MAXIMIZING THE LIFETIME OF CLUSTERS WITH SLEPIAN-WOLF CODING

Tianqi Wang, Wendi Heinzelman, Alireza Seyedi, and Azadeh Vosoughi

Electrical and Computer Engineering Department, University of Rochester

ABSTRACT

In this paper, we propose an iteration-free algorithm to find the optimal configuration, including transmit power and source coding rates, to maximize the lifetime of a cluster utilizing Slepian-Wolf source coding of data sent to a fusion center. Exact closed form solutions are derived when the fusion center is not energy constrained. When the fusion center is energy constrained, a near optimum solution is provided. Numerical results demonstrating the performance of the proposed algorithms are also provided.

Index Terms— Distributed source coding, Wireless sensor networks, Energy efficiency, Lifetime

I. INTRODUCTION

In general, there are two ways to reduce energy consumption in wireless sensor networks (WSNs): efficient communication and data compression. Among the data compression schemes, distributed source coding (DSC) in particular has great potential in WSNs [1]-[3]. DSC exploits the spatial correlation commonly found in sensed signals. With only the knowledge of the spatial correlation between their signals, the sensor nodes can compress their data using DSC without communicating each other.

Slepian-Wolf coding, which is concerned with lossless DSC of two correlated discrete sources, has recently been utilized to address the energy efficient gathering of correlated data in WSNs using mathematical optimization techniques [2][4][5]. Researchers have abstracted the wireless media as a graph with fixed cost per information bit, and then jointly optimized the data gathering tree and the rate allocation across the source nodes. In practice, however, different communication constraints (e.g., network topology, energy distribution) can significantly alter the communication cost per information bit. Consequently, when considering minimization of consumed energy, it is essential that the communication techniques are considered in conjunction with the compression techniques.

In this paper, we propose an iteration-free solution to maximize the amount of data gathered during the cluster lifetime with Slepian-Wolf coding. The optimization is conducted on both the source coding and the communication sides. The closed-form solutions for joint optimal transmit powers and Slepian-Wolf coding rates are provided.
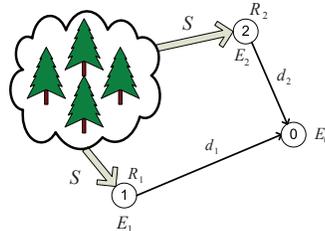
Fig. 1. The 2-node Slepian-Wolf coding scheme

II. PROBLEM DEFINITION

Fig. 1 shows a typical Slepian-Wolf coding scheme with 2 sensor nodes that send data to a fusion center. During the lifetime of this cluster, each sensor senses S samples. The lifetime of the cluster is defined as the time duration elapsed from the beginning to the time when the fusion center or both sensor nodes die.

We denote the source coding rate at node i with R_i. The Slepian-Wolf theorem states that if $R_1 \geq H(X_1|X_2), R_2 \geq H(X_2|X_1), R_1 + R_2 \geq H(X_1, X_2)$, X_1 and X_2 are reconstructible perfectly at the fusion center. Here $H(\cdot)$ represents the entropy function, and X_i is the source random variable (sample) sensed at node i [6].

Assume that sensor node i can transmit for time duration T_i, and it transmits SR_i data bits to the fusion center. Since both sensors observe the same number of samples, we have

$$BT_i \log(1 + \frac{P_i}{d_i^n P_N}) \geq SR_i, \qquad (1)$$

where B is the bandwidth, P_i is the transmit power of node i, d_i is the transmission distance from node i to the fusion center, n is the path loss exponent and P_N is the additive white gaussian noise (AWGN) power. The inequality of the above equation comes from Shannon's channel capacity theorem [6]. Therefore, the transmit power of node i is bounded by

$$P_i \geq (2^{\frac{SR_i}{BT_i}} - 1)d_i^n P_N. \qquad (2)$$

That is, the lower bound on the transmit power of node i is determined by its time duration T_i and SR_i.

The energy constraints at the sensors are $T_i(P_i + P_{CT}) \leq E_i$, where E_i denotes the residual energy at node i and P_{CT} is the circuit power consumption at the sensor nodes. To facilitate our analysis, we utilize (2) to relax the energy

constraints. That is, we assume the lower bounds on transmit powers are achievable. Therefore, the sensor nodes' energy constraints become $T_i[(2^{\frac{SR_i}{BT_i}} - 1)d_i^n P_N + P_{CT}] \leq E_i$. The energy constraint at the fusion center is $P_{CR}(T_1 + T_2) \leq E_0$, where E_0 is the energy available at the fusion center. P_{CR} is the circuit power consumption at the fusion center and represents the power consumption for receiving data.

Under the energy and Slepian-Wolf coding constraints, the maximization of the total gathered samples during the lifetime of a cluster can be modeled as

$$
\begin{aligned}
\min \quad & -S \\
\text{s.t.} \quad & C_0: \ T_i > 0, S > 0, \\
& C_1: \ P_{CR}(T_1 + T_2) \leq E_0, \\
& C_2: \ T_i[(2^{\frac{SR_i}{BT_i}} - 1)d_i^n P_N + P_{CT}] \leq E_i, \\
& C_3: \ R_1 \geq H(X_1|X_2), \\
& C_4: \ R_2 \geq H(X_2|X_1), \\
& C_5: \ R_1 + R_2 \geq H(X_1, X_2),
\end{aligned}
\tag{3}
$$

where C_1 and C_2 are energy constraints at the fusion center and sensor nodes, respectively, and C_3, C_4 and C_5 are Slepian-Wolf rate constraints. Our goal is, for a given residual energy and communication environment, to find the optimum transmission time T_i, coding rates R_i, and transmit powers P_i that maximize S.

III. MAXIMIZING LIFETIME WITHOUT ENERGY CONSTRAINT AT THE FUSION CENTER

The number of samples S is a monotonically decreasing function of R_1 and R_2. Therefore, R_1^* and R_2^* should lie on the boundary $R_1 + R_2 = H(X_1, X_2)$. Thus, we always have $R_1^* + R_2^* = H(X_1, X_2)$. This observation implies that constraint C_5 is always active. When there is no energy constraint at the fusion center, the sensor nodes are then free to choose the most energy-efficient transmission method by only considering constraints $C_2 - C_5$. Let us define

$$
D_i^* = \max \left\{ T_i \log_2(1 + \frac{E_i - P_{CT}T_i}{T_i d_i^n P_N}) \right\}.
\tag{4}
$$

Eq. (4) represents the unconstrained maximum information bits that node i can send, given that both sensors use up their energy.

Proposition 1. Without loss of generality, let us assume D_{1*} is acheivable, then

$$
\begin{aligned}
R_1^* &= \max \left\{ H(X_1|X_2), \frac{H(X_1,X_2)}{1 + D_2^*/D_1^*} \right\}, \\
R_2^* &= H(X_1, X_2) - \max \left\{ H(X_1|X_2), \frac{H(X_1,X_2)}{1 + D_2^*/D_1^*} \right\}, \\
T_1^* &= \frac{E_1}{\frac{\ln 2(P_{CT} - d_1^n P_N)}{W\left(\frac{\ln 2}{2d_1^n P_N}(P_{CT} - d_1^n P_N)\right)} + P_{CT} - d_1^n P_N}, \\
T_2^* &= \arg \left\{ T_2 \log_2(1 + \frac{E_2 - P_{CT}T_2}{T_2 d_2^n P_N}) = D_1^* \frac{R_2^*}{R_1^*} \right\},
\end{aligned}
\tag{5}
$$

where $W(\cdot)$ is the Lambert function.

Proof: Assume that node 1 sends $\hat{D}_1 = SR_1^*$ bits, and node 2 sends $\hat{D}_2 = SR_2^*$ bits. Then we have

$$
\begin{aligned}
\hat{D}_1 + \hat{D}_2 &= S(R_1^* + R_2^*) \\
\overset{a}{\Longrightarrow} S &= (\hat{D}_1 + \hat{D}_2)/H(X_1, X_2), \\
\overset{b}{\Longrightarrow} S &= \frac{\hat{D}_1(1 + \frac{R_2^*}{R_1^*})}{H(X_1, X_2)} = \frac{\hat{D}_2(1 + \frac{R_1^*}{R_2^*})}{H(X_1, X_2)},
\end{aligned}
\tag{6}
$$

Step (a) is from the fact that $R_1^* + R_2^* = H(X_1, X_2)$. Step (b) is from the fact that $\hat{D}_1/\hat{D}_2 = R_1^*/R_2^*$. Step (a) implies, in the case where there is no fusion center energy constraint, we need to maximize $\hat{D}_1 + \hat{D}_2$, in order to maximize the number of gathered samples. Since there is no constraint from the fusion center side, at least one of $\hat{D}_1$ and $\hat{D}_2$ should be the unconstrained maximum value from (4). Without loss of generality, we assume $\hat{D}_1 = D_1^*$ and $\hat{D}_2 \leq D_2^*$.

When $\hat{D}_1 = D_1^*$ and $\hat{D}_2 \leq D_2^*$, from step (b), we have $S = D_1^*(1 + \frac{R_2^*}{R_1^*})/H(X_1, X_2)$. To maximize S, R_2^*/R_1^* needs to be the largest possible value. That is R_1^* has to be the smallest feasible value and R_2^* is the largest feasible value. We have

$$
\begin{aligned}
D_1^* \frac{R_2^*}{R_1^*} \leq D_2^* &\implies \frac{H(X_1, X_2) - R_1^*}{R_1^*} \leq \frac{D_2^*}{D_1^*}, \\
&\implies R_1^* \geq \frac{H(X_1, X_2)}{1 + D_2^*/D_1^*}.
\end{aligned}
\tag{7}
$$

Also considering the constraint $R_1^* \geq H(X_1|X_2)$, we have $R_1^* = \max \left\{ H(X_1|X_2), \frac{H(X_1, X_2)}{1 + D_2^*/D_1^*} \right\}$, and $R_2^* = H(X_1, X_2) - R_1^*$.

When node 1 sends D_1^*, the optimal T_1^* can be found from (4). The corresponding transmit time duration T_2^*, however, is not unique. One possible analytical solution of T_2^* is

$$
T_2^* = \arg \left\{ T_2 \log_2(1 + \frac{E_2 - P_{CT}T_2}{T_2 d_2^n P_N}) - D_1^* \frac{R_2^*}{R_1^*} = 0 \right\}. \square
$$

Proposition 1 gives the exact optimal solution of $T_1, T_2, R_1,$ and R_2 when there is no energy constraint at the fusion center. The transmit powers can be obtained by assuming both sensor nodes use up their energy. Since the optimal solution in this case is not unique, this assumption does not affect the optimality of the solution provided by proposition 1.

IV. MAXIMIZING LIFETIME WITH ENERGY CONSTRAINT AT THE FUSION CENTER

When the fusion center is deployed in the field with limited energy supply, based on the amount of available energy at the fusion center, there are two possibilities.

If the residual energy at the fusion center is larger than the critical value $\hat{E}_0 = P_{CR}(T_1^* + T_2^*)$, where T_1^* and T_2^* are given by (5), then this scenario is equivalent to the case when there is no energy constraint at the fusion center, as considered in Section III.

On the other hand, if $E_0 < \hat{E}_0$, the total transmission time duration of the sensor nodes is limited by the fusion center energy constraint. Due to the limited transmission time budget, the sensor nodes have to send their data in

a non-optimal way. As the available energy at the fusion center increases, thereby the total transmission time budget increases, the sensor nodes will be able to choose a more efficient transmission method. Correspondingly, the gathered data at the fusion center will increase. In this section, we focus on the optimization of a cluster where the fusion center energy constraint is the limiting factor of the cluster, or a *fusion-center-limited cluster*.

Given a limited transmission time budget, to send as much data as possible, the sensor nodes will use as much transmit power as possible. Therefore, both sensor nodes and the fusion center will use up their energy. This observation leads us to the following propositions:

Proposition 2. In fusion-center-limited clusters, we have

1) If $\frac{E_1}{d_1^n} > \frac{E_2}{d_2^n}$, then $R_1^* = H(X_1)$, and $R_2^* = H(X_2|X_1)$.

2) If $\frac{E_1}{d_1^n} < \frac{E_2}{d_2^n}$, then $R_1^* = H(X_1|X_2)$, and $R_2^* = H(X_2)$.

3) If $\frac{E_1}{d_1^n} = \frac{E_2}{d_2^n}$, R_1^* and R_2^* can be any points on the line segment $R_1 + R_2 = H(X_1, X_2), R_1 > H(X_1|X_2), R_2 > H(X_2|X_1)$.

Proof: Since the fusion center is the limiting factor, the fusion center and both sensor nodes should use up their energy. Therefore, both the fusion center and the sensor nodes' energy constraints are active:

$$P_{CR}(T_1 + T_2) = E_0, \tag{8}$$

$$T_i[(2^{\frac{SR_i}{BT_i}} - 1)d_i^n P_N + P_{CT}] = E_i. \tag{9}$$

From (9), we have

$$\log_2\left(\frac{\frac{E_i}{T_i} - P_{CT}}{d_i^n P_N} + 1\right) = \frac{SR_i}{BT_i},$$
$$\overset{a}{\Longrightarrow} T_i \log_2\left(\frac{E_i}{T_i d_i^n P_N}\right) \approx \frac{SR_i}{B},$$
$$\Longrightarrow T_i \log_2\left(\frac{E_i}{d_i^n P_N}\right) - T_i \log_2(T_i) \approx \frac{SR_i}{B}, \tag{10}$$
$$\overset{b}{\Longrightarrow} T_i \approx \frac{SR_i/B}{\log_2\left(\frac{E_i}{d_i^n P_N}\right) + \frac{6}{\ln 2}},$$

where step (a) assumes that the signal to noise ratio (SNR) is much greater than 1, and $E_i \gg P_{CT}$. Step (b) employs the approximation $T_i \log_2(T_i) \approx -T_i \frac{6}{\ln 2}$, which becomes accurate when $T_i \to 0$. Define $U_i = \log_2\left(\frac{E_i}{d_i^n P_N}\right) + \frac{6}{\ln 2}$, then we have $T_i = \frac{SR_i}{BU_i}$. Correspondingly, we have

$$T_1 + T_2 = \frac{S}{B}\left(\frac{R_1}{U_1} + \frac{R_2}{U_2}\right), \tag{11}$$

$$\frac{T_1}{T_2} = \frac{R_1 U_2}{R_2 U_1}. \tag{12}$$

The fusion center energy constraint implies

$$P_{CR}(T_1 + T_2) = E_0. \tag{13}$$

Solving equations (11) and (13) for S, we have

$$S = \frac{BE_0}{P_{CR}}\left(\frac{R_1}{U_1} + \frac{R_2}{U_2}\right)^{-1}. \tag{14}$$

Maximizing (14) under the constraint $R_1 + R_2 = H(X_1, X_2)$, we directly have the following results: when $U_1 > U_2$, then $R_1^* = H(X_1)$, and $R_2^* = H(X_2|X_1)$; when $U_1 < U_2$, then $R_1^* = H(X_1|X_2)$, and $R_2^* = H(X_2)$; when $U_1 = U_2$, R_1^* and R_2^* can be any points on the line segment $R_1 + R_2 = H(X_1, X_2), R_1 \geq H(X_1|X_2), R_2 \geq H(X_2|X_1)$. $\square$

Proposition 3. In fusion-center-limited clusters, the closed-form near-optimal transmit time durations are:

$$\begin{aligned} T_1^* &= \frac{E_0}{P_{CR}}\frac{R_1^* U_2}{R_1^* U_2 + R_2^* U_1}, \\ T_2^* &= \frac{E_0}{P_{CR}}\frac{R_2^* U_1}{R_2^* U_1 + R_1^* U_2}, \end{aligned} \tag{15}$$

where R_1^* and R_2^* are determined from proposition 2.

Proof: By solving equations (12) and (13) for T_1 and T_2, we directly have proposition 3. $\square$

V. NUMERICAL RESULTS

We assume that two sensor nodes are uniformly placed within a disk, with a radius of 100m, centered at the fusion center. The path loss exponent is $n = 4$. Also, we set $H(X_1) = H(X_2) = 1$ and $H(X_1|X_2) = H(X_2|X_1) = h$. We compare the following cases: (1) $R_1 = H(X_1), R_2 = H(X_2)$ (No Slepian-Wolf coding), optimized T_i; (2) $R_1 = H(X_1|X_2), R_2 = H(X_2)$, optimized T_i; (3) $R_1 = R_2 = H(X_1, X_2)/2$, optimized T_i; (4) Jointly optimized T_i and R_i (analytical results P.1, P.2 and P.3.); (5) Jointly optimized T_i and R_i (numerical).
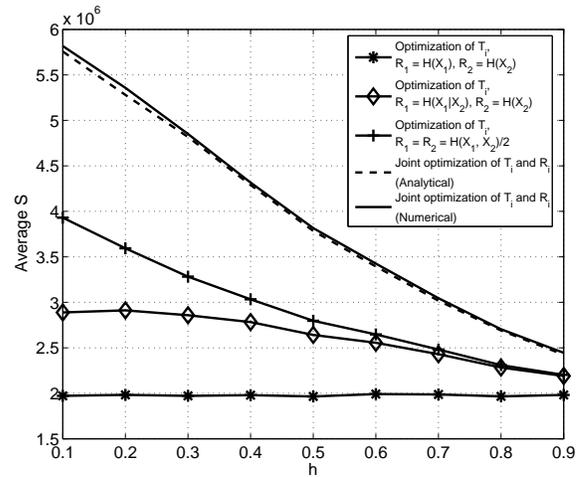
Fig. 2. Comparison of the performance in clusters with sufficient energy at the fusion center as a function of conditional entropy.

First let us consider the case where the fusion center is not the limiting factor. We fix the sensor nodes' residual energy $E_1 = E_2 = 0.1$ J, and the fusion center residual energy $E_0 = 100$ J. In Fig. 2, we compare the average maximum number of samples with various conditional entropies in the above cases. Case 1, where no Slepian-Wolf coding

is applied, performs the worst. Case 3 outperforms case 2 because the evenly distributed burden in case 3 avoids the possible early drain of energy in case 2 where there is always one fixed node sending more data than the other. However, the performance gap between cases 1, 2, and 3 diminishes as h increases, since as h increases, the difference between $H(X_1), H(X_2), H(X_1|X_2), H(X_2|X_1)$ and $H(X_1, X_2)/2$ becomes smaller.

The numerical optimization of T_1, T_2, R_1, and R_2 provides the best performance. However, an iterative numerical optimization in the $\mathbb{R}^4$ domain is exceedingly complicated in most WSN applications, where energy is a limited resource. Moreover, when the distribution of the source and/or the profile of the channel are time-varying, periodic updates on T_1, T_2, R_1, and R_2 make iterative numerical optimizations even more costly. The analytical results in (5) allow us to jointly optimize T_1, T_2, R_1 and R_2 at a considerably lower computational cost.
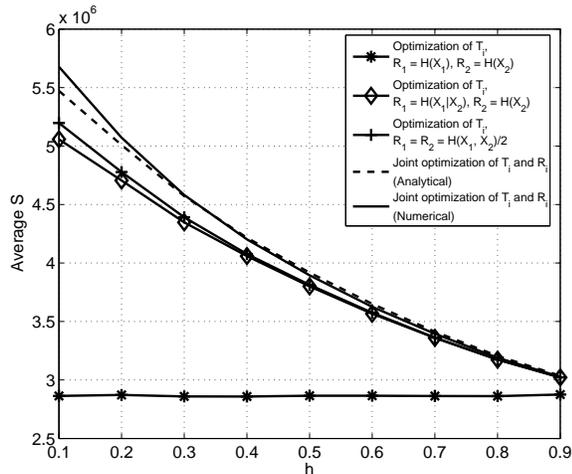
Fig. 3. Comparison of the performance in fusion-center-limited clusters.

To look at the fusion-center-limited case, we assume that the sensor nodes' residual energy are $E_1 = E_2 = 100$ J, and the fusion center residual energy $E_0 = 0.1$ J. Fig. 3 shows the comparison of the performance of the five cases. The analytical results provide a performance that is almost as good as the numerical optimizations, and there is a performance gain over the fixed source coding rate cases (cases 1-3), especially at small h where X_1 and X_2 are highly correlated. Our analytical approximations provide a near-optimal performance in a fusion-center-limited cluster.

Comparing Fig. 3 and Fig. 2, the optimization of R_1 and R_2 in in clusters with no fusion center energy constraint is more important than in fusion-center-limited clusters. As shown in Fig. 2, in clusters with no fusion center energy constraint, the analytical optimization provides a gain of as much as 101% on the gathered data over fixed-rate Slepian-Wolf coding cases (cases 2 and 3); while in fusion-center-limited

clusters, as shown in Fig. 3 the most gain that the analytical optimization can achieve is merely 8%. This is because, in clusters with no fusion center energy constraint, the total available transmission time duration is determined by the sensor nodes, and an inefficient Slepian-Wolf coding scheme would accelerate the drain of energy at the sensor nodes and thereby reduce the total available transmission time duration. However, in fusion-center-limited clusters, the total available transmission time duration is mainly determined by the fusion center energy storage, and it is not affected by the adopted Slepian-Wolf coding scheme as much as in clusters with no fusion center energy constraint. Thus, when the sensor nodes' energy is limited while the fusion center energy storage is abundant, an optimized distributed source coding scheme becomes crucial.

VI. CONCLUSIONS

In this paper, we provide analytical results to maximize the lifetime of a 2-source Slepian-Wolf coding cluster with energy constraints and Slepian-Wolf coding constraints. The optimization parameters are source coding rates and transmit durations of both source nodes. The optimal transmit powers of source nodes are also found. For the case where there is no active energy constraint at the fusion center, we have found an exact optimal solution. For the case where there is an energy constraint at the fusion center, near-optimal solutions are obtained. The optimization in the case where there is no energy constraint at the fusion center results in a very significant gain in lifetime.

VII. REFERENCES

[1] A. Majumdar, K. Ramachandran, and I. Kozintsev, "Distributed Coding for Wireless Audio Sensors," *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, Oct. 2003.

[2] J. Chou, D. Petrovic, and K. Ramachandran, "A Distributed and Adaptive Signal Processing Approach to Reducing Energy Consumption in Sensor Networks," *Proceedings of the IEEE Twenty-Second International Annual Joint Conference (INFO-COM)*, 2003.

[3] F. Oldewurtel, M. Foks, and P. Mahonen, "On a Practical Distributed Source Coding Scheme for Wireless Sensor Networks," *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, May 2008.

[4] Stephen P. Boyd and Lieven Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, first edition, 2004.

[5] R. Cristescu, B. Beferull-Lozano, and M. Vetterli, "Networked Slepian-Wolf : Theory, Algorithms, and Scaling Laws," *IEEE Trans. on Information Theory*, vol. 51, no. 12, pp. 4057–4073, Dec. 2005.

[6] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*, Wiley-Interscience, NJ, second edition, 2006.