

IMPACT OF PHYSICAL LAYER DESIGN ON THE DEPLOYMENT OF A SENSOR NETWORK WITH CORRELATED OBSERVATIONS

T. Banerjee[†], A. A. Kherani[†], M. Debbah^{*}, A. Roumy^{**}

[†] Dept. of Comp. Sc. and Engg., IIT Delhi, India,

^{*} SUPELEC, Gif-sur-Yvette, France,

^{**} IRISA, Rennes, France.

ABSTRACT

We consider a sensor network which observes a spatially correlated random field. The Slepian-Wolf scheme is known to help reduce the transmission requirement in such a scenario. We seek a “good” operating point for this network, in order to maximize the information gathered with a constraint on the network budget (which includes the number of nodes and their transmission powers). We present a comparative study of the system performance obtained under various transmission/reception schemes and fading coefficient statistics. Finally, we study the optimal density of the sensor network.

1. INTRODUCTION

We consider a bounded region on which sensor nodes are to be spread in order to observe some stochastic process (a random field over the region, a standard example is temperature at different points on this region) [1]. The function of these sensor nodes is to make an observation and convey it to the fusion center¹. Because of the spatial correlation of the observed random field, the readings of sensors located close-by are not independent of each other. Thus, in order to be able to track (or, estimate) the random field, one may allow for the possibility that the sensor nodes convey less information than their individual measurements, and that the fusion center performs some processing to get the complete information.

Clearly such a scheme, if practical, would allow gain in the performance of the network. It was shown in [6] that such a scheme, using Slepian Wolf encoding [10] is indeed possible and further, *this can be done without explicit (per-measurement-basis) coordination among the sensor nodes*. The gain in using such a scheme is clearly the increased lifetime of individual sensors and the network as a whole.

When looking at the system from an information theoretic point of view, the collection of sensors across the field is

This work is supported in part by the Network of Excellence in Wireless Communications (NEWCOM), E. C. Contract no. 507325 and by the INRIA-ARC project InFormAtioN theory (IFANY).

¹We are assuming that there is only one fusion center, and that all the sensor nodes are able to convey their information to this fusion center directly.

gathering a certain amount of information. Thus one would like to deploy as many sensors as possible to increase the information captured. But because of the spatial dependence of observations, the joint entropy does not increase linearly with the number of sensors. On the other hand, the cost of transmitting the data may increase linearly with the number of sensors (monetary cost, power). Hence, a combination of the rate of increase of Joint Entropy with number of sensors and the cost incurred per sensor dictates the actual number of sensors used.

Apart from the source coding technique used (e.g. distributed source coding (DSC) or Slepian-Wolf [10]), there are other physical layer aspects that affect the choice of the number of sensors used in the region of observation. These are,

(1) The choice of the multiple access scheme used, orthogonal or interfering channels,

(2) The statistics of the fading coefficient (deterministic or random).

The case of deterministic orthogonal channels is tackled in [6]. In the present paper, we consider more setups. More precisely for various types of transmission schemes (orthogonal or interfering like CDMA based access) and also various types of receivers for CDMA access we find the optimal rate allocation or the optimal power allocation for each link. The optimality is with respect to minimizing the power consumption in the network while using Slepian-Wolf encoding at the sources. Finally, we show the effect of these physical layer elements on the choice of the number of sensors deployed in the region of observation. Note that optimal density of a sensor network is also studied in [4]. The context is however different from ours since the sensor network has a task of distributed decision making, rather than estimation (as is the case in our work).

2. MATHEMATICAL MODEL

Consider a region where N sensors are deployed to observe some phenomena and send the observations to a single sink or fusion center. Let us model the data gathered by each sensor $i \in \mathcal{N} = \{1, \dots, N\}$ as an instance of a discrete random

variable X_i . R_i denotes the rate (in bits per source symbol) at which node i sends data. The transmit power P_i used by node i represents the cost of sending R_i bits. Sensors are placed in the bounded area and the position of each node is known and fixed.

The discrete random variables $\{X_i\}$ (as observed by different sensors) are correlated and their joint distribution is a function of the inter-node distances. Let X^N [resp., x^N] be the random vector (X_1, \dots, X_N) [resp., the observations vector (x_1, \dots, x_N)] and let $p(x^N) = \mathbb{P}(X^N = x^N)$ be the joint probability mass function of their observations. Then, the joint entropy $H(X^N)$ of the observation variables is given by

$$H(X^N) = - \sum_{x^N} p(x^N) \log_2 p(x^N).$$

The joint entropy is the total information (in bits per source symbol) which needs to be transmitted using the sensors.

Property 1. *Adding more sensors to the field by keeping the position of existing ones fixed can only increase $H(X^N)$.*

Proof. $H(X^{N+1}) = H(X^{N+1}|X^N) + H(X^N) \geq H(X^N)$. Thus, $H(X^N)$ is an increasing function of N . \square

It is this non-decreasing nature of the total information gained by increasing the number of sensor nodes that motivates us to use as many sensor nodes as possible. However, increasing the number of sensor nodes may not always provide us with a proportional return because of the following considerations:

1. Because the sensor measurements are correlated, the total information $H(X^N)$ gathered by N sensor nodes is assumed to increase sublinearly with N . Further intuitively (see example in section 5) $H(X^N)$ is a *non-decreasing* concave function of N , thus implying *decreasing returns* on increasing N .
2. Each sensor node comes with its own cost having either monetary or other interpretation. The cost of buying N sensors is $r(N)$ where $r(\cdot)$ is an increasing concave function such that $\lim_{N \rightarrow \infty} r(N) = \infty$. A particular example is per-sensor cost of b units so that the cost of buying N sensors is $r(N) = bN$ units.
3. For interfering channels, the more sensors are deployed, the more interferences are generated.

The total power consumed to send $H(X^N)$ is $\sum_{i=1}^N P_i$. We assume that the sensor network has a bound P on the total power usage to be consumed. We define a reward function for the power consumed by $f(P - \sum_i P_i)$. Here $f(\cdot)$ is a *concave* increasing function with $\lim_{x \rightarrow 0} f(x) = -\infty$ while $f(P)$ is a finite positive quantity. In a sensor network with renewable energy sources [7], $P.T$ can be thought of as the expected total energy available to be harvested in between successive transmissions of duration T .

We can now formulate the following optimization problem. If our net gain in terms of information is $g(N)$ where $g(\cdot)$ is an increasing concave function (it could be $H(X^N)$ for example), then the optimization problem would be.

$$\max_{N, R_i, P_i} \left[g(N) + f\left(P - \sum_i P_i\right) - r(N) \right], \quad (1)$$

subject to reliable communication

where the condition reliable communication depends on the transmission scenario, as detailed below.

2.1. Reliable communication

We consider two communication scenarios: orthogonal and interfering channels. In the first case and for distributed encoding, source channel separation holds [2], whereas it may not hold in the second case [5]. However, since there exists no practical distributed joint source channel coding schemes for more than 2 sources (for 2 sources see [13, 8]), we assume separated source channel coding schemes. Therefore, for reliable communication, the set of rates and powers $\{R_i, P_i\}$ should lie in the intersection of the Slepian Wolf region and of the capacity region.

Slepian Wolf region: constraint due to DSC. The Slepian Wolf region gives the set of achievable rates for distributed compression without cooperation between the sources ([10] for 2 sources and [12] for N sources). It is given by the following linear constraints,

$$\sum_{i \in \Sigma} R_i \geq H(X_\Sigma | X_{\Sigma^c}),$$

for any subset Σ of the set of nodes in the network \mathcal{N} (Σ^c is the set of nodes not in the set Σ).

Proposition 2. *The optimum occurs on the dominant face of the Slepian Wolf region such that $\sum_i R_i = H(X^N)$.*

Proof. The total information available is $H(X^N)$ and if the individual sensor nodes transmit at rates higher than $H(X^N)$, then the receiver will get the same information, i.e., $H(X^N)$, although at a higher network resource cost. \square

Proposition 3. *Separation of optimization. The joint optimization problem (1) can be carried out:*

- 1/ for each N , compute $H(X^N)$
- 2/ compute R_i, P_i to achieve this $H(X^N)$ that minimizes the sum power, under error free transmission:

$$\{R_i^*, P_i^*\} = \arg \min_{\{R_i, P_i\}} \sum_i P_i \quad (2)$$

- 3/ finally optimize the global cost function over N

Proof. From Proposition 2, R_i does not influence the net gain $g(N) = H(X^N)$ (as long as $\sum_i R_i = H(X^N)$). Therefore

(1) is equivalent to

$$\max_N \left[g(N) - r(N) + \max_{\{R_i, P_i\}} f\left(P - \sum_i P_i\right) \right].$$

Moreover $f(\cdot)$ is an increasing function, and

$$\max_{\{R_i, P_i\}} f\left(P - \sum_i P_i\right) = f\left(P - \min_{\{R_i, P_i\}} \sum_i P_i\right). \quad \square$$

With the above model in place, we now study various possibilities of multiple access schemes and find a rate allocation and expression for the total power consumed.

3. ORTHOGONAL MAC

In this section we find an operation point in the Slepian-Wolf region, for an orthogonal MAC (Multiple access channel) scheme.

This orthogonality can be achieved through protocols (CSMA) or multiple access techniques (such as TDMA, FDMA or orthogonal CDMA).

3.1. Slow fading

The capacity of the additive white Gaussian channel (AWGN) (between node i and the sink) with transmit power P_i , channel gain $\nu_i \in \mathbb{R}^+$ and noise variance σ^2 is

$$C_i(P_i) \triangleq \log_2\left(1 + \frac{\nu_i P_i}{\sigma^2}\right)$$

We further assume that the channel gains $\{\nu_i\}$ are fixed quantities, known by the sink. To achieve error free transmission, each rate should satisfy:

$$R_i \leq C_i(P_i)$$

Property 4. *The optimum occurs on the boundary such that $R_i = C_i(P_i)$.*

Proof. Proof by contradiction, similar as for Proposition 2. \square

This proposition allows us to reduce the number of variables and the optimization problem (2) now reads

$$\{R_i^*\} = \arg \min_{\{R_i\}} \sum_i (2^{R_i} - 1) \frac{\sigma^2}{\nu_i}$$

subject to error free compression (i.e. the rates R_i must lie in the Slepian Wolf region), and $P_i^* = C_i^{-1}(R_i^*)$

This optimization problem admits a closed form solution if we don't take into account the Slepian Wolf constraints. More precisely, due to the convexity of the exponential function, the optimal unconstrained rate allocation is:

$$\tilde{R}_i = \frac{1}{N} \left(H(X^N) + \log_2 \frac{\nu_i^{N-1}}{\prod_{j \neq i} \nu_j} \right) \quad (3)$$

Since this point may not lie in the Slepian Wolf region, we now propose a greedy algorithm to find a feasible point in

the admissible region. The idea is to choose a vertex of the Slepian Wolf region with a rate as close as possible as to the unconstrained optimal solution (3) in order to get small sum power. It leads to the following algorithm:

Algorithm 5. For all $i_1 \in \mathcal{N}$,

- $R_{i_1} = H(X_{i_1})$ and $P_{i_1} = (2^{R_{i_1}} - 1) \frac{\sigma^2}{\nu_{i_1}}$

- for $j = 2 : N$,

$$i_j = \arg \min_{k \neq i_1, \dots, i_{j-1}} H(X_k | X_{i_1}, \dots, X_{i_{j-1}}) - \frac{1}{N} \log_2 \frac{\nu_k^{N-1}}{\prod_{l \neq k} \nu_l} \quad (4)$$

$$= \arg \min_{k \neq i_1, \dots, i_{j-1}} H(X_{i_1}, \dots, X_{i_{j-1}}, X_k) - \log_2 \nu_k \quad (5)$$

$$R_{i_j} = H(X_{i_j} | X_{i_1}, \dots, X_{i_{j-1}}) \text{ and } P_{i_j} = (2^{R_{i_j}} - 1) \frac{\sigma^2}{\nu_{i_j}}$$

- Compute $\sum_{j=1}^N P_{i_j}$.

Finally, choose $\min_{i_1} \sum_{j=1}^N P_{i_j}$.

(5) follows from the fact that (4) can be rewritten as $H(X_{i_1}, \dots, X_{i_{j-1}}, X_k) - \log_2 \nu_k - H(X_{i_1}, \dots, X_{i_{j-1}}) + \frac{1}{N} \log_2 \prod_l \nu_l$, where the two last terms are common for all $k \neq i_1, \dots, i_{j-1}$.

The output of algorithm 5 is a rate allocation denoted \hat{R}_i with sum power

$$\sum_i \hat{P}_i = \sum_i \frac{\sigma^2}{\nu_i} (2^{\hat{R}_i} - 1) \quad (6)$$

3.2. Fast fading

We now consider fast fading channels and assume that the channel state is available at the receiver only. In this case, the capacity of this fast fading AWGN channel is [3]:

$$C(P_i) \triangleq \mathbb{E}_\nu \left[\log_2 \left(1 + \frac{\nu P_i}{\sigma^2} \right) \right]$$

We further assume that all the channel fading have the same distribution. As an example, we consider a channel distribution with two mass points:

$$p(\nu) = \frac{1}{2} \delta(\nu) + \frac{1}{2} \delta(\nu - a) \quad (7)$$

with $a \in \mathbb{R}^+$. This channel can be seen as a rough approximation of a Rayleigh channel with two states. For such a channel, the capacity is

$$C(P_i) = \frac{1}{2} \log_2 \left(1 + \frac{a P_i}{\sigma^2} \right)$$

Algorithm 5 can also be applied in order to find a feasible rate allocation. The only changes are: $\nu_i = a, \forall i$ and 2^{R_i} is changed into 2^{2R_i} .

4. CDMA WITH RANDOM CODES

We consider now interfering channels and more precisely, a CDMA scheme with random codes. CDMA linear receivers are used in order to separate the users. Therefore, each user sees an equivalent channel with noise variance equal to the signal to interference plus noise ratio (SINR) at the output the CDMA linear receiver. For this channel, single user capacity achieving codes are used. We use the asymptotic analysis of Linear MMSE receiver (MMSE) and Matched filter (MF) from [11] to obtain the SINR. Although the results in [11] are said to be asymptotic, they are seen to be valid even for small values of N . Hence for the sake of comparison, we assume that the results for MMSE and MF used are valid for all N .

4.1. Matched filter receiver

For the matched filter, the output SINR is given by

$$\frac{P_i \nu_i}{\sigma^2 + c}$$

where $c = \frac{1}{N} \sum_{j \neq i} P_j \nu_j \approx \frac{1}{N} \sum_j P_j \nu_j$. The rates are chosen in order to minimize the sum power given by

$$\sum_i P_i^{\text{MF}} = \frac{1}{1 - \frac{1}{N} \sum_j (2^{R_j} - 1)} \sum_i \frac{\sigma^2}{\nu_i} (2^{R_i} - 1) \quad (8)$$

Interestingly, for the matched filter, communication is not always possible. The rates should be not too large and verify: $1 - \frac{1}{N} \sum_j (2^{R_j} - 1) \geq 0$.

At low SNR, $\frac{1}{N} \sum_j (2^{R_j} - 1)$ is negligible in (8) s.t. the same rate allocation $\{\hat{R}_i\}$ as in Algorithm 5 can be applied. At high SNR, $\frac{1}{N} \sum_j (2^{R_j} - 1)$ dominates and Algorithm 5 can be applied with $\nu_i = 1$.

The case of a fading channel with distribution (7) gives similar sum power with $\nu_i = a$ and 2^{R_i} changed into 2^{2R_i} . The rate allocation is given by Algorithm 5 with $\nu_i = a$.

4.2. MMSE receiver

For the MMSE receiver, the output SINR is given by $P_i \nu_i \beta$ where

$$\frac{1}{\beta} = \sigma^2 + \frac{1}{N} \sum_{j \neq i} \frac{\nu_j P_j}{1 + \nu_j P_j \beta} \approx \sigma^2 + \frac{1}{N} \sum_j \frac{\nu_j P_j}{1 + \nu_j P_j \beta}$$

The sum power is given by

$$\sum_i P_i^{\text{MMSE}} = \frac{1}{\frac{1}{N} \sum_j 2^{-R_j}} \sum_i \frac{\sigma^2}{\nu_i} (2^{R_i} - 1) \quad (9)$$

As for the matched filter, at low SNR the rate allocation $\{\hat{R}_i\}$ of Algorithm 5 is applied, whereas at high SNR, Algorithm 5 is applied with $\nu_i = 1$.

The case of a fading channel with distribution (7) gives similar sum power with $\nu_i = a$ and 2^{R_i} changed into 2^{2R_i} . The rate allocation is given by Algorithm 5 with $\nu_i = a$.

5. COMPARATIVE STUDY

We consider here the entropy-based model of [9] that approximates well the measurements of the daily rainfall precipitation. So, for an N sensor network, where all nodes are located on a line equally spaced by d , the joint entropy is given by:

$$H(X_1^N) = H(X_1) + (N - 1) \left(1 - \frac{1}{\frac{d}{c} + 1} \right) H(X_1) \quad (10)$$

where c captures the correlation between the sources (the smaller c , the bigger the correlation). We further assume that the area of the network is $L \in \mathbb{R}^+$. All sensors are assumed to be equally spaced, s.t. adding one more sensor to the network impose to move all the sensors. Moreover the sensors range the whole interval of length L (a sensor is placed at each extremity of the interval). Thus, the distance d_N between 2 consecutive sensors depends on the number of sensors and is

$$\forall N \geq 2, d_N = \frac{L}{N - 1}$$

From (10) we finally get the joint entropy as

$$H(X_1^N) = H(X_1) + \frac{(N - 1)L}{L + c(N - 1)} H(X_1) \quad (11)$$

The growth of this joint entropy is sublinear. It is even bounded since

$$\lim_{N \rightarrow +\infty} H(X_1^N) = H(X_1) \left(1 + \frac{L}{c} \right)$$

The growth of the joint entropy versus the number of sensors deployed in the network is shown on Fig. 1.

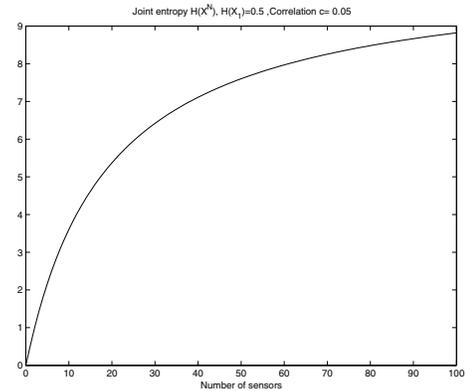


Fig. 1. Joint entropy. $H(X_1) = 0.5$, $L = 1$, $c = 0.1$.

We can now compare the variation of the total power consumed with the number of sensors for orthogonal MAC (6), MF (8) and MMSE (9) receivers. Figure 2 gives the desired plot for a fading AWGN channel with distribution (7). Simulations show that the suboptimal but feasible rate allocation $\{\hat{R}_i\}$ of Algorithm 5 gives sum powers close from the

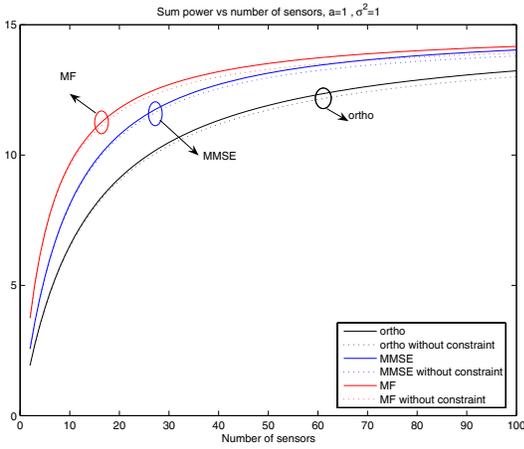


Fig. 2. The total power consumed in the sensor network as a function of the number of sensors. 2-mass point fading channel (7) with $a = 1$, $\sigma^2 = 1$.

optimal but unconstrained solution, that may not be feasible. Therefore the suboptimal approach gives little loss.

Finally, we compare the optimal density for the 3 access schemes in Figure 3. The utility function is

$$\frac{H(X^N)}{\sum_i P_i + 0.02N}$$

Therefore the optimal density corresponds to the number of sensors that maximizes the amount of received data while minimizing the power and the cost (monetary for instance).

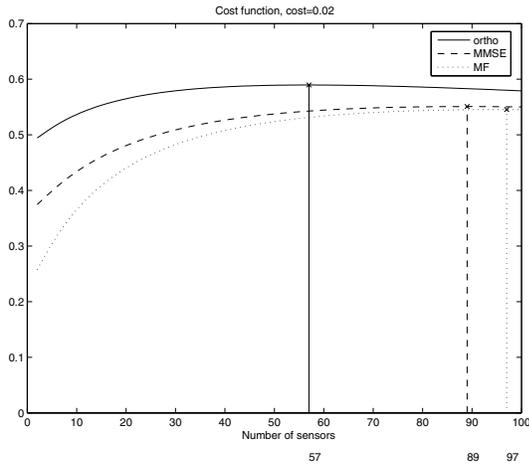


Fig. 3. The system utility as a function of the number of nodes. The utility function is $H(X^N)/(\sum_i P_i + 0.02N)$. The sensors are placed at equal distances from each other.

6. CONCLUSION

In this paper we studied the impact of the physical layer on a sensor network, where Slepian Wolf encoding is used in order to compress the data. We derived the optimal rate allocation that minimizes the sum power without Slepian Wolf constraint and proposed a greedy rate/power allocation algorithm to minimize the sum power under the Slepian Wolf constraint. Simulations showed that both solutions are very similar. Finally, we derived the optimal density of the sensor network that maximizes the amount of received data per unit cost.

7. REFERENCES

- [1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. "Wireless sensor networks: a survey." *Computer Networks*, 38(4):393422, April 2002.
- [2] J. Barros, S. D. Servetto, "Network Information Flow with Correlated Sources," *IEEE Trans. on Information Theory*, Vol. 52(1), pp. 155-170, 2006.
- [3] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. on Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct 1998.
- [4] J.-F. Chamberland and V.V. Veeravalli. "How Dense Should a Sensor Network be for Decentralized Detection with Correlated Observations?." *IEEE Trans. on Information Theory*, 52(11):5099-5106, Nov. 2006.
- [5] T. Cover, A. El Gamal and M. Salehi, "Multiple Access Channels with Arbitrarily Correlated Sources," *IEEE Trans. on Information Theory*, Vol. IT-26, No. 6, pp. 648-657, Nov. 1980.
- [6] R. Cristescu, B. Beferull-Lozano, M. Vetterli, "Networked Slepian-Wolf: theory, algorithms, and scaling laws", *IEEE Trans. on Information Theory*, 51(12):4057- 4073, Dec. 2005.
- [7] K. Kar, A. Krishnamurthy, N. Jaggi, "Dynamic node activation in networks of rechargeable sensors" *IEEE/ACM Trans. in Networking*, Vol. 14 , Issue 1, pp. 15 - 26, Feb. 2006.
- [8] A. Lapidoth, S. Tinguely, "Sending a Bi-Variate Gaussian Source over a Gaussian MAC", *ISIT*, 2006.
- [9] S. Patten, B. Krishnamachari, R. Govindan, "The impact of spatial correlation on routing with compression in wireless sensor networks," in *IPSN 2004*.
- [10] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources", *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 471-480, July 1973.
- [11] D. N. C. Tse and S. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Inform. Theory*, vol. 45, pp. 641657, Mar. 1999.
- [12] J. K. Wolf, "Data reduction for multiple correlated sources," in *Proc. 5th Colloquium Microwave Communication Budapest, Hungary*, pp. 287-295, June 1973.
- [13] W. Zhong and J. Garcia-Frias, "LDGM Codes for Channel Coding and Joint Source-Channel Coding of Correlated Sources", *EURASIP Journal on Applied Signal Processing*, pp. 942-953, May 2005.