Maximizing Sample Rate for Distributed Source Coding over Multiple Access Channels

Tianqi Wang, Wendi Heinzelman, Alireza Seyedi, and Azadeh Vosoughi
Department of Electrical and Computer Engineering, University of Rochester
Email: {tiwang,wheinzel,alireza, azadeh}@ece.rochester.edu

Abstract—We investigate the problem of maximizing the sample rate at the transmitters in a wireless network using distributed source coding (DSC) and lossless transmission. With the consideration of circuit power consumption, average power constraints and DSC rate constraints, we study different orthogonal multiple access channels (MACs) and prove that the optimal orthogonal MAC is either TDMA or CDMA, depending on the joint entropy of the correlated sources, average power constraints, circuit power consumption, noise power and path loss.

I. INTRODUCTION

In wireless sensor networks (WSNs), sensors are usually densely deployed, and the information gathered by adjacent nodes is highly correlated in the space domain [1], [2]. Efficiently exploiting this correlation between the observed sources is critical in improving the energy efficiency of WSNs [3], [4]. As a compression technique for correlated random variables, distributed source coding (DSC) can be directly applied in WSNs. Given the knowledge of the joint distribution of correlated sources at the nodes, DSC enables distributed compression of these sources without requiring direct communication between the nodes [5]. The collaborative nature of WSNs greatly facilitates the adoption of DSC.

In terms of providing the highest sum rate, the optimality of orthogonal multiple access channels (MAC), including code division multiple access (CDMA), time division multiple access (TDMA) and frequency division multiple access (FDMA), has been well studied in the literature [6], [7]. In this paper, we will show that neither CDMA nor TDMA can guarantee optimality in terms of supporting the maximum sum rate at the source nodes when considering DSC and circuit power consumption; rather, depending on the system parameters, either CDMA or TDMA may provide the best performance. While joint source-channel coding with correlated sources has also been studied in the literature [8]–[11], in this work we separate source and channel coding. Though this separation of source and channel coding is not optimal from an information theoretic perspective [6], this is a reasonable assumption because there has been rapid progress over the past several years in constructing practical Slepian-Wolf codes for compression of correlated discrete sources, and in developing conventional capacity achieving channel coding, while constructive low complexity joint source-channel codes for correlated sources is absent. Hence, we consider separate source and channel coding in our work.

In this paper, information describing the environment of interest is measured in samples per second, or sample rate. We say that a WSN has a high performance if the network can support all of its sensors having a high sample rate. Moreover, we assume DSC is used in the sensor network and each source node has the same sample rate. Also, the circuit power of each sensor is viewed as non-negligible. The goal of this work is to find the maximum sample rate at the source nodes that can be supported by a WSN that utilizes lossless Slepian-Wolf coding, when considering different MACs.

II. SYSTEM MODEL

Fig. 1 shows a two-source cluster where the source nodes (nodes 1 and 2) gather information of the environment of interest and send the information to the fusion center (node 0). Both source nodes sample the environment at a sample rate of $S$ (samples/second). The observed samples at node $i$ is denoted by $X_i$, which is a discrete random variable where the samples from the two source nodes are correlated with each other. At the source nodes, Slepian-Wolf coding is used, and the coding rate at node $i$ is denoted by $R_i$ (bits/sample). After source coding, the sensor nodes need to perform channel coding based on the resulting bit streams. The resulting random variable after channel coding at node $i$ is denoted by $W_i$, and the rate of the link between node $i$ and the fusion center is denoted by $r_i$ (bits/second). The outputs at the channel decoder and the DSC decoder are denoted by $(W_1, W_2)$ and $(\hat{X}_1, \hat{X}_2)$, respectively. The average power constraint of node $i$ is $P_i$ and the distance between node $i$ and the fusion center is $d_i$. Additive white Gaussian noise (AWGN) channels are assumed in this paper. $Y$ is the received signal at the fusion center and can be expressed by

$$Y = \frac{W_1}{d_1^2} + \frac{W_2}{d_2^2} + Z,$$

(1)
where $Z$ is a Gaussian random variable with zero mean and variance $P_N$. $d_n$ denotes the path loss effect where $n$ is the path loss exponent. Our goal is to maximize the sample rate $S$ at the source nodes given the average power constraints $P_i$.

In this paper the source coding and the channel coding are considered separately, i.e., the codewords for $X_1$ and $X_2$ are independently generated and the sequences after DSC are simply used as indices of the codewords. Thus, to maximize the sample rate, we consider the capacity region of the 2-source MAC and the Slepian-Wolf coding rate region individually [5], [6]. The capacity region of the 2-source MAC is

$$
\begin{align*}
  r_1 &\leq B Z(W_1;Y|W_2), \\
  r_2 &\leq B Z(W_2;Y|W_1), \\
  r_1 + r_2 &\leq B Z(W_1,W_2;Y),
\end{align*}
$$

(2)

where $Z(\cdot,\cdot)$ is the mutual information, and $B$ is the channel bandwidth. The Slepian-Wolf coding rate region is

$$
\begin{align*}
  R_1 &\geq H(X_1|X_2), \\
  R_2 &\geq H(X_2|X_1), \\
  R_1 + R_2 &\geq H(X_1,X_2),
\end{align*}
$$

(3)

where $H(\cdot)$ is the entropy. The channel and source coding rates can be related to each other by

$$
  r_1 = S R_1, r_2 = S R_2.
$$

(4)

From (3) and (4), we have

$$
\begin{align*}
  r_1 &\geq S H(X_1|X_2), \\
  r_2 &\geq S H(X_2|X_1), \\
  r_1 + r_2 &\geq S H(X_1,X_2).
\end{align*}
$$

(5)

Eventually, a general optimization problem for maximizing the sample rate $S$ of the 2-source system under the multiple access channel and the Slepian-Wolf coding rate constraints is

$$
\begin{align*}
 \min & -S \\
 \text{s.t.} & C_1 : r_1 \leq B Z(W_1;Y|W_2), \\
 & r_2 \leq B Z(W_2;Y|W_1), \\
 & r_1 + r_2 \leq B Z(W_1,W_2;Y), \\
 & C_2 : r_1 \geq S H(X_1|X_2), \\
 & r_2 \geq S H(X_2|X_1), \\
 & r_1 + r_2 \geq S H(X_1,X_2),
\end{align*}
$$

(6)

where the constraint $C_1$ is the capacity region of the 2-source MAC and the constraint $C_2$ is the Slepian-Wolf coding rate constraint.

III. OPTIMAL MULTIPLE ACCESS SCHEME

A. The optimality criterion for multiple access channels

We first find the relationship between the sample rate at the source nodes and the sum rate of a MAC by proposition 1.

**Proposition 1** Given any feasible sample rate $S$ in (6), there exists a set of channel rates $r_1$ and $r_2$ such that $(r_1 + r_2)/H(X_1,X_2) = S$.

**Proof:** From the definition (6), we know that any feasible sample rate $S$ is the minimum value of the following sequence

$$
\left\{ \frac{r_1}{H(X_1|X_2)} + \frac{r_2}{H(X_2|X_1)} + \frac{r_1 + r_2}{H(X_1,X_2)} \right\},
$$

(7)

where $r_1$ and $r_2$ are in a convex set defined by (2). First, note that $(r_1 + r_2)/H(X_1,X_2)$ cannot be the largest element of sequence (7). This is because, we have

$$
\frac{r_1 + r_2}{H(X_1,X_2)} = \frac{r_1}{H(X_1|X_2)} \leq \frac{r_2}{H(X_2|X_1)},
$$

(8)

which implies either

$$
\frac{r_1}{H(X_1|X_2)} \leq \frac{r_2}{H(X_2|X_1)} \leq \frac{r_1 + r_2}{H(X_1,X_2)}
$$

or

$$
\frac{r_2}{H(X_2|X_1)} \leq \frac{r_1 + r_2}{H(X_1,X_2)}.
$$

If $r_1/H(X_1|X_2) < (r_1 + r_2)/H(X_1,X_2)$ and $r_2/H(X_2|X_1) < (r_1 + r_2)/H(X_1,X_2)$, then we must have

$$
\frac{r_1 + r_2}{H(X_1,X_2)} \leq \frac{r_2}{H(X_2|X_1)} \leq \frac{r_1}{H(X_1|X_2)},
$$

(9)

which is contradictory to the assumption that $r_2/H(X_2|X_1) < (r_1 + r_2)/H(X_1,X_2)$.

To prove proposition 1, we need to analyze three cases: $(r_1 + r_2)/H(X_1,X_2)$, $r_1/H(X_1,X_2)$, or $r_2/H(X_2|X_1)$ is the minimum of the set in (7), respectively. If $(r_1 + r_2)/H(X_1,X_2)$ is the minimum of the set in (7), proposition 1 is trivial. Otherwise, if

$$
\frac{r_1}{H(X_1|X_2)} \leq \frac{r_2}{H(X_2|X_1)} \leq \frac{r_1 + r_2}{H(X_1,X_2)},
$$

(10)

due to the convexity of the channel rate region and the fact that $(r_1 + r_2)/H(X_1,X_2)$ cannot be the maximum of the set in (7), we can always fix $r_1$ and find a $r_2^* < r_2$ so that

$$
\frac{r_1}{H(X_1|X_2)} = \frac{1}{H(X_1,X_2)} \geq \frac{r_1 + r_2^*}{H(X_2|X_1)} \leq \frac{r_2^*}{H(X_2|X_1)}.
$$

(11)

Thus, proposition 1 is true. Similarly, proposition 1 is also true when $r_2/H(X_2|X_1)$ is the minimum of the set. $\square$

From proposition 1, it is straightforward to see that if $S^*$ is the maximum sample rate of problem (6), then there must be a set of channel coding rates $r_i^*$ such that $\sum_{i=1}^2 r_i^*/H(X_1,X_2) = S^*$. This is better illustrated by Fig. 2. As shown in Fig. 2, the bottom left region is defined by the multiple access channel in (2) and the upper right

![Fig. 2. An illustration of the multiaccess channel rate regions (2 source nodes).](image-url)
region is defined by the Slepian-Wolf region in (3). Given a MAC technique, the region defined by (2) is fixed. Under this circumstance, for a sample rate \( S \) to be feasible, the two regions defined by (2) and (3) must have intersecting area. As \( S \) increases, the region defined by (3) will shift to the upper right direction as shown in the figure. Thus, for a sample rate \( S \) to be optimal (denoted by \( S^* \)), these two regions must be tangent as shown in Fig. 2, and the channel rates \( r^*_1 \) on the tangent point must satisfy \( \sum_{i=1}^{2} r^*_i / H(X_1, X_2) = S^* \).

Based on the Slepian-Wolf coding region, we find that the optimal channel coding rates \( r^*_1 \) and \( r^*_2 \) always have
\[
\frac{H(X_1 | X_2)}{H(X_2)} \leq r^*_1 \leq \frac{H(X_1)}{H(X_2 | X_1)},
\]
(12)
Therefore, a necessary and sufficient condition for a multiple access channel scheme to be optimal in the sense of maximizing the sample rate is that it provides the highest sum rate among all possible multiple access channel techniques within channel rate region \( H(X_1 | X_2) / H(X_2) \leq r_1 / r_2 \leq H(X_1) / H(X_2 | X_1) \).

**B. The optimality of TDMA and CDMA**

After obtaining the above condition, we further evaluate the performance of different MACs. We consider the case when each source node has an average power constraint \( P_i \) and has non-negligible circuit power consumption \( P_{CT} \).

TDMA provides higher sum rate than FDMA when \( P_{CT} > 0 \) in any channel rate region. This can be easily shown as follows. The sum rate of a multiple access channel using TDMA with an average power constraint is
\[
r_{1, \text{TDMA}} + r_{2, \text{TDMA}} = \sum_{i=1}^{2} \theta_i B \log_2 \left( 1 + \frac{P_i}{d_i^n P_N} \right),
\]
(13)
where \( P_i \) is the transmit power of node \( i \) and \( \theta_i \) is the portion of the time resource assigned to node \( i \). The constraints are
\[
\theta_1 + \theta_2 \leq 1, \theta_i (P_i + P_{CT}) \leq P_i.
\]
(14)
On the other hand, the sum rate of a multiple access channel using FDMA is
\[
r_{1, \text{FDMA}} + r_{2, \text{FDMA}} = \sum_{i=1}^{2} \theta_i B \log_2 \left( 1 + \frac{P_i}{d_i^n P_N \theta_i} \right),
\]
(15)
where \( \theta_i \) is the portion of the frequency resource assigned to node \( i \). The constraints are
\[
\theta_1 + \theta_2 \leq 1, (P_i + P_{CT}) \leq P_i.
\]
(16)
Therefore,
\[
\sum_{i=1}^{2} r_{i, \text{FDMA}} = \sum_{i=1}^{2} \theta_i B \log_2 \left( 1 + \frac{P_i}{d_i^n P_N \theta_i} - \frac{P_{CT}}{d_i^n P_N \theta_i} \right) < \sum_{i=1}^{2} \theta_i B \log_2 \left( 1 + \frac{P_i}{d_i^n P_N \theta_i} - \frac{P_{CT}}{d_i^n P_N \theta_i} \right) = \sum_{i=1}^{2} r_{i, \text{TDMA}}.
\]
(17)
That is TDMA can provide higher sum channel rate than FDMA in any channel rate region, i.e., TDMA is always superior to FDMA. Thus, during the following analysis, FDMA is no longer considered a candidate for MAC.

Next we show that based on the system parameters, either CDMA or TDMA may be superior. Although it is proven in [12] that TDMA provides higher maximum sum rate than CDMA when \( P_{CT} > 0 \), TDMA cannot guarantee a higher sum rate than CDMA in any channel rate region. For instance, when using CDMA, we have the maximum channel rate for node 2 is
\[
r_{2, \text{max, CDMA}} = B \log_2 \left( 2 + \frac{P_2}{d_2^n P_N} \right).
\]
Using TDMA, we have the maximum channel rate for node 2 is
\[
r_{2, \text{max, TDMA}} = \max \theta B \log_2 \left( 1 + \frac{P_2 - P_{CT}}{d_2^n P_N} \right),
\]
(18)
the resulting \( \theta \) that maximizes the channel rate for node 2 is
\[
\theta^* = \begin{cases} \alpha(d_2) \bar{P}_2, & \alpha(d_2) \leq \frac{1}{P_2} \\ 1, & \text{otherwise} \end{cases}
\]
(19)
where
\[
\alpha(d_i) = \frac{1}{P_{CT} - d_i^n P_N} - \frac{1}{W(2 \log_2 \left( \frac{d_i^n P_N + P_2 - P_{CT}}{P_{CT}} \right) - 1)},
\]
(20)
and \( W(\cdot) \) is the Lambert function [13]. This is a direct result from setting the first order derivative of the objective function to zero [12]. Thus, when \( \alpha(d_2) > 1/P_2 \), and \( r_1 = 0 \), we have \( r_{2, \text{max, CDMA}} = r_{2, \text{max, TDMA}} \). Furthermore, in the case when \( \alpha(d_2) < 1/P_2 \), if we set for both TDMA and CDMA that \( 0 < r_1 = \epsilon < B \log_2(1 + (P_1 - P_{CT})/(d_1^n P_N + P_2 - P_{CT})) \), then for CDMA, we have the maximum channel rate for node 2 is
\[
r_{2, \text{max, CDMA}} = B \log_2 \left( 1 + \frac{P_2 - P_{CT}}{d_2^n P_N} \right) = r_{2, \text{max, CDMA}}.
\]
(21)
For TDMA, since \( r_1 > 0 \) and \( \theta \) cannot be set as 1, which is optimal for node 2 when \( \alpha(d_2) > 1/P_2 \), the maximum channel rate for node 2 is now
\[
r_{2, \text{max, TDMA}} = \max_{0 \leq \theta < 1} \left\{ \theta B \log_2 \left( 1 + \frac{P_2 - P_{CT}}{d_2^n P_N} \right) \right\}
\]
\[
< r_{2, \text{max, CDMA}} = r_{2, \text{max, CDMA}},
\]
(22)
Therefore, as an example when \( \alpha(d_2) > 1/P_2 \) and
\[
\frac{r_1}{r_2} = \left( \frac{P_1 - P_{CT}}{d_1^n P_N + P_2 - P_{CT}} \right) / \left( \frac{P_1 - P_{CT}}{d_2^n P_N} \right),
\]
(23)
CDMA provides higher sum rate than TDMA. That is TDMA cannot provide higher sum rate than CDMA in all channel rate regions. This can be illustrated by Fig. 3. The transmission distances are set as \( d_1 = 30 \) m and \( d_2 = 50 \) m. For both nodes, the circuit power consumption is \( P_{CT} = 100 \) mW, and average power constraint is \( P_i = 120 \) mW. The path loss exponent is \( n = 4 \), and the noise power is \(-116.49 \) dBmW. Signal bandwidth is \( B = 100 \) KHz. From the figure, it is shown that if
\[
\begin{cases}
H(X_1) & \frac{r_1}{r_2} < \frac{H(X_1)}{H(X_2)} \\
H(X_2) & \frac{H(X_1)}{H(X_2)} \leq \frac{r_1}{r_2} \leq \frac{H(X_1)}{H(X_2)} \in \text{Region 1},
\end{cases}
\]
(24)
CDMA, which we refer to as $\theta_2$ while if support higher sample rate at the source nodes than CDMA; TDMA can provide higher sum channel rate and thereby can support higher sample rate at the source nodes. As a result, to maximize the sample rate at the source nodes, both TDMA and CDMA should be considered as possible multiple access channel techniques. Which multiple access technique is superior is determined by the entropy of the source random variables, the average power constraints, the circuit power consumptions, the noise power and the path loss.

C. Optimization problem formulation

A general modeling of a MAC using TDMA and CDMA is illustrated in Fig. 4. The parameter $\theta_{12}$ is the portion of the time resource assigned for CDMA, $\theta_1$ is the portion of the time resource assigned for only node 1 to communicate, and $\theta_2$ is the portion of the time resource assigned for only node 2 to communicate. That is, when $\theta_{12} = 0$, the MAC only uses TDMA; when $\theta_1 + \theta_2 = 0$, the MAC only uses CDMA; when $\theta_1 + \theta_2 \neq 0$ and $\theta_{12} \neq 0$, the MAC uses both TDMA and CDMA, which we refer to as hybrid TDMA/CDMA.

If the transmit power of node $i$ is denoted by $P_i$, then the average power constraints for nodes 1 and 2 are

$$(\theta_1 + \theta_{12})(P_1 + P_{CT}) \leq \bar{P}_1, (\theta_2 + \theta_{12})(P_2 + P_{CT}) \leq \bar{P}_2, \quad (26)$$

The rates of node 1 and node 2 can be expressed as

$$\begin{align*}
\theta_1 B \log_2 (1 + \frac{P_1}{\sigma_1^2 P_N}) + r_{1,\text{CDMA}} & = r_1, \\
\theta_2 B \log_2 (1 + \frac{P_2}{\sigma_2^2 P_N}) + r_{2,\text{CDMA}} & = r_2,
\end{align*} \quad (27)$$

where $r_{1,\text{CDMA}}$ and $r_{2,\text{CDMA}}$ are the channel rates of nodes 1 and 2 when using the CDMA MAC, and

$$\begin{align*}
r_{1,\text{CDMA}} + r_{2,\text{CDMA}} & \leq \theta_{12} B \log_2 \left( 1 + \sum_{i=1}^{2} \frac{P_i}{\sigma_i^2 P_N} \right), \\
r_{1,\text{CDMA}} & \leq \theta_{12} B \log_2 \left( 1 + \frac{P_1}{\sigma_1^2 P_N} \right), \\
r_{2,\text{CDMA}} & \leq \theta_{12} B \log_2 \left( 1 + \frac{P_2}{\sigma_2^2 P_N} \right).
\end{align*} \quad (28)$$

Built on the general MAC model, the maximization of the sample rate at the source nodes can be established as follows:

$$\begin{align*}
\min \quad & -S \\
\text{s.t.} \quad & C_0: \quad S > 0, \theta_i \geq 0, \\
& C_1: \quad (\theta_1 + \theta_{12})(P_1 + P_{CT}) \leq \bar{P}_1, \\
& \quad (\theta_2 + \theta_{12})(P_2 + P_{CT}) \leq \bar{P}_2, \\
& C_3: \quad \theta_1 + \theta_2 + \theta_{12} \leq 1, \\
& C_4: \quad r_1 \geq SH(X_1|X_2), \quad r_2 \geq SH(X_2|X_1), \\
& \quad r_1 + r_2 \geq SH(X_1, X_2), \\
& C_5: \quad r_{1,\text{CDMA}} + r_{2,\text{CDMA}} \leq \theta_{12} B \log_2 \left( 1 + \sum_{i=1}^{2} \frac{P_i}{\sigma_i^2 P_N} \right), \\
& \quad r_{1,\text{CDMA}} \leq \theta_{12} B \log_2 \left( 1 + \frac{P_1}{\sigma_1^2 P_N} \right), \\
& \quad r_{2,\text{CDMA}} \leq \theta_{12} B \log_2 \left( 1 + \frac{P_2}{\sigma_2^2 P_N} \right).
\end{align*} \quad (29)$$

The above optimization problem has optimization variables $\theta_1$, $\theta_{12}$, $\theta_2$, $r_{1,\text{CDMA}}$, and $r_{2,\text{CDMA}}$. Regarding the optimal solutions to the above optimization problem, we have the following proposition:

Proposition 2 To maximize the sample rate, the optimal policy is to choose either TDMA or CDMA, or equivalently either $\theta_{12} = 0$ or $\theta_1 + \theta_2 = 0$ in (29).

Proof: For any ratio $\beta$ within $H(X_1|X_2)/H(X_2) \leq \beta \leq H(X_1)/H(X_2|X_1)$, if only TDMA is used, the resulting channel rates are $r_{1,\text{TDMA}}$ and $r_{2,\text{TDMA}}$, while if only CDMA is used, the resulting channel rates are $r_{1,\text{CDMA}}$ and $r_{2,\text{CDMA}}$, where

$$r_{1,\text{TDMA}}/r_{2,\text{TDMA}} = r_{1,\text{CDMA}}/r_{2,\text{CDMA}} = \beta.$$ 

Thus, the sum rate using the hybrid TDMA/CDMA is

$$\begin{align*}
(\theta_1 + \theta_{12})(r_{1,\text{TDMA}} + r_{2,\text{TDMA}}) + \theta_{12}(r_{1,\text{CDMA}} + r_{2,\text{CDMA}}).
\end{align*} \quad (30)$$

Since the sum rate of the hybrid TDMA/CDMA is a linear combination of the sum rates of TDMA and CDMA, to maximize the sum rate of the hybrid TDMA/CDMA with the rate ratio $\beta$, we have $\theta_1 + \theta_2 \to 0$ when $(r_{1,\text{TDMA}} + r_{2,\text{TDMA}}) \leq (r_{1,\text{CDMA}} + r_{2,\text{CDMA}})$, or otherwise $\theta_{12} \to 0$. That is, the optimal policy is to choose either TDMA or CDMA, whichever provides the higher sum rate at the rate ratio $\beta$. The rate ratio $\beta$ is only determined by the joint entropy of the observed random variables. □
HDOM variables are fixed to rate using the general MAC based on the optimization model. The transmission distances in Fig. 5(a) are $d_1 = d_2 = 100$ m, and the transmission distances in Fig. 5(b) are $d_1 = 40$ m, $d_2 = 160$ m. In Fig. 5(a), TDMA is optimal in terms of maximizing the sample rate at the source nodes, while in Fig. 5(b) CDMA is optimal. In both cases the general MAC provides the highest sample rate, and the performance of the general MAC technique always overlaps with either TDMA or CDMA. That is, it is demonstrated that the optimal policy is to choose either TDMA or CDMA. Another interesting observation is that asymptotically CDMA is the optimal in terms of maximizing the sample rate as the average power increases. As shown in Fig. 5(a), the resulting maximum sample rate using only CDMA approaches that using TDMA as average power increases. This is because as $P$ increases, the impact of circuit power will gradually diminish. Eventually, the MAC channel problem with the consideration of circuit power consumption will converge to a traditional MAC channel problem without circuit power where CDMA always provides higher sum rate than TDMA [6].

V. CONCLUSIONS

We have investigated the joint optimal rate allocation of multiple access channels and Slepian-Wolf coding in a WSN, with the consideration of circuit power consumption and average power constraints. We established and solved the optimization problem to maximize the sample rate at the source nodes based on a general MAC model. The results show that the optimal policy is to choose either TDMA or CDMA based on the entropy of the source variables, average power constraints, circuit power consumption, noise power and path loss.

REFERENCES