ABSTRACT

We consider a sensor network involving sensors that are placed in specific locations. A point phenomenon is being detected and tracked by the activated sensors. The latter collect data characterizing parameters of the phenomenon; possibly compress it and transport it to a central node. The central node processes the received data to derive an estimate of the phenomenon’s parameters. It is essential that the estimate reproduced at the center reflects the parameters characterizing the phenomenon at a sufficiently high fidelity level. Our sensing stochastic process models account for distance dependent observation noise perturbations as well as location dependent correlations between observation noise components, and assume sample mean estimates to be employed at the processing center. As such, they are distinctly different than corresponding models presented in the literature. We develop computationally efficient algorithms for determining the specific set of sensors to be activated, so that a sufficiently low reproduction distortion level can be attained. Corresponding algorithms are also derived for a system that operates under communications capacity constraints. For those sensor selection problems that are NP-hard, we introduce computationally efficient heuristic algorithms. We develop computationally efficient algorithms for determining the specific set of sensors to be activated, so that a sufficiently low reproduction distortion level can be attained. Corresponding algorithms are also derived for a system that operates under communications capacity constraints. For those sensor selection problems that are NP-hard, we introduce computationally efficient heuristic algorithms.

The models and results presented in this paper consist of a number of sensor network configurations. For each configuration, we define the underlying optimization problem (aiming to minimize the involved distortion function), and present an algorithm for determining the identity of sensors to be activated. Such configurations include: a). The noise processes perturbing sensor observations at the different sensors are assumed to be uncorrelated. We derive an activation scheme that is proven to be optimal. b). The above mentioned noise processes are assumed to be correlated with the correlation function being dependent upon nodal locations. The related optimization problem is shown to be NP-hard. We proceed to develop an effective sensor activation heuristic algorithm. c). A system as defined for case (b) is assumed except that a single hop multiple access wireless communications channel is used to transport each sensor’s data to the processing center. The latter is assumed to be capacity limited. d). A system as defined for case (b) is assumed except that a multi-hop network with limited capacity is used to transport each sensor’s data to the processing center. Scalable low complexity heuristic algorithms are developed for solving problems (c)-(d). We use our algorithms to present illustrative system performance results. We demonstrate that the activation of sensors that belong to a critical set of sensors can provide a distinct reduction in the distortion measure, while the activation of additional sensors outside such a set may not lead to further distinct estimate fidelity improvements.

1. INTRODUCTION

We consider a sensor network involving sensors that have been placed in specific locations. A point phenomenon is being detected and tracked by the activated sensors. The latter proceed to collect data characterizing parameters of the detected phenomenon, possibly compress it and transport it to a central node. The latter collects the received data to derive an estimate of the phenomenon’s parameters. The estimate must well reflect the parameters characterizing the detected phenomenon at a sufficiently high fidelity level. Our models are different than those considered in other studies. We characterize the observations at each sensor node to be perturbed by noise processes whose variance levels depend on the distance of the underlying sensor from the phenomenon. Nodes that are located closer to the location of the point phenomenon are assumed to gather information that describes the estimated parameter in a more faithful fashion. Furthermore, data readings at sensors that are located close to each other may observe be highly correlated. Consequently, our models account for such features, so that data readings collected at distinct sensors can be statistically independent or correlated with a correlation function that depends on the relative locations of the involved sensors.

We develop computationally efficient algorithms for determining the number and the identities of the sensors to be selected for activation, to yield sufficiently low prescribed reproduction distortion levels. For those sensor selection problems that are NP-hard, we describe computationally efficient heuristic algorithms.
tive system performance results. For example, we illustrate the variation of the distortion measure with the iterative progression of the selection algorithm. We demonstrate that the activation of sensors that belong to a critical set of sensors can provide a distinct reduction in the distortion measure, while the activation of additional sensors outside such a set may not lead to further distinct estimate fidelity improvements.

Different sensing and data gathering approaches have been investigated in recent years in solving related problems. Some recent research works have focused on jointly optimizing source coding and routing in wireless sensor networks where the data readings by different sensors are correlated. These approaches aim to reduce the total transmission cost of transporting data readings from sensor nodes to a sink node. In [2], the Slepian-Wolf joint coding mechanism has been studied, under which the optimal coding is complex and transmission optimization is simple. In [1-3], various data gathering mechanisms with in-network data fusion capabilities are investigated. Such mechanisms have high complexity joint data fusion and routing operations required at the involved network nodes.

Recent research works presented in [4, 5] investigate mechanisms for sequentially activating sensor nodes for the purpose of collaborative target localization and tracking. The expected conditional entropy function of the posterior target location distribution is used as the objective function for minimization. In [6], a mathematical framework is presented for the selection of sensor nodes to sense a point phenomenon so that the distortion between the actual phenomenon’s attribute and the estimate produced at a processing center is minimized. Data readings of different sensors are assumed there to have same entropy levels. In contrast with the results presented in this paper, the latter paper does not present algorithms for the selection of the sensors; neither is the network capacity limitation factor involved in the framework posed there.

The remainder of this paper is structured as follows. In Section 2, we present the models used to describe the point phenomenon and the observed sensing processes. In Sections 3 and 4, we develop algorithms to calculate the optimal number of sensors to be activated without and with the capacity resources constraint respectively. In Section 5, we give several examples to demonstrate the effectiveness of our algorithms. Conclusions are drawn in Section 6.

2. MODELS OF THE PHENOMENON AND SENSING PROCESSES

Assume the point phenomenon to be located at position \( z_0 \). We assume that the signal characterizing the state of the sensed phenomenon is modeled as a Gaussian random variable, denoted by \( s \sim N(m, \sigma^2) \).

Assume that \( N \) sensors have been placed in specific locations \( \{z_1, z_2, ..., z_N\} \). We assume the data reading at any sensor \( i \) to be modeled as the random variable \( s_i = s + n_i \), where \( n_i \) represents the noise process perturbing the data readings. We assume that \( n_i \) can be modeled as a white Gaussian random variable so that \( n_i \sim N(0, \sigma^2) \).

We set the following parameters: \( \sigma^2 \) represents the normalized variance constant of the noise process; \( d_{i0} \) denotes the distance between the \( i \)th sensor and the phenomenon, where \( d_{i0} = |z_i - z_0| \). We assume that the noise power level characterizing \( n_i \) increases with the distance between the sensor and the point phenomenon. Hence, \( a(d_{i0}) \) increases with \( d_{i0} \), \( 0 \leq a(d_{i0}) \leq 1 \). Such an \( a(d_{i0}) \) function is illustrated as \( a(d_{i0}) = 1 - e^{-d_{i0}/c_0} \), where \( d_0 > 0 \) is identified as the noise distance factor constant.

Assume that the noise processes at two sensors that are located at different locations \( \{z_i, z_j\} \) are spatially correlated, having the correlation function \( c(d_{ij}) \). An example of \( c(d_{ij}) \) is shown as \( c(d_{ij}) = e^{-d_{ij}/c_0} \), where \( c_0 > 0 \) is identified as the noise correlation factor constant.

For \( i, j \in \{1, 2, ..., N\} \), we obtain:

\[
E(n_n) = \sigma^2 c(d_{ij}) a(d_{i0}) a(d_{j0}) \tag{1}
\]

\[
E(s_s) = E(s) = (m^2 + \sigma^2) \tag{2}
\]

\[
E(s_s) = \sigma^2 c(d_{ij}) a(d_{i0}) a(d_{j0}) (m^2 + \sigma^2) \tag{3}
\]

Assume \( M \) out of \( N \) sensors to be activated to sense the point phenomenon. The activated sensors proceed to collect data concerning the detected phenomenon, possibly data compress it and transport it across a communications network system to a processing center. Assume a processing center to be located at \( z_p \). The latter receives \( M \) messages per unit time from each one of the activated sensors. Assume the center to calculate an estimated state of the sensed phenomenon by computing the corresponding sample mean value, \( y = \frac{1}{M} \sum_{i=1}^{M} s_i \).
We define an illustrative distortion measure by using a mean squared error function. Assume that the signals received at the activated M sensors, \( s_1, s_2, \ldots, s_M \), are delivered to the processing center without any compression distortion or errors contributed by the communication processes. The resulting distortion function is then expressed as follows:

\[
D(s_1, s_2, \ldots, s_M) = E \left\{ \left( \frac{1}{M} \sum_{i=1}^{M} s_i - s \right)^2 \right\}
\]

\[
= E \left\{ \left( \frac{1}{M} \sum_{i=1}^{M} s_i \right)^2 \right\} - E \left\{ \left( \frac{2}{M} \sum_{i=1}^{M} s_i \right) s \right\} + E\{s^2\}
\]

\[
= \frac{1}{M^2} \left( \sum_{i=1}^{M} \sum_{j=1}^{M} \sigma_i^2 c_i d_i a(d_{i0}) \sigma_j^2 c_j d_j a(d_{j0}) \right)
\]

When the noise processes perturbing the data readings at different sensors are statistically independent and identically distributed, the distortion measure is reduced to zero as the number of independent samples M increases. When the underlying noise processes are statistically independent, but are characterized by having different noise power levels, the resulting distortion measure does not necessarily decrease as we activate a larger number M of sensor nodes. Thus, the activation of an excessive number of sensors may lead to noise accumulation, causing degradation in the estimation quality.

3. OPTIMAL NUMBER OF SENSORS TO BE ACTIVATED

3.1 Sensor Selections under Uncorrelated Noise Processes

In this section, we present an algorithm that is used to calculate the optimal number of sensors and the related sensor set to be activated so that the distortion measure is minimized under the assumption that the noise processes perturbing the data readings at different sensors are uncorrelated (e.g., \( c_{ij} = \delta_{ij} \)). The correlated case is studied in the subsequent Section.

We define the indicator variable \( x_i = \{0, 1\} \), \( i = 1, 2, \ldots, N \) to identify the state of the ith sensor. We set \( x_i = 1 \), if sensor \( i \) is activated. Otherwise, we set \( x_i = 0 \). The resulting distortion value is calculated as follows.

\[
D(x_1, x_2, \ldots, x_N) = E \left\{ \left( \frac{\sum_{i=1}^{N} x_i s_i}{\sum_{i=1}^{N} x_i} - s \right)^2 \right\}
\]

\[
= \frac{1}{\sum_{i=1}^{N} \sum_{j=1}^{N} x_i x_j} \sum_{i=1}^{N} \sum_{j=1}^{N} (x_i x_j) \sigma_i^2 \delta_{ij} \sigma_j^2 c_i d_i a(d_{i0}) a(d_{j0})
\]

The resulting distortion function is then expressed as:

\[
D(x_1, x_2, \ldots, x_N) = \sigma_N^2 \sum_{i=1}^{N} x_i a(d_{i0}) = c^T x
\]

where \( c = \sigma_N^2 [a(d_{10}) \ a(d_{20}) \ \ldots \ a(d_{N0})]^T \)

and \( H = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & \vdots \\ \vdots & \vdots & \ddots & 1 \\ 1 & \vdots & \ldots & 1 \end{bmatrix} \), so that that \( x^T H x \) denotes the square of the number of activated sensor nodes.

We prove that the following iterative selection process leads to the activation of a set of sensor nodes that yields the lowest distortion level.

Algorithm 1:

1. Order the list of sensors \( U=\{1, 2, \ldots, N\} \) in accordance with their noise variances values \( \sigma_i^2 a(d_{i0}) \).
2. \( k=1, U=U \setminus \{1\}, x^*=A_1=\{1\}, D(A_1)=\sigma_N^2 a(d_{10}) \).
3. \( k=k+1, A_k = A_{k-1} \cup \{k\} \).
4. \( D(A_k) = D(D(A_{k-1}))(k-1)^2 + \sigma_N^2 a(d_{10}) \) / \( k^2 \)
5. If \( D(A_k) < D(x^*) \), then \( x^* = A_k \) and \( M^* = k \).
6. \( U = U \setminus \{k\} \). If \( U \neq \emptyset \), go to Step 3.
7. Activate sensor set \( x^* \).

Under this iterative sensor selection process, at each step, a set of sensor nodes is composed. Let \( A_k \) denote the set of sensors selected for consideration at the \( k^{th} \) step iteration. At step \( k+1 \), the set \( A_{k+1} \) is constructed by adding a new sensor to \( A_k \). Denote by \( U \) the set of sensors that are not selected at the \( k^{th} \) step. New sensor \( i \) is added into set \( A_k \) if \( a(d_{i0}) \leq a(d_{j0}) \), \( \forall i, j \in U \). The process stops after all N sensors have been selected to form the set \( A_N \). The sensor set that yields the lowest distortion level is tracked. The sensor set formed at the end of process identifies the sensors that need to be activated. The optimality of this algorithm is proven in the following.

Theorem 1: Algorithm 1 yields an optimal set of sensors whose activation results in the lowest distortion level.

Proof: Assume \( x \) to denote the optimal set of sensors that yields the minimal distortion level. We show in the following, by contradiction, that any sensor included in \( x \) must have a noise variance value that is not higher than that characterizing any sensor that is not included in \( x \). Assume that \( x \) contains sensor \( i \), where its data reading has a noise variance \( \sigma_N^2 a(d_{i0}) \) that is higher than the noise
variance \( \sigma_s^2 a(d_{ij}) \) of sensor \( j \) that is not included in the set \( x \). We then replace sensor \( i \) by sensor \( j \) to form a new set of sensors, \( y \). We have: \( x^T H x = y^T H y \) and \( c^T x > c^T y \). Hence, \( \frac{c^T x}{x^T H x} > \frac{c^T y}{y^T H y} \), so that the new set yields a lower distortion value. This result thus contradicts the assumption that \( x \) is an optimal set of sensors. Hence, it is sufficient to consider only the \( N \) sets described by the above defined iterative algorithm. ■

The computational complexity of the algorithm is of the order \( O(N) \). Thus, it is highly scalable and efficient.

### 3.2 Sensor Selections under Spatially Correlated Noise Processes

In this section, we present an algorithm that is used to calculate the optimal set of sensors to be activated so that the distortion measure is minimized under the assumption that the noise processes perturbing the data readings at different sensors are spatially correlated. The correlation function is described in Section 2.

The distortion function is calculated as follows:

\[
D(x_1, x_2, \ldots, x_N) = E \left[ \sum_{i=1}^{N} \sigma_s^2 c(d_{ij}) \sqrt{a(d_{ij}) a(d_{ij})} \right] = \frac{x^T Q x}{x^T H x} \tag{6}
\]

where \( Q = \sigma_s^2 \left[ c(d_{ij}) \sqrt{a(d_{ij}) a(d_{ij})} \right]^{N \times N} \).

We identify the following feasibility problem (determining whether a given vector \( x \) yields a distortion level that is no higher than value \( t \)):

\[
x^T Q x - t x^T H x = x^T P(t) x \leq 0 \tag{7}
\]

where \( P(t) = Q - tH \).

The problem of minimizing \( x^T P(t) x \) is an unconstrained binary quadratic programming problem. This problem has been the focus of a considerable amount of research in recent years, including exact and heuristic solutions [7-11]. It has been shown to be NP-hard. Clearly, if the binary quadratic feasibility problem included in this algorithm is solved in an exact manner, the algorithm will not yield a scalable procedure. It will not scale well for applications that involve more than 200 variables. Therefore, we present a scalable heuristic algorithm that provides suboptimal solution to the above problem.

**Algorithm 2:**

1. Find sensor \( i \) that has the lowest \( \sigma_s^2 a(d_{ij}) \).
2. \( k=1, \ U=\{1, 2, \ldots, N\} \setminus i \), \( x^*=A_1=\{i\} \), \( M^*=1 \), \( D(A_1)=\sigma_s^2 a(d_{ij}) \)
3. \( k=k+1 \).
4. \( A_k = A_{k-1} \cup i \), so that \( \forall i, j \in U \)
   \[
   \sum_{i=1}^{k} c(d_{ij}) \sqrt{a(d_{ij}) a(d_{ij})} + a(d_{ij})
   \]
   \[
   < 2 \sum_{i=1}^{k} c(d_{ij}) \sqrt{a(d_{ij}) a(d_{ij})} + a(d_{ij})
   \]
5. \( U = U \setminus \{i\} \).
6. \( D(A_k) = \frac{k^2}{D(A_{k-1})(k-1)^2 + 2 \sigma_s^2 \sum_{i=1}^{k} c(d_{ij}) \sqrt{a(d_{ij}) a(d_{ij})} + \sigma_s^2 a(d_{ij})} \)
7. If \( D(A_k) < D(A_{k-1}) \), then \( x^*=A_k \) and \( M^*=k \).
8. If \( k < N \), go to Step 3.
9. Activate sensor set \( x^* \).

Algorithm 2 iteratively constructs a set of sensors by adding a new sensor whose addition results in the lowest distortion level among all other possible sensor additions. We note that the computational complexity of the algorithm is of the order \( O(N^3) \).

### 4. OPTIMAL SET OF SENSORS TO BE ACTIVATED UNDER COMMUNICATIONS CAPACITY RESOURCES

#### 4.1 Sensor Rate Allocation under a Single-hop Network Configuration

When the network capacity is limited, every sensor would have to quantize its data reading and then send the quantized data reading to the processing center. The quantization of each data reading is determined by the capacity allocated for the underlying sensor. In this Section, we assume that every sensor is located one-hop away from the processing center and study the selective sensor activation problem under the network capacity constraint.

Assume that we assign data rate \( R_i \) for any sensor \( i \). The capacity constraint is \( \sum_{i=1}^{N} R_i \leq R \), where \( R \) denotes the overall available capacity to access the processing center. Denote the quantized data reading at sensor \( i \) by \( \hat{s}_i \). We have
\[ \hat{s}_i = s_i + q_i, \] where \( q_i \) is a random variable that represents the quantization error. We define the quantization distortion associated with the quantized data reading as follows:

\[ \hat{D}_i = (\hat{s}_i - s_i)^2 = q_i^2 \] (8)

Recall that the variance of the data reading at sensor \( i \) is equal to \( \sigma^2 + \sigma_i^2 a(d_{io}) \). From rate distortion theory, the minimum rate for sensor \( i \) that could achieve a quantization distortion \( \hat{D}_i \) for the quantized data reading is calculated as \( R_i = \max \left\{ \frac{1}{2} \log_2 \frac{\sigma^2 + \sigma_i^2 a(d_{io})}{\hat{D}_i}, 0 \right\} \).

We first assume \( M \) sensors to be activated. The distortion measure is thus calculated as follows:

\[ D((R_i)) = E \left[ \left( \frac{1}{M} \sum_{i=1}^{M} \hat{s}_i - s \right)^2 \right] = E \left[ \left( \frac{1}{M} \sum_{i=1}^{M} \hat{s}_i - \frac{1}{M} \sum_{i=1}^{M} s_i + \frac{1}{M} \sum_{i=1}^{M} - s \right)^2 \right] \]

\[ = E \left[ \left( \frac{1}{M} \sum_{i=1}^{M} \hat{s}_i - s \right)^2 + \left( \frac{1}{M} \sum_{i=1}^{M} s_i - s \right) \left( \frac{1}{M} \sum_{i=1}^{M} q_i + \left( \frac{1}{M} \sum_{i=1}^{M} q_i \right)^2 \right) - \frac{1}{M} \sum_{i=1}^{M} \sum_{j=1}^{M} \sigma_i^2 c(d_{ij}) a(d_{ij}) a(d_{ij}) + \frac{1}{M} \sum_{i=1}^{M} \sum_{j=1}^{M} \sigma_i^2 a(d_{ij}) \right]\] (9)

subject to \( \sum_{i=1}^{M} R_i \leq R \).

In considering single-hop sensor network configuration involving \( M \) activated sensor nodes, our objective here is to determine the data rate level to be assigned for the data compression scheme at each sensor, so that the aggregate data rate is lower than the available communications capacity level and the overall distortion measure is minimized.

**Theorem 2:** Under the optimal configuration for the above defined problem, activated sensors are configured to yield a fixed quantization distortion level, when considering such sensors whose data reading distortion values (variances) are lower than this fixed level. The optimal rate allocated for sensor \( i \), \( i=1, 2, \ldots, M \), is calculated as follows:

\[ R_i = \begin{cases} \frac{1}{2} \log_2 \frac{\sigma^2 + \sigma_i^2 a(d_{io})}{\frac{\lambda M^2}{2 \ln 2}}, & \text{if } \sigma^2 + \sigma_i^2 a(d_{io}) > \frac{\lambda M^2}{2 \ln 2} \\ 0, & \text{if } \sigma^2 + \sigma_i^2 a(d_{io}) \leq \frac{\lambda M^2}{2 \ln 2} \end{cases} \] (10)

where \( \lambda \) is a constant determined by \( \sum_{i=1}^{M} R_i = R \).

**Proof:** Using a Lagrange multiplier \( \lambda \), we define the following function:

\[ \Lambda(R_1, R_2, \ldots, R_M) = \frac{1}{2} \sum_{i=1}^{M} \sigma_i^2 c(d_{ij}) a(d_{ij}) a(d_{ij}) + \frac{1}{M} \sum_{i=1}^{M} \sigma_i^2 a(d_{io}) \]

\[ + \lambda \left( \sum_{i=1}^{M} R_i - R \right) \]

\[ \frac{\partial \Lambda(R_1, R_2, \ldots, R_M)}{\partial R_i} = -\sigma^2 - \sigma_i^2 a(d_{io}) \left( \frac{1}{2^{2R_i}} \right) \ln 2 (2^{2R_i}) + \lambda \]

Using Kuhn-Tucker conditions, we choose \( \lambda \) so that

\[ \frac{\partial \Lambda(R_1, R_2, \ldots, R_M)}{\partial R_i} = \begin{cases} \frac{2^{2R_i}}{\lambda M^2} (\sigma^2 + \sigma_i^2 a(d_{io})), & \text{if } \sigma^2 + \sigma_i^2 a(d_{io}) > \frac{\lambda M^2}{2 \ln 2} \\ R_i = 0, & \text{if } \sigma^2 + \sigma_i^2 a(d_{io}) \leq \frac{\lambda M^2}{2 \ln 2} \end{cases} \]

\[ \Rightarrow \begin{cases} \hat{D}_i = \frac{\lambda M^2}{2 \ln 2}, & \text{if } R_i > 0 \\ \hat{D}_i = \sigma^2 + \sigma_i^2 a(d_{io}), & \text{if } R_i = 0 \end{cases} \]

Note that, when the data reading of sensor \( i \) has a variance that is lower than the common quantization distortion level, we set \( R_i = 0 \). Thus, it is effectively unnecessary for such sensor \( i \) to send its data reading to the processing center. However, as shown in Eq. (9), the data reading of sensor \( i \) should still count for calculating the sample mean. In doing so, the processing center can hypothetically pick a data reading for sensor \( i \) according to its data reading distribution, and then quantize it using the desired quantization level. Effectively, sensor \( i \) will actually send infrequently its data reading level.

**4.2 Sensor Selection under a Single-hop Sensor Network Configuration**

In Section 4.1, we have assumed that \( M \) sensors are activated. We have then determined the optimal rate allocation scheme. In this section, we determine the optimal set of sensors to be activated, among all \( N \) sensor nodes, so that the overall realized distortion level is minimized, under the limited communications transport capacity level.

Using the result presented in the previous section, we conclude that the activated sensor nodes contribute equal distortion levels. Denote the latter common quantization distortion value by \( \hat{D} \). The joint optimization problem is then formulated as follows:

\[ \min_{\hat{D}, x_1, x_2, \ldots, x_N} \hat{D} \left( \hat{D}; x_1, x_2, \ldots, x_N \right) = E \left[ \frac{\sum_{i=1}^{N} x_i^2}{\sum_{i=1}^{N} x_i} - s \right]^2 \]
subject to $B^T x \leq R$, where
\begin{equation}
B = \frac{1}{2} \begin{bmatrix}
\log_2(\sigma_i^2 + \sigma_{x_i}^2 a(d_{i,j})) & \ldots & \log_2(\sigma_i^2 + \sigma_{x_i}^2 a(d_{i,N}))
\end{bmatrix} - \left(\frac{1}{2} \log_2 \hat{D}\right) f.
\end{equation}

As a binary programming problem, this optimization problem is NP-hard even when it involves a fixed $\hat{D}$ level. In the following, we present a heuristic algorithm that employs an iterative routine to find a sub-optimal solution, in selecting a set of sensors to be activated.

**Algorithm 3:**

1. Find sensor $i$ that has the lowest $\sigma_N^2 a(d_{i,i})$.
2. $U = \{1, 2, \ldots, N\} \setminus i$, $k=1$, $x^*=A_1=\{i\}$, $M^*=1$.
3. $D(A_1) = \sigma_N^2 a(d_{i,i})$, $\hat{D}(A_1) = (\sigma^2 + \sigma_N^2 a(d_{i,i}))/2^{2\sigma}$.
4. $k = k+1$.
5. For each $i \in U$, do the following
   a. Set the rate vector as:
   $$R_j(i) = \max \left\{ \frac{1}{2} \log_2 \frac{\sigma^2 + \sigma_N^2 a(d_{i,j})}{\lambda(i)}, 0 \right\}, j \in A_{k-1} \cup \{i\}$$
   b. by solving for $\lambda(i) > 0$ so that $\sum_{j \in A_{k-1} \cup \{i\}} R_j(i) = R$
6. End For.
7. Select sensor $i$, so that $\forall i,j \in U$
   we have $2 \sum_{i \in A_{k-1}} c(d_{i,j}) \sqrt{a(d_{i,j})} a(d_{i,i}) + a(d_{i,j}) + k\lambda(i)/\sigma_N^2$
   $$< 2 \sum_{i \in A_{k-1}} c(d_{i,j}) \sqrt{a(d_{i,j})} a(d_{i,j}) + a(d_{i,j}) + k\lambda(j)/\sigma_N^2.$$
8. $U = U \setminus \{i\}$, $A_k = A_{k-1} \cup \{i\}$.
9. If $\exists j \in A_k$, $R_j(i) \leq 0$, go to step 11.
10. $D(A_{k-1})(k-1)^2 + 2\sigma_N^2 \sum_{i \in A_{k-1}} c(d_{i,j}) \sqrt{a(d_{i,j})} a(d_{i,i}) + \sigma_N^2 a(d_{i,j})$
11. $D(A_k) = \frac{\lambda(i)}{k}$.
12. If $D(A_k) + \hat{D}(A_k) < D(A_{k-1}) + \hat{D}(A_{k-1})$, then $x^* = A_k$ and $M^* = k$.
13. If $k < N$, go to Step 3.

Algorithm 3 iteratively constructs a set of sensors by adding a new sensor whose addition results in the lowest distortion level among all other possible sensor additions. At each step of the process, the capacity to be allocated for each sensor is calculated. We note that the computational complexity of the algorithm is of the order $O(N^3)$.

### 4.3 Sensor Rate Allocation under a Multi-hop Network Configuration

In this section, we assume that sensors might be located multi-hop away from the processing center. We assume that a given set of $M$ sensors has been activated. Hence, the transport load constraint for these $M$ activated sensors is
\begin{equation}
\sum_{i=1}^{M} \frac{d_{i,j}}{d_j} \leq SRF \times R.
\end{equation}

where $d_j$ represents the nominal communication range, $d_{i,j} > 0$ denotes the distance between sensor $i$ and the processing center, and SRF denotes the averaged spatial reuse factor under given routes between the $M$ sensors and the processing center.

We present the following theorem for the minimization of the overall distortion measure under the multi-hop sensor network configuration.

**Theorem 3:** For the above stated problem under which the best quantization level at each sensor is determined to minimize the overall system distortion level, the optimal quantization distortion level assigned to each one of the sensors is linearly proportional to the distance between the sensor and the processing center, provided this quantization distortion value is lower than the variance level of the data reading. The optimal rate allocated to sensor $i$ is calculated as follows:
\begin{equation}
R_i = \begin{cases}
\frac{\lambda M \sigma^2}{2 \ln 2} \frac{d_{w}}{d_j}, & \text{if } \sigma^2 + \sigma_N^2 a(d_{i,i}) > \frac{\lambda M^2}{2 \ln 2} \frac{d_{w}}{d_j} \\
0, & \text{if } \sigma^2 + \sigma_N^2 a(d_{i,i}) \leq \frac{\lambda M^2}{2 \ln 2} \frac{d_{w}}{d_j}
\end{cases}
\end{equation}

where $\lambda$ is a constant determined by $\sum_{i=1}^{M} R_i (d_{i,j}/d_j) = SRF \times R$.

We omit the proof here because of the limited space.

### 4.4 Sensor Selection under a Multi-hop Network Configuration

We consider in this section the joint selection of the number and identities of the sensors to be activated under a multi-hop network configuration. To minimize the overall distortion obtained at the processing center, we need to activate the optimal set of sensor nodes. However, as noted in Section 4.1, this joint optimization problem is NP-hard. In the following, we present a heuristic algorithm that employs an iterative routine to find a sub-optimal solution in
deriving a set of sensors to be activated, with each sensor proceeding to compress its observed data at a calculated quantization level.

**Algorithm 4:**

1. Find sensor $i$ that has the lowest $\sigma_N^{-2} a(d_{i0})$.
2. $U=\{1, 2, \ldots, N\} \setminus i$, $k=1$, $x^*=A_1=\{i\}$, $M^*=1$, $D(A_1)=\sigma_N^{-2} a(d_{i0})$, $\hat{D}(A_1)=\left(\sigma^2 + \sigma_N^{-2} a(d_{i0})\right)/2$.
3. $k=k+1$.
4. For each $i \in U$, do the following
5. Set the rate vector as, by solving $\lambda(i) > 0$:
   
   $$R_j(i) = \max\left\{\frac{1}{2} \log_2 \frac{\sigma^2 + \sigma_N^{-2} a(d_{i0})}{\lambda(i) d_{ij}} \right\}, j \in A_{k-1} \cup\{i\}$$
   
   so that $\sum_{j \in A_{k-1} \cup\{i\}} R_j(i) d_{ij} = SRF \times R$
6. End For.
7. Select sensor $i$, so that $\forall j, i \in U$
   
   
   $$\sum_{j \in A_{k-1}} c(d_{ij}) a(d_{ij}) a(d_{i0}) + a(d_{ij}) + \frac{\lambda(i)}{\sigma_N^{-2}} \sum_{m \notin j, i \in A_{k-1}} d_{jm}$$
   
   $$<2 \sum_{j \in A_{k-1}} c(d_{ij}) a(d_{ij}) a(d_{i0}) + a(d_{ij}) + \frac{\lambda(i)}{\sigma_N^{-2}} \sum_{m \notin j, i \in A_{k-1}} d_{jm}$$
   
   $U = U \setminus \{i\}$, $A_k = A_{k-1} \cup i$.
8. If $\exists j \in A_k$, $R_j(i) \leq 0$, go to step 11.
9. $D(A_k)(k-1)^2 + 2\sigma_N^{-2} \sum_{j \in A_k} c(d_{ij}) a(d_{ij}) a(d_{i0}) + \sigma_N^{-2} a(d_{i0})$
10. $D(A_k) = \frac{\lambda(i)}{k^2} \sum_{j \in A_k} d_{jm}/d_{ij}$.
11. $\hat{D}(A_k) = \frac{\lambda(i)}{k^2} \sum_{j \in A_k} d_{jm}/d_{ij}$.
12. If $D(A_k) + \hat{D}(A_k) < D(A_{k-1}) + \hat{D}(A_{k-1})$, then $x^* = A_k$ and $M^* = k$.
13. Activate sensor set $x^*$.

Algorithm 4 adopts a computation procedure that is similar to that used in Algorithm 3, except that the rate allocation vector for the selected sensors has been modified to account for the multi-hop distances between the involved sensors and the central node.

6. ILLUSTRATIVE EXAMPLES

In this section, we present illustrative configurations to demonstrate the effectiveness of the algorithms presented above. For illustration purposes, 400 sensors are deployed in a grid manner over a $950m \times 950m$ operational area. We assume that a point phenomenon is detected and is located at $(475m, 475m)$. The parameter to be monitored in characterizing the point phenomenon is assumed to be governed by the statistics of a Gaussian random variable $N(0, 1)$. We assume the noise processes perturbing the data readings at sensor nodes to have the following parameters defined in the model presented in Section 2: the normalized noise variance is $\sigma_N^{-2} = 1$, the noise distance factor is equal to $c_0 = 200m$.

We first assume that the available network capacity is sufficiently high to support distributions of the data readings to the processing center. Algorithm 2 is employed to select a sub-optimal number and set of sensors to be activated.

![Figure 1. Sensor Activation under Energy Constraints](image-url)
vated is noted to be equal to 13. It is observed to be a lower number of activated sensors than that determined by using Algorithm 2 when no communications capacity constraints are imposed.

For a multi-hop networking scenario, when the network capacity is insufficient for sensors to send their original data readings to the processing center, we employ Algorithm 4. We assume a nominal communications range of $d_i=220\text{m}$. The processing center is assumed to be located at (975m, 975m). Thus, a sensor node is located at most 7 hops away from the processing center. Note that for communications purposes, a node that may not be elected to participate in the sensing process may have to be activated to relay messages sensed by other nodes. Performance results, showing sensor activation and rate allocation results, are depicted in Fig. 3. Using Algorithm 4, we find the number of sensors to be activated to be equal to 12.

![Figure 2. Sensor Activation with Capacity Constraints under Single-hop Sensor Networks](image)

![Figure 3. Sensor Activation with Capacity Constraints under Multi-hop Sensor Networks](image)

**7. CONCLUSIONS**

We consider a sensor network in which activated sensors sense a parameter of a detected point phenomenon and report their observations to a processing center. The data readings at each sensor are perturbed by noise processes. The latter are location dependent. The noise processes perturbing sensors positioned at different locations could be spatially correlated. The center calculates an estimate of the parameter of interest by using a sample mean computation. The resulting mean square distortion level is computed. Our models are distinctly different than other published models. We present computationally efficient heuristic algorithms that calculate the numbers and identities of the sensor nodes to be activated, aiming to yield low distortion levels. These algorithms are further extended to account for communications capacity resource limitations. We illustrate the performance sensitivity of the fidelity measure on the identities of the activated sensors, as well as on the available levels of communications capacity resources.

**ACKNOWLEDGMENT**

This work was supported by the National Science Foundation (NSF) under Grant No. ANI-0087148, by University of California/Conexant MICRO Grant No. 04-100, and by University of California/Nokia MICRO Grant No. 05-054.

**REFERENCES**


