
Xiaolu Zhang, Meixia Tao and Chun Sum Ng
Department of Electrical and Computer Engineering
National University of Singapore
Singapore 117576
Email: {zhangxiaolu, mxtao, elengcs}@nus.edu.sg.

Abstract—We study transmission schemes for lifetime maximization in wireless sensor networks. Specifically we consider the network where all the sensors observe uncorrelated signals and forward the quantized observations to a common fusion center. We demonstrate how the minimum node lifetime is maximized under individual estimation accuracy constraints through adapting transmission times and powers assigned to different sensor nodes. A partially distributed algorithm is proposed to implement joint time sharing and power control, for which each sensor node only needs local information and little common information. In addition, we obtain a necessary and sufficient condition for convergence of the algorithm. Simulation results show a significant lifetime gain over existing schemes, especially when the sensing environment becomes more heterogeneous and the number of nodes increases.

I. INTRODUCTION

Energy limitation is one of the greatest distinctions between a wireless sensor network (WSN) and other wireless data networks such as cellular networks and wireless local area networks (WLAN). A natural problem in this area is the maximization of network lifetime, i.e., the time that it takes for the first node to die. The power in WSN is consumed for transmitting and receiving, data processing, sensing, and switching between active mode and sleep mode. For simplicity, we consider only the power required for transmission which is dominant in total power consumption as in [1] and [2].

Adaptive resource allocation design is a potent technique for improving wireless data network performance under application-specific constraints [3]. The concept is to adapt the transmission parameters, such as transmission power, transmission time and channel bandwidth, to the underlying channel, interference, and system preferences. The work [4] shows that adaptation can obtain up to 20 dB power saving over nonadaptive transmission for a single communication link. In multi-user systems, significant system performance gain can also be achieved by optimal transmission scheduling and power control. For instance, in a multi-access channel, the authors in [5] proposed a water-filling based power control algorithm to maximize the sum-of-rate capacity subject to the average transmission power constraint of each user.

The benefit to wireless data networks from adaptive resource allocation motivates us to apply the adaptation in wireless sensor networks. The optimal power control problem for the decentralized estimation of a noise-corrupted deterministic signal in an inhomogeneous sensor network is studied in [2]. The proposed power control scheme suggests that the sensors with bad channel should decrease quantization resolutions for total energy saving. Yao and Giannakis in [6] propose to minimize the total energy consumption by varying the transmission times assigned to different sensor nodes under the individual rate constraint for each node. However, minimizing the total power consumption does not guarantee maximum lifetime for a network since the objective, total power minimization, always results in some of nodes with non-rechargeable batteries running out of energy quickly [7].

We consider the lifetime maximization for a cluster-based WSN. Sensor nodes are organized into clusters. Each cluster consists of multiple sensor nodes and one common fusion center (FC), and is responsible for monitoring a certain geographical area. The focus of this paper is on single-hop data collection from sensor nodes to the FC in a particular cluster. The responsibility of sensor nodes is to sense the surrounding environment, quantize the observation and report to the FC. We assume that FC is less energy-constrained than sensor nodes.

A joint time sharing and power control (JTPC) policy is proposed in a time-framed WSN. In a predefined frame duration, each sensor node should transmit the observed information using sufficient quantization resolutions so as to guarantee a certain estimation accuracy. At any time instant only one node can access the common channel, but each time frame is shared by all the nodes in an adaptive time-division fashion. We are interested in maximizing the network lifetime, the time taken for the first node to run out of energy, by varying the transmission times and transmission powers. Besides channel quality as considered in most of the literature, we also take into account the fact that different nodes may have different initial energy supplies and observation qualities. Since the fusion center only has limited computational capacity and the network cannot afford large control signalling overhead, a partially distributed algorithm is introduced where each node only needs to know the local information and little common
information. The convergence behavior of the algorithm is also explored. Numerical results illustrate significant advantages of JTPC over existing energy-efficient schemes, especially when the sensing environment becomes more heterogeneous.

The rest of the paper is organized as follows. Section II describes the system model. In Section III, we formulate the problem of optimal JTPC and derive the sufficient-and-necessary condition of network lifetime maximization. A partially distributed algorithm is developed and its convergence is proven in Section IV. Section V presents and evaluates the simulation results and, finally, the paper concludes in Section VI.

II. SYSTEM MODEL

Consider a transmission scheduling problem in a cluster of WSN with a FC and a set of $K$ sensor nodes deployed sparsely. The data fusion procedure is illustrated in Fig. 1. At sensor node $k$, the observation corrupted by additive noise, $\omega_k$, is denoted by

$$x_k = \theta_k + \omega_k, \quad k \in K,$$

(1)

where $K$ denotes the set of nodes, and $\theta_k$ denotes the signal to be observed and is uncorrelated for different $k$’s. The noise $\omega_k$, $k \in K$, is zero-mean independent random variable with variance $\sigma^2_k$. Due to bandwidth limitation between the sensor nodes and FC, each sensor needs to perform local quantization using $L_k$ bits of their observations to obtain a discrete message $m_k(x_k, L_k)$. During a time frame, the quantized message of each node is transmitted consecutively in an adaptive time sharing fashion through a fading channel with additive white Gaussian noise (AWGN). Without loss of generality, the time-frame duration is assumed to be normalized. The channel gain between FC and sensor nodes is modelled as $g_k$. We assume that $g_k$ changes very slowly and remains approximately constant during the network lifetime. Let $p = [p_1, p_2, \ldots, p_K]$ denote the time-sharing policy, where $p_k$ represents the fraction of the frame duration allocated to node $k$. It is assumed that the transmission is error-free and the information rate is given by the mutual information. Thus we have:

$$L_k = \left[p_k \log_2 \left( 1 + \frac{g_k p_k}{N_0} \right) \right], \quad \forall k \in K,$$

(2)

where $\lfloor x \rfloor$ represents the largest integer that is not greater than $x$, $N_0$ is the noise power, and $p = [p_1, p_2, \ldots, p_K]$ is the transmission power vector. Hence the energy consumed by node $k$ in a time frame is $\rho_k T_k p_k$. Without loss of generality, the noise power is assumed to be one throughout this paper.

Suppose that the signal range that each sensor can observe is $[-W, W]$, and we quantize the observation $x$ by uniformly dividing this range into $2^L$ levels. Lemma 1 shows the relationship between the quantization scheme and the estimation accuracy (as denoted by (b) in Fig. 1).

Lemma 1 (See [2]): Let $m(x, L)$ be an $L$-bit quantization of $x \in [-W, W]$. Then the mean square error (MSE), $D$, of an unbiased estimator satisfies

$$D \triangleq \mathbb{E}(|m(x, L) - \theta|^2) \approx \frac{W^2}{(2^L - 1)^2} + \sigma^2$$

(3)

where the expectation $\mathbb{E}(\cdot)$ is taken with respect to both the sensor observation noise and quantization noise. The approximation becomes sufficiently accurate when $L \gg 1$.

At the beginning of the transmission, FC needs to broadcast some common information to all the sensor nodes. Combining individual local information (available energy, channel state information and observation noise) with the common information, each sensor is to decide the transmission parameters (transmission power and time sharing fraction), which is discussed in the next section. Note that the quantization scheme is related to the transmission parameters by equation (2).

III. OPTIMAL JTPC

Under the assumption that the time sharing fraction can take arbitrary value between 0 and 1, we investigate JTPC policy when each node observes a spatially uncorrelated signal. Our goal is to prolong the minimum lifetime among all the nodes under individual estimation accuracy constraints. We choose the MSE to quantify the estimation accuracy. Then the optimal JTPC policy can be formulated as the problem, $\textbf{P1}$.

$$\textbf{P1:} \quad \max_{\rho, p} \min_k T_k$$

s.t. $D_k \leq \bar{D}_k, \quad \forall k \in K$ \hspace{1cm} (4)

$T_k p_k = E_k, \quad \forall k \in K$ \hspace{1cm} (5)

$$\sum_{k=1}^{K} \rho_k = 1$$

$$\rho_k \geq 0, \quad \forall k \in K,$$

where $T_k$ denotes the lifetime of node $k$, $E_k$ is the initial energy, and $D_k$ is the target MSE.

Using Lemma 1 and (2), we can express $D_k$ as a convex function of $g_k$, given by

$$D_k = w^2 [(1 + g_k p_k)^2 - 1]^{-2} + \sigma^2_k.$$
Substituting (6) into (4) results in

\[ p_k \geq \frac{1}{g_k} \left( \frac{a_k}{2} - 1 \right), \tag{7} \]

where

\[ R_k = \log_2 \left( \frac{W}{D_k - \sigma_k^2} + 1 \right), \quad \forall k \in K. \]

Applying (7) in the constraint (5), we obtain the inverse of node \( k \)’s lifetime as

\[ u_k(\rho_k) := \frac{1}{T_k} = \frac{\rho_k}{E_k g_k} \left( \frac{a_k}{2} - 1 \right). \tag{8} \]

The quantity \( u_k(\rho_k) \) can be interpreted as the normalized energy consumption of each sensor node in one time frame with respect to its initial energy. Thus far, the power optimization variables \( \{\rho_k\} \) and the constraints (4) (5) are all absorbed. The problem \( P1 \) is now equivalent to minimizing the maximum of \( u_k(\rho_k) \) subject to the constraints that \( \sum_{k=1}^{K} \rho_k = 1 \) and \( 0 \leq \rho_k \leq 1 \), \( \forall k \in K \). Due to the convexity of \( u_k(\rho_k) \) in \( \rho_k \) as can be verified easily, the optimal solution to \( P1 \) satisfies that \( u_k(\rho_k) = u_i(\rho_i), \quad \forall k, i \in K \).

We can now transform \( P1 \) into a new problem, \( P2 \), which includes the time-sharing variable vector \( \rho \) only,

\[
P2: \min_{\rho} \quad \sum_{k=1}^{K} (u(\rho) - u_k(\rho_k))^2 \\
\text{s.t.} \quad \sum_{k=1}^{K} \rho_k = 1, \quad 0 \leq \rho_k \leq 1, \quad \forall k \in K,
\]

where \( u(\rho) := \frac{1}{K} \sum_{k=1}^{K} u_k(\rho_k) \). The optimal solution of \( P2 \) obviously achieves the maximum network lifetime. To implement it in a distributed way, instead of solving the problem \( P2 \) at the FC, we add a penalty function to relax the constraint. The equivalent problem can then be expressed as:

\[
P3: \min_{\rho} \quad f(\rho) = \sum_{k=1}^{K} [u(\rho) - u_k(\rho_k)]^2 + \mu \left( 1 - \sum_{k=1}^{K} \rho_k \right)^2 \\
\text{s.t.} \quad 0 \leq \rho_k \leq 1, \quad \forall k \in K
\]

where \( \mu \) is a penalty factor and can take any arbitrary positive value.

**Theorem 1:** Problems \( P2 \) and \( P3 \) have the same and unique optimal solution. The unique optimal solution is \( \rho^* \) if and only if \( \nabla f(\rho^*) = 0 \).

**Proof:** According to the discussion above, if \( \rho^* \) is an optimal solution of Problem \( P2 \), it must satisfy the following condition:

\[
\left\{ \begin{array}{l}
  u_i(\rho^*_i) = u_j(\rho^*_j), \quad \forall i, j \in K \\
  \sum_{k=1}^{K} \rho_k^* = 1
\end{array} \right.
\]

We now prove the existence and uniqueness of the above solution \( \rho^* \).

**Existence and uniqueness:** when \( x \to 0, \sum_{k=1}^{K} u_k^{-1}(x) \to +\infty \) (here, \( u_k^{-1}(\cdot) \) is the inverse function of \( u_k(\cdot) \)), and when \( x \to +\infty, \sum_{k=1}^{K} u_k^{-1}(x) \to 0 \). Meanwhile, \( \sum_{k=1}^{K} u_k^{-1}(x) \) is a strictly decreasing function of \( x \). Thus there exists a unique \( x^* \) satisfying \( \sum_{k=1}^{K} u_k^{-1}(x^*) = 1 \). It is easy to see that a unique solution \( \rho_k^* = u_k^{-1}(x^*) \) satisfies (9), and condition (9) results in the minimum of the objective function of Problem \( P2 \) and \( P3 \).

The following shows that condition (9) holds if and only if \( \nabla f(\rho^*) = 0 \). When \( K \) is sufficiently large and/or the step size mentioned in Section IV is sufficiently small, \( \bar{u}(\rho) \) can be viewed as a constant at each updating \( \rho_k \). Hence, we have

\[
\nabla_k f(\rho^*) = -2 \left\{ u_k'(\rho_k) \bar{u}(\rho) - u_k(\rho_k) + \mu \left( 1 - \sum_{k=1}^{K} \rho_k \right) \right\},
\]

for all \( k \in K \). When \( \nabla f(\rho^*) = 0 \), if we assume that \( 1 - \sum_{k=1}^{K} \rho_k > 0 \), then it holds \( \bar{u}(\rho) - u_k(\rho_k) < 0, \forall k \in K \). However, it contradicts with the definition of \( \bar{u}(\rho) := 1/K \sum_{k=1}^{K} u_k(\rho_k) \). Similarly, it can be proven that \( 1 - \sum_{k=1}^{K} \rho_k < 0 \) does not hold when \( \nabla f(\rho^*) = 0 \). When condition (9) is satisfied, obviously \( \nabla f(\rho^*) = 0 \). Therefore, \( \rho^* \) is the unique optimal solution if and only if \( \nabla f(\rho^*) = 0 \).

**IV. Partially Distributed Transmission Scheme Adaptation**

Using gradient projection, the time sharing fractions are adjusted in the opposite direction to the gradient \( \nabla f(\rho) \), i.e.,

\[
\rho(t+1) = \rho(t) - \gamma \nabla f(\rho(t))_0,
\]

where \( \gamma \) is a stepsize, and \( [y]_0^0 = \min\{\max\{y, a, b\}\} \). As seen in (10), the time sharing adaptation can be implemented at each sensor node individually in a distributed manner, and the only common information required is \( \bar{u}(\rho) \) and \( \sum_{k=1}^{K} \rho_k \). Thus, it has a good scalability.

**Theorem 2:** There exists a constant \( \gamma_0 > 0 \) such that if \( 0 < \gamma < \gamma_0 \), then \( \lim_{t \to \infty} \nabla f(\rho(t)) = 0 \). That is, given an arbitrary initial point, \( 0 < \rho(0) < 1 \), the time sharing policy \( \rho \) generated by the iterative function (10) converges to the unique optimal solution \( \rho^* \).

The proof of this theorem uses the following Lemmas.

**Lemma 2 (Lipschitz Continuity of \( \nabla f(\rho) \)):** The function \( f \) is continuously differentiable and there exists a constant \( B \) such that \( || \nabla f(\rho) - \nabla f(\rho') ||_2 \leq B || \rho - \rho' ||_2 \).

**Lemma 3 (Descent Lemma in [8]):** If \( \nabla f(\rho) \) is Lipschitz continuous then \( f(\rho + \rho') \leq f(\rho) + (\rho')^T \nabla f(\rho) + 0.5B \| \rho' \|_2^2 \).

Define the error term \( s_k(t) := [\rho_k(t+1) - \rho_k(t)]/\gamma \). Using Lemma 2 and 3, we can obtain

\[
f(\rho(t+1)) \leq f(\rho(t)) + \gamma s_k^T(t) \nabla f(\rho(t)) + \gamma^2 B / 2 \| s(t) \|_2^2 \leq f(\rho(t)) - \gamma \left( 1 - \frac{B^2}{2} \right) \| s(t) \|_2^2.
\]
Summing the above inequalities from $t = 0$ yields
\[
\begin{align*}
  f(\rho(t+1)) & \leq f(\rho(0)) - \gamma \left( 1 - \frac{B \gamma}{2} \right) \sum_{t=0}^{\infty} \| s(t) \|_2^2 \\
                         &= f(\rho(0)) - \gamma \left( 1 - \gamma \gamma_0 \right) \sum_{t=0}^{\infty} \| s(t) \|_2^2,
\end{align*}
\]
where $\gamma_0 = 2/B$. Since $f(\rho(t+1)) \geq 0, \forall t$, we obtain
\[
\sum_{t=0}^{\infty} \| s(t) \|_2^2 \leq \frac{1}{\gamma \left( 1 - \gamma \gamma_0 \right)} f(\rho(0)) < \infty.
\]
This implies that $\lim_{t \to \infty} s(t) = 0$.

Combining with Theorem 1, we can achieve the convergence of $\rho(t)$ generated by (10) to the unique optimal point.

**Proposition 1:** When $R_k = R$ ($\forall k \in K$), the optimal time sharing fraction only relates to the product of available energy and channel gain. More time sharing fraction will be given to the user with smaller product.

Proposition 1 suggests that the sensors with bad channel and/or energy condition will take longer time for transmission. This is in agreement with the idea of “lazy scheduling” [9], which achieves energy saving by lowering the transmission power and hence taking longer transmission time.

**V. NUMERICAL RESULTS**

This section compares the network lifetime of our proposed JTPC policy with the following two transmission schemes:

- **Uniform TDMA with power control** (UTP): it can be viewed as a suboptimal solution of JTPC by fixing $\rho_k = 1/K, \forall k \in K$. The optimal transmission power $p_k$ under the constraint (4) can be obtained from (7).

- **Inverse-log scheduling** (ILS) [6]: the total energy in a sensor network is minimized by varying the transmission times assigned to different sensor nodes. The time-sharing fraction in the ILS scheme has an inverse-log form,

\[
\rho_k' = \frac{R_k}{\log(\lambda g_k)}, \forall k \in K,
\]

where $\lambda$ is determined by the constraint $\sum_k \rho_k' = 1$, and it can be obtained by numerical search.

The simulation is carried out under various network and channel configurations by varying path losses, available energies, and observation noise variances. The channel gain is modelled as $g = ad^{-n}$, where $d^{-n}$ is the effect of path loss, and $a$ is an attenuation constant. Here, the distance between sensor node $k$ and the FC, $d_k$, is randomly generated following a uniform distribution in the region $[1,10]$. Similar to [10], observation noise variances $\sigma_k^2$ ($\forall k \in K$) are generated according to the distribution $0.05(1 + a Y^2)$, where $Y^2$ follows the exponential distribution. In addition, we set available energies $E_k = E_{\text{min}}(1 + b Y^2)$, where $E_{\text{min}}$ is the minimum energy of all nodes. Through adjusting the parameters $n$, $a$, and $b$, we generate $g_k$, $E_k$ and $\sigma_k^2$ for all $k \in K$ to model different sensing environments. In the simulation, we set the target MSE, $D_k = 2\sigma_k^2 \forall k \in K$. We use the normalized deviation $\sqrt{\text{Var}(t)/E(t)}$ defined in [10] to measure the heterogeneity of a variable $t$, where $\text{Var}(t)$ denotes the variance of $t$ and $E(t)$ the mean of $t$.

Fig. 2 compares the network lifetime in time frames of JTPC, UTP and ILS by choosing the simulation parameters $n = 3$, $a = 1$, and $b = 0$, and varying the number of sensor nodes from 10 to 20. It is shown that the network lifetimes decrease exponentially with the number of sensor nodes. This is because the increase of the number of sensor nodes reduces the average transmission times of each node when the duration of a time frame is fixed. Correspondingly, the energy consumption increases dramatically. This phenomenon suggests that partitioning a WSN into several subnetworks could prolong the network lifetime at the expense of increasing the number of FCs. The performance of ILS is a little worse than that of UTP. It is because minimizing the total energy consumption in ILS leads to some nodes running out of energy quickly. It is also observed that the lifetime gains of JTPC over UTP and ILS increase exponentially with the number of nodes. Here, we define the lifetime gain of one scheme over another as the ratio of network lifetimes obtained by the two schemes. The following figures illustrate the lifetime gains obtained by the proposed scheme under different sensing environments.

Fig. 3 plots the lifetime gain of JTPC over the UTP and ILS schemes versus normalized deviation of channel path losses by varying the path loss exponent $a$. Both $\sigma_k^2$ and $E_k$ are fixed by letting $a = b = 0$. Fig. 4 is plotted by fixing $d_k = 5.5$, $\forall k \in K$, $b = 0$ and varying $a$. It is observed that the lifetime gains increase when observation noise variances become more heterogeneous. In addition, for fixed normalized deviation of noise variances, the lifetime gain rises as the number of sensors increases. The relationship between the lifetime gains and the normalized deviation of initial energies is shown in Fig. 5. A trend similar to that of Figs. 3 and 4 is observed.
VI. CONCLUSION

We have considered a joint time sharing and power control policy to prolong the lifetime of cluster-based WSNs under the average estimation distortion constraints. Our transmission scheme focused on single-hop data collection from sensor nodes to a common FC in a particular cluster where each node observes a spatially uncorrelated signal. Considering the practical limited initial energy and computational capacity of FC, we developed a partially distributed algorithm where the JTPC policy can be calculated at each node with the local information of the path loss, available energy and observation quality. Numerical examples show that significant lifetime gain can be achieved when compared with the UTP and ILS schemes, and the gain becomes more significant when sensing environment becomes more heterogeneous and the number of nodes increases.

A challenge of the future work would be to extend JTPC policies to a more generic distributed estimation framework discussed in [10] where the observed signals could be partially correlated. Also the definition of network lifetime considered in this paper is very simple. In a real WSN, the death of some nodes may not affect the functioning of the whole network. More useful definitions of network lifetime, for example, the time when the $L^{th}$ ($L < K$) node dies, are needed.

REFERENCES