Title

On Scalability and Source/Channel Coding Decoupling in Large Scale Sensor Networks

Authors

Ameesh Pandya
Greg Pottie

Publication Date

2003

On Scalability and Source/Channel Coding Decoupling in Large Scale Sensor Networks

Ameesh Pandya and Greg Pottie

Department of Electrical Engineering

University of California, Los Angeles

{ameesh, pottie}@ee.ucla.edu

This material is based upon work supported by the National Science Foundation (NSF) under Cooperative Agreement #CCR-0121778.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).

Abstract

Prior results have shown that for ad hoc networks with uniform source-destination probabilities, where each node generates traffic, the transport capacity for each node in the network declines with the network size n [1]. In this paper, we show that the finite per-node throughput and hence scalability is achievable with the adjustment of source-destination pair distributions.

We also explore the rate-distortion bound in the context of sensor networks. Considering a network over a finite region, with finite Gaussian point sources and densely distributed sensors, the otherwise difficult to obtain data rate region under a fidelity criterion, reduces to a partial side information problem for Gaussian sources. The key concept in proving this is to consider source, sensor, and communications relay densities as separate quantities.

Index Terms

sensor network, scalability, source-destination pair distribution, rate-distortion bound, source separation, partial side information.

I. INTRODUCTION

Large scale sensor networks consisting of hundreds or thousands of nodes will link the physical world to global communication networks for a broad set of applications [2], [3], [4]. Individual nodes will have some combination of sensing, signal processing and communications capability and may self-organize for a variety of cooperative sensing and communication tasks, subject to resource constraints such as energy and bandwidth [2].

Basic information theoretic questions for such networks include the minimum resources required to extract information about some source to some level of fidelity - a rate distortion problem [5], and whether the network capacity per node is scalable subject to bandwidth constraints [1]. Even given that source-channel coding separability fails for network information theory problems, there remain many long-standing unsolved problems for source and channel coding treated individually. However, it is possible to derive rate regions in order to answer questions about whether the required information can be extracted or the network can scale [6], [7] given resource constraints.

The question of the capacity region for ad hoc network has recently been addressed by Gupta and Kumar [1]. Their key result is that given n nodes in the unit disk and a uniform traffic pattern, one obtains the per-node capacity of $O\left(1/\sqrt{n}\right)$ bit-meter per second (transport capacity), assuming simple point-to-point coding. The results in [1] suggest that static ad hoc networks are inherently non-scalable, i. e. per-node capacity $\rightarrow 0$ as $n \rightarrow \infty$.

After Gupta and Kumar's capacity results [1] for wireless networks, a considerable amount of work has been done in this area [8], [9], [10], [11]. Grossglauser and Tse in [12] modified the model in [1] to explicitly include mobility. They allowed for unbounded delay and used only one hop relaying, taking advantage of mobility. They showed that $O\left(n\right)$ throughput for a mobile ad hoc network is achievable. Recently, Bansal and Liu in [13] proposed a routing algorithm which exploits the patterns in mobility of nodes to provide guarantees on delay. Here we show that scalability is possible even for static networks.

In this paper, we study the scalability issue and show that the appropriate choice of source-destination pair distribution can lead to finite per node capacity in ad hoc networks. Applications of this result include sensor networks and hierarchical communication networks.

We also consider the rate-distortion region characterization for a sensor network with finite point sources and under a fidelity criterion. We note that the model of the sensor network and the objectives are critical to the conclusions. Servetto [6] for example poses the problem as a many-to-one correlated coding problem and for his model concludes that the information can be extracted at the desired level of fidelity for dense networks. Neuhoff, et. al. [7] on the other hand have a significantly different set of objectives and conclude these cannot be met. The model considered in [7] has a data-gathering wireless sensor network in which densely deployed sensors take periodic samples of the sensed field, and then quantize, losslessly encode and transmit them to a single receiver/central controller where snapshot images of the sensed field are reconstructed. For this particular model, [7] shows that the efficiency with which the sensor network functions, degrades with the increase in the density of the sensors. By contrast, in this paper, the sensor network problem consists of extracting information about sources in some region to some desired level of fidelity, and transmitting this information to some gateway(s). We assume point sources to be Gaussian for analytical simplicity. We show that the question of feasible rates for such a network can be made into a relatively easy problem to deal with by allowing the network to be dense, i.e. letting the number of sensors and communication relays be much larger than the number of sources. A sub-optimal decoupling of source and channel coding can be shown to be sufficient for achieving scalability in this context. We also explore how the issues of scalability, source separation, and information extraction can be dealt with by altering the relative densities of sources, communication relays and the sensors.

The rest of the paper is organized as follows. In Section II we discuss some early work that motivated our formulation. In Section III we present the S-D pair distribution for achieving scalability in sensor networks. The consequences of decoupling source, communication relay and sensor densities for sensor networks are discussed in Section IV. The paper concludes in Section V with the discussion of future work.

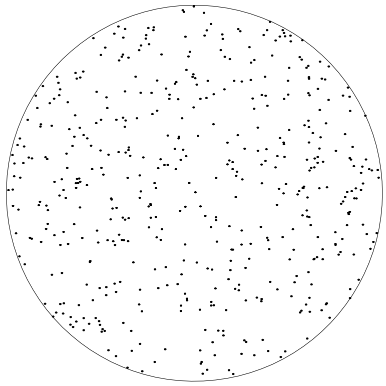

Fig. 1. n nodes located in the disk of unit area with uniform traffic pattern

II. Prior Scalability Results

A. Point-to-Point Coding Model [1]

In [1], Gupta and Kumar devised a rule for the achievable global transmission rate of an ad hoc network. The rate is measured in bit-meters per second (the transport capacity). In their analysis they used a simple *point-to-point coding* model. At any given time, a receiver decodes a message from only one sender, considering simultaneous transmissions purely as noise, and similarly at any given time, a sender transmits information to only one receiver. Consider n nodes located in the disk of unit area (Figure 1) with a uniform traffic pattern. Each node transmits at W bits per second over a common wireless channel. Packets are sent from source to destination in a multi-hop fashion. Radios that are sufficiently distant can transmit concurrently; the total amount of data that can be simultaneously transmitted for one hop increases linearly with the total area of the ad hoc network.

If node density is constant, this means that the total one hop capacity is $O(n)$. However, as the network grows larger, the number of hops between each source and destination may also grow larger, depending on communication patterns. One might expect the average path length to grow with the spatial diameter of the network, or equivalently the square root of area, or $O(\sqrt{n})$. With this assumption, the total end-to-end capacity is roughly $O\left(n/\sqrt{n}\right)$, and end-to-end throughput available to each node is [1], [11]

$$O\left(\frac{1}{\sqrt{n}}\right)$$

The inverse square root of n behavior can be intuitively understood as follows: Every bit has to travel at least the distance that separates its source from its destination. It may travel this distance either through a single direct transmission, or through multiple transmissions via relay nodes [12].

Gupta and Kumar also demonstrated the existence of a global scheduling scheme achieving $\Omega\left(1/\sqrt{n\log n}\right)$ for a uniform random network with a random traffic pattern.

The application of throughput result of [1] to a relay traffic pattern, where there is only one source/destination pair while the rest of the network is at their service, would result in $O\left(\sqrt{n}\right)$ bits per second or a more careful application would yield $O\left(1\right)$. In fact, Gastpar and Vetterli in [14], [15] showed that the capacity for such a traffic pattern is $O\left(\log n\right)$ as the number of nodes in the network, n, goes to infinity. The extension of results in [14], [15] to the relay traffic pattern with multiple transmissions is not known.

To summarize, the results in [1] indicates the lack of scalability in static ad hoc networks.

III. Scalability in Sensor Networks by Source-Destination Pair Distribution

In this section, we demonstrate that some choices of the source-destination pair distribution in networks can indeed achieve finite per node capacity and hence scalability. This concept of source-destination pair distribution for achieving the scalability is motivated by the work in [11].

A. *Scalability for Point-to-Point Coding Model*

Successful communication is still possible with almost constant node bandwidth, if the distribution of the source-destination (S-D) pairs is such that the average hops per communication is small enough.

We consider the same 2-D framework as in [1]. Servetto in [6] observes that the scalability in sensor networks of arbitrary size is achievable as long as the rate at which nodes generate information decays faster then the throughput of the network. However, we note that it is not the fact of correlated sources which is most fundamental to this result, but rather the S-D pair distribution.

Observation III.1: For the 2-D geometric model defined in [1], the average distance between source and destination should be $O(\frac{1}{\sqrt{n \log n}})$.

Clearly, in order for the average throughput per node to be constant, the average number of hops between source and destination should grow as $O(1)$. This follows immediately by observing that this criterion is necessary for transport capacity meeting the upper bound of $\sqrt{n \log n}$.

1) Example of a distribution that achieves the non-increasing average number of hops with n: Consider a network with n nodes randomly distributed over an unbounded area with some particular distribution as follows(Figure 2). The x and y coordinates of the node locations are independent Gaussian distributions with zero mean and variance σ^2. Defining R as the transmission range of each node, the probability density function (pdf) of the link distance, r, between any two arbitrary nodes is given by:

$$p_r(r) = \frac{r}{2\sigma^2} e^{-r^2/4\sigma^2} \tag{1}$$

The derivation for the pdf of the link distance can be found in [16]. [16] also considers the case when x and y coordinates have different variances.

The probability of a 2-hop connection between an arbitrary source and destination pair is given

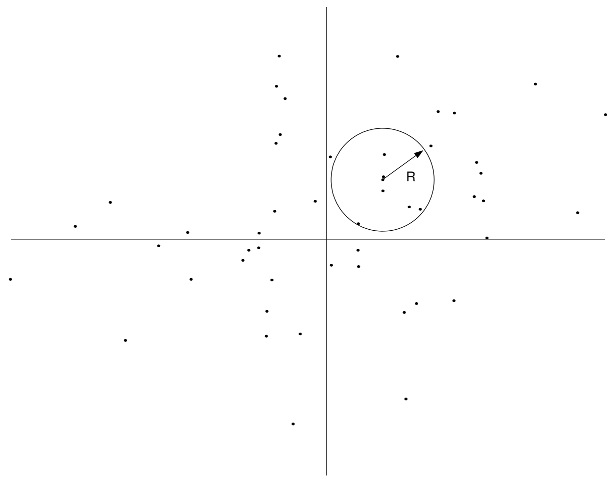

Fig. 2. Random Network with dispersion parameter σ, transmission range R.

by [17]:

$$\begin{aligned} P_2 &= Pr\{1 \rightarrow 2 \text{ in 2 hops}\} \\ &= Pr\{R < r < 2R \text{ and at least one other node in the area of intersection} \\ &= \underbrace{\iiint}_{R<r<2R} p_{x,y}(x_1, y_1, x_2, y_2) \\ &\quad \times \left[1 - \left[1 - \underbrace{\iint}_{A(x_1,y_1,x_2,y_2)} p_{x,y}(x_3, y_3) dx_3 dy_3 \right]^{n-2} \right] dx_1 dy_1 dx_2 dy_2 \end{aligned} \quad (2)$$

As $n \rightarrow \infty$, (2) can be approximated with the following upper bound [17]:

$$P_2 < P_{2\infty} = \int_{R^2/4\sigma^2}^{R^2/\sigma^2} e^{-\nu} d\nu P_{2\infty} = e^{-R^2/4\sigma^2} - e^{-R^2/\sigma^2} \quad (3)$$

Similarly, the asymptotic probability of an m-hop connection is given by:

$$P_m < P_{m\infty} = e^{(m-1)^2 R^2/4\sigma^2} - e^{m^2 R^2/4\sigma^2} \quad (4)$$

Miller [17] then showed that the average number of hops between node pair can be calculated as,

$$E\{h\} = \sum_{m=1}^{n-1} mP_m < \sum_{m=1}^{\infty} mP_{m\infty} \tag{5}$$

$$E\{h\} < \sum_{m=0}^{\infty} e^{-\frac{m^2}{4\gamma^2}} \doteq \overline{h}_+(\gamma) \tag{6}$$

where γ is the mobile dispersion given by $\gamma = \sigma/R$. Using the non-linear regression techniques and subsequent linearization, the asymptotic average number of hops for random source destination pair is,

$$\overline{h}_+(\gamma) \approx 0.5 + 1.772(\sigma/R) \tag{7}$$

In terms of actual distance, we simply multiply by the transmission range of each node R, to upper bound the average hop distance per transmitted bit by

$$E\{h\} < R\left(0.5 + 1.772\frac{\sigma}{R}\right) \tag{8}$$

The average hop distance in (8) is for an unbounded disk. Since the nodes in the scenario described earlier are zero-mean Gaussian distributed, 99.7% of all the nodes are expected to lie within a 3σ radius of the center. The resulting area of the disk of radius 3σ is $9\pi\sigma^2$. We therefore scale the result in (8) by this factor. Also note, that the average hop distance is consistent with the framework of Gupta and Kumar [1],

$$H = \frac{R}{9\pi\sigma^2}\left(0.5 + 1.772\frac{\sigma}{R}\right) \tag{9}$$

which is independent of n.

Hence, the zero-mean truncated Gaussian is one of the many distributions that achieve the finite per node capacity for the *geometric model* of [1].

Our approach is based on results in [6], but one can find the similar results with the alternative approach in [11]. J. Li, et. al. in [11] show that the traffic pattern determines whether an ad hoc network's per-node capacity will scale to large networks. In particular, they show that for the total capacity to scale up with network size the average distance between source and destination nodes must remain small as the network grows. If local communication predominates, path lengths could remain nearly constant as the network grows leading to a constant per-node available throughput [18].

In practice, scalability can be achieved in two basic ways:

1) Local cooperative processing to produce decisions (e. g. in sensor networks).
2) Adding communications hierarchy so that communications in each level is local (e. g. telecommunication network).

The latter of course requires additional resources, but typically also provides latency benefits.

IV. Consequences Of Decoupling Source, Sensor And Relay Densities

Along with the ways to approach network scalability, another interesting question is the feasible rate region under a fidelity criterion. We now find the rate region for a wireless sensor network with finite point sources and under a fidelity criterion and discuss the consequences of spatial source separation in dense networks.

In sensor networks, the basic problem is to extract measurements of some physical phenomenon, to some desired level of fidelity, subject to constraints on energy consumption and bandwidth (resources). Nodes may also have explicit limits on signal processing and storage, which we will neglect here. By considering source, sensor and communications relay densities separately, we show that extraction of such information can be easily achieved without the requirement of

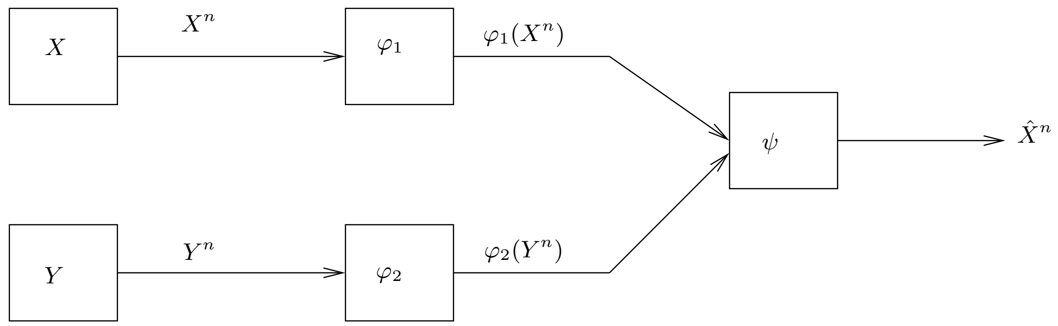

Fig. 3. The coding system with partial side information at the decoder

complicated joint source-channel coding schemes, in the limit of high sensor and relay densities. Further, this formulation admits simple classification of a broad set of network information theory problems.

Before we cite our main results, we shall define some preliminaries required for our formulation.

A. Preliminaries

1) Coding System with Partial Side Information at the Decoder [19]: Consider the system as shown in figure 3.

X and Y are correlated memoryless Gaussian sources with squared distortion measures. For analytical simplicity, the observation data is assumed to be Gaussian. The challenge is to determine appropriate codes and data rates such that the gateway/data-fusion center can reproduce the data from the main node using the other node as a source of partial node information, subject to some distortion criterion. At the decoder we reproduce only X and do not care about Y. Rather, Y acts as a helper to reproduce X by providing side information at the data fusion node with some distortion criterion. This is the so-called partial side information problem.

2) Rate Distortion for the partial side information problem for Gaussian Source [19]: The main source, X, and a partial side information correlated source, Y (see Figure 3), $\{(X_t, Y_t)\}_{t=1}^{\infty}$

are stationary Gaussian memoryless sources, for each observation time, $t = 1, 2, 3, \ldots$. Let the random pair (X_t, Y_t) take values in $\mathcal{X} \times \mathcal{Y}$. The probability density function (pdf) $p_{XY}(x, y)$ is given by

$$p_{XY}(x, y) = \frac{1}{2\pi|\Lambda|^{1/2}} \exp\left\{-\frac{1}{2}\vec{x}\Lambda^{-1T}\vec{x}\right\}$$

where $\vec{x} = (x, y) \in \mathcal{X} \times \mathcal{Y}$ and Λ is a covariance matrix defined by

$$\Lambda = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}, \quad -1 < \rho < 1 \tag{10}$$

Let us denote independent copies of $\{X_t\}_{t=1}^{\infty}$ as $X^n = \{X_1, X_2, \ldots, X_n\}$ and that of $\{Y_t\}_{t=1}^{\infty}$, as $Y^n = \{Y_1, Y_2, \ldots, Y_n\}$. Consider a coding system where data sequences X^n, Y^n are separately encoded to $\varphi_1(X^n)$, $\varphi_2(Y^n)$ and sent to an information processing/data fusion center. The decoder function $\psi = (\psi_1, \psi_2)$ observes $(\varphi_1(X^n), \varphi_2(Y^n))$ and estimates $\left(\hat{X}^n, \hat{Y}^n\right)$. Let $\mathcal{F}_{n,\delta}(R_1, R_2)$ denote the set of all such coding and decoding schemes, $(\varphi_1, \varphi_2, \psi)$, which can exist with the properties mentioned above. Let,

$$d_1 : \mathcal{X}^2 \to [0, \infty), \quad d_2 : \mathcal{Y}^2 \to [0, \infty) \tag{11}$$

define the distortion measures, which in our case is the squared distortion measure. The average distortions Δ_1 and Δ_2 are given by,

$$\Delta_1 = \mathbf{E}\left\{\frac{1}{n}\sum_{t=1}^{n} d_1\left(X_t, \hat{X}_t\right)\right\}$$

$$\Delta_2 = \mathbf{E}\left\{\frac{1}{n}\sum_{t=1}^{n} d_2\left(Y_t, \hat{Y}_t\right)\right\}$$

Then for given positive numbers D_1 and D_2, a rate pair (R_1, R_2), is admissible if for any $\delta > 0$, $n \geq n_0(\delta)$, there exists a triple

$$(\varphi_1, \varphi_2, \psi) \in \mathcal{F}_{n,\delta}(R_1, R_2)$$

such that $\Delta_i \leq D_i + \delta$, $i = 1, 2$.

Note, that D_2 can be large as we do not care about the reproduction of Y. For an encoding system using Y as a helper, the rate distortion region is given by:

$$\mathcal{R}\left(D_1\right) = \left\{\left(R_1, R_2\right) : \left(R_1, R_2\right) \text{is admissible}\right\} \tag{12}$$

For the special case of the correlated Gaussian sources, the following theorem applies [19]:

Theorem IV.1: Consider the following encoding functions:

$$\varphi_1 : X^n \rightarrow \mathcal{C}_1 = \left\{1, \ldots, C_1\right\},$$

$$\varphi_2 : Y^n \rightarrow \mathcal{C}_2 = \left\{1, \ldots, C_2\right\},$$

to be such that the rate constraints being satisfied are

$$\frac{1}{n} \log C_i \leq R_i + \delta, i = 1, 2$$

then for an admissible rate (R_1, R_2) and for some $D_2 > 0$, the partial side information at the decoder coding system data rates for correlated Gaussian sources can be fused to yield an effective data rate (with respect to source X) satisfying the following lower bound

$$\mathcal{R}_1\left(D_1\right) \geq \frac{1}{2} \log^+ \left[\frac{\sigma_X^2}{D_1}\left(1 - \rho^2 + \rho^2 2^{-2R_2}\right)\right] \tag{13}$$

where $\log^+ x = \max\left\{\log x, 0\right\}$. This is the desired rate region.

The proof of the Theorem IV.1 can be found in [19]. Note that when there is no helper, (13) collapses to the classic Gaussian rate-distortion expression [20].

B. Spatial Source Separation

Consider any two dimensional region, S, with m Gaussian point sources (Figure 4) where $m < \infty$. The point sources are spread over a region S with some random distribution.

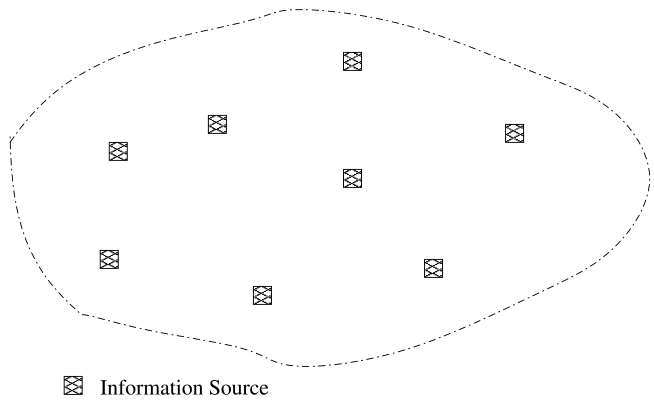

Fig. 4. A region S with Gaussian point sources

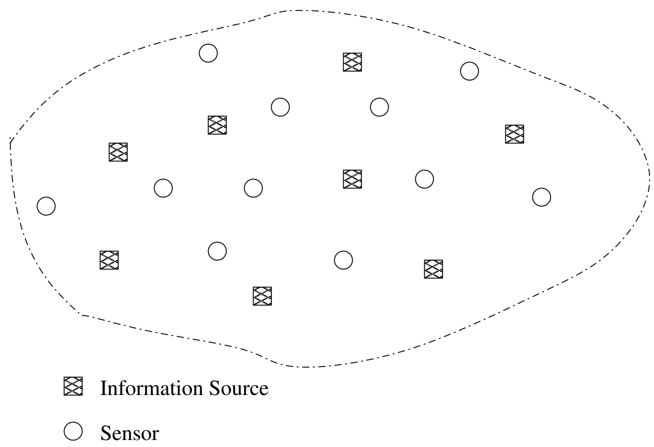

Fig. 5. A region S with Gaussian point sources and sensors of $O(m)$

Consider $p < \infty$ randomly distributed sensors in the region S to fulfill the functions of gathering information. Also, the sensors are assumed to be *iid* in location. Furthermore, the distribution by which the sensors are deployed is assumed to have positive density at every point over S (Figure 5). We also assume that power decays with distance.

The challenge is to find the rate distortion bound of such a network. In other words, we need to know the achievable distortion $D(R)$, for the above described sensor network model. The distortion measure is assumed to be squared error as in Section IV-A.

When the number of sensor nodes in the network is of $O(m)$, where m is the number of finite

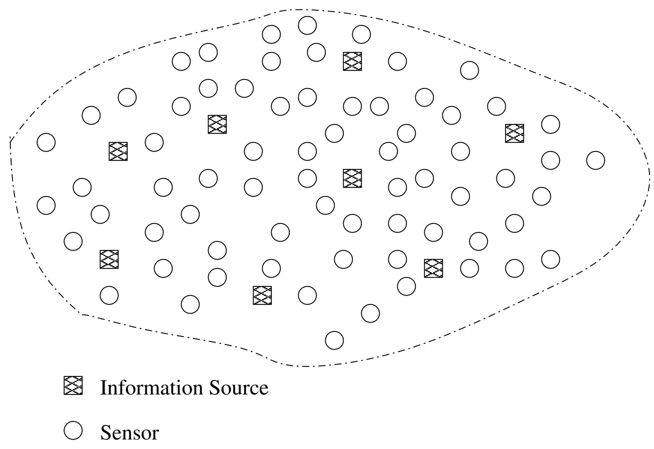

Fig. 6. A region S with Gaussian point sources and sensor nodes $\rightarrow \infty$

Gaussian point sources, not all D are achievable regardless of the rate constraints, since sensors may not be close enough to sources. Also, not all the rates are achievable because capacity may not be sufficient. Here, there is an interference among information streams and in the signals received by each sensor and so a joint source-channel coding approach would be needed to achieve a large fraction of the rate region. There is little prospect of actually implementing such a system for large m, although clearly it is a rich regime for future research. The concept of joint source-channel coding is well described in [15].

For the network in Figure 5, the rate distortion bound is a hard problem to solve even for the case of Gaussian point sources. But, it can be readily converted to a simple formulation by increasing the density of sensor nodes i.e. allowing sensor nodes, p, to go to infinity ($p \rightarrow \infty$). This is depicted in Figure 6.

As $p \rightarrow \infty$, we will have at least one sensor node in very close vicinity to the point source. In this scenario, the rate distortion bound for the network reduces to the individual rate distortion bounds, for each point source. Hence, we have the separation of point sources in the network. The clear distinction of our approach from [6], [7] is that we do not need every node to gather information, only the closest. We could, for lower distortion, use partial side information at the decoder from a correlated sensor node for the rate-distortion coding locally to each point source. This is sufficient to achieve desired distortion level D.

Proposition IV.2: For the network in Figure 6, as the number of sensors, p, goes to infinity, there exists at least one sensor node in a very close vicinity of the point sources with probability 1. Here, we have an assumption that sensor nodes are *iid* with distribution, F, having positive density at every point on region S.

Proof: Since F has positive density at every point in S, for any $\delta > 0, \exists B_1$ s.t. $P\left(||X_i - B_1|| < \delta\right) = \epsilon > 0$, where X_i is the sensor node and B_1 is the point source.

$$\begin{aligned} P\left(||X_i - B_1|| < \delta \; \forall i\right) &= 1 - P\left(||X_i - B_1|| \geq \delta \; \forall i\right) \\ &= 1 - P\left(||X_1 - B_1|| \geq \delta\right)^p \\ &\rightarrow 1 \text{ as } p \rightarrow \infty \end{aligned} \tag{14}$$

■

Theorem IV.3—(Source Separation): A network with finite Gaussian point sources, say m, and number of sensor nodes, p, going to infinity can be considered as a network with m separate coding systems with partial side information at the decoder. We assume that power decays at least as the square of the distance.

Proof: The proof follows from Proposition IV.2. Since $p \rightarrow \infty$, the frequency re-use distance goes to zero and thus there is no interference between information pathways through the network. Hence, a simple relay is adequate for carrying the traffic even if it is not capacity achieving. For the nodes that are in very close vicinity to the Gaussian point sources, SNR for those nodes goes to infinity. So, any value of distortion, D, will be achievable and there is no substantial interference among sources. Hence, local processing will be sufficient. In the scenario considered here, each

Gaussian point source can be separated from the rest and source coding can be posed as a partial side information problem. This is because Gaussian point sources are independent of each other. Since we have m point sources in the network of Figure 6, it can be seen as m separate coding system with partial side information at the decoder for the correlated Gaussian sources.

The assumption that the signal decays with distance faster than some rate is required to avoid interference growing without bound for large fields of sources. This assumption is quite reasonable for typical deployments of sensor networks for many physical phenomena of interest. ■

From Theorem IV.3 and limiting cooperation to only two nodes per source, it is evident that the data rate, R_{X_i}, $i = 1, 2, , \ldots, m$, associated with each point source is given by

$$\mathcal{R}_{X_i}\left(D_{X_i}\right) \geq \frac{1}{2}\log^{+}\left[\frac{\sigma^2_{X_i}}{D_{X_i}}\left(1-\rho^2+\rho^2 2^{-2R_{Y_i}}\right)\right] \tag{15}$$

where X_i is the main source and Y_i is a helper such that $(i = \{1, 2, \ldots, m\})$. The rate distortion bound for the network will be the ensemble over the position of all the point sources. Clearly, extending to the q-helpers per source $(q > 1)$, lower distortions would be achievable for a given density of sensors.

In terms of achievable distortion for a point source we have,

$$D_{X_i}\left(R_{X_i}\right) \geq \frac{\sigma^2_{X_i}}{2^{2R_{X_i}}}\left[\frac{\sigma^2_{X_i}}{D_{X_i}}\left(1-\rho^2+\rho^2 2^{-2R_{Y_i}}\right)\right] \tag{16}$$

for $i = \{1, 2, \ldots, m\}$.

The achievable distortion for the network will be $D = \sum_{i=1}^{m} D_{X_i}$. For a point source, if we fix the node density and allow an algorithm to select a node closest to the source while making the rest of the nodes inactive, it is less likely to achieve the desired distortion. But this approach is a practical way to deal with achieving the desired distortion, at the cost of increased density.

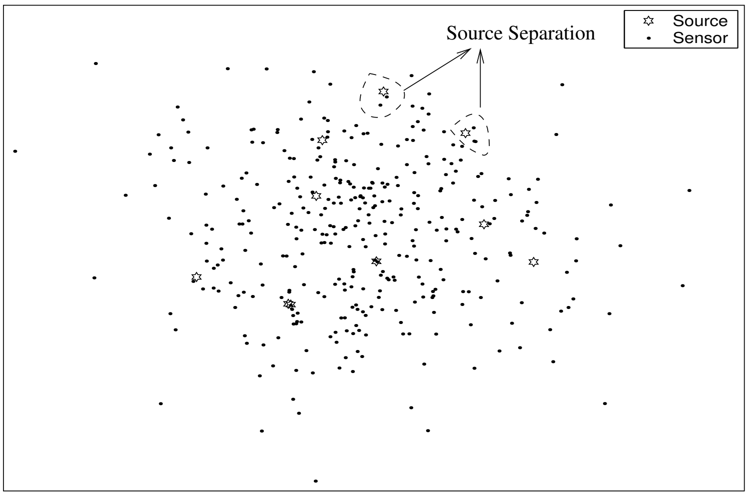

Fig. 7. Gaussian distributed sources and sensor nodes

Note that we consider only those sensor nodes that are near to the source and it is sufficient to have only partial side information at the decoder. Consider a source and node distribution as shown in Figure 7. Consider the case of determining the achievable rate for a desired distortion level 0.4 a partial side information at the decoder. The covariance matrix is given by,

$$\begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}$$

The correlation between the main source X and helper Y is $\rho_{XY} = 1/\sqrt{3}$. The minimum rate at which Y needs to transmit the information is $R_Y = 1.5$. From (15), the rate at which X needs to transmit to achieve the desired distortion level is $R_X \geq 1.205$.

To illustrate the source separation, consider the multi terminal system shown in Figure 8. A portion of a distributed cluster of sensor nodes is observing a phenomenon and generating source data. Algorithms exist which can determine which nodes in the proximity of the phenomenon need to be activated and which can remain dormant [21], [20]. On the completion of a boot up process, one data node acts as the main data source (e. g. that which is closest to the phenomenon), and a

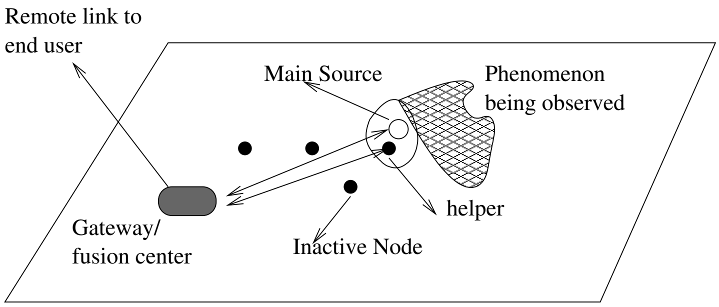

Fig. 8. Data fusion for Wireless Networked Sensor System

node close to main data source generates correlated data. Hence, the other node acts as a source of partial information at the decoder. Here, we consider an example of only one source but this can be easily extended for m sources.

Now it is also possible that for given values of m, p and D, the capacity of the network may be inadequate. However, by allowing the number of communication relays $n \gg p$ then the information can be extracted. Large over-provisioning will enable decoupling of source and channel coding.

V. Conclusion and Future Work

In this paper, we have discussed the issues of scalability, and extraction of the measurements of a physical phenomenon to a desired level of fidelity by sensor networks. We show that these issues can be tackled by considering the relative densities of sources, communication relays and sensors. The results are summarized in Table I.

Scalability can be controlled by making the S-D pair distribution peaked to local. When this occurs, the local interactions will dominate resources. Thus, the cooperative signal processing and communication problems are most profitably considered in these local domains - the typical interactions (that are application specific) thus may involve relatively small numbers of nodes.

TABLE I

SUMMARY OF MAIN RESULTS

Question Asked	**Source Density**	**Relay Density**	**Sensor Density**	**Solution**
Scalability?	Same	Same	Same	only if we make S-D pair distribution local
Information Extraction?	Fixed	Same	Same	only if correlation increases fast enough [6]
Scalability and Information Extraction?	Fixed	Traffic Dependent	Distortion Dependent	Yes by Source Separation

Further, even though the optimization may be intricate, it is feasible because of the small numbers. The adjustment in S-D pair distribution for scalability can be done directly through cooperative signal processing to bias high volume traffic to local destinations. Average delay may also be controlled by adjusting the S-D pair distribution by biasing towards closer nodes, but this does not help with peak delay.

For sensor networks, we believe the main objective is information extraction to some level of fidelity. We show that the objective of scalability and information extraction under a distortion criterion is attainable by the independent play of the relative densities of sensors, sources and relays. For this we consider a network of m Gaussian point sources, p sensors and n communication relays over a finite region and show that the rate-distortion bound for such a network can be found by a set of q-helper problems as $p \to \infty$. The case $n \gg p \gg m$ is easy to deal with. For such cases, we consider the decoupling of source, relays and sensor densities to achieve any distortion level D as $p \to \infty$. This is the idea of source separation. The source separation leads to the useful abstraction that the scale for specialized communications and signal processing for linkage to the physical world need not be large in the limit of high sensor density relative to the source density. At the networking and higher layers, the standard approaches for large scale networks can be utilized.

The cases where $n \approx p \approx m$ are much more difficult problems. This is because as we approach the critical sampling density (such as Nyquist sampling), larger scale interactions are required, and the communication and source coding become tightly coupled. If the number of sensors is limited, certain distortion values are not achievable regardless of the data rate available. Notice that both distortion and capacity must be considered jointly since the data rate must also be achievable, and if we use some efficient scheme, we might use fewer resources in extracting the necessary information. Such problems remain both open and interesting.

There also remain a large set of open problems with varying ratios of n, p and m. For example, given a distortion level (spatial and value), what is the minimum density of sensors required? There are many such resource optimization problems that will differ in character according to communication resources and source densities. Among the optimization parameters are energy, bandwidth and latency. The extension to non-Gaussian sources may also be challenging.

While in this paper we consider point sources, we believe that very similar results may be obtained for distributed sources subject to fidelity constraints.

ACKNOWLEDGMENT

The authors would like to acknowledge helpful discussions with Huiyu Luo.

REFERENCES

[1] P. Gupta and P. R. Kumar, “The capacity of wireless networks”, *IEEE Transactions on Information Theory*, 46(2):388-404, March 2000.

[2] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, “Instrumenting the world with wireless sensor networks”, *In Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 2033, Salt Lake City, UT.

[3] J. M. Kahn, R. H. Katz and K. S. J. Pister, “Emerging challenges: Mobile Networking for Smart Dust”, *Journal of Communications And Networks*, vol. 2(3), pp. 188-196, September 2000.

[4] D. Estrin, et. al., “Embedded, Everywhere: A Research Agenda for Networked Systems of Embedded Computers”, *Computer Science and Telecommunications Board (CSTB) Report*, National Academy Press 2001.

[5] M. Ahmed, “Decenteralized Information Processing in Wireless Peer-to-Peer Networks”, *Ph.D. Thesis*, UCLA 2003.

[6] S. D. Servetto, “On the Feasibility of Large-Scale Wireless Sensor Networks”, *In the Proceedings of the 40th Annual Allerton Conference on Communication, Control, and Computing*, Urbana, IL, October 2002.

[7] D. Marco, E. Duarte-Melo, M. Liu and D. L. Neuhoff, “On the Many-to-One Transport Capacity of a Dense Wireless Sensor Network and the Compressibility of Its Data”, *In Proceedings of International Workshop on Information Processing in Sensor Networks (IPSN)*, April 2003.

[8] P. Gupta and P. R. Kumar, “Internets in the Sky: The Capacity of Three Dimensional Wireless Networks”, *Communications in Information and Systems*, 1(1):39–49, 2001.

[9] P. Gupta and P. R. Kumar, “Towards an information theory of large networks: An achievable rate region”, submitted to *IEEE Transactions on Information Theory*, 2001.

[10] L-L. Xie and P. R. Kumar, “A network information theory for wireless communication: Scaling laws and optimal operation”, Submitted to *IEEE Transactions on Information Theory*, April 2002.

[11] Jinyang Li, Charles Blake, Douglas S. J. De Couto, Hu Imm Lee, and Robert Morris, “Capacity of ad hoc wireless networks”, *In Proceedings of the 7th ACM International Conference on Mobile Computing and Networking*, pp. 61-69, Rome, Italy, July 2001.

[12] M. Grossglauser and D. Tse, “Mobility increases the capacity of ad-hoc wireless networks”, *In Proceedings INFOCOM*, Anchorage, April 2001.

[13] Nikhil Bansal and Zhen Liu, “Capacity, Mobility and Delay in Wireless Ad hoc Networks”, *In Proceedings INFOCOM* 2003.

[14] M. Gastpar and M. Vetterli, “On the capacity of wireless networks: The relay case”, *In Proceedings INFOCOM*, New York, NY, June 2002.

[15] M. Gastpar, “To Code or not to Code”, *Ph.D. Thesis*, EPFL, Laussane, 2002.

[16] Leonard E. Miller, “Distribution of link distances in a wireless network”, *Journal of Research of the National Institute of Standards and Technology*, 106(2), March-April 2001.

[17] Leonard E. Miller, “Probability of a two-hop connection in a random mobile network”, *John Hopkins University 2001 Conference on Information Sciences and Systems*, March 2001.

[18] Chih-fan Hsin, “Ad Hoc Networks and Challenges”, *Paper Review in University of Michigan*, Fall 2001.

[19] Yasutada Oohama, “Gaussian Multiterminal Source Coding”, *IEEE Transactions on Information Theory*, 43(6):1912-1923, November 1997.

[20] T. M. Cover and J. A. Thomas, “Elements of Information Theory”, Wiley, New York, NY, 1991.

[21] A. El Gamal and T. Cover, “Multiple User Information Theory”, *Proceedings of the IEEE*, 68(12):1466–1483, December 1980.