On the relation between Source and Channel Coding and Sensor Network Deployment

Iordanis Koutsopoulos[1], Stavros Toumpis[2†] and Leandros Tassiulas[1]
[1]Department of Computer Engineering and Communications, University of Thessaly, Greece, and Center for Research and Technology Hellas (CERTH), emails: {jordan,leandros}@uth.gr
[2]Telecommunications Research Center Vienna (ftw.), email: toumpis@ftw.at

***Abstract*— In this paper, we identify analogies between the themes of source and channel coding and some problems that arise in the context of wireless sensor networks. Our aim is to establish a framework within which well-known methods from the former two areas are used in tackling problems associated to sensor networks. First, we address an important tradeoff between required precision and transmission rate in a class of sensor networks by using rate-distortion theory. We then address the problem of optimal placement of super-sensors in an area covered by sensors with the objective to minimize a generic cost factor that captures several special cases, and we show the analogy to a vector quantization problem. Based on the analogy of the resulting system with a discrete source emitting symbols, we show that a traffic load balancing problem in the sensor network can be reduced to an entropy maximization problem. Finally, we consider a hierarchical sensor coverage problem that involves deploying a set of sensors with sophisticated sensing capabilities over a grid of ordinary sensors. We cast the problem in the framework of channel coding by defining appropriate analogies under the common denominator of redundancy.**

I. INTRODUCTION

Sensor networks consist of several generally stationary battery-powered sensors, which are constrained in physical size and processing capabilities and form wireless networks with the objective of monitoring a certain area and communicating information created there at a central site[1]. Sensor networks have received an unprecedented amount of interest recently from the research community as well as the industry due to their wide range of applications. Such applications include military communications, disaster recovery, surveillance or intrusion detection, environmental or other process monitoring, health-care and industrial quality or inventory control, to name a few.

Sensor networks should be deployed and operated so that the following objectives are met: (i) the sensors cope with bandwidth limitations especially in the case of transfer of bandwidth-demanding information, (ii) they must use their limited energy resources judiciously and thus extend the network lifetime, (iii) they must monitor or measure and represent a physical process subject to some specified accuracy, and (iv) the network must be able to handle a certain amount of sensor failures by creating a fault-tolerant system.

Work partly supported by **K***plus* funding for the ftw. project I0 "Signal and Information Processing", and by the European Network of Excellence on Wireless Communications (NEWCOM).

†Corresponding author, email: toumpis@ftw.at

The authors of [2] discuss distributed signal compression by exploiting the inherent spatial or temporal correlation among sensor output data and indirectly address objective (i). Regarding objective (ii), in [3], the problem of routing traffic from monitoring nodes to gateways so as to maximize network lifetime is considered. The solution involves balancing energy consumption among monitoring nodes. The work in [4] tackles the multi-objective problem of improving data quality in terms of distortion and increasing network lifetime by performing iteratively the steps: (a) for fixed sensor placement, find flows (i. e., a routing discipline) from sensors to the sink [3], (b) for fixed flows, find a new sensor placement that improves lifetime and does not increase distortion and (c) find a sensor placement that improves distortion and does not increase lifetime. Objective (iii) per se is the topic of [5] and objective (iv) is studied in [6].

Clustering is another technique that is oftentimes employed in sensor networks in view of the aforementioned objectives. The hierarchical structure imposed by clustering techniques leads to significant reduction of consumed energy, since nodes have to relay traffic to their corresponding cluster-heads and do not need to employ long-range transmissions. A direct consequence of clustering is the elongation of network lifetime. Several centralized or distributed clustering approaches have been proposed in the literature (see [7] and references therein). Clustering methods often aim at creating clusters such that nodes within a cluster are within one [8] or $d > 1$ hops [9] from their cluster-heads, while attempting to balance the load among cluster-heads and reduce the number of clusters in the network.

This paper constitutes an attempt to view some fundamental problems that arise in sensor networks within a source- or channel-coding framework. The idea is first to draw the analogies among the research areas and then to utilize well-known methods and theory from the latter two areas in order to provide solution to problems related to sensor networks.

To this end, we first consider a flat sensor network architecture where all nodes belong to the same class in the sense of having the same (limited) processing capabilities. We discuss the tradeoff of obtaining accurate measurements versus reducing the number of bits that need to be communicated and show its analogy to the information-theoretic tradeoff of rate-distortion theory.

Next, we consider the situation where the available sensors are of two types: ordinary monitoring nodes of limited size, energy and processing capabilities and more powerful sensors, so called "super-sensors". An interesting arising issue is the problem of optimally placing a set of super-sensors over a deployed topology of ordinary sensors with the objective to minimize a generic cost that captures energy cost for communication between ordinary sensors and super-sensors. We show that this problem can be viewed in the context of vector quantization, where the communication cost has similar role to that of distortion resulting from quantization. We also explain how the problem of maximizing source entropy can be mapped to an associated problem in the context of sensor networks.

Finally, we consider an application of channel coding in a problem that arises in sensor networks, where each sensor is subject to failure with some probability. The question is to identify the amount of redundancy that needs to be added to the system (in the sense of additional sensors) so as to achieve a certain level of performance, which in this case is described by the probability of error in measurements. In Section II we present an overview of vector quantization and in Sections III, IV and V we discuss the above mentioned problems. Section VI concludes our paper.

II. Overview of Vector Quantization

Vector quantization can be viewed as a generalization of scalar quantization. Given a set of n-dimensional continuous-valued vectors distributed based on some joint probability density function (p.d.f.) $p(\mathbf{x})$, the problem is to find quantization vector levels $\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_L$ so as to minimize the expected distortion between the continuous-valued vectors and quantization vector levels. The distortion between the actual point $\mathbf{x}$ and quantized point $\hat{\mathbf{x}}$ is $d(\mathbf{x}, \hat{\mathbf{x}})$, where $d(\cdot,\cdot)$ is the distortion measure, which could be, for example, the ℓ_p-norm, with p a positive integer. For illustrative purposes we restrict ourselves to the case $n = 2$. Let $Q(\cdot)$ denote a function such that $Q(\mathbf{x}) = \hat{\mathbf{x}}$ is the vector quantization level to which the point $\mathbf{x}$ is mapped. As shown in Fig. 1, vector quantization results in a division of the two-dimensional plane into cells C_k, $k = 1, \dots, L$, so that $C_k \triangleq \{\mathbf{x} : Q(\mathbf{x}) = \hat{\mathbf{x}}_k\}$.

The probability that $\hat{\mathbf{x}}_k$ is the output of quantization is $p_i \triangleq P(\mathbf{x} \in C_k) = \int_{C_k} p(\mathbf{x})\, d\mathbf{x}$, and the average distortion is

$$
\begin{aligned}
D &= \sum_{k=1}^{L} P(\mathbf{x} \in C_k) E[d(\mathbf{x}, \hat{\mathbf{x}}_k) | \mathbf{x} \in C_k] \\
&= \sum_{k=1}^{L} \left(\int_{C_k} p(\mathbf{x})\, d\mathbf{x} \right) \int_{\mathbf{x} \in C_k} d(\mathbf{x}, \hat{\mathbf{x}}_k) p(\mathbf{x}) d\mathbf{x}. \qquad (1)
\end{aligned}
$$

It can be shown [10] that the mapping of the continuous-valued points $\mathbf{x}$ into cells that minimizes the expected distortion is performed based on the nearest-neighbor rule, that is $Q(\mathbf{x}) = \hat{\mathbf{x}}_k$ if and only if $d(\mathbf{x}, \hat{\mathbf{x}}_k) \leq d(\mathbf{x}, \hat{\mathbf{x}}_j)$, for $k \neq j$, $j = 1, \dots, L$. In addition, the vector quantization level for a cell should be selected so that, assuming the boundaries of the cell are fixed, the average distortion in the cell C_k is minimized.

The optimal vector quantization problem has been studied extensively in the literature, and several algorithms have been

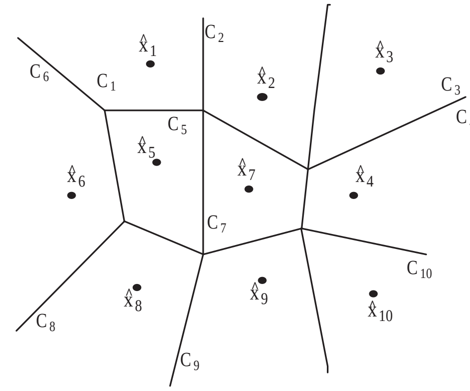

Fig. 1. Quantization points $\hat{\mathbf{x}}_k$ and quantization regions C_k.

proposed. The so called Lloyd algorithm is simple and intuitive and its main steps are outlined as follows [10]:

Step 1: Set iteration number $i = 0$, $D(0) = \infty$, and arbitrarily pick a set of output vectors $\hat{\mathbf{x}}_k(i)$.

Step 2: Set $i = i + 1$. Create the regions $C_k(i)$ by mapping each point $\mathbf{x}$ to the closest $\hat{\mathbf{x}}_k(i)$.

Step 3: Select new quantization points $\hat{\mathbf{x}}_k(i)$. Each point $\hat{\mathbf{x}}_k(i)$ is chosen so that average distortion is minimized, under the constraint that its corresponding region C_k remains fixed and equal to $C_k(i)$.

Step 4: Calculate the average distortion $D(i)$ by using (1) with the quantization points $\hat{\mathbf{x}}_k(i)$ and corresponding regions $C_k(i)$. If $D(i-1) - D(i) > \epsilon$, go to Step 2, else terminate algorithm.

The algorithm is guaranteed to converge to a local minimum of the distortion. By repeated applications of the algorithm with different initializations, it is possible to arrive at the global minimum of distortion.

III. Rate-distortion tradeoff in a flat sensor network architecture

In this section, we discuss the fundamental problem of information representation in a sensor network and associate it to rate-distortion theory. As shown in Fig. 2, we consider a wireless sensor network deployed over a large region whose task is to sense the occurrence of a sequence of events. We limit ourselves to the case where the information pertains to location and time data (e.g as in the cases of monitoring a forest for the event of fire and battlefield monitoring for intruders).

In order to preserve their limited energy resources, it is important for the nodes to transmit the information with as few bits as possible, since energy consumption is proportional to the number of transmitted bits. Another reason that motivates the use of few bits for sensor transmission is the fact that communication between the sensors and the central site may often take place over a common wireless link of certain capacity. That is, a certain maximum number of bits can be transmitted from all sensors through the link. The bandwidth of the link is shared among transmitting sensors based on some scheduling principle. In addition, due to the time-varying link quality, link capacity fluctuates and in some time periods when

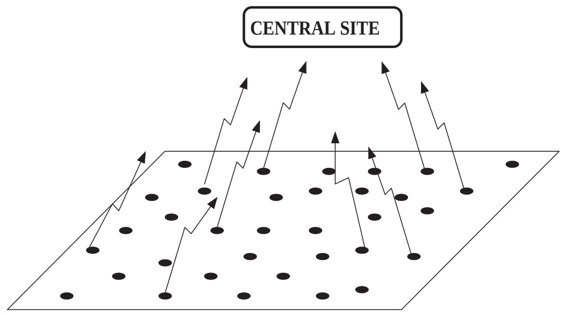

Fig. 2. Our setting: A wireless sensor network is deployed over a region, with the aim of sensing a sequence of events, and reporting their location to a central site. Communication is over a common wireless link of limited capacity.

channel quality is deteriorated, the number of bits that can be accommodated in the link is even more limited. Lastly, sensor devices are inherently weak, and minimizing the number of bits they must transmit simplifies their design.

However, the attempt of sensors to convey information with few bits will affect the precision of transmitted data, since representation of data with fewer bits brings along the notion of distortion. Namely, a tradeoff exists between the need to preserve energy resources and that of transmitting information with certain accuracy. The situation is analogous to the basic issue raised in rate distortion theory [11], where the problem is to convey the output of a random source of information by using as few information bits as possible, and with introducing as little distortion as possible. The basic question of rate distortion theory is: given a maximum acceptable distortion, find the minimum rate representation of the source that achieves it. Or equivalently, for a given rate, find the minimum amount of achievable distortion.

We now specify our setting in detail. The arrival of events is modeled as a point process $\{\mathbf{X}_n\}$ over a two-dimensional region $\mathcal{A}$. Each realization of the process, $\{\mathbf{x}_n\}$ is a sequence of points that corresponding to event occurrences. Events appear independently of each other at periodic time intervals and an event appears in an area of size $d\mathbf{x}$ around point $\mathbf{x}$ with probability $p(\mathbf{x})d\mathbf{x}$, where $\int_{\mathcal{A}} p(\mathbf{x})\, d\mathbf{x} = 1$, with $p(\cdot)$ representing a event occurrence spatial probability density function (p.d.f). We assume that the region is covered with a large number of sensors that can determine the exact location of the event with high accuracy. In order to preserve energy by reducing the number of transmitted, a sensor may report as $\hat{\mathbf{x}}_k$ the location of the k-th event whose actual location is $\mathbf{x}_k$. Then, a distortion $d(\mathbf{x}, \hat{\mathbf{x}})$ is incurred, where $d(\cdot,\cdot)$ is the distortion measure, which for example can be the squared-error distortion or the ℓ_p-norm distortion, with p a positive integer. Let$\{\hat{\mathbf{X}}_n\}$ be the point process reported by the sensors. The distortion between the actual and reported process over a block of K events is

$$d(\mathbf{X}_n, \hat{\mathbf{X}}_n) = \frac{1}{K} \sum_{k=1}^{K} d(\mathbf{x}_k, \hat{\mathbf{x}}_k) \quad (2)$$

and since $d(\mathbf{X}_n, \hat{\mathbf{X}}_n)$ is a random variable, its expected value is defined as the distortion D.Assume now that a sensor uses R bits per event on average. A rate distortion pair (R, D) is called *achievable* if it is possible to represent event locations using on the average R bits per event while incurring average distortion D. The closure of the set of all achievable rate distortion pairs is the rate distortion region. The *rate distortion function* is defined as the minimum of rates R so that pair (R, D) is in the rate distortion region for a given distortion D. Similarly, the *distortion rate function* $D(R)$ is defined as the minimum of distortions D such that (R, D) is in the rate distortion region for a given rate R. These notions capture the optimal tradeoff between the number of bits used by a sensor in order to represent an event and the precision in the description.

Under this formulation, our problem becomes the calculation of $R(D)$ (equivalently $D(R)$) for a given distribution $p(\cdot)$ and a given distortion function $d(\cdot,\cdot)$, and the determination of the strategy that achieves them. This problem has already been studied extensively in the past, in the context of rate distortion theory [11]. In that setting, it is typically assumed that the process $\{\mathbf{X}_n\}$ is a signal created, and optimally distorted, locally. In our case, the creation, and subsequent distortion of the signal, happen on a network wide basis, and this differentiation must be taken into account.

In general, the rate distortion curve is achieved only by the use of a **rate distortion code** [11] that codes a group of symbols $\{\mathbf{X}^k\}$, $k = 1, \ldots, K$, to a group of symbols $\{\hat{\mathbf{X}}^k\}$, $j = 1, \ldots, K$, with an average distortion that approaches D, and a bit rate that approached R, as $K \to \infty$. In our setting however, as already discussed, the input symbols of $\{\mathbf{X}_n\}$ will have to be distorted locally. Therefore, the rate distortion function will not be in general achievable, but will represent an upper bound to the performance of the wireless sensor network.

In its task to represent the required information with the minimum number of necessary bits *locally*, a sensor can apply quantization. As discussed in Section II, this entails the mapping of the actual location of an event $\mathbf{x}$ into one of L discrete values $\hat{\mathbf{x}}_\ell$, $\ell = 1, \ldots, L$, the quantization levels. Assuming that the spatial p.d.f. of the process $\{\mathbf{X}_n\}$ is known, the optimum quantizer results in a rate $R = \lceil \log_2 L \rceil$ bits per event and distortion $D(R)$. The corresponding distortion-rate curve is the uppermost one in Fig. 3).

However, more efficient encoding can be performed, in the sense that we can further reduce the number of transmitted bits per event for the same distortion, or equivalently decrease distortion for the same rate (see middle curve in Fig. 3. Indeed, quantization produces a discrete information source $\{\hat{\mathbf{x}}_n\}$, with $\hat{\mathbf{x}}_n \in \{\hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_L\}$, where each quantization level $\hat{\mathbf{x}}_i$ has probabilities of appearance p_i, which comes as an output of the quantization procedure. By applying source coding techniques that are known as entropy coding, a sensor can reduce the number of transmitted bits. The underlying principle of entropy coding techniques, such as Huffman coding is the following: in order to reduce the average number of bits per symbol (event in our case), fewer information bits should be used to describe symbols (in our case events) that have higher probability of occurrence. Note that the bottom curve corresponds to the distortion-rate function $D(R)$. As

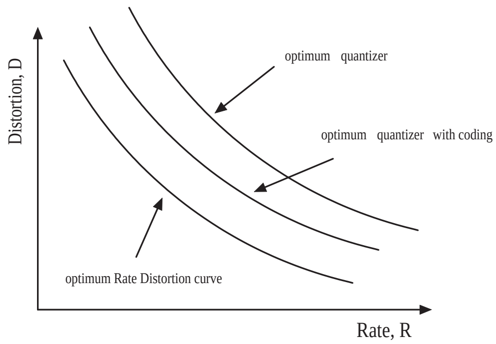

Fig. 3. The rate distortion curve for quantization, quantization followed by entropy coding and for a Gaussian continuous-amplitude source.

discussed, this curve represents the information theoretic lower bound on distortion for a given rate, or equivalently the lower bound on required rate for a given distortion.

It should be noted that the framework described above refers to the tradeoff between transmission rate and accuracy of transmitted data for the case of one measuring sensor. In the more general case, several sensors sense an event and each sensor i is limited to a transmission rate of r_i bps. Each sensor does not collaborate with others and transmits its observation with different rate to a central site, which subsequently needs to reconstruct the observed event process within the prescribed distortion bounds. The objective is to find a sensor transmission vector with minimum total rate subject to a maximum distortion constraint. This brings upon an information-theoretic multi-terminal source coding problem which is referred to as the CEO problem [12]. In the CEO problem, a source is observed by several observers and independent messages are sent by them to another agent, the CEO. This latter agent attempts to recover the source to meet a fidelity constraint. It is also interesting to investigate the opportunity of distributed selection of sensors that will transmit the event process subject to the aforementioned constraints.

IV. Optimal Placement of super-sensors as a Vector Quantization problem

We now consider a sensor architecture that consists of two types of sensors. The first one includes cheap, ordinary sensors of small physical size having stringent energy constraints and limited processing and communication capabilities. Their only task is to sense the environment. A function $p(\mathbf{x})$ captures the information generation density (in terms of bps/m^2) from these sensors as a function of position $\mathbf{x} \in \mathcal{A}$, where $\mathcal{A}$ is the coverage area. The second type of sensors is referred to as super-sensors and includes devices of larger physical dimensions with higher processing and communication capabilities and more energy resources. We are given a certain number of super-sensors. Instead of transmitting information traffic to the operations center, each sensor relays information to its closest super-sensor, which has a larger transmission range and is responsible for forwarding the traffic further. Sensors are much more densely deployed that super-sensors.

Let $\mathbf{x}_i$ be the location of the i-th super-sensor and consider a sensor in position $\mathbf{x}$. Let $c(|\mathbf{x}_i - \mathbf{x}|)$ be the cost of sending a transmission from position $\mathbf{x}_i$ to position $\mathbf{x}$. The cost function $c(\cdot)$ can capture several kinds of communication cost between a sensor and a super-sensor. For example, if $c(|\mathbf{x}_i - \mathbf{x}|) = |\mathbf{x}_i - \mathbf{x}|^\alpha$, where α is a constant, it captures the energy consumed for the single-hop information transfer from a sensor to a super-sensor. To see this, assume that the propagation mode between a sensor and a super-sensor includes only path loss with attenuation exponent α. Let simultaneous transmissions from several sensors towards the same super-sensor be resolved with a scheduling protocol that allocates transmission turns to sensors, so that collisions at the super-sensors are avoided. Also, let us neglect the interference from transmissions of sensors to neighboring super-sensors. In this admittedly simplistic model, the SNR at the receiver of the super-sensor i at distance $|\mathbf{x}_i - \mathbf{x}|$ from the sensor is $P/(N|\mathbf{x}_i - \mathbf{x}|^\alpha)$, where P is the transmission power and N is the receiver thermal noise power. Clearly, in order to achieve a given SNR requirement at the super-sensor, the sensor should transmit with power proportional to $|\mathbf{x}_i - \mathbf{x}|^\alpha$. If the cost function takes the form $c(|\mathbf{x}_i - \mathbf{x}|) = |\mathbf{x}_i - \mathbf{x}|$, then it is proportional to the distance between a sensor and a super-sensor and can roughly denote the number of hops in which a sensor reaches a super-sensor. Finally, if the cost is exponential in the number of hops, i.e $c(|\mathbf{x}_i - \mathbf{x}|) = \left[1 - (1-w)^{|\mathbf{x}_i - \mathbf{x}|}\right]$, it captures the probability of error in packet transmission from the sensor to the super-sensor, where w is the probability of packet transmission error over one hop and errors occur independently in different hops.

The problem then is as follows. Given the information density function $p(\mathbf{x})$ and a number of super-sensors L, find the locations of super-sensors $\mathbf{x}_1, \ldots, \mathbf{x}_L$ that minimize total communication cost between sensors and super-sensors. The problem can be seen to be equivalent to that arising in the context of vector quantization, where the objective is to identify vector quantizer levels that minimize distortion. The outcome of the vector quantization is the division of area $\mathcal{A}$ in L areas (or clusters) C_i, $i = 1, \ldots, L$, so that the total cost

$$C = \sum_{i=1}^{L} \int_{C_i} c(|\mathbf{x}_i - \mathbf{x}|) p(\mathbf{x}) d\mathbf{x} \tag{3}$$

is minimized.

As was mentioned in Section II, the Lloyd algorithm results in local minima in the optimization problem above, while repeated runs with different starting points can lead to closer approximations of global minima as well. Another methodology that can aid in avoiding local minima is that of deterministic annealing [13]. The quantization distortion as outcome of the algorithm can be directly mapped to the total cost required in order to communicate messages from sensors to super-sensors. At the end of the algorithm, a partition of the area has been achieved and each super-sensor is allocated with some percentage q_i of the information rate of the network, where

$$q_i = \int_{C_i} p(\mathbf{x}) d\mathbf{x}. \tag{4}$$

Clearly, this is a measure of the total rate load that super-sensor i will receive from sensors that belong to its cluster. A percentage vector $\mathbf{q} = (q_1, \ldots, q_L)^T$ corresponds to the set of L super-sensors. This vector can be considered to be the probability mass function of a discrete memoryless source emitting L possible source symbols, where the probability of occurrence of i-th symbol is q_i. The entropy of this source is,

$$H(\mathbf{q}) = -\sum_{i=1}^{L} q_i \log q_i \tag{5}$$

We now proceed to another analogy. The portion of rate q_i that corresponds to super-sensor i reflects the intensity of traffic load that super-sensor i will receive from sensors. A meaningful objective in that case would be to find an allocation of sensors to super-sensors such that the load is balanced among super-sensors. This would come at the expense of a larger communication cost. Balancing the load of super-sensors can be pursued in order to prevent overload events or even preserve energy resources of super-sensors. Balancing the load of super-sensors is equivalent to maximizing the entropy of a symbol-emitting source, which is achieved when all the q_i are equal. It would be interesting to devise an algorithm that solves this optimization problem with the additional constraints $q_1 = \ldots = q_L$. It would also be worthwhile to study the tradeoff between balancing the load of super-sensors and reducing energy consumption. We are currently investigating these research issues.

V. A CHANNEL CODING FRAMEWORK FOR A SENSOR NETWORK PROBLEM

We now focus on another problem that arises in the context of sensor network deployment. Consider a two-dimensional $N \times N$ grid representing an area to be covered with sensors. An event occurring in a certain square (i, j), namely in the i-th row and j-column of the grid is modeled by a Bernoulli random variable $X_{i,j}$. Therefore, $X_{i,j}$ equals 1 or 0, with respective probabilities q and $1-q$ which may or may not be known. The set of all events occurring in the area is modeled by a two-dimensional matrix process $\{\mathbf{X}(n)\} = \{X_{i,j}(n) : i, j = 1, \ldots, N\}$, where n is the time index. A sensor is placed at the center of each square with the aim to track $X_{i,j}(n)$. The sensing range of the sensor is assumed to be exactly the required one so as to sense the square where the sensor is appointed.

Due to the fact that sensor devices are prone to failures, certain sensors covering some squares may malfunction and thus will not able to provide accurate information about the local event in this square. Sensor malfunction may happen occasionally (i.e. at specific time instances) or may be the result of permanent failure. Let p denote the probability of malfunction of a sensor. The objective of sensor deployment is to accurately track the event process that takes place within the area. To this end and in order to overcome the sensor malfunction problem, a solution would be to utilize additional sensors. For instance, if additional sensors were allocated with the task of sensing each square, the probability of correct detection of the event in the square would increase due to the introduced redundancy. However, this solution would incur the use of a significantly larger number of sensors than what was originally planned for the area.

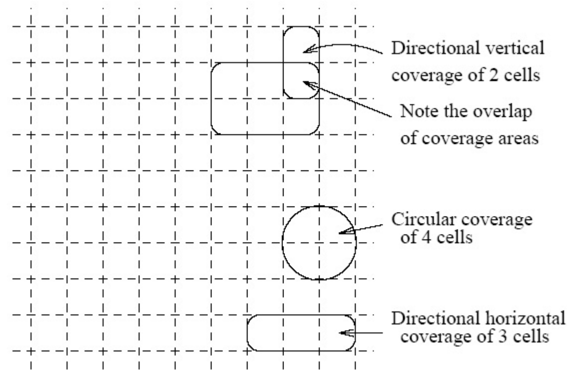

Fig. 4. A two-dimensional grid with circular or directional coverage areas of sophisticated sensors.

Consider now the alternative of using a few additional sensors which are more sophisticated in terms of their sensing abilities. For instance, a sensor with greater receiver sensitivity would have greater sensing range and would be able to sense a square area consisting of more squares, e.g 4, 9, or more. Alternatively, a sensor could be equipped with directional sensing capabilities, so that it could sense more than one squares in the horizontal or vertical direction, as shown in Fig. 4. This is often the case for movement detection with surveillance cameras or sound detection scenarios.

We refer to the type of coverage provided by such a sophisticated sensor as *coverage pattern* (i, j) if the sensor can cover i and j consecutive cells in the horizontal and vertical direction respectively. In general, a sophisticated sensor k may have a collection of n_k coverage patterns $\mathcal{C}_k = \{(i_\ell, j_\ell) : \ell = 1, \ldots, n_k\}$ to choose from, depending on its level of sophistication.

The output signal of a sophisticated sensor that covers more than one cells is some modulo-2 operation on events in the individual squares such as ”OR”, ”AND” or ”XOR”, since the sensors cannot distinguish among individual events. Usually, there exists a specification on a maximum allowable probability of error in detecting the events, P^*, which reflects the required reliability of the detection mechanism. Assume that sophisticated sensors perform a certain modulo-2 operation and consider a certain deployment in the grid, which is characterized by the identities of cells that are covered by each sophisticated sensor. A placement is called *feasible* with respect to P^* if the probability of error in event detection does not exceed P^*. In this setting, there exist several challenging questions, such as:

- Given the probability of sensor malfunction q, a number of sophisticated sensors K, each with its set of coverage patterns $\mathcal{C}_k$, for $k = 1, \ldots, K$, and given a maximum allowable error probability P^*, does there exist a feasible placement?
- In the problem setup above, identify a sensor placement that minimizes the probability of error in the detection of events in the entire grid.

We now proceed to the analogy with channel coding. The values of events that emanate from the grid form the set of information bits that are about to be transmitted through a channel. Now, we can distinguish the following kinds of malfunction of a sophisticated sensor that covers a square of the grid:

- A sensor can measure an event correctly with probability $1 - q$ and incorrectly with probability q.
- A sensor can measure an event correctly with probability $1 - q$ but cannot provide any measurement with probability q.

Sensors are caused to malfunction independently of one another. The first kind of malfunction can be seen to be analogous to the situation encountered by bits passing through a *binary symmetric channel* with probability of error q, while the second kind of malfunction can be identified as equivalent to the case when bits pass through a *binary erasure channel* with probability of erasure q [11].

In addition, the identification of a placement of K sophisticated sensors to an $N \times N$ grid clearly adds redundancy to the measurement bits. Therefore, each placement corresponds to a code, in which K parity (redundancy) bits are appended to N^2 information bits and form the codeword that is transmitted through the channel. Since each sophisticated sensor performs a modulo-2 operation on bits of events in covered squares, the code is obviously systematic, namely each parity bit is a combination of some information bits. However, an additional constraint dictated by the definition of the coverage pattern comes into play: the required codeword should be formed so that each parity bit includes the combination of *only consecutive bits in the horizontal and/or the vertical directions*.

The theory of block codes (see for example Chap.8 in [10]) can be applied in order to derive efficient encoding and decoding schemes, so that the probability of correct detection of events is improved. The encoding operation is performed by a encoder which can be represented by a matrix operation on the information bit vector in order to generate the codeword bit vector. As was mentioned earlier, the encoder should take into account the additional constraints on the parity bits that stem from the nature of coverage patterns. The design of such an encoder in the setting of a sensor network is an interesting open issue.

Furthermore, decoding pertains to reconstruction of the original set of information bits that describe the events so that the minimum number of errors occur. For the binary symmetric channel case, the optimum decoder in the sense of minimizing the probability of a codeword error is based on the concept of minimum distance decoding. According to this principle, the decoder compares the received codeword to all possible transmitted ones and decides in favor of the one that is closest (in the Hamming distance sense) to the received one. However, the decoding rule for the binary erasure channel needs more investigation. In either case, the channel coding framework offers fertile ground for the derivation of algorithms and performance bounds in the framework of a sensor network. The ultimate objective is to specify the tradeoff between the level of required redundancy and quality of measurements captured by probability of detection error.

VI. Discussion

In this paper, we considered some problems that arise in the context of sensor networks and studied them from the viewpoint of well-known problems from the area of source and channel coding. The paper constitutes a first step towards understanding the nature of such problems and could serve as a guideline for obtaining algorithms and performance bounds that have their origins in the areas of source and channel coding.

There exist several directions for future study and we list a few of them. With respect to the first problem, the analogy with rate-distortion theory will establish performance tradeoffs between distortion and rate for the case of one observing sensor. For the case of several observing sensors, the CEO problem offers an interesting framework. The study for the second problem should focus on design of algorithms for clustering schemes that minimize communication cost. The tradeoff between load balancing at cluster-heads and increase in communication cost should also be quantified. An interesting issue pertains to the case where the information generation density function cannot be assumed to be known. For the last problem, the design of efficient encoding and decoding rules as well as fundamental performance limits in terms of probability of event detection warrant further attention. Finally, we note that the last problem can be cast in the framework of compression and source coding, if the objective is to deploy the sophisticated sensors such that event information bits are conveyed with a small number of bits and tolerable distortion.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramanian, and E. Cayirci, “A survey on sensor networks,” *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.

[2] Z. Xiong, A. D. Liveris, and S. Cheng, “Distributed source coding for sensor networks,” *IEEE Signal Processing Mag.*, vol. 21, no. 5, pp. 80–94, Sep. 2004.

[3] J.-H. Chang and L. Tassiulas, “Energy conserving routing in wireless ad-hoc networks,” in *Proc. IEEE INFOCOM*, vol. 1, Tel Aviv, Israel, Mar. 2000, pp. 22–31.

[4] V. Sharma, E. Frazzoli, and P. G. Voulgaris, “Improving lifetime data gathering and distortion for mobile sensing networks,” in *Proc. IEEE SECON*, Santa Clara, CA, Oct. 2004.

[5] D. Ganesan, R. Cristescu, and B. Beferull-Lozano, “Power-efficient sensor placement and transmission structure for data gathering under distortion constraints,” in *Proc. IEEE IPSN*, Berkeley, CA, Apr. 2004, pp. 142–150.

[6] B. Carbunar, A. Grama, J. Vitek, and O. Carbunar, “Coverage preserving redundancy elimination in sensor networks,” in *Proc. IEEE SECON*, Santa Clara, CA, Oct. 2004.

[7] S. Bandyopadhyay and E. J. Coyle, “An energy efficient hierarchical clustering algorithm for wireless sensor networks,” in *Proc. IEEE INFOCOM*, vol. 3, San Francisco, CA, Mar.-Apr. 2003, pp. 1713–1723.

[8] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-efficient communication protocol for wireless micro-sensor network,” in *Proc. IEEE HICSS*, Maui, HI, Jan. 2000.

[9] A. D. Amis, R. Prakash, T. H. P. Vuong, and D. T. Huynh, “Max-min d-cluster formation in wireless ad-hoc networks,” in *Proc. IEEE INFOCOM*, vol. 1, Tel Aviv, Israel, Mar. 2000, pp. 32–41.

[10] J. G. Proakis, *Digital Communications*, 3rd ed. New York, NY: McGraw Hill, 1995.

[11] T. Cover and J. Thomas, *Elements of Information Theory*, 1st ed. New York: Wiley Interscience, Aug. 1991.

[12] T. B. J. Chen, X. Jiang and S. B. Wicker, “An upper bound on the sum-rate distortion function and its corresponding rate allocation schemes for the CEO problem,” *IEEE Journal on Select. Areas Commun.*, vol. 22, no. 6, pp. 977–987, Aug. 2004.

[13] C. Rose, “Write or radiate? [information flow],” in *Proc. IEEE VTC-Fall*, vol. 4, Orlando, FL, Oct. 2003, pp. 2322–2327.