COOPERATIVE MODELS FOR SYNCHRONIZATION, SCHEDULING AND TRANSMISSION IN LARGE SCALE SENSOR NETWORKS: AN OVERVIEW

Anna Scaglione, Yao-Win Hong

School of Electrical and Computer Engineering
Cornell University

ABSTRACT

What is the difference between classical remote sensing and sensor networks? What kind of data models that one can assume in the context of sensor networks? Can the sensors in the network concurrelty contribute to the sensing objective, without creating network conflicts? It is becoming apparent that methodologies designed to resolve network resource allocation conflicts in the communications among open systems have several bottlenecks when applied to sustain networkign among concurrent sensing nodes. Can we structure the network activities so that they are always directly beneficial to the sensing task?

The goal of this paper is to articulate these questions and indicate how some resource allocation conflicts can be removed embracing colaborative networking approaches among the sensors.

1. INTRODUCTION

Sensor networks are generally considered as suffering the curse of size. The capacity per square meter (i.e. the number of bits that is sent and successfully received over a square meter range) of a multi-hop network vanishes as an $O(\sqrt{N})$ as the number of nodes N becomes large [1]. Centralized architectures with a fusion center have even worse scaling laws, such as the $O(1/N)$ scaling performance derived in [6] or the $O(logN/N)$ attainable by utilizing antenna sharing schemes indicated in [7]. Thus, using non cooperative models in large scale sensors networks creates such a burden that it is crucial to reduce the number of sensors to the minimum needed.

Of course, mother nature often poses problems of great complexity and with such a large description that they are inherently ill posed, no matter how good is our design and for these problems we cannot expect good scaling laws. However, peculiarities of the sensor networks applications are that: 1) sensors are not end users, their identity is merely an attribute of the measurement - this implies that *fairness* is not a real constraint in the resource allocation; 2) if there is spatial redundancy in the system and sensors can *cooperate* rather than contend for the medium to deliver critical information; 3) the aggregate entropy/complexity of the data in space and time is typically low. In fact, sensor networks are deployed to record unexpected, dispersed events, sudden but localized changes, inconsistencies or discontinuities that occur rarely.

Henceforth, here we argue that *the curse of scale* is largely due to the ad-hoc approach followed in dividing functionalities, mostly that of communication and data processing, and the tendency of organizing activities serially rather than in parallel, vertically (or hierarchically) rather then horizontally. For all those data collection or inference problems that present a *sparse* structure, there is a great opportunity of turning the abundance of sensors as a resource to simplify the problem rather than complicating it and the key to do that is to allow cooperation. Next we make an attempt to characterize the type of problems that are of interest in the sensor network arena, putting them into perspective compared with the traditional approaches followed in the remote sensing literature and highlighting the new problems and the breakthroughs needed to have robust and efficient sensor network technology.

1.1. Remote sensing or remote sensors networks?

Remote sensing is hardly a new problem: it is a subject that had its breakthroughs in the early 1960's [2]. The remote sensing literature has put most of its emphasis on the data processing with limited attention to the issue of deploying sensors and getting data from them. For a long time the central problem has been that of extracting good estimates or inferences from few noisy sensors. Today, sensors are cheap, in some cases disposable and it is easier to deploy a large quantity of them. It is not atypical to have thousands of sensors and much fewer time snapshots to use for the lack of stationarity of the underlying process. At the other end of the spectrum is the sensor networking problem, which often ignores the final objective of the system. If a difference in the actual models can be pointed out is that in sensor networks' papers the sensing is often not assumed to be *remote* and, in fact, more and more often it is the end user that is

WORK SUPPORTED IN PART BY THE NATIONAL SCIENCE FOUNDATION UNDER GRANT CCR-0514243.

remote, not the sensors. Hence, in a nutshell, the difference between sensor networking and remote sensing is in the relative distance of the sensors from the end user and from the sensed environment. Interesting question that one may want to answer are as follows: "Is it better to send many small sensors to Mars that communicate back to Earth or to observe Mars with powerful sensors on Earth?"

1.2. Models for sensor networks

It is useful to make a coarse classification of the problem faced in the data collection. We denote by $(t, \ )$ the observation recorded at time t and location .

1- Fields with sparse representation: In this class of problems the observation recorded $(t, \ )$ are random or deterministic signals, which can be scalar observations (e.g. a temperature field) or vector observations (e.g. images); the key to address tractable data collection problems is to consider cases where the observation recorded have a sparse representation over some appropriate basis:

$$(t, \ ) = \sum_{k=1}^{M} C_k \ \ _k(t, \ ). \tag{1}$$

The expansion in (1) can be valid in the exact sense when assuming a deterministic model or can be interpreted as being equal to the signal in the mean square sense for a random model. Using a Gaussian random model and rate distortion theoretic arguments, for example, it was shown in [4] that the number of bits necessary to represent the sensor field can be shown to grow logarithmically in the number of sensors if the total distortion is fixed.

2- Fields with few sources in noise The case where $(t, \ )$ originates from a finite number of sources embedded in noise, i.e.:

$$(t, \ ) = \sum_{k=1}^{K} \ _k(t, \ ) + \ \ (t, \ ), \tag{2}$$

is the classical array processing setup where, assuming that the signals are narrowband, the problem is often reduced as follows:

$$(t, \ ) = A(\ ) \ (t) + \ \ (t, \ ), \tag{3}$$

with $A(\ )$ representing the so called array manifold matrix and (t) being the vector of source signals. The aim is the identification of the array manifold to localize the sources by separating the mixture of signals.

3- Fields with conditionally independent/Markov-process observations This model is mostly embraced for technical reasons and has two versions: 1) the first and simplest one postulate the independence of the random observations $(t, \ )$ with respect to the position, given a certain hypothesis H_i, $i = 1, \ldots, M$ (this is a classical model used in the distributed detection literature); 2) the second one introduces an ordering among the sensor positions and postulates that the observations form conditionally a Markov chain or, often, they are samples of a Gauss-Markov process.

4- Fields with random dynamical sources This model is considered when tracking sources that move in the sensor field. For simplicity, let us address the case of one source: the position of the target is considered equivalent to a state variable (t) and observation $(t; \ )$ such that

$$(t{+}1) = f(t, \ _t) + \ _k(t), \quad (t{+}1; \ ) = g(t, \ _{t+1}, \ ) + \ _k(t). \tag{4}$$

Under bandwidth and energy constraint there are two approaches that can be taken: 1) in the data collection process the transmissions map the data $(t, \ )$ into a representation $\hat{\ }(t, \ )$ that is used to perform some inference; 2) in the data processing the transmissions map $(t, \ )$ into an estimate or a decision $\hat{\ }$ on some parameters embedded in the model. The sensor network question is essentially how demanding this is in terms of network resources and if it can be done efficiently. Clearly, to fully answer the question one should quantify the performance loss associated with the lack of direct access to $(t, \ )$. The answers lie at the intersection of statistical signal processing, signal compression and communication/information theory. One emerging trend is that of introducing collaboration among the sensor nodes to speed up and render more power efficient the data gathering. Although the interplay between cooperation among nodes and performance gain is not fully understood, as discussed in the introduction, there are clear indications that the traditional networking is too limiting and not scalable.

In the following two sections we discuss how cooperation can lead to very rapid network synchronization and scheduling that achieves cooperative gains in multi-hop broadcast networks.

2. COOPERATIVE SYNCHRONIZATION

One of the key issues in any collaborative activity is that the system has to be able to rapidly achieve a common sync. This allows the sensors to pace their activities so as to not interfere with each other. Most synchronization protocols proposed for networks operate at the packet-level, where the time information is exchanged explicitly among nodes, and they require the processing of these messages to calibrate the time difference between their local clock. Because these schemes are evolutions of synchronization protocols used over the Internet, they use point-to-point transmissions instead of capitalizing on the broadcast nature of the wireless channel. There are essentially three families of approaches: 1) Centralized methods, utilizing the Global Positioning System (GPS) [8] to synchronize the network to an external timescale. The accuracy is in the order of 200nsec but the cost of the hardware is high and it needs a line-of-sight to a GPS satellite, which is unrealistic in many practical settings; 2) Master-slave methods, such as the Mills Network Time Protocol (NTP) [9] widely used over the Internet. In the NTP, a hierarchy is constructed among nodes in the network with multiple roots synchronized to external global time sources (e.g. the GPS) and all the other nodes synchronized to its corresponding parent-node; in NTP the accuracy is strongly affected by the symmetry of the delays in the transmission paths between the transmitter and receiver node and it also degrades significantly when the level of the node is low within the hierarchy and these problems are retained in the wireless adaptations of the NTP protocol such as in [10]; 3) Decentralized methods, such as the Reference Broadcast Synchronization (RBS) scheme [10], where each node sends reference beacons to neighbors using a physical-layer broadcast. The nodes that received the beacon exchange the arrival time of the beacon relative to the local clock and obtain a time difference matrix and clock skew information. RBS eliminates the contribution of the random processing time spent at the transmitter node due to the protocol processing and the variable delays of the operating system and therefore is more accurate than NTP. However, the RBS scheme requires a large amount of data exchange which makes it non-scalable.

A network synchronization protocol that is totally decentralized, and where is no estimate or exchange of times stamps or multiple access control involved since it operates exclusively at the physical layer is that proposed in [12]. We are not aware of any other alternative that shares the scalability of this method. It essentially emulates a mechanism that explains the ability to sync observed in several biological networks. The synchronization is not achieve by estimating clock timing from other nodes, it is the result of the dynamics of a set of *coupled* non linear differential equation that evolve concurrently in the nodes. In the ideal system $x_i(t)$ is the state variable of node i and the system dynamics are regulated by:

$$\frac{dx_i(t)}{dt} = -\gamma x_i(t) - \frac{\delta(x_i(t)-1)}{|\ddot{x}_i(t)|} + \epsilon \sum_{j \neq i} \frac{\delta(x_j(t)-1)}{|\ddot{x}_j(t)|} \quad (5)$$

the term $\epsilon\delta(x_j(t)-1)/|\ddot{x}_j(t)|$ is valid for $j \neq i$ and is called the coupling. The concurrent activity of these equations has only a stable point, which is that where all states $x_i(t)$ evolve identically. It is easy to see that when the states are the same, the coupling has no effect since it is concurrent with the contribution to the equation that resets the state to zero $\frac{\delta(x_i(t)-1)}{|\ddot{x}_i(t)|}$.

In a sensor system the access to the other node state variable can be only through a noisy observations of their signals and, while the [12] partially addressed this issue, it is important to derive design principles that can be applied to more general synchronization objectives and transmission interfaces. In particular, it is of great interest: 1) to develop robust mechanisms, in continuous or discrete time, that provide the same effect of (5); 2) to investigate the fundamental performance limits and means square error of decentralized synchronization systems where the coupling is affected by random noise. It is, in fact, not straightforward to compare the performance of this decentralized clock generation system with formal Cramér Rao bounds for time estimation.

3. COOPERATIVE DATA DRIVEN TRANSMISSION

How difficult is to extract information which is distributed? The concept of communication complexity is formally defined in as the problem of minimizing the binary messages that are involved in the computation of a function $f(X,Y)$ where X and Y are known originally by only one of the two terminals. While wondering if functions were computable with a minimal sequence of messages among terminals, Orlisky and El Gamal [13] noticed that, quote: *"Considerable insight has been obtained by noting that sequences are compressible to their entropy, notwithstanding the fact that most sequences are incompressible because their entropy is maximal. Does a similar property hold for communication complexity?"*. Most of the theoretical work on communication complexity and data querying considers strategies that are applicable to "wired" parallel computers but that are not suitable for a wireless sensor network. On the other hand, it is well-known [3] that the separation of source and channel coding is not optimal and that feedback improves the capacity of the multiple access channel.

Data driven communication is a keyword that is seldom used in the literature on physical layer transmission. It essentially means that the data type that the sensor node acquires automatically assigns a channel to transmit. In [11] we argued that to reach minimum complexity in the data collection it is necessary to formulate the sensor communication problem as the construction of a *guessing game* where each query activates the transmission of a code from collaborating sensors and the answer to the query is the feedback that allows to formulate the next most likely guess. The optimization objectives in designing the communication protocols can be that of minimizing the number of queries to groups of nodes (time efficiency) or invoking the answers from a minimum number of sensors (space efficiency) or requiring the minimum total energy for the acquisition (energy efficiency). This approach would naturally lead to a data and computation driven multiple access scheme. A special case of this class of strategies is what we refer to as the Group Testing Multiple Access (GTMA) [11]. In general, let S_l be the time slot interval at which the l-th test is performed and let $B_{l,i} = \mathcal{B}[T_l, X_i, Z_{l-1}, \ldots, Z_1]$ be the ith node answer, function of the test T_l the sensor data X_i and the previous feedback $Z_{l-1}, \ldots, Z_1$. The short-

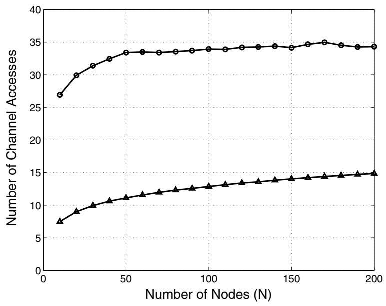

Fig. 1. The rate-distortion lower bound (△) and the average number of tests (•) for our data driven data collection algorithm. Each sensor sample is quantized into a 8 bit message where the quantizer contains two 4-level uniform quantizers, within the range $[-2\sigma^2, 2\sigma^2]$, that are biased by the amount $\sigma^2/2$ (see [11] for details). The vector of samples   has mean 0, is spatially stationary and has covariance matrix $R_V(\xi,\nu) = sinc(\frac{\pi\xi}{N})sinc(\frac{\pi\nu}{N})$ where $\xi,\nu = -N+1,\cdots,0,1,\cdots,N-1$.

est slot duration $|S_l|$ is the Nyquist limit $1/W$ where W is the signals bandwidth, although in general $|S_l|W \geq 1$. The baseband complex equivalent model of the received signal in multi-path (convolution channel $h_i(t)$) and additive Gaussian noise $w(t)$ is in general a mixture of signals, with individual powers P_i, that can be written as $x_l(t) = \sum_{i\in U_l} h_i(t) * \sqrt{P_i} s_i(t; B_{l,i}) + w(t)$, $t \in S_l$. The discrete-time representation of the received signal in is the vector ($_i$ are convolution matrices and $_l \sim \mathcal{N}(\mathbf{0}, N_0\ )$):

$$_l = \sum_{i\in U_l} \ _i\sqrt{P_i}\ \ _i(B_{l,i}) + \ \ _l \qquad (6)$$

The feedback is $Z_l = \mathcal{D}(\ _l|Z_{l-1},\ldots,Z_1)$ where $\mathcal{D}(.)$ is the decision rule. The design of the optimum set of encoded signals $s_i(t; B_{l,i})$ is not a classical multiple access problem. If the signals were designed to allow the receiver to resolve all X_i without the feedback Z_k from the previous tests and without the selection of a specific group U_l, the scheme would be equivalent to polling the sensors individually. In our preliminary results in [11] we considered a rudimentary, yet effective, memoryless coding technique and *linear* modulation, which can be summarized saying that 1) the answer $B_{l,i} = \mathcal{B}[T_l, X_i]$ has no memory of the previous feedback and it is a simple boolean answer (yes/no $\rightarrow$ $B_{l,i} = 0/1$); 2) $s_i(t; B_{l,i}) = B_{l,i}p(t) \rightarrow \ _i(B_{l,i}) = B_{l,i}$ and, 3) the detection rule uses $_l$ to detect the presence or absence of at least one positive answer $B_{l,i} = 1$, i.e. in this case $\mathcal{D}(\ _l; Z_{l-1},\ldots,Z_1) = \mathcal{D}(\ _l) = \{||\ _l||^2 \geq \tau\}$, where τ is an energy detection threshold. Nodes that reply $B_{l,i} = 1$ indicate to the receiver that the test was a wrong guess and that the search needs to proceed. Fig. 1 gives the average number of tests required by one instance of the family of algorithms we just described [11] and the corresponding the rate-distortion lower-bound (we assume noise free reception). The sensors are spread on a uniform lattice in a unit area. We observe that the average number of tests (channel uses) for the algorithm does not increase linearly with N which is the number of sensors. In fact, the growth rate increases similar to that of the rate-distortion bound, although a degradation in performance is observed due to its sub-optimal structure.

4. REFERENCES

[1] P. Gupta and P. R. Kumar. The Capacity of Wireless Networks. *IEEE Trans. Inform. Theory*, 46(2):388–404, 2000.

[2] Harry L. Van Trees, *Detection, Estimation, and Modulation Theory: Radar-Sonar Signal Processing and Gaussian Signals in Noise*, Wiley-Interscience; Reprint edition.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 1991.

[4] A. Scaglione and S. Servetto, "On the interdependence of routing and data compression in multi-hop sensor networks," in *Proc. of Mobicom 2002*, Atlanta, GA, Sept. 2002.

[5] Feng Zhao, Jie Liu, Juan Liu, Leonidas Guibas, and James Reich. Collaborative signal and information processing. Proc. IEEE, 91(8):11991209, August 2003

[6] 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, ISPN 2003, Palo Alto, CA, Apr. 2003.

[7] H. E. Gamal, On the scaling laws of dense wireless sensor networks, to appear in IEEE Trans. on Inform. Theory, 2005.

[8] J. Mannermaa, K. Kalliomaki, T. Mansten, and S. Turunen, "Timing performance of various GPS receivers," in *Proc. of the 1999 IEEE IFCS*, Apr. 1999, pp. 287–290.

[9] D. L. Mills, "Internet time synchronization: The network time protocol," *IEEE Trans. Commun.*, vol. 39, no. 10, pp. 1482–1493, Oct. 1991.

[10] J. Elson, L. Girod, and D. Estrin, "Fine-grained network time synchronization using reference broadcasts," in *Proc. of the 5th OSDI*, Dec. 2002.

[11] Y.-W. Hong and A. Scaglione, "Content-based multiple access: Combining source and multiple access coding for sensor networks," in *Proc. MMSP 2004*, Siena, Italy, Sept. 2004.

[12] ——, A Scalable Synchronization Protocol For Large Scale Sensor Networks And Its Applications, IEEE JSAC, Vol. 23, Issue 5, May 2005 pp. 1085 - 1099.

[13] A. Orlitsky and A. E. Gamal, "Average and randomized communication complexity," IEEE Trans. Inform. Theory, vol. 36, no. 1, pp. 3-16, Jan. 1990.