{yxchen,qzhao}@ece.ucdavis.edu

Vikram Krishnamurthy, Dejan Djonin

University of British Columbia
Vancouver, BC V6T 1Z4, Canada
{vikramk,ddjonin}@ece.ubc.ca

ABSTRACT

This paper addresses optimal sensor scheduling for maximizing network lifetime. We formulate this problem as a stochastic shortest-path multi-armed bandit problem. The optimal transmission scheduling policy is thus to choose the sensor with the largest Gittins index. Exploiting the underlying structure of the sensor scheduling problem, we derive a closed-form expression for the Gittins index. We show that choosing the sensor with the most residual energy is an optimal strategy when the channel fading is independently and identically distributed across sensors.

1. INTRODUCTION

A wireless sensor network (WSN) consists of a large number of low-cost, low-power, energy-constrained sensors. Sensors monitor a certain phenomenon and transmit their measurements to the access point (AP). One of the key problems arising in WSNs is sensor scheduling: which sensors should be scheduled for transmission to achieve an optimal performance specified by the underlying network applications (see [1–3] and references therein). In [4], a decentralized suboptimal sensor scheduling protocol is proposed for lifetime maximization. Exploiting the local channel state information (CSI) of each sensor, this protocol requires every sensor to estimate its channel realization in each data collection. Recently, an optimal transmission scheduling is obtained via stochastic shortest path (SSP) formulation in [5]. This optimal protocol requires centralized implementation in which sensors have to send pilot signals to the AP for global CSI acquisition. The basic idea of these two protocols is to exploit channel diversity among sensors. In practice, however, the transmission power of sensors may be limited to a small range due to hardware implications. This may exclude the possibility of fully exploiting channel diversity. As a consequence, the extra energy consumed in channel acquisition and centralized scheduling may override the benefit of using CSI.

We are thus motivated to exploit the channel statistics rather than realizations in sensor scheduling[1]. We formulate the resulting problem as an SSP multi-armed bandit problem which is special class of Markov decision process with a non-discounted reward and a finite but random stopping time. The optimal transmission scheduling for network lifetime maximization is thus given by an indexable policy that chooses the sensor with the largest Gittins index. Although a closed-form expression for the Gittins index does not exist for a general multi-armed bandit problem, we derive a closed-form expression for the Gittins index by exploiting the rich structure of the transmission scheduling problem. We also show that choosing the sensor with the most residual energy is optimal when the channel is independently and identically distributed (i.i.d.) across sensors. In this case, the optimal scheduling policy does not even require the knowledge of channel statistics.

2. NETWORK MODEL AND LIFETIME DEFINITION

We consider a WSN with N sensors, each powered by a non-rechargeable battery with initial energy E_0. In each data collection, one of these N sensors is selected to transmit its measurement encoded in a packet with fixed size to the AP through a wireless fading channel. The channels between the AP and sensors follow the block fading model with the block length equal to the transmission time of one packet. We also assume that the channel fading is i.i.d. across data collections and independently distributed across sensors.

In each data collection, the minimum energy $E(c_n)$ required for sensor n to transmit a packet successfully to the AP is a random variable depending on its current channel state c_n:

$$E(c_n) = E_c + \frac{\overline{E}}{c_n} \qquad (1)$$

where E_c is the energy consumed in transmitter circuitry and

This work is prepared through collaborative participation in the Communications and Networks Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011.

[1]Note that although channel realizations are not used in scheduling, local CSI of the scheduled sensor is used to adapt its transmission power.

$\overline{E}$ is the energy required to achieve an acceptable received SNR at the AP when the channel gain is one. In practice, sensors can only transmit at a finite number L of power levels due to their hardware limitations. To conserve energy, every sensor chooses, according to its channel gain, the lowest power level that meets the SNR threshold for successful reception at the AP. Hence, the energy w_n consumed by sensor n in transmitting its packet is a random variable depending on its current channel state c_n:

$$w_n = \begin{cases} \min\limits_k \{\varepsilon_k : \varepsilon_k \geq E(c_n)\}, & \text{if } c_n \geq \frac{\overline{E}}{\varepsilon_L - E_c}, \\ \infty, & \text{otherwise}, \end{cases} \tag{2}$$

where ε_k is the energy consumed in transmitting a packet at power level k and $0 < \varepsilon_1 < \ldots < \varepsilon_L < \infty$. Note that if the channel experiences deep fading ($c_n < \frac{\overline{E}}{\varepsilon_L - E_c}$), sensor n will fail to transmit in this data collection since its energy requirement exceeds the maximum sensor energy consumption: $E(c_n) > \varepsilon_L$. For simplicity, we write $w_n = \infty$ in this case.

The distribution of the energy consumption requirement w_n of sensor n is thus determined by its channel distribution:

$$\begin{aligned} &\Pr\{w_n = \varepsilon_1\} = \Pr\left\{c_n \geq \frac{\overline{E}}{\varepsilon_1 - E_c}\right\}, \\ &\Pr\{w_n = \varepsilon_k\} \\ &\qquad = \Pr\left\{\frac{\overline{E}}{\varepsilon_k - E_c} \leq c_n < \frac{\overline{E}}{\varepsilon_{k-1} - E_c}\right\}, \quad 2 \leq k \leq L, \\ &\Pr\{w_n = \infty\} = \Pr\left\{c_n < \frac{\overline{E}}{\varepsilon_L - E_c}\right\}. \end{aligned} \tag{3}$$

Since the sensor energy consumption is restricted to the set $\{\varepsilon_k\}_{k=1}^L$, the residual energy e_n of sensor n at the beginning of a data collection belongs to the set $\mathcal{E}$:

$$\mathcal{E} = \left\{e : e = E_0 - \sum_{k=1}^{N} \alpha_k \varepsilon_k \geq 0 \text{ for } \alpha_k \geq 0 \text{ and } \alpha_k \in \mathbb{Z}\right\}.$$

We define the network lifetime $\mathcal{L}$ as the number of data collections until any sensor dies or the first transmission failure[2]. In other words, the network is considered dead if any sensor's residual energy drops below the minimum energy consumption (*i.e.,* $e_n < \varepsilon_1$ for any n) or the scheduled sensor fails to transmit, whichever occurs first.

In this paper, we seek the answer to the following question: given the residual energy profile $\mathbf{e} = (e_1, \ldots, e_N)$ at the beginning of each data collection and the channel distributions, which sensor should be scheduled for transmission so that the network lifetime is maximized.

3. SSP MULTI-ARMED BANDIT FORMULATION

An SSP multi-armed bandit process consists of N independent Markov systems, each of which has an inevitable terminating state. Based on the current states of all systems, we select one system to work on and receive a non-discounted reward. Only the state of the chosen system evolves according to its transition probabilities; the states of other systems remain fixed. The goal of an SSP multi-armed bandit is to maximize the total expected reward until any system reaches its terminating state.

To formulate the optimal transmission scheduling as an SSP multi-armed bandit process, we model the sensor network by N independent Markov systems, each of which represents the state evolution of a sensor in the network. In each data collection, we characterize the state of a sensor by its residual energy. Since the current channel realizations are unknown, the scheduled sensor may not have enough energy for transmission. In this case, a failure in data collection occurs and the network lifetime terminates. We introduce state t to represent this situation and define the state space $\mathcal{S}$ of each Markov system as

$$\mathcal{S} \triangleq \{\text{state } i = e : e \in \mathcal{E}\} \cup \{\text{state } t\}. \tag{4}$$

The terminating state space $\mathcal{S}_t \subset \mathcal{S}$ is defined as

$$\mathcal{S}_t \triangleq \{e : e < \varepsilon_1\} \cup \{\text{state } t\}, \tag{5}$$

where states in the set $\{e : e < \varepsilon_1\}$ indicate the death of the sensor and state t indicates a failure in data collection. By the lifetime definition, the network is considered dead if any sensor reaches a terminating state in $\mathcal{S}_t$.

Next, we define the action space, the reward, and the transition probabilities for the Markov system representing sensor n. The action space consists of two elements: $\mathcal{A} \triangleq \{0, 1\}$. Action $a = 0$ indicates that sensor n is not chosen; the state of sensor n does not change. Action $a = 1$ indicates that sensor n is scheduled for transmission; the state of sensor n transits from i to j with probability $p_{ij}^{(n)}$ and a reward $r_{ij}^{(n)}$ is earned. The transition probability $p_{ij}^{(n)}$ depends on the channel distribution of sensor n, which can be obtained as

$$p_{ij}^{(n)} = \begin{cases} \Pr\{w_n = \varepsilon_k\} \mathbb{I}_{[e'_n = e_n - \varepsilon_k]}, & j = e'_n, \\ \Pr\{w_n > e_n\}, & j = t, \end{cases} \tag{6}$$

where state $i = e_n$ and $\mathbb{I}_{[x]}$ is the indicator function: $\mathbb{I}_{[x]} = 1$ if x is true and $\mathbb{I}_{[x]} = 0$ otherwise. The reward $r_{ij}^{(n)}$ obtained when the state of sensor n transits from i to j is defined as

$$r_{ij}^{(n)} = \mathbb{I}_{[j \neq t]}. \tag{7}$$

Hence, the reward indicates whether the transmission of the chosen sensor is successful. Since only one sensor is scheduled in each data collection, the total reward obtained until the network dies (*i.e.,* any sensor reaches one of its terminating states) represents the total number of successful data collections, which is the network lifetime. Our goal is to find the optimal policy that specifies which sensor to choose in each data collection for network lifetime maximization.

[2] Similar analysis can be carried out for a more general lifetime definition where the network is considered dead when any sensor dies or the number of transmission failures exceeds a threshold.

4. OPTIMAL TRANSMISSION SCHEDULING

In the last section, we have formulated the optimal transmission scheduling as an SSP multi-armed bandit process. It is shown in [6] that the optimal policy is given by an indexable strategy that chooses the sensor whose current state has the largest Gittins index in each data collection. In this section, we address the implementation of this optimal policy and derive a closed-form expression for the Gittins index. We also calculate the optimal expected network lifetime.

4.1. Gittins index

For completeness, we first give a brief description of the Gittins index. We consider the single sensor case and modify the Markov system associated with sensor n in the following way [6]: under action $a = 0$, sensor n transits to terminating state t with probability 1 and a terminating reward g is received. Our goal here is to maximize the total expected reward before sensor n reaches any terminating state in (5).

Applying the Bellman's equation, the optimal expected reward $V(i)$ of the modified system starting from state i is given by

$$V(i) = \max\left\{g, R_n(i) + \sum_{j\in\mathcal{S}} p_{ij}^{(n)} V(j)\right\}, \qquad (8)$$

where $R_n(i)$ is the expected reward in state i associated with sensor n under action $a = 1$:

$$R_n(i) = \sum_{j\in\mathcal{S}} r_{ij}^{(n)} p_{ij}^{(n)} = \Pr\{w_n \leq e_n\}. \qquad (9)$$

Note that the optimal reward $V(i) = 0$ for all $i \in \mathcal{S}_t$. The Gittins index $\gamma_n(i)$ of state i is defined as the smallest value of terminating reward g at which action $a = 0$ is optimal:

$$\gamma_n(i) = \min_g \{g : V(i) = g\}. \qquad (10)$$

4.2. Optimal Policy

Since the Gittins index associated with each sensor only depends on its own state and is independent of other sensors' states, the optimal transmission scheduling policy can be implemented in a distributed fashion via opportunistic carrier sensing [4]. Specifically, in each data collection, every sensor chooses a backoff time inversely proportional to its Gittins index and transmits with its chosen backoff delay if the channel is available before its backoff time expires. Since a sensor with larger Gittins index has a shorter backoff time, this opportunistic carrier sensing scheme ensures that the sensor with the largest Gittins index seizes the channel and transmits if the channel propagation delay is negligible.

The Gittins index is the key to the optimal transmission scheduling. For a general N-armed bandit problem, a closed-form expression for the Gittins index does not exist, and the complexity of the most efficient algorithm to calculate the Gittins index is cubic in the number of non-terminating states [7]: $\mathcal{O}(M^3)$ where $M = |\mathcal{S}\backslash\mathcal{S}_t|$. Exploiting the rich structure of sensor scheduling problem, we derive a closed-form expression for the Gittins index and hence reduce its computational complexity to linear $\mathcal{O}(M)$.

Theorem 1: For sensor scheduling problem, the Gittins index $\gamma_n(i)$ of state i associated with sensor n is given by

$$\gamma_n(i) = \gamma_n(e_n) = \frac{\Pr\{w_n \leq e_n\}}{\Pr\{w_n > e_n - \varepsilon_1\}}, \qquad (11)$$

where the distribution of w_n is determined by the channel distribution in (3).

Proof: See Appendix A for details □□□

Corollary 1: When channel fading is i.i.d. across sensors, choosing the sensor with the most residual energy is optimal in maximizing network lifetime.

Proof: Since the channel distribution is identical across sensors and the Gittins index $\gamma_n(e_n)$ given in (11) increases with the residual energy e_n, Corollary 1 follows. □□□

Corollary 1 shows that when the channel fading is i.i.d. across sensors, the optimal transmission scheduling can be implemented even when the channel distribution is unknown.

4.3. Optimal Lifetime

Applying the Bellman's equation, we obtain the optimal lifetime $V^*(\mathbf{e})$ starting from the network residual energy profile $\mathbf{e} = (e_1, \ldots, e_N)$ as

$$V^*(\mathbf{e}) = \max_n \left\{R_n(e_n) + \sum_{e'_n\in\mathcal{S}} p_{e_n e'_n}^{(n)} V^*(\mathbf{e}')\right\}, \qquad (12)$$

where $\mathbf{e}' = (e_1, \ldots, e_{n-1}, e'_n, e_{n+1}, \ldots, e_N)$ denotes the energy profile to which $\mathbf{e}$ transits when sensor n is chosen and $R_n(e_n)$ is the expected reward in state e_n given in (9). Calculating (12) in an increasing order of the total energy $\sum_{n=1}^N e_n$ in the network, we can readily obtain the optimal lifetime in one iteration. Furthermore, applying the Gittins strategy, we can also simplify (12) as

$$V^*(\mathbf{e}) = \sum_{k=1}^{L} \Pr\{w_a = \varepsilon_k\} \mathbb{I}_{[e_a - \varepsilon_k \geq 0]} \left[1 + V^*(\mathbf{e}'_k)\right] \qquad (13)$$

where $\mathbf{e}'_k = (e_1, \ldots, e_{a-1}, e_a - \varepsilon_k, e_{a+1}, \ldots, e_N)$ and $a = \max_n\{\gamma_n(e_n)\}$ is the index of the sensor whose current state e_n has the largest Gittins index.

5. NUMERICAL EXAMPLE AND CONCLUSION

Fig. 1 compares the expected network lifetime of the optimal transmission scheduling policy, the scheme that chooses the sensor with the most residual energy, and the scheme that

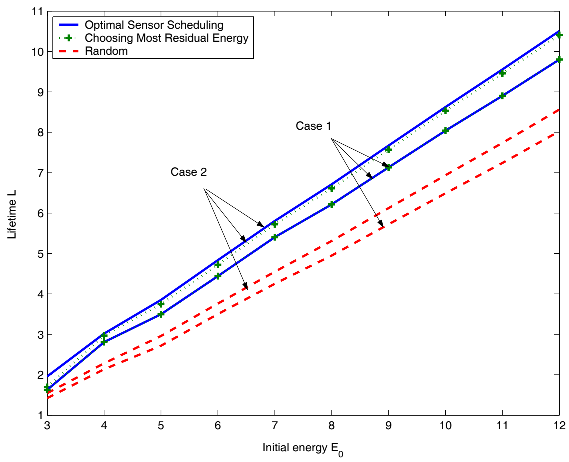

Fig. 1. Lifetime comparison. $N = 2$, $L = 3$. In case 1, channels are i.i.d.: $\Pr\{w_n = 1\} = \Pr\{w_n = 2\} = \frac{1}{4}, \Pr\{w_n = 3\} = \frac{1}{2}$ for $n = 1, 2$. In case 2, channels are non-identically but independently distributed: $\Pr\{w_1 = 1\} = \Pr\{w_1 = 2\} = \frac{1}{4}, \Pr\{w_1 = 3\} = \frac{1}{2}, \Pr\{w_2 = 1\} = \Pr\{w_2 = 3\} = \frac{1}{4}, \Pr\{w_2 = 2\} = \frac{1}{2}$.

randomly schedules a sensor. Exploiting channel distribution, the optimal policy outperforms the random protocol without increasing the implementation overhead since the Gittins index can be pre-calculated according to the channel distribution. The performance gain of the optimal policy over the random protocol increases with the initial energy E_0. When the channel is i.i.d. across sensors (case 1), choosing the sensor with the most residual energy has the optimal lifetime performance, which confirms Corollary 1. Clearly, when the channel is not identically distributed (case 2), this strategy is no longer optimal, but the performance degradation is small.

6. CONCLUSION

In this paper, we formulated the optimal sensor scheduling as an SSP multi-armed bandit problem. The rich structure of the scheduling problem enables us to derive a closed-form expression for the Gittins index. We also showed that choosing the sensor with the most residual energy is optimal when channel is i.i.d. across sensors[3].

7. APPENDIX: PROOF OF THEOREM 1

To derive the closed-form Gittins index (11), we consider the modified Markov system associated with sensor n that has been introduced in Section 4.1. Let $\mathcal{U}$ be the set of states whose Gittins index has been calculated. Initiate $\mathcal{U} = \phi$.

Step 1: *Calculate the largest Gittins index in the set $\mathcal{S}\backslash\mathcal{U}$.* Suppose state $i^* = e_n^*$ has the largest Gittins index and the modified Markov system has a terminating reward $g = \gamma_n(i^*)$. It has been shown in [6] that the optimal policy for the modified Markov system is to choose action $a = 0$ whenever the Gittins index of the current state is smaller than the terminating reward g. Hence, the optimal expected reward starting from any non-terminating state $j \in \mathcal{S}\backslash\mathcal{S}_t\backslash\mathcal{U}$ is $V(j) = g = \gamma_n(i^*)$. From the definition of Gittins index, we obtain the optimal total expected reward starting from state i^* as

$$\begin{aligned} V(i^*) =& \gamma_n(i^*) = R_n(i^*) + \sum_{j \in \mathcal{S}\backslash\mathcal{U}} p_{i^*j}^{(n)} V(j) \\ =& R_n(i^*) + F_n(i^*)\gamma_n(i^*) \end{aligned} \tag{14}$$

where $F_n(i^*) = \Pr\{e_n^* - w_n \geq \varepsilon_1\} = \Pr\{w_n \leq e_n^* - \varepsilon_1\}$ is the probability that state i^* does not transit to a terminating state in one step. Hence, the largest Gittins index in $\mathcal{S}\backslash\mathcal{U}$ is given by

$$\gamma_n(i^*) = \frac{R_n(i^*)}{1 - F_n(i^*)} = \frac{\Pr\{w_n \leq e_n^*\}}{\Pr\{w_n > e_n^* - \varepsilon_1\}}. \tag{15}$$

Step 2: *Determine the state i^* that has the largest Gittins index and remove it from the Markov system.* From (15), it can be readily shown that state i^* maximizes $\frac{R_n(i)}{1-F_n(i)}$ over all $i \in \mathcal{S}\backslash\mathcal{U}$, *i.e.,*

$$i^* \in \arg\max_{i \in \mathcal{S}\backslash\mathcal{U}} \frac{R_n(i)}{1 - F_n(i)}. \tag{16}$$

If there are more than one states that achieve the largest Gittins index (15), we will choose the one with the most residual energy. Since $\Pr\{w_n \leq x\}$ is a non-decreasing function in x, state i^* has the most residual energy in $\mathcal{S}\backslash\mathcal{U}$. Hence, state i^* is not reachable from any remaining state $i \in \mathcal{S}\backslash\mathcal{U}$, *i.e.,* $p_{ii^*}^{(n)} = 0$. We can thus remove state i^* from the system without changing the remaining states.

Step 3: *Let $\mathcal{U} = \mathcal{U} \cup \{i^*\}$.* Goto Step 1 until $\mathcal{U} = \mathcal{S}\backslash\mathcal{S}_t$.

Following the above procedure, we find that the Gittins index can be computed using (15), which is the same as (11).

8. REFERENCES

[1] A. Logothetis and A. Isaksson, "On sensor scheduling via information theoretic criteria," in *Proc. of the American Control Conferences*, (San Diego, CA), pp. 2402 – 2406, June 1999.

[2] Y. He and E. K. P. Chong, "Sensor scheduling for target tracking in sensor networks," in *Proc. of 43rd IEEE Conference on Decision and Control*, pp. 743 – 748, Dec. 2004.

[3] V. Krishnamurthy, "Algorithms for optimal scheduling and management of Hidden Markov model sensors," *IEEE Trans. Signal Processing*, vol. 50, pp. 1382 – 1397, June 2002.

[4] Y. Chen and Q. Zhao, "Distributed transmission protocol for lifetime maximization in sensor networks," in *Proc. of IEEE 6th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 895 – 899, June 2005.

[5] Y. Chen, Q. Zhao, V. Krishnamurthy, and D. Djonin, "Transmission Scheduling for Optimizing Sensor Network Lifetime: A Stochastic Shortest Path Approach" *Submitted to IEEE Transactions on Signal Processing*, Jan. 2006.

[6] I. Dumitriu, P. Tetali, and P. Winkler, "On playing golf with two balls," *SIAM J. Discrete Math.*, vol. 18, pp. 604 – 615, July 2003.

[7] A. -K. Katta and J. Sethuraman, "A note on bandits with a twist," *SIAM J. Discrete Math.*, vol. 18, pp. 110 – 113, July 2004.

[3] The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U. S. Government.