Energy-efficient transmission and bit allocation schemes in wireless sensor networks

Xiaoxia Zhang*, Liang-Liang Xie and Xuemin Shen

Department of Electrical and Computer Engineering,
University of Waterloo,
Waterloo, ON, Canada
Email: x79zhang@uwaterloo.ca
Email: llxie@uwaterloo.ca
Email: sshen@uwaterloo.ca
*Corresponding author

Abstract: Energy-efficient transmission and bit allocation schemes are investigated in multi-source single-sink Wireless Sensor Networks (WSNs). For transmission over Additive White Gaussian Noise (AWGN) channels with path loss, this work shows that the overall energy consumption can be minimised if each sensor transmits with the minimum power and cooperates with others in Time-Division Multiple Access (TDMA) mode. From the efficient correlated source coding perspective, the Slepian–Wolf coding theorem is applied. Jointly considering the two aspects, we propose a closed form bit allocation scheme to minimise the overall energy consumption. The underlying idea is to assign more bits to nodes with better channel conditions. Additionally, based on the definition of network lifetime as the time before the first sensor fails, we further maximise the network lifetime by developing a heuristic algorithm to balance energy consumption among sensors. The superiority of the proposed scheme is validated by both analytical and simulation results.

Keywords: WSN; wireless sensor network; bit allocation; energy efficiency; network lifetime; Slepian–Wolf coding.

Reference to this paper should be made as follows: Zhang, X., Xie, L-L. and Shen, X. (2012) 'Energy-efficient transmission and bit allocation schemes in wireless sensor networks', *Int. J. Sensor Networks*, Vol. 11, No. 4, pp.241-249.

Biographical notes: Xiaoxia Zhang is presently working towards the PhD degree at Department of Electrical and Computer Engineering in University of Waterloo, ON, Canada. She received her BE degree from Beijing University of Posts and Telecommunications, Beijing, China, in 2008 and MASc degree in Electrical and Computer Engineering from the University of Waterloo, Waterloo, ON, Canada, in 2010. Her current research interests include network information theory and wireless sensor networks.

Liang-Liang Xie is presently a Professor at the Department of Electrical and Computer Engineering in University of Waterloo, Waterloo, ON, Canada. He received his BS degree in Mathematics from Shandong University, Jinan, China, in 1995 and the PhD degree in Control Theory from the Chinese Academy of Sciences, Beijing, in 1999. He did his Postdoctoral research with the Automatic Control Group, Linkoping University, Linkoping, Sweden, during 1999–2000 and with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, during 2000–2002. His research interests include wireless networks, information theory, adaptive control and system identification.

Xuemin (Sherman) Shen is a Professor and University Research Chair, Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research focuses on resource management, wireless network security, vehicular ad hoc and sensor networks. He served as the Technical Program Committee Chair for IEEE VTC'10 Fall, and the Technical Program Committee Chair for IEEE Globecom'07. He also serves/served as the Editor-in-Chief for *IEEE Network*, *Peer-to-Peer Networking and Application*, and *IET Communications*. He is a registered Professional Engineer of Ontario, Canada, an IEEE Fellow, an Engineering Institute of Canada Fellow, and a Distinguished Lecturer of IEEE Vehicular Technology Society and Communications Society.

1 Introduction

Sensors that are capable of sensing, data processing and communicating have enabled the realisation of Wireless Sensor Networks (WSNs) (Akyildiz et al., 2002). In a WSN, a large number of nodes are deployed in an area to monitor an environment by measuring some physical parameters such as temperature, humidity and pressure (Knaian et al., 2000; Mainwaring et al., 2002; Cook and Das, 2005; Lynch and Loh, 2006; Patel et al., 2008; Li and Liu, 2009; Cheng et al., 2011). However, generally, wireless sensor nodes carry very limited irreplaceable power sources. Inspite of the significant advances in processor computing, little improvement has been achieved for the battery recharging and replacing in many scenarios. Consequently, two primary concerns in WSNs are to save the overall energy consumption and to prolong the network lifetime, namely the time when all the nodes are functional.

Tremendous research efforts have been dedicated to the problem of maximising energy efficiency in low-power WSNs from several perspectives including data collecting, source coding and transmission. In particular, on correlated data gathering and source coding, Cristescu et al. (2004) proposed a closed form optimal rate allocation scheme to minimise some transmission cost function which is proportional to the distance from the source node to the sink, thus to decrease the overall transmitting power. As for the transmission (channel coding) in a wireless environment, Knopp and Humblet (1995) showed that in order to maximise the aggregate capacity with fixed power, only the user having the best channel condition is allowed to transmit at any given time, which is essentially the same as Time-Division Multiple Access (TDMA). Hou et al. (2008) provide an algorithm to maximise the sum-rate of all the nodes under the constraint of satisfying a given network lifetime requirement. Mergen et al. (2006) examine the trade-off between improving the sum-rate and saving energy in many aspects of designing the sensor networks. Gastpar and Vetterli (2003) considered both source and channel coding to optimise the rate distortion function under a fixed power constraint. Finally, Roumy and Gesbert (2007) studied the combination of source and channel coding to provide an algorithm to minimise the overall power consumption and to maximise network lifetime; nevertheless, it only considers a special scenario where one sensor is only correlated to another sensor.

In this paper, both source and channel coding are jointly considered towards the energy efficiency problem in WSNs where correlation exists among all sensors.[1] We first consider the single transmission in Additive White Gaussian Noise (AWGN) channels with path loss. The capacity is known as $\frac{1}{2}\log\left(1+\frac{P/d^{\gamma}}{N}\right)$, where P is the transmitting power, d is the distance from the source node to the sink node, γ is the path loss exponent and N is the noise variance. We focus on the sensor network scenario where the environment and data do not change rapidly, so that there is enough transmitting time, and the transmitting rate is not a major concern. We demonstrate the optimal strategy for the single source is to employ minimum transmitting power in terms of improving the energy efficiency. Then we extend the model to consider the multi-source single-sink WSNs. For transmission part, it is further shown that when all the sensor nodes transmit in the TDMA mode to the common sink node, the overall energy consumption can be minimised.

Besides the transmitting power P, the overall energy consumption also depends on the transmitting time. When the transmission rate is fixed, transmitting time is proportional to the total number of transmission bits $\mathcal{B}$. Since in TDMA mode, the channel can be regarded as a single user channel at any given time, source-channel separation theorem (Cover and Thomas, 1991) holds. For the source coding part, the distributed source coding theorem (Slepian–Wolf coding) (Slepian and Wolf, 1973) can be applied in order to further improve the resource efficiency. Therefore, we propose a bit allocation scheme to minimise the overall energy consumption without loss of any information. Based on this criterion, we further develop an algorithm to adjust the optimal bit allocation scheme to maximise the network lifetime.

The remainder of the paper is organised as follows. In Section 2, we present an energy-efficient transmitting strategy for a single sensor. If there are multiple sensors, Section 3 proves that TDMA transmission mode is optimal in terms of the overall energy saving. Section 4 presents the background of correlated source coding. The optimisation problem of minimising the overall energy consumption is formulated and a closed form bit allocation solution is obtained in Section 5. Section 6 extends the optimisation objective to include lifetime prolonging and provides an algorithm which aims at achieving the two aforementioned goals. Section 7 contains some numerical results.

2 Energy-efficient method for single-sensor single-sink transmission

We first consider the case where there is only one source node and one sink node. Suppose the sensor collects data from its environment during one time period and transmits the encoded data in the next time period. The encoded data of one time period is represented by $\mathcal{B}$ bits where $0 < \mathcal{B} < \infty$.

Generally, the transmitting power of each sensor is both lower bounded by $P_{\min}$ and upper bounded by $P_{\max}$. Still, there are many different levels of transmitting power the sensor can adopt. It can use either lower power but longer time, or higher power but shorter time to send the same amount of bits. The question is; which way is better if energy efficiency is the primary concern?

It is well known that for an AWGN channel, the reliable transmission rate is bounded by its capacity: $R \le \frac{1}{2}\log(1+SNR)$, where SNR is the received signal-to-noise ratio. Besides, in a wireless environment, the channel capacity is affected by path loss. Denote the distance from the source to the sink node by d and path loss exponent is a constant γ. Then the channel capacity is:

$$R \le \frac{1}{2}\log\left(1+\frac{P/d^{\gamma}}{N}\right), \tag{1}$$

where, P is transmitting power and N is the noise variance.

In order to fully utilise the power, the node should transmit at a rate as close to the capacity as possible. Thus, in the following analysis, without loss of generality, we assume that $R = \frac{1}{2}\log\left(1+\frac{P/d^{\gamma}}{N}\right)$.

Theorem 1: *In the single source case, given that the transmitting power can be chosen arbitrarily within the bound $P_{\min} \le P \le P_{\max}$, in order to transmit $\mathcal{B}$ bits in total, using P_{min} consumes the least overall energy.*

Proof: Denote the transmission time by T. Then,

$$\mathcal{B} = RT = \frac{1}{2}\log\left(1+\frac{P/d^{\gamma}}{N}\right)T, \tag{2}$$

where $\mathcal{B}$ and N are constants. The overall energy consumption in this transmission is:

$$E = PT = P\frac{\mathcal{B}}{\frac{1}{2}\log\left(1+\frac{P/d^{\gamma}}{N}\right)}. \tag{3}$$

Then the power that minimises the overall energy consumption can be obtained simply by taking derivative with regard to P, i.e.

$$\begin{aligned}\frac{dE}{dP} &= 2\mathcal{B}\left[\frac{P}{\log\left(1+\frac{P/d^{\gamma}}{N}\right)}\right]' \\ &= \frac{2\mathcal{B}}{\left[\log\left(1+\frac{P/d^{\gamma}}{N}\right)\right]^2\left(1+\frac{P/d^{\gamma}}{N}\right)} \\ &\quad \left[\left(1+\frac{P/d^{\gamma}}{N}\right)\log\left(1+\frac{P/d^{\gamma}}{N}\right)-\frac{P/d^{\gamma}}{N}\log e\right].\end{aligned}$$

Let $f(x) = (1+x)\log(1+x) - x\log e$, where $x = \frac{P/d^{\gamma}}{N}$.

Since $P > 0, N > 0, d > 0, x > 0$,

$$f'(x) = \log(1+x) \ge 0, \forall x > 0. \tag{4}$$

Taking into account the fact that $f(0) = 0$, we can attain

$$f(x) > 0, \forall x > 0, \tag{5}$$

which implies that

$$\frac{dE}{dP} > 0, \forall P > 0, d > 0 \text{ and } N > 0. \tag{6}$$

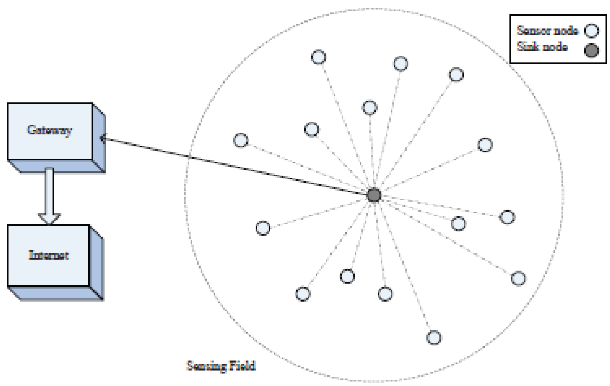
Therefore, the energy consumption function is monotone increasing with transmitting power P. In other words, using minimum power $P_{\min}$ consumes the least energy. This result is in consistence with our observation. Since $\log(1+x)$ approaches a linear function only when $x \to 0$, the smaller power we use, the more efficient is the transmission.

3 Energy-efficient method for multi-sensor single-sink transmission

For multi-source single-sink transmission, there are two types of transmission modes: source nodes transmitting one after another, i.e. every time there is only one source-sink connection TDMA and more than one source-sink connections existing at some specific time (multiple access channel).

Consider a network with n nodes all transmitting to a single-sink node as depicted in Figure 1. Let $\mathcal{N}$ be the set of sensor indices: $\mathcal{N} = \{1,\dots,n\}$. Then for each node $i \in \mathcal{N}$, it uses its minimum power $P_{i,min}$ and the total bits it needs to transmit is some fixed value $\mathcal{B}_i$. Suppose the distance from node i to the sink node is d_i and path loss exponent γ is a constant for every path.

Figure 1 A sensor network with n sensors and one sink node (see online version for colours)



Theorem 2: *For multi-source single-sink transmission, given that each node uses its minimum power and has finite total bits to transmit, using TDMA among all the nodes is more energy efficient.*

Proof: The overall energy consumption can be calculated as $\sum_{i=1}^{n} P_{i,min}T_i$, where $T_i = \mathcal{B}_i / R_i$ is the transmission time,

$$E = \sum_{i=1}^{n} P_{i,min}\frac{\mathcal{B}_i}{R_i}. \tag{7}$$

1 Using TDMA, the instant rates of all source nodes can reach their separate channel capacities, i.e.

$$R_i = \frac{1}{2}\log\left(1+\frac{P_{i,min}/d_i^{\gamma}}{N}\right), \forall i \in \mathcal{N}. \tag{8}$$

Notice that traditionally, the concept 'rate' usually refers to the average rate in a time frame. However, here we adopt the 'instant rate' exclusively for the convenience of calculation of the overall energy consumption.

2 Suppose at some specific time, a set $S \subseteq \mathcal{N}$ with $|S| \geq 2$ all connect to the sink node, then the rate vector should lie in multiple access channel capacity region. For any subset $S_1 \subseteq S$ (We use $R_i^{'}$ to differentiate):

$$\sum_{i \in S_1} R_i^{'} \leq \frac{1}{2} \log \left(1 + \frac{\sum_{i \in S_1} P_{i,min} / d_i^{\gamma}}{N} \right)$$
$$< \sum_{i \in S_1} \frac{1}{2} \log \left(1 + \frac{P_{i,min} / d_i^{\gamma}}{N} \right) = \sum_{i \in S_1} R_i.$$

So,

$$R_i^{'} < R_i, \forall i \in S.$$

Since both $P_{i,min}$ and $\mathcal{B}_i$ are fixed, the energy consumption function is reversely proportional to the instant rate. Thus, the overall energy consumption in Case 2 is greater than that in Case 1. Therefore, transmitting using TDMA is more energy-efficient.[2]

4 Source coding for sensors gathering correlated data

In general, the measurements of sensors, e.g. temperature, humidity, sound, etc., are highly correlated. When encoding those correlated sources, Slepian and Wolf (1973) showed that a total rate of $R = \sum_{i=1}^{n} R_i = H(W_1, W_2, \ldots, W_n)$, which is the joint entropy of information from all sensors, is sufficient even if these nodes are not able to communicate with each other.

More precisely, if there are multiple n sources $W_1, W_2, \ldots, W_n$ drawn *i.i.d* according to $p(w_1, w_2, \ldots, w_n)$ that are encoded separately, and there is one sink node that can decode them together, then the set of achievable source coding rate vectors must lie in Slepian–Wolf region (here, we use R_S to represent the source coding rate):

$$R_S(U) > H(W(U) \mid W(U^C)), \tag{9}$$

for all $U \subseteq \{1, 2, \ldots, n\}$, where

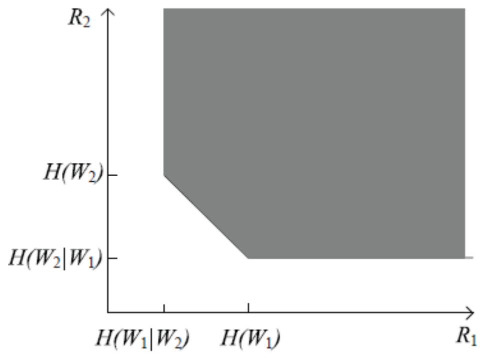
$$R_S(U) = \sum_{j \in U} R_{Sj}, \tag{10}$$

and

$$W(U) = W_j : j \in U. \tag{11}$$

Figure 2 shows the Slepian–Wolf region for the case of two sources.

Figure 2 Slepian–Wolf region for two correlated sources



Suppose one time period for all the sensors to gather data is represented by T_S and $T_S > \sum_{i=1}^{n} T_i$. This is because the measurements of physical phenomena that are sensed by sensors usually do not vary rapidly, e.g. the temperature in a day. In other words, the source rate is much smaller than the instant transmitting rate, so T_S is greater than the transmitting time. Under this assumption, all the sensors gather data in the first period T_S, with rate R_{Si}, then in the next period, they cooperate using TDMA to transmit the encoded data to the sink node, with larger transmitting rate R_i.

The total information generated by sensor i is $R_{Si}T$, which should be equal to the bits that are transmitted by sensor i, i.e.

$$R_{Si} T = \mathcal{B}_i. \tag{12}$$

So $\mathcal{B}_i$ must satisfy:

$$\sum_{i \in U} \mathcal{B}_i > H(W(U) \mid W(U^C))T, \forall U \subseteq \{1, 2, \ldots, n\}. \tag{13}$$

This sophisticated result has a simple interpretation. The information collected by each sensor contains two parts: the unique information of this sensor and some common information that can also be conveyed by other sensors. When encoding these sources, each sensor must first encode its unique information; then this sensor cooperates with other sensors to divide the task of the common part.

5 Source-channel communication in minimisation of energy consumption

In this section, we consider the problem of how to allocate information bits, i.e. $\mathcal{B}_i$ among the sensor nodes for transmission in order to minimise the overall energy consumption.

Intuitively, if a sensor is very far away from the destination, channel condition is severely degraded. The transmitting rate is very slow even if high power is used, which is a waste of energy. Thus we tend to assign less task (bits) to this channel. For example, we only use this channel to transmit those information that is uniquely generated in this sensor and use other good channels to transmit those common information.

The overall energy consumption is:

$$E = \sum_i P_{i,min} \frac{\mathcal{B}_i}{R_i} = \sum_i P_{i,min} \frac{\mathcal{B}_i}{\frac{1}{2}\log\left(1+\frac{P_{i,min}/d_i^\gamma}{N}\right)}$$
$$= \sum_i \frac{P_{i,min}}{\frac{1}{2}\log\left(1+\frac{P_{i,min}/d_i^\gamma}{N}\right)} \mathcal{B}_i.$$

Since the coefficient of $\mathcal{B}_i$ is only related to i and when $P_{i,min}$ and d_i are fixed, this coefficient is a constant. Let a_i represent $\frac{P_{i,min}}{\frac{1}{2}\log\left(1+\frac{P_{i,min}/d_i^\gamma}{N}\right)}$, this problem can be formulated as follows:

$$\begin{aligned} &\min \sum_i a_i \mathcal{B}_i \\ &\text{subject to} \sum_{i\in U} \mathcal{B}_i > H(W(U) \mid W(U^c))T, \forall U \subseteq \mathcal{N}. \end{aligned} \tag{14}$$

Theorem 3: *Without loss of generality, let* $a_1 > a_2 > \ldots > a_n$*, then the solution to the above optimisation problem is* $\mathcal{B}^*$*:*

$$\begin{aligned} \mathcal{B}_1^* &= H(W_1 \mid W_2 \cdots W_n)T \\ \mathcal{B}_2^* &= H(W_2 \mid W_3 \cdots W_n)T \\ &\vdots \\ \mathcal{B}_{n-1}^* &= H(W_{n-1} \mid W_n)T \\ \mathcal{B}_n^* &= H(W_n)T. \end{aligned} \tag{15}$$

Proof: We prove Theorem 3 in two steps:

1 First, the solution (15) satisfies all the constraints in optimisation problem (14).

2 Second, this solution can minimise the overall energy consumption.

Proof of Step 1: Define set $S_j = \{j, j+1, \ldots, n\}$, where $1 \le j \le n$, so $S_1 = \mathcal{N}$.

Then for any subset $U \subseteq \mathcal{N}$,

$$\begin{aligned} \sum_{j\in U} \mathcal{B}_j^* &= \sum_{j\in U} H(W_j \mid W(S_{j+1}))T \\ &= \sum_{j\in U} H(W_j \mid W(S_{j+1} \cap U), W(S_{j+1} \cap U^c))T \\ &\overset{(a)}{\ge} \sum_{j\in U} H(W_j \mid W(S_{j+1} \cap U), W(U^c))T \\ &\overset{(b)}{=} H(W(U) \mid W(U^c))T, \end{aligned} \tag{16}$$

where (a) follows from conditioning reduces entropy and (b) follows from the chain rule for entropy.

Proof of Step 2: We prove this part by induction. We start from $\mathcal{B}_n^* = H(W_n)T$.

$\mathcal{B}_1^* + \ldots + \mathcal{B}_n^* = H(W_1, \ldots, W_n)T$, which has already achieved the Slepian–Wolf bound. If there exists another set of solution, say, $\mathcal{B}_1^{'}, \ldots, \mathcal{B}_n^{'}$, this set of solution must achieve the bound as well. Since $\mathcal{B}_n^*$ is already the largest value that $\mathcal{B}_n$ can be, let $\mathcal{B}_n^{'} = H(W_n)T - \delta$, where δ is a positive small value that does not affect the Slepian–Wolf conditions. Then this δ must be added to some $\mathcal{B}_i^*$, where $1 \le i \le n-1$. Then

$$\begin{aligned} E^{'} &= a_1\mathcal{B}_1^{'} + \ldots + a_n\mathcal{B}_n^{'} \\ &= a_1\mathcal{B}_1^* + \ldots + a_{i-1}\mathcal{B}_{i-1}^* + a_{i+1}\mathcal{B}_{i+1}^* + \ldots + a_{n-1}\mathcal{B}_{n-1}^* \\ &\quad + a_i\mathcal{B}_i^{'} + a_n\mathcal{B}_n^{'} \\ &= E^* + (a_i - a_n)\delta. \end{aligned}$$

Since $a_i \ge a_n$, $(a_i - a_n)\delta \ge 0$ and thus $E^{'} \ge E^*$. So $\mathcal{B}_n^*$ is optimum.

Given $\mathcal{B}_n^*$, $\mathcal{B}_{n-1}^* = H(W_{n-1} \mid W_n)T$ is the largest value that $\mathcal{B}_{n-1}$ can choose. By the same argument, it is obvious that $\mathcal{B}_{n-1}^*$ is the optimal value.

Similarly, we can prove that $\mathcal{B}^*$ is the optimum solution. Actually, the interpretation of this solution is quite straight-forward. The coefficient a_i is the energy consumption per bit. Thus, it is the most efficient if we transmit more bits to smaller a_i and less bits to larger a_i.

6 Lifetime maximisation

In WSNs, energy consumption and network lifetime are the two primary concerns. In previous section, we presented a bit allocation scheme that can minimise the overall energy consumption. However, if there is one sensor that consumes more energy than the rest of the sensors, this sensor is the first to die. And when any sensor dies, the whole network must lose a part of the information, so we say the network dies. Notice that, network lifetime has different definitions in various application scenarios, here we adopt the most commonly used notion that network lifetime is the time before the first node fails (Dietrich and Dressler, 2009). Therefore, in this section, we are looking for a scheme that can average the energy consumption distribution among all the sensors, thus prolong the network lifetime, while at the same time, minimising the overall energy consumption.

The problem can be formulated by the following lexicographic optimisation:

$$\begin{aligned} &lex\min(\max_i E_i, \sum_i E_i) \\ &\text{subject to} \sum_{i\in U} \mathcal{B}_i > H(W(U) \mid W(U^c))T, \forall U \subseteq \mathcal{N}, \end{aligned} \tag{17}$$

where $E_i = a_i\mathcal{B}_i$.

This optimisation problem has a hierarchical structure, i.e. the first objective has higher priority than the second objective (Ehrgott, 2005). Under the condition that the highest energy consumption has been minimised, we try to minimise the overall energy consumption to achieve lexicographic optimality.

To minimise E_i for some i, we only need to reduce its corresponding $\mathcal{B}_i$, i.e. assign less bits that need to be transmitted by this sensor. Since we still desire the least overall energy consumption, we start from the point $\mathcal{B}^*$, find the node that consumes the most energy, reduce the bits of this node and reallocate these bits among the rest nodes, until the highest node's information cannot be reduced anymore. This algorithm can achieve lexicographical optimality.

Suppose $E_k, 1 \le k \le n$ is the largest, so $\mathcal{B}_k$ needs to be reduced. It can be seen that $\mathcal{B}_1^* + \mathcal{B}_2^* + \ldots + \mathcal{B}_k^* = H(W_1, W_2, \ldots, W_k)T$, which is already the Slepian–Wolf bound. Thus, $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_{k-1}$ need to increase the same bits in total. However, $\mathcal{B}_k^* + \mathcal{B}_{k+1}^* + \ldots + \mathcal{B}_n^* > Slepian-Wolf\ bound$, So $\mathcal{B}_{k+1}, \ldots, \mathcal{B}_n$ may not need to change, as will be proved later.

Suppose we reduce $\mathcal{B}_k$ by 1 bit, so this bit needs to be added to $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_{k-1}$. Since $a_1 > a_2 > \ldots > a_{k-1}$, adding this 1 bit to node $k-1$ will consume the least extra energy, thus is the most efficient choice. Therefore, when we need to transfer some bits from node *k* to the nodes before it, we prefer to add those bits to the node just before it first, and then two positions before it, and so on till the first node.[3] How many bits can we transfer at most to node $l, 1 \le l \le k-1$ without violating any of the Slepian–Wolf conditions?

Lemma 1: *If node k consumes the most energy, $\mathcal{B}_k^*$ needs to be reduced, and we add those bits to $\mathcal{B}_{k-1}^*$ until $\mathcal{B}_{k-1}$ is full, then to $\mathcal{B}_{k-2}^*$ until $\mathcal{B}_{k-2}$ is full, , and so on till $\mathcal{B}_1$ is full.*[4] *In this process, we can transfer to $\mathcal{B}_l, 1 \le l \le k-1$ at most $I(W_l; W_k \mid W_{l+1}, \ldots, W_{k-1}, W_{k+1}, \ldots, W_n)T$ bits before it is full.*

Proof: We only need to prove the case when $l = k-1$. For the rest *l*s, this result can be easily extended by induction.

Suppose after transferring bits from $\mathcal{B}_k^*$ to $\mathcal{B}_{k-1}^*$ and $\mathcal{B}_{k-1}$ being full, $\mathcal{B}_k^*$ has been reduced to $\hat{\mathcal{B}}_k$ and $\mathcal{B}_{k-1}^*$ has been added to $\hat{\mathcal{B}}_{k-1}$ and all the rest of $\mathcal{B}_i^*$ s remain the same. The only possibility that Slepian–Wolf condition might not hold is that the addition of any subset of $\{\mathcal{B}_1^*, \mathcal{B}_2^*, \ldots, \mathcal{B}_{k-2}^*\}$ and $\hat{\mathcal{B}}_k$ might not be greater than the Slepian–Wolf bound. Thus, we are looking for the largest $d_k^{k-1} = \mathcal{B}_k^* - \hat{\mathcal{B}}_k$, which is the bits we can transfer at most.

Since $\forall V \subseteq \{1, \ldots, k-2\}$,

$$\begin{aligned}\sum_{i\in V} \mathcal{B}_i^* + \mathcal{B}_k^* &= \sum_{i\in V} H(W_i \mid W_{i+1}, \ldots, W_n)T \\ &\quad + H(W_k \mid W_{k+1}, \ldots, W_n)T.\end{aligned} \tag{18}$$

However, after $\mathcal{B}_k^*$ has been reduced to $\hat{\mathcal{B}}_k$, according to Slepian–Wolf bound,

$$\begin{aligned}\sum_{i\in V} \mathcal{B}_i^* + \hat{\mathcal{B}}_k &= H(W(V), W_k \mid W(V^c), \\ &\quad W_{k-1}, W_{k+1}, \ldots, W_n)T.\end{aligned} \tag{19}$$

Thus,

$$\begin{aligned}d_k^{k-1} &= \mathcal{B}_k^* - \hat{\mathcal{B}}_k \\ &= \sum_{i\in V} H(W_i \mid W_{i+1}, \ldots, W_n)T \\ &\quad + H(W_k \mid W_{k+1}, \ldots, W_n)T \\ &\quad - H(W(V), W_k \mid W(V^c), W_{k-1}, W_{k+1}, \ldots, W_n)T\end{aligned} \tag{20}$$

$$\begin{aligned}&\overset{(a)}{=} \sum_{i\in V} H(W_i \mid W_{i+1}, \ldots, W_n)T \\ &\quad + H(W_k \mid W(V^c), W_{k-1}, W_{k+1}, \ldots, W_n)T \\ &\quad + I(W_k; W(V^c), W_{k-1} \mid W_{k+1}, \ldots, W_n)T \\ &\quad - H(W(V), W_k \mid W(V^c), W_{k-1}, W_{k+1}, \ldots, W_n)T\end{aligned} \tag{21}$$

$$\begin{aligned}&\overset{(b)}{\ge} H(W(V) \mid W(V^c), W_{k-1}, \ldots, W_n)T \\ &\quad + H(W_k \mid W(V^c), W_{k-1}, W_{k+1}, \ldots, W_n)T \\ &\quad - H(W(V), W_k \mid W(V^c), W_{k-1}, W_{k+1}, \ldots, W_n)T \\ &\quad + I(W_k; W(V^c), W_{k-1} \mid W_{k+1}, \ldots, W_n)T\end{aligned} \tag{22}$$

$$\overset{(c)}{=} I(W_k; W(V^c), W_{k-1} \mid W_{k+1}, \ldots, W_n)T \tag{23}$$

$$\overset{(d)}{\ge} I(W_k; W_{k-1} \mid W_{k+1}, \ldots, W_n)T, \tag{24}$$

where (*a*) follows from $I(X;Y) = H(X) - H(X \mid Y)$, (*b*) follows from equation (16), (*c*) follows from $H(X) + H(Y \mid X) = H(X,Y)$ and (*d*) follows from the chain rule of mutual information.

Notice that expression (24) is for the case when $V = \{1, \ldots, k-2\}$, which is the tightest bound. Thus, d_k^{k-1} is at most $I(W_k; W_{k-1} \mid W_{k+1}, \ldots, W_n)T$, and this completes the proof.

There are two remarks. First, after transferring $I(W_k; W_{k-1} \mid W_{k+1}, \ldots, W_n)T$ bits from $\mathcal{B}_k^*$ to $\mathcal{B}_{k-1}^*$, $\mathcal{B}_k$ becomes

$$\begin{aligned}&H(W_k \mid W_{k+1}, \ldots, W_n)T - I(W_k; W_{k-1} \mid W_{k+1}, \ldots, W_n)T \\ &= H(W_k \mid W_{k-1}, W_{k+1}, \ldots, W_n)T,\end{aligned} \tag{25}$$

and $\mathcal{B}_{k-1}$ becomes

$$H(W_{k-1} \mid W_{k+1}, \ldots, W_n)T, \tag{26}$$

which means that $\mathcal{B}_k^*$ and $\mathcal{B}_{k-1}^*$ have switched their positions in solution (15). If $\mathcal{B}_{k-1}$ is full and node *k* still consumes the highest energy, we need to further transfer bits from $\mathcal{B}_k$ to $\mathcal{B}^*_{k-2}$, which will result in that $\mathcal{B}^*_k$ and $\mathcal{B}^*_{k-2}$ switch positions, and so on until $\mathcal{B}^*_k$ has been moved to the first place.

Second, bits that can be reduced at most from $\mathcal{B}^*_k$ are

$$\begin{aligned} &H(W_k \mid W_{k+1},\dots,W_n)T - \\ &H(W_k \mid W_1,\dots,W_{k-1},W_{k+1},\dots,W_n)T \\ &= I(W_k;W_1,\dots,W_{k-1} \mid W_{k+1},\dots,W_n)T, \end{aligned} \tag{27}$$

which exactly equals to $d_k^{k-1} + d_k^{k-2} + \dots + d_k^1$ by the chain rule.

Lemma 2: *Even if* $\mathcal{B}_k$ *has been reduced to the bound,* $H(W_k \mid W_1,\dots,W_{k-1},W_{k+1},\dots,W_n)T$ *,* $\mathcal{B}_{k+1},\dots,\mathcal{B}_n$ *can remain the same and satisfy Slepian–Wolf conditions.*

Proof: Among all the summations between $\mathcal{B}^*_k$ and any subset of $\{\mathcal{B}^*_{k+1},\dots,\mathcal{B}^*_n\}$, $\mathcal{B}^*_k + \mathcal{B}^*_{k+1}$ is the closest to the bound. Thus, if $\mathcal{B}^*_k$ is reduced the most, $\mathcal{B}_k + \mathcal{B}^*_{k+1}$ is still above the bound, so are the remaining summations.

$$\begin{aligned} &\mathcal{B}_{k\,min} + \mathcal{B}^*_{k+1} \\ &= H(W_k \mid W_1,\dots,W_{k-1},W_{k+1},\dots,W_n)T + \\ &\quad H(W_{k+1} \mid W_{k+2},\dots,W_n)T \\ &\overset{(a)}{\geq} H(W_k \mid W_1,\dots,W_{k-1},W_{k+1},\dots,W_n)T + \\ &\quad H(W_{k+1} \mid W_1,\dots,W_{k-1},W_{k+2},\dots,W_n)T \\ &\overset{(b)}{=} H(W_k,W_{k+1} \mid W_1,\dots,W_{k-1},W_{k+2},\dots,W_n)T, \end{aligned} \tag{28}$$

where (*a*) follows from conditioning reduces entropy and (*b*) follows from the chain rule of entropy. Since equation (28) is the Slepian–Wolf bound, the proof is complete.

The above explanations can be summarised in the procedures of Algorithm 1.

Notice that when there are multiple nodes consuming the highest energy, we prefer to check the one with the smallest index first. This is because of the chance that this index is 1. In this case, reducing bits of other nodes will only consume more overall energy without prolonging lifetime since the first node is always the network's bottleneck.

The value of δ is chosen based on equations (29)–(31), where the first and second terms are bits that $\mathcal{B}_{k_1}$ can be reduced at most until this sensor's energy consumption is the same as the second highest one, and the third term is the total reduced bits when $\mathcal{B}_{k_1}$ can be decreased to the Slepian–Wolf bound.

The outcome derived from the network lifetime maximisation algorithm is as the following. The most energy consumption node, i.e. the first node that fails, is either the one that has the least bits to transmit, i.e. $\mathcal{B}_i = H\left(W_i \middle| W_1,\dots,W_{i-1},W_{i+1},\dots,W_n\right)T$, or the first node. In either case, the network lifetime has achieved maximum.

Algorithm 1: Algorithm to maximise network lifetime

1 Calculate the energy consumption for each node $i = 1,2,\dots,n$ to find the highest one, whose index is denoted by k_1. If there are multiple nodes consuming highest energy, let k_1 be the smallest index. Let k_2 denote the node index with the second highest energy consumption;

2 if $k_1 = 1$, stop since $\mathcal{B}^*_1$ is already the smallest and cannot be reduced anymore;

3 else if $k_1 \neq 1$, and

$\mathcal{B}_{k_1} > H(W_{k_1} \mid W_{k_1-1},W_{k_1+1},\dots,W_n)T,$

let

$$\begin{aligned} \delta = \min\{&\mathcal{B}_{k_1} - \frac{a_{k_2}}{a_{k_1}}\mathcal{B}_{k_2}, \frac{a_{k_1}\mathcal{B}_{k_1} - a_{k_1-1}\mathcal{B}_{k_1-1}}{a_{k_1} + a_{k_1-1}}, \\ &\mathcal{B}_{k_1} - H(W_{k_1} \mid W_{k_1-1},W_{k_1+1},\dots,W_n)T\}, \end{aligned} \tag{29}$$

transfer δ bits from $\mathcal{B}_{k_1}$ to $\mathcal{B}_{k_1-1}$;

4 else if

$H(W_{k_1} \mid W_{k_1-2},W_{k_1-1},W_{k_1+1},\dots,W_n)T$
$< \mathcal{B}_{k_1} \leq H(W_{k_1} \mid W_{k_1-1},W_{k_1+1},\dots,W_n)T,$

let

$$\begin{aligned} \delta = \min\{&\mathcal{B}_{k_1} - \frac{a_{k_2}}{a_{k_1}}\mathcal{B}_{k_2}, \frac{a_{k_1}\mathcal{B}_{k_1} - a_{k_1-2}\mathcal{B}_{k_1-2}}{a_{k_1} + a_{k_1-2}}, \\ &\mathcal{B}_{k_1} - H(W_{k_1} \mid W_{k_1-2},W_{k_1-1},W_{k_1+1},\dots,W_n)T\}, \end{aligned} \tag{30}$$

transfer δ from $\mathcal{B}_{k_1}$ to $\mathcal{B}_{k_1-2}$;

⋮

5 else if

$H(W_{k_1} \mid W_1,\dots,W_{k_1-1},W_{k_1+1},\dots,W_n)T$
$< \mathcal{B}_{k_1} \leq H(W_{k_1} \mid W_2,\dots,W_{k_1-1},W_{k_1+1},\dots,W_n)T,$

let

$$\begin{aligned} \delta = \min\{&\mathcal{B}_{k_1} - \frac{a_{k_2}}{a_{k_1}}\mathcal{B}_{k_2}, \frac{a_{k_1}\mathcal{B}_{k_1} - a_1\mathcal{B}_1}{a_{k_1} + a_1}, \\ &\mathcal{B}_{k_1} - H(W_{k_1} \mid W_1,\dots,W_{k_1-1},W_{k_1+1},\dots,W_n)T\}, \end{aligned} \tag{31}$$

transfer δ from $\mathcal{B}_{k_1}$ to $\mathcal{B}_1$;

6 else if

$\mathcal{B}_{k_1} = H(W_{k_1} \mid W_1,\dots,W_{k_1-1},W_{k_1+1},\dots,W_n)T,$

stop since $\mathcal{B}_{k_1}$ is already the smallest;

7 Go to step 1.

7 Numerical results

The improvement of the overall energy consumption and the network lifetime by our proposed schemes depends significantly on the geometric and initial information distribution among all the sensors.[5] In this section, we consider two specific sensor network scenarios, sensors distributed on a line and in a square area to illustrate the optimality of our proposed scheme. First, suppose all the sensor nodes are uniformly distributed on a line and the sink node is at one end. Suppose the data collected by each sensor is composed of two parts: $W_i = W + Z_i$, where $W \sim \mathcal{N}(0, N)$ is the common information and $Z_i \sim \mathcal{N}(0, N_i)$ is the noise, where $N_i = d_i^2$.

We compare the results obtained from our proposed scheme with an average bit allocation scheme, where each sensor transmits its unique information and $1/n$ of the common information. Figure 3 shows the overall energy consumption of the proposed scheme and the average bit allocation scheme. It can be seen that the proposed scheme always achieves lower overall energy consumption. When the number of nodes is 50, the proposed scheme is approximately 30% more energy efficient.

Figure 3 Comparison of the scheme proposed in Theorem 3 and the average bit allocation scheme in terms of the overall energy consumption when sensors are deployed on a line

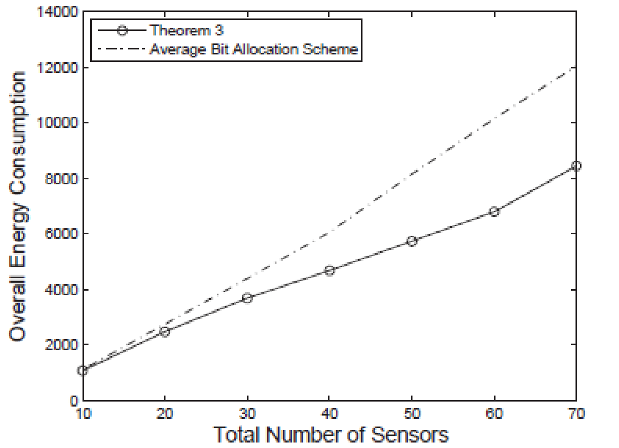


Figure 4 compares the peak energy consumption of Algorithm 6 and the average bit allocation scheme. It can be seen that the proposed algorithm effectively reduces the maximum energy consumption and hence prolongs the network life time. When the number of nodes is 50, the proposed algorithm reduces the peak energy consumption by approximately 20%.

Second, consider the WSN environment where sensors are uniformly deployed in a unit area and the sink node is located at the centre. The area is relatively small such that all the measurements encounter the same disturbance. Specifically, the data collected by each sensor is $W_i = W + Z$, where $W \sim \mathcal{N}(0, N)$ is the common information and $Z \sim \mathcal{N}(0, N_1)$ is the noise. Comparing with the average bit allocation scheme, the results are depicted in Figures 5 and 6. It can be seen that our proposed scheme and algorithm always outperform the average bit allocation scheme significantly.

Figure 4 Comparison of the maximum energy consumption of Algorithm 1 and the average bit allocation scheme when sensors are deployed on a line

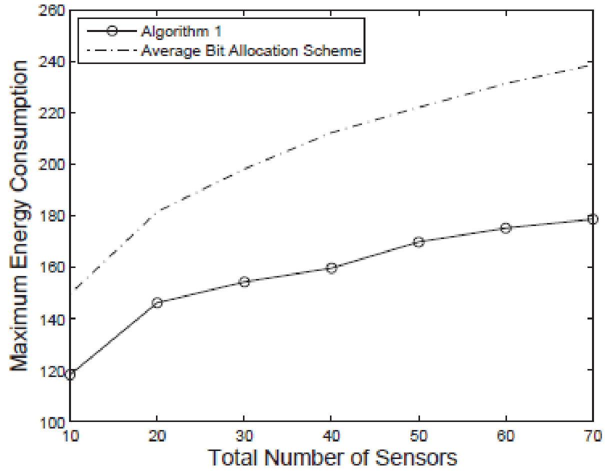


Figure 5 Comparison of the scheme proposed in Theorem 3 and the average bit allocation scheme in terms of the overall energy consumption when sensors are deployed in a unit area

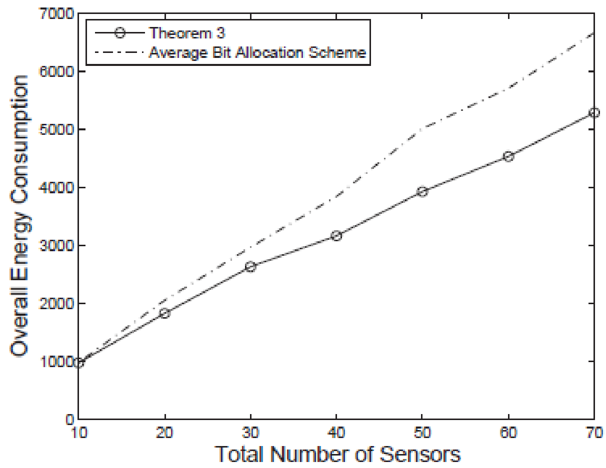
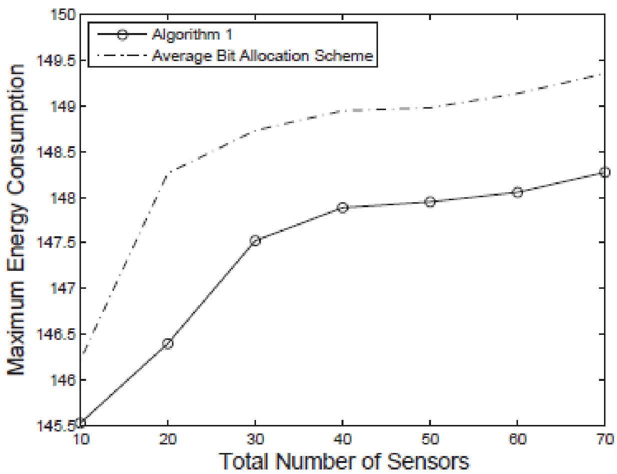


Figure 6 Comparison of the maximum energy consumption of Algorithm 1 and the average bit allocation scheme when sensors are deployed in a unit area



8 Conclusions

In this paper, we have studied the energy efficiency issue for multi-source single-sink WSNs. For transmission part, we have proved that employing minimum transmitting power for a single sensor is optimal in terms of energy efficiency. In the multi-source scenario, transmission in the fashion of TDMA among all sources is most energy efficient. Exploiting source correlation, by Slepian–Wolf coding, we have provided a closed form bit allocation scheme and a heuristic algorithm for minimising the overall energy consumption and for maximising the network lifetime, respectively.

References

Akyildiz, I.F., Su, W., Sankarasubramaniam, Y. and Cayirci, E. (2002) 'Wireless sensor networks: a survey', *Computer Networks*, Vol. 38, No. 4, pp.393-422.

Cheng, P., Cao, X., Chen, J., Cao, K., Sun, Y. and Shen, X. (2011) 'A data processing and analysis testbed for WSNs: design and implementation', *3rd International Conference on Communications and Mobile Computing (CMC)*, 18–20 April, Qingdao, pp.475-480.

Cook, D. and Das, S. (2005) *Smart Environments: Technologies, Protocols, and Applications*, Wiley-Interscience, Hoboken, New Jersey.

Cover, T.M. and Thomas, J.A. (1991) *Elements of Information Theory*, Wiley, New York.

Cristescu, R., Beferull-Lozano, B. and Vetterli, M. (2004) 'On network correlated data gathering', *23rd AnnualJoint Conference of the IEEE Computer and Communications Societies, INFOCOM 2004*, 7–11 March, Vol. 4, pp.2571–2582.

Dietrich, I. and Dressler, F. (2009) 'On the lifetime of wireless sensor networks', *ACM Transactions on Sensor Networks (TOSN)*, Vol. 5, No. 1, pp.1-39.

Ehrgott, M. (2005) *Multicriteria Optimization*, Springer, Verlag.

Gastpar, M. and Vetterli, M. (2003) 'Source-channel communication in sensor networks', *Lecture Notes in Computer Science*, Springer, pp.162-177.

Hou, Y., Shi, Y. and Sherali, H. (2008) 'Rate allocation and network lifetime problems for wireless sensor networks', *IEEE/ACM Transactions on Networking*, Vol. 16, No. 2, pp.321-334.

Knaian, A. (2000) *A Wireless Sensor Network for Smart Roadbeds and Intelligent Transportation Systems*, Master's Thesis, MIT.

Knopp, R. and Humblet, P. (1995) 'Information capacity and power control in single-cell multiuser communications', *IEEE International Conference on Communications--Gateway to Globalization: ICC'95 Seattle*, 18–22 June, Seattle, WA, USA, Vol. 1, pp.331–335.

Li, M. and Liu, Y. (2009) 'Underground coal mine monitoring with wireless sensor networks', *ACM Transactions on Sensor Networks (TOSN)*, Vol. 5, No. 2, p.10.

Lynch, J. and Loh, K. (2006) 'A summary review of wireless sensors and sensor networks for structural health monitoring', *Shock and Vibration Digest*, Vol. 38, No. 2, pp.91-130.

Mainwaring, A., Culler, D., Polastre, J., Szewczyk, R. and Anderson, J. (2002) 'Wireless sensor networks for habitat monitoring', *WSNA'02: Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications*, ACM, New York, NY, USA, pp.88-97.

Mergen, G., Zhao, Q. and Tong, L. (2006) 'Sensor networks with mobile access: energy and capacity considerations', *IEEE Transactions on Communications*, Vol. 54, No. 11, pp.2033-2044.

Patel, S., Lorincz, K., Hughes, R., Huggins, N., Growdon, J., Standaert, D., Dy, J., Welsh, M. and Bonato, P. (2008) 'A body sensor network to monitor parkinsonian symptoms: extracting features on the nodes', *5th International Workshop on Wearable Micro and Nanosystems for Personalised Health, pHealth2008*, pp.21-23.

Roumy, A. and Gesbert, D. (2007) 'Optimal matching in wireless sensor networks', *IEEE Journal of Selected Topics in Signal Processing*, Vol. 1, No. 4, pp.725-735.

Slepian, D. and Wolf, J. (1973) 'Noiseless coding of correlated information sources', *IEEE Transactions on Information Theory*, Vol. 19, No. 4, pp.471-480.

Zhang, X., Xie, L. and Shen, X. (2010) 'An energy-efficient bit allocation scheme in wireless sensor networks', *Proceedings of IEEE Globecom'10*, 6–10 December, Miami, Florida, USA, pp.1–5.

Notes

1 Part of this work was presented at IEEE Globecom'10 (Zhang et al., 2010).

2 For the multiple access channels, it is well known that CDMA is optimal in achieving the maximum rates with fixed power (Cover and Thomas, 1991). However, we are interested in the problem of sending maximum bits with fixed energy consumption, and it is shown that TDMA outperforms CDMA.

3 Node order is based on the index.

4 Here, $\mathcal{B}_l$ is 'full' means that it achieves its largest value without violating any Slepian–Wolf condition.

5 However, this does not impair the optimality of our proposed schemes.