2006

Spatial Correlation-Based Collaborative Medium Access Control in Wireless Sensor Networks

Mehmet C. Vuran
IEEE, mcvuran@cse.unl.edu

Ian F. Akyildiz
IEEE

Spatial Correlation-Based Collaborative Medium Access Control in Wireless Sensor Networks

Mehmet C. Vuran, *Member, IEEE*, and Ian F. Akyildiz, *Fellow, IEEE*

Abstract—Wireless Sensor Networks (WSN) are mainly characterized by dense deployment of sensor nodes which collectively transmit information about sensed events to the sink. Due to the spatial correlation between sensor nodes subject to observed events, it may not be necessary for every sensor node to transmit its data. This paper shows how the spatial correlation can be exploited on the Medium Access Control (MAC) layer. To the best of our knowledge, this is the first effort which exploits spatial correlation in WSN on the MAC layer. A theoretical framework is developed for transmission regulation of sensor nodes under a distortion constraint. It is shown that a sensor node can act as a representative node for several other sensor nodes observing the correlated data. Based on the theoretical framework, a distributed, spatial Correlation-based Collaborative Medium Access Control (CC-MAC) protocol is then designed which has two components: *Event MAC (E-MAC)* and *Network MAC (N-MAC)*. E-MAC filters out the correlation in sensor records while N-MAC prioritizes the transmission of route-thru packets. Simulation results show that CC-MAC achieves high performance in terms energy, packet drop rate, and latency.

Index Terms—CC-MAC, energy efficiency, medium access control, spatial correlation, wireless sensor networks.

I. INTRODUCTION

WIRELESS SENSOR NETWORKS (WSN) are event based systems that rely on the collective effort of several microsensor nodes observing a physical phenomenon. Typical WSN applications require spatially dense sensor deployment in order to achieve satisfactory coverage [11]. As a result, several sensor nodes record information about a single event in a sensor field. Due to the high density in the network topology, the sensor records may be spatially correlated subject to an event. The degree of spatial correlation increases with the decreasing internode separation.

Exploiting spatial correlation in the context of collaborative nature of the WSN can lead to significant performance improvement of communication protocols. Possible approaches to utilize spatial and temporal correlation in WSN are already investigated in [20]. Intuitively, data from spatially separated sensor nodes are more useful for the sink than the highly correlated data from closely located sensor nodes. Hence, it may not be necessary for every sensor node to transmit its data to the sink;

instead, a smaller number of sensor data might be adequate to transmit certain event features to the sink within a certain distortion constraint.

In this paper, we show how this spatial correlation can be exploited on the Medium Access Control (MAC) layer. The shared wireless channel between sensor nodes and the energy considerations make the Medium Access Control a crucial part in WSN. The existing MAC protocols cannot be applied here directly because of the event-based traffic properties and the energy constraints in WSN. Also, the channel access must be coordinated such a way that possible collisions are minimized or eliminated. These requirements can be satisfied by intelligent management of transmission attempts among nodes by exploiting the spatial correlated nature of the event information. Based on the theoretical framework for spatial correlation in WSN, we develop a distributed, spatial Correlation-based Collaborative MAC (CC-MAC) protocol that regulates medium access and prevents redundant transmissions from closely located sensors.

The remainder of the paper is organized as follows. We summarize the recent work on spatial correlation in WSN in addition to existing MAC protocols in Section II. In Section III, we develop a theoretical framework for spatial correlation in WSN and derive a distortion metric for reliable detection/estimation of physical phenomenon. Based on this framework, we develop a centralized node selection algorithm for correlation filtering in Section IV. Details of the distributed CC-MAC protocol that regulates medium access to reduce the number of nodes transmitting information are presented in Section V. CC-MAC performance analysis and simulation results are presented in Section VI and the paper is concluded in Section VII.

II. RELATED WORK

There exists some research to study the correlation in WSN [6], [13], [17] in recent years. In these papers, the information theoretical aspects of the correlation are explored in depth. In other words, these studies aim to find the optimum rate to compress redundant information in the sensor observations. More recently, the relation between distortion, spatio-temporal bandwidth and power for large sensor networks is investigated [7]. However, no correlation (spatial or temporal) between sensor observations is considered in this work. Moreover, none of the above solutions develop communication network protocols.

In [22], spatial and temporal correlation is exploited to eliminate the acknowledgment in the communication. While the number of acknowledgments is considerably reduced, the number of redundant packets is still high in the network. The joint routing and source coding is introduced in [17] to reduce the amount of traffic generated in dense sensor networks with
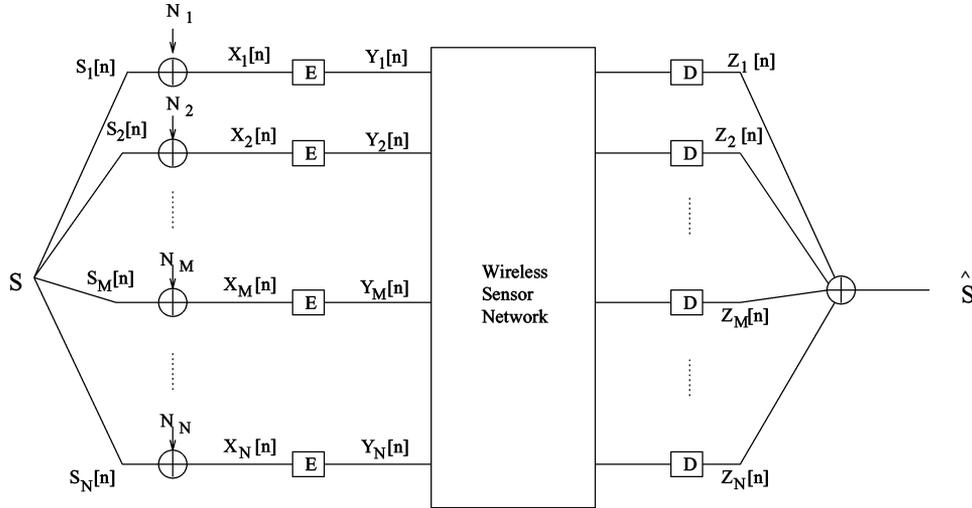
Fig. 1. Correlation model and architecture.

spatially correlated records. While joint routing and source coding reduces the number of transmitted bits, the number of transmitted packets remains unchanged from the network point of view. In our opinion, the number of transmitted packets can be further minimized by regulating the network access based on the spatial correlation between the sensor nodes. Moreover, the relation between spatial and temporal sampling rate on the overall network delay and energy consumption is studied in [3]. However, the spatial and temporal correlation between sensor observations is not investigated.

Current studies on medium access in WSN focus mainly on the energy-latency tradeoffs. S-MAC [21] aims to decrease the energy consumption by using sleep schedules with virtual clustering. T-MAC [5], a variant of S-MAC, incorporates variable sleep schedules to further decrease the energy consumption. However, in both protocols, sensor nodes keep sending redundant data with increased latency due to periodic sleep durations. In [1], an energy-aware TDMA-based MAC protocol is presented where the sensor network is assumed to be composed of clusters and gateways. Each gateway acts as a cluster-based centralized network manager and assigns slots in a TDMA frame. The protocol assumes a cluster-based topology, which requires significant additional processing complexity and overhead in the overall sensor network. An energy-efficient collision-free MAC protocol, which is based on a time-slotted structure, is presented in [15]. Each node determines its own time slot using a distributed election scheme based on traffic requirements of its every two-hop neighbor. Although the protocol achieves high delivery ratio with tolerable delay, the performance of the protocol depends on the two-hop neighborhood information in each node. Since this information is collected through signaling, in the case of high density sensor networks, the signaling cost increases significantly resulting in either incomplete neighbor information due to collisions or high energy consumption due to signaling costs.

So far, the existing MAC solutions focus on decreasing energy consumption by modifying known medium access techniques. In our opinion, event-based MAC protocols exploiting topology and traffic properties of WSN need to be developed. In

[20], we introduced the characteristics of spatial and temporal correlation in WSN. Based on [20], we develop the spatial correlation-based collaborative MAC (CC-MAC) protocol which aims to *reduce the energy consumption of the network by exploiting spatial correlation in WSN without compromising the channel access latency and the distortion achieved.*

III. THEORETICAL FRAMEWORK

The objective of the theoretical framework is to investigate the effect of spatial correlation on the reliability of event estimation at the sink in WSN. In particular, we investigate the relation between the positions of sensor nodes in the event area and the event estimation reliability. Consequently, we introduce a node selection algorithm that achieves acceptable estimation reliability by using a small subset of all the sensor nodes in the event area. Based on the theoretical framework, we develop the spatial CC-MAC protocol that exploits the spatial correlation between sensor nodes to reduce the energy consumption of the network without compromising the estimated reliability achieved at the sink.

A. Architecture and Spatial Correlation Model for WSN

The model for the information collected by N sensors in an event area is illustrated in Fig. 1. The sink is interested in estimating the event S in the sensor field, according to the observations of the sensor nodes, n_i, in the event area.[1]

Each sensor node n_i observes $X_i[n]$, the noisy version of the event information, $S_i[n]$, which is spatially correlated to the event source, S. In order to communicate this observation to the sink through the WSN, each node has to encode its observation. The encoded information, $Y_i[n]$, is then sent to the sink through the WSN. The sink, at the other end, decodes this information to obtain the estimate, $\hat{S}$, of the event S from the sensor field. The

[1]Note that although throughout our theoretical analysis we consider point source S for the physical phenomenon, a moving source such as in a tracking application can also be considered. In any given time period, the source can be approximated as a point source.

encoders and the decoders are denoted by E and D in Fig. 1, respectively.

The main objective of WSN is to estimate the event, S, from collective observations of the sensor nodes within a reliability measure required by the application. This reliability measure is chosen as the distortion between the event source, S, and its estimation, $\hat{S}$, at the sink. This event distortion is given by

$$D_E = E\left[d_E(S, \hat{S})\right] \tag{1}$$

where $d_E(S, \hat{S})$ is the distortion measure. In this paper, we refer to the maximum distortion allowed by the sensor application as the *distortion constraint*. Our goal is to exploit spatial correlation without compromising the distortion constraint.

The event source in interest can be modeled as a random process $s(t, x, y)$ as a function of time t and spatial coordinates (x, y). Each observed sample, $X_i[n]$, of sensor node, n_i, at time $t = t_n$ [2] is represented as

$$X_i[n] = S_i[n] + N_i[n], \quad i = 1, \ldots, N \tag{2}$$

where the subscript i denotes the spatial location of the sensor node n_i, i.e., (x_i, y_i), $S_i[n]$ is the realization of the space-time process $s(t, x, y)$ at $(t, x, y) = (t_n, x_i, y_i)$, and $N_i[n]$ is the observation noise. $\{N_i[n]\}_n$ is a sequence of i.i.d Gaussian random variables of zero mean and variance σ_N^2. We further assume that the noise each sensor node encounters is independent of each other, i.e., $N_i[n]$ and $N_j[n]$ are independent for $i \neq j$ and $\forall n$.

Since we only consider the spatial correlation between nodes, we can assume that the samples are temporally independent. Hence, by dropping the time index n, (2) can be restated as

$$X_i = S_i + N_i, \quad i = 1, \ldots, N \tag{3}$$

where the observation noise N_i of each sensor node n_i is modeled as i.i.d. Gaussian random variable of zero mean and variance σ_N^2. The event information, S_i, at each point of the event area is modeled as joint Gaussian random variables (JGRVs) as

$$E\{S_i\} = 0, \quad var\{S_i\} = \sigma_S^2, \quad i = 1, \ldots, N$$
$$\rho_{(i,j)} = K_\vartheta\left(d_{(i,j)}\right) = \frac{E[S_i S_j]}{\sigma_S^2} \tag{4}$$

where $\rho_{(i,j)}$ and $d_{(i,j)} = \|\mathbf{s}_i - \mathbf{s}_j\|$ denote the correlation coefficient and the distance between nodes n_i and n_j located at coordinates $\mathbf{s}_i$ and $\mathbf{s}_j$, respectively, and $K_\vartheta(\cdot)$ is the correlation model.

Note that the event source, S, is also a JGRV with the same properties. Since it is of special interest, the correlation coefficient, and the distance between the event source S and the node n_i are denoted by $\rho_{(s,i)}$ and $d_{(s,i)}$, respectively, throughout our discussions.

[2]Note that we use a discrete-time model since each node is assumed to sample the event information synchronously after the initial wake-up.

The covariance function, $K_\vartheta(\cdot)$, used in (4), models the relation between the correlation coefficient between the sensor observations, $\rho_{(i,j)}$, and the distance, $d_{(i,j)}$, between the nodes n_i and n_j. The covariance function is assumed to be nonnegative and decrease monotonically with the distance $d = \|\mathbf{s}_i - \mathbf{s}_j\|$, with limiting values of 1 at $d = 0$ and of 0 at $d = \infty$. Generally, covariance models can be classified into four standard groups as *Spherical*, *Power Exponential*, *Rational Quadratic*, and *Matérn* [4]. Although our results about the distortion function apply to all the covariance models, we use the *Power Exponential* model in this paper because the physical event information such as, electromagnetic waves, is modeled to have an exponential autocorrelation function [18].

Hence, the covariance function, $K_\vartheta(\cdot)$, is given as

$$K_\vartheta^{PE}(d) = e^{(-d/\theta_1)^{\theta_2}}, \quad \text{for } \theta_1 > 0, \quad \theta_2 \in (0, 2] \tag{5}$$

where for $\theta_2 = 1$, the model becomes exponential.

The parameter, θ_1, controls the relation between the distance, $d_{(i,j)}$, and the correlation coefficient, $\rho_{(i,j)}$, between the nodes n_i and n_j. The parameter θ_1 is important since the spatial correlation between the sensor observations depend on the distance between the sensor nodes. We investigate the effects of θ_1 on the distortion in Section III-B.

As each sensor node n_i observes an event information X_i, this information is encoded and then sent to the sink through the WSN. In traditional point-to-point communication, the optimum performance is obtained by compressing the information according to the source statistics and then adding redundant information to accommodate the errors introduced in the wireless channel. This technique is known as the *separation principle*. In WSN, where multiple nodes try to send information about the same event, however, it is known that the joint source-channel coding outperforms separate coding [6], [14]. In addition, for Gaussian sources, if the source is Gaussian and the cost on the channel is the encoding power, then the uncoded transmission is optimal for point-to-point transmission [8]. Furthermore, for sensor networks with finite number of nodes, the uncoded transmission outperforms any approach based on the above *separation principle*, leading to the optimal solution for infinite number of nodes [6]. More specifically, *separation principle* may incur an exponential penalty in terms of communication resources, as a function of the number of sensors [7].

Hence, we adopt the uncoded transmission for sensor observations in this paper. Each node n_i sends to the sink, a scaled version, Y_i, of the observed sample X_i according to encoding power constraint P_E, i.e.,

$$Y_i = \sqrt{\frac{P_E}{\sigma_S^2 + \sigma_N^2}} X_i, \quad i = 1, \ldots, N \tag{6}$$

where σ_S^2 and σ_N^2 are the variances of the event information S_i and the observation noise N_i, respectively.

The sink needs to calculate the estimation of each event information, S_i, in order to estimate the event S from the sensor field. Since the uncoded transmission is used, it is well-known that the minimum mean square error (MMSE) estimation is the

optimum decoding technique [6]. Hence, the estimation, Z_i, of the event information S_i is simply the MMSE estimation of Y_i, which is given by

$$Z_i = \frac{E[S_i Y_i]}{E[Y_i^2]} Y_i. \tag{7}$$

Note that the estimated values of Z_i's are spatially correlated since the actual event information S_i's are spatially correlated. This spatial correlation results in redundancy in each event information sent to the sink. Although the sink is interested in estimating the event S from the sensor field, with a distortion constraint, intuitively, this constraint can still be met by using a subset of sensor nodes rather than all the nodes in the event area.

In order to investigate the distortion achieved when a subset of nodes send information, we assume that only M out of N nodes send information to the sink, where N is the total number of sensor nodes in the event area. Since the sink decodes each Y_i using the MMSE estimator, the event source can simply be computed by taking the average of all the event information received at the sink. Then, $\hat{S}$, the estimate of S, is given as

$$\hat{S}(M) = \frac{1}{M} \sum_{i=1}^{M} Z_i. \tag{8}$$

The event distortion achieved by using M packets to estimate the event S is given as

$$D_E(M) = E\left[\left(S - \hat{S}(M)\right)^2\right] \tag{9}$$

where we use the mean-squared error as the distortion metric in (1). Using (3), (6), and (7), the estimate Z_i of each event information S_i can be written as

$$Z_i = \frac{E[S_i Y_i]}{E[Y_i^2]} \sqrt{\frac{P}{\sigma_S^2 + \sigma_N^2}} (S_i + N_i) = \frac{\sigma_S^2}{\sigma_S^2 + \sigma_N^2} (S_i + N_i). \tag{10}$$

Using (8) and (10) in (9), the distortion function $D_E(M)$ is found to be

$$D_E(M) = \sigma_S^2 - \frac{\sigma_S^4}{M(\sigma_S^2 + \sigma_N^2)} \left(2 \sum_{i=1}^{M} \rho_{(s,i)} - 1\right)$$
$$+ \frac{\sigma_S^6}{M^2(\sigma_S^2 + \sigma_N^2)^2} \sum_{i=1}^{M} \sum_{j \neq i}^{M} \rho_{(i,j)}. \tag{11}$$

$D_E(M)$ shows the event distortion achieved at the sink as a function of the number of sensor nodes M that send information to the sink and correlation coefficients $\rho_{(i,j)}$ and $\rho_{(s,i)}$ between nodes n_i and n_j, and the event S from the sensor field and the sensor node n_i, respectively.

In order to gain more insight regarding the distortion function, $D_E(M)$, shown in (11), we present a case study for a sample sensor network in Section III-B.
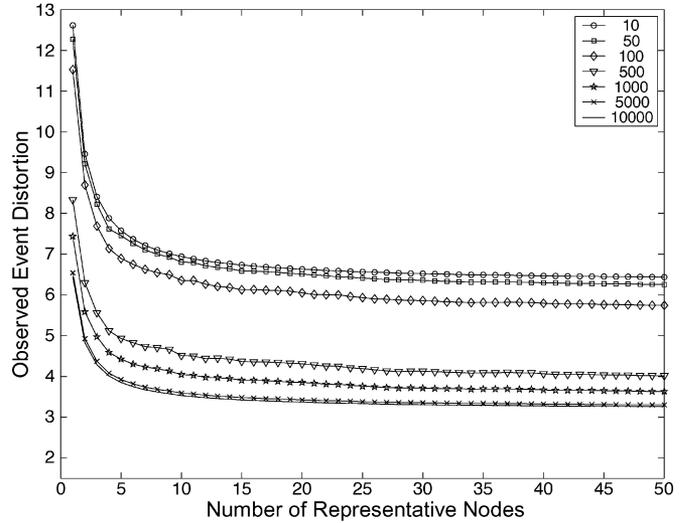

Fig. 2. Average distortion for different θ_1 values according to changing representative node number.

B. Case Study

In a 500 by 500 grid, we deployed 50 sensor nodes randomly. We use the *Power Exponential* model with $\theta_2 = 1$ and $\theta_1 = \{10, 50, 100, 500, 1000, 5000, 1000\}$ as the covariance model. For each value of θ_1 we calculated the distortion function (11) by varying the number of sensor nodes sending information. Starting from 50 nodes, the nodes that send event information to the sink is decreased. We refer to these nodes that send information to the sink as the *representative nodes*.

The simulations were performed in a fixed topology with 1000 trials for each number of representative nodes. Representative nodes are selected randomly among the 50 nodes for each trial and the distortion function is calculated according to the locations of the nodes. The average distortion calculated from these simulations is shown in Fig. 2.

As shown in Fig. 2, the achieved distortion stays relatively constant when the number of representative nodes is decreased from 50 to 15. This behavior is due to the highly redundant data sent by the sensor nodes that are close to each other. Based on the results shown in Fig. 2 and (11), we conclude the following about the achieved distortion at the sink:

Remark 1: The minimum distortion is achieved when all the sensor nodes in the event area send information to the sink. However, the achieved distortion at the sink can be preserved even though the number of the representative nodes decreases. As a result, significant energy savings are possible by allowing less number of sensor nodes send information about an event to the sink.

Remark 2: Based on (11), there are two factors effecting the distortion other than the number of representative nodes.

1) The correlation coefficient, $\rho_{(s,i)}$, between a sensor node n_i sending the information and the event source S, affects the distortion function negatively. The distortion increases as the distance between the event source S and the node n_i increases. Intuitively, if a representative sensor node is chosen apart from the source, it observes inaccurate data resulting in higher distortion at the sink.
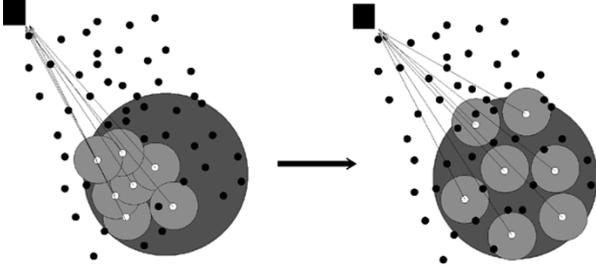
Fig. 3. Spatial re-usage in sensor networks.

2) The correlation coefficient, $\rho_{(i,j)}$, between each representative node n_i and n_j affects the distortion positively. As the distance between sensor nodes increases, the distortion decreases. Since further apart sensor nodes observe less correlated data, the distortion is decreased if these nodes are chosen as the representative nodes.

Consequently, due to the spatial correlation between sensor observations, significant energy savings can be achieved by choosing representative nodes among the sensor nodes in the event area without degrading the achieved distortion at the sink. It is clear that the reduced number of sensor nodes transmitting information decreases contention in the wireless medium resulting in decreased energy consumption due to both transmission of packets and collisions.

As a result, our aim is to find the minimum number of representative nodes that achieve the distortion constraint given by the sensor application. This minimum number is given as

$$M^* = \arg\min_M \{D_E(M) < D_{\max}\}$$

where $D_{\max}$ is the maximum distortion allowed by the sensor application.

It is important to note that the minimum number of representative nodes, M^*, depends on the locations of the representative nodes. It follows from our previous discussions that, for a fixed number of representative nodes, the minimum distortion can be achieved by choosing these nodes such that *(i)* they are located as close to the event source S as possible and *(ii)* are located as farther apart from each other as possible. As an example, as illustrated in Fig. 3, choosing representative nodes such that they are spread over the event area results in a decrease in distortion, due to less redundant data sent by these nodes. Note that such a formation also improves the medium access success during the transmission of the information. Since the representative nodes are not located close to each other, the probability of collision in the wireless medium decreases. As a result, *exploiting spatial correlation not only improves the distortion but also utilizes the wireless channel* due to the spatial reuse property of the wireless medium.

In order to apply our results on the distortion function $D_E(M)$, a node selection technique is required in order to select the representative nodes resulting in minimum energy expenditure. For this purpose, we introduce the *Iterative Node Selection (INS)* algorithm in Section IV. The INS algorithm is a control agent, running at the sink, which determines the

minimum number of representative nodes as well as their locations, based on the distortion constraint, $D_{\max}$. Accordingly, the average distance between the representative nodes is determined and each sensor node in the WSN is informed of this average distance value. Each node, then, performs the spatial CC-MAC protocol, explained in Section V, which exploits spatial correlation distributively.

IV. THE ITERATIVE NODE SELECTION (INS) ALGORITHM

According to the results obtained in Section III, we introduce the *Iterative Node Selection (INS)* algorithm to find the number and location of the representative nodes in WSN. INS tries to find the ideal locations of representative sensor nodes such that the required distortion can be maintained at the sink. Based on the INS Algorithm results, the CC-MAC protocol is performed distributively by each sensor node to achieve the required performance. The INS algorithm requires only the statistical properties of the node distribution as input and provides a correlation radius value for distributed operation as output. As pointed out in Section III, the locations of the representative sensor nodes should be chosen such that the redundancy between event information is decreased. The selection of locations of correlated points based on a distortion constraint has been investigated with the well-established Vector Quantization (VQ) methods in coding theory [10]. Hence, these methods can be exploited by suitable mapping of our problem. We first give an overview about the VQ design problem and then introduce our INS algorithm.

The VQ design can be stated as follows: Given a vector source with its statistical properties known, given a distortion constraint, and given the number of codevectors, the VQ algorithm tries to find a codebook and a partition which result in the smallest average distortion. More specifically, the VQ algorithm aims to represent all possible codewords in a code space by a subset of codewords, i.e., the codebook, within the distortion constraint. Hence, the VQ solutions fit well with our problem, where we want to represent all the sensor nodes in an event area with a smaller number of representative nodes. If we choose two dimensional codevectors, the code-space in the VQ approach can be mapped to the network topology with the node places as the codeword spaces. Once the codebook is determined, the VQ algorithm uses Voronoi regions to determine the partition of a code such that any information in this partition is represented by the codevector. The Voronoi regions determine the areas closest to the points representing the area. Applying the VQ algorithms, e.g., [10], to our representative node selection problem, the codebook and the partitions can be found. The codebook represents the locations of the representative nodes, while the partitions represent the areas of which the representative nodes are responsible for.

Since the VQ algorithms require only the statistical properties of the code space, for the selection of representative nodes, only the statistical properties of the topology are required at the sink. These properties constitute the density of the network and the type of node distribution, e.g., uniform, Gaussian, or Poisson. Therefore, the INS algorithm does not require the exact locations of the nodes to be collected at the sink. It is assumed that the statistical properties the WSN topology is known by the INS

```
INS()
    M = N;
    Calculate  D_E(M)
    while  (D_E(M) ≤ D_max)
        M = M - k
        RunVQ()
        Calculate  ρ_(s,i)  and  ρ_(i,j)   ∀ i,j
        Calculate  D_E(M)
    end
end;
```

Fig. 4.   Iterative Node Selection Algorithm. RunVQ() performs VQ algorithm and finds the places of representative nodes.
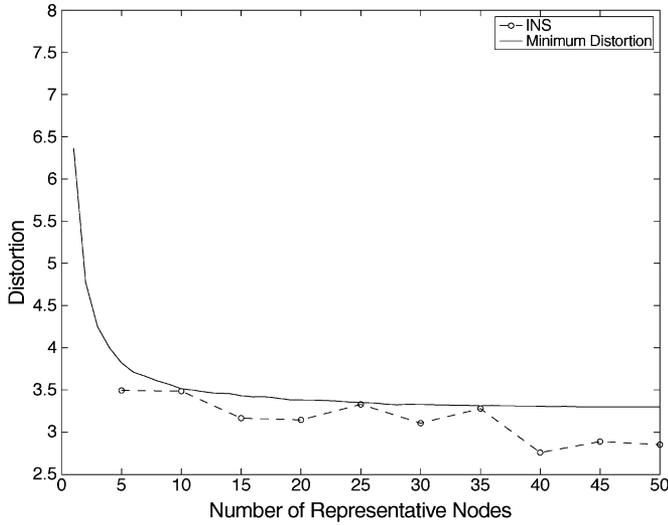

Fig. 5.   Comparison between minimum distortion among 1000 random trials and the distortion found by the INS solution.
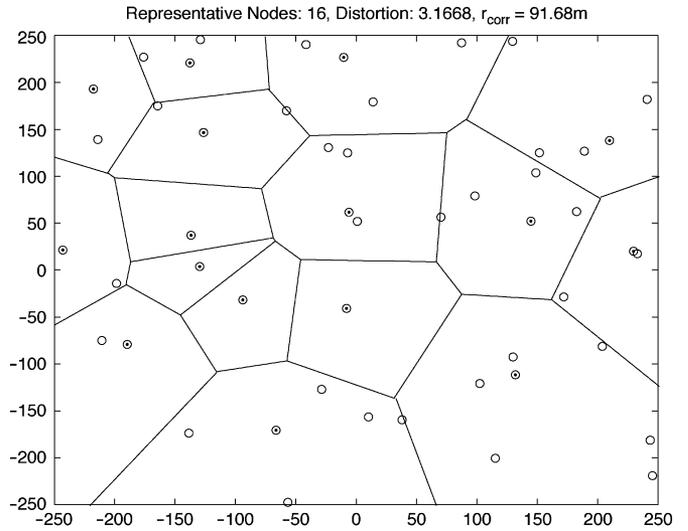

Fig. 6.   16 Representative nodes chosen by VQ algorithm and the Voronoi regions representing the correlation regions.

node is shown by a circle and a dot. The representative nodes are the only nodes that transmit their event information to the sink, while the other nodes are suppressed from sending information. The Voronoi regions determine the areas where only one node is allowed to transmit its event information to the sink.

As explained before, the INS algorithm creates a sample topology for the sensor network to exploit spatial correlation and filter correlation between the nodes. Fig. 6 shows that, if a node transmits data, then the nodes in its proximity are not required to send data. We call this area, specified by the INS algorithm, as the *correlation region* of the representative node. Based on these observations, we introduce the following definitions for clarification.

- *Correlation Radius* (r_{corr}): The radius of the correlation region is called the *correlation radius*, r_{corr}. The INS algorithm determines this value from the average radius of Voronoi regions. Nodes with internode distance, $d_{(i,j)}$, smaller than the correlation radius, r_{corr} are assumed to contain highly correlated data. This distance is assumed to be known *a priori* from the exchange of local information during network initialization or to be estimated from the received control signal strength as discussed in [16].

- *Correlation Neighbor:* A node n_j is said to be the *correlation neighbor* of node n_i if its distance, $d_{(i,j)}$, to the node n_i is smaller than the correlation radius, r_{corr}. The correlation neighbors are shown as circles in Fig. 6.

When the INS algorithm converges, the average radius of the Voronoi regions, i.e., the correlation radius, r_{corr}, is calculated and the distributed CC-MAC protocol, as explained in Section V, is then performed according to the correlation radius, r_{corr}. In order to exploit the spatial correlation between sensor nodes and to improve the performance of the WSN, our MAC protocol tries to create the correlation regions distributively. Note that the INS algorithm determines the representative nodes that would achieve the minimum distortion given the number of representative nodes. However, since this centralized information is not suitable for distributed control, only

algorithm.[3] Based on the statistical properties, the INS algorithm first forms a sample topology. Then, as shown in Fig. 4, the algorithm starts with selecting all the nodes in the event region as representative nodes and iteratively decreases the number of representative nodes, M. For each value of M, the positions of the nodes are found such that the distortion $D_E(M)$ is minimized.

The INS algorithm decreases the number of representative nodes until the distortion constraint, $D_{\max}$, is met. The VQ solution is used to find the positions of the representative nodes for each iteration using the sample topology created from the statistical properties of the network. The distortion, $D_E(M)$ is then calculated using (11). Once the maximum allowed distortion is met, the algorithm terminates.

Using a VQ toolbox [19], we simulate the INS algorithm using the same topology used in Section III-B. The locations of representative nodes are determined for 5 to 50 representative nodes. The distortion found from these selections are shown in Fig. 5. Moreover, we show the minimum distortion found among the 1000 random trials in Section III-B for $\theta_1 = 5000$. Note that, by choosing the representative node locations based on the VQ algorithm rather than the random selection, significant improvement in the achieved event distortion is possible as shown in Fig. 5. In Fig. 6, we show the locations of the selected nodes for 16 representative nodes. Each representative

---

[3]Note that the density and node distribution can be evident from the initial deployment of the WSN.

the correlation radius, r_{corr}, is informed to the individual nodes, so that they try to form the correlation regions in a distributed manner and choose the representative nodes, accordingly. Since the INS algorithm resides at the sink and requires no location information, no additional energy consumption is introduced at the sensor nodes which perform only the CC-MAC protocol.

During the lifetime of the network, the network topology can change due to failure or battery drain of sensor nodes. However, since the distortion depends on the physical phenomenon, such a change should not effect the distortion achieved at the sink unless the number of sensors decreases significantly. In such a case, the event information cannot be captured at the desired distortion level even if all the nodes send information, i.e., $r_{corr} = 0$. Hence, new nodes can be deployed. If new nodes are deployed in the sensor field in order to increase the spatial resolution or to improve the connectivity, the CC-MAC operation is not affected since the desired distortion is achieved at the sink.

Overall, the goals of the CC-MAC protocol is to determine representative nodes without any explicit internode communication, to create the correlation regions, and to prevent the correlation neighbors from transmitting its event information.

V. SPATIAL CORRELATION-BASED COLLABORATIVE MAC (CC-MAC) PROTOCOL

Based on the results presented in Section III and Section IV, we propose the spatial CC-MAC protocol that aims to collaboratively regulate sensor node transmissions. It follows from our earlier discussion in Section III that the distortion constraint can be achieved even though the number of nodes sending information about an event is decreased. Furthermore, by intelligently selecting the locations of representative nodes, the distortion $D_E(M)$ can be further reduced. In order to achieve these goals, the INS algorithm, which resides at the sink, determines the correlation radius, r_{corr}, for a distortion constraint, D_{max}, as explained in Section IV. This information is then broadcast to each sensor node during the network setup. The CC-MAC protocol, which is implemented at each sensor node, then performs MAC distributively. CC-MAC exploits spatial correlation in the MAC layer by using the correlation radius, r_{corr}, to suppress the redundant information from being injected into the WSN.

We now present the principles of CC-MAC protocol in detail. When a specific source node, n_i, transmits its event record to the sink, all of its correlation neighbors have redundant information with respect to the distortion constraint, D_{max}. This redundant information, if sent, increases the overall latency and contention within the correlation region, as well as wasting scarce WSN energy resources. Our proposed spatial CC-MAC protocol aims to *prevent the transmission of such redundant information and in addition, prioritize the forwarding of filtered data to the sink*.

In WSN, the sensor nodes have the dual functionality of being both data originators and data routers. Hence, the medium access is performed for two reasons:

- *Source Function:* Source nodes with event information perform medium access in order to transmit their packets to the sink.
- *Router Function:* Sensor nodes perform medium access in order to forward the packets received from other nodes to the next destination in the multi-hop path to the sink.
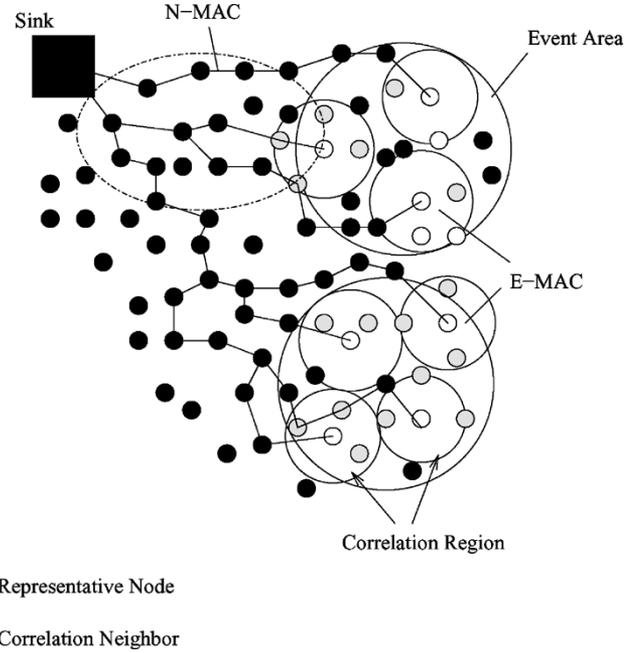

Fig. 7. E-MAC and N-MAC. The representative node transmits its record on behalf of the entire correlation region, while all correlation neighbors suppress their transmissions.

According to the spatial correlation between observations in WSN, the medium access attempts due to the source function of the sensor nodes should be coordinated such that the transmission of the redundant information to the sink is collaboratively prevented. However, once a packet is injected into the network, it has to be reliably transmitted to the sink since the correlation has now been filtered out. Hence, the route-thru packet is more valuable at an intermediate node than its own generated data packet.

In order to address these two different contention attempts in WSN, our spatial CC-MAC protocol contains two components corresponding to the source and router functionalities. *Event MAC (E-MAC)*, filters out the correlated records and *Network MAC (N-MAC)* ensures prioritization of route-thru packets. More specifically, a node performs E-MAC when it has to transmit its sensor reading to the sink, while N-MAC is performed when a node receives a packet and tries to forward it to the next hop. A typical WSN with the E-MAC and N-MAC application areas are shown in Fig. 7.

Since centralized medium access is not preferred in WSN, we use a distributed protocol to determine the representative nodes. Both E-MAC and N-MAC use a CSMA/CA based medium access control with appropriate modifications and enhancements. The information about correlation formation is embedded inside the RTS/CTS/DATA/ACK packets. Each node is informed about the correlation information about a node using these packets. As a result, additional signalling is not required for our CC-MAC protocol. We explain the packet structure and the principles of both E-MAC and N-MAC in the following subsections.

A. Packet Structure

In order to address the unique characteristics of WSN, i.e., spatially correlated information and higher priority route-thru

RTS Packet Structure

Frame Control	Duration	RA	TA	*FH*	FCS
16	16	48	48	1	32

Bits

CTS Packet Structure

Frame Control	Duration	RA	*FH*	FCS
16	16	48	1	32

Bits

DATA Packet Structure

MAC Header	*FH*	Frame Body	FCS
240	1	0–18496	32

Bits

Fig. 8.   Structures for RTS, CTS and DATA packets.

Fig. 9.   Two cases for E-MAC. The figure shows two cases for correlation region, r_{corr}, and transmission region, R.

packets, a bit in the reserved space of RTS, CTS and DATA packet structures is used as a new field called *First Hop (FH) Field* as shown in Fig. 8. The *FH field helps the sensor nodes to differentiate the type of packet*, i.e., newly generated packet (source functionality) or a route-thru packet (router functionality), and perform E-MAC or N-MAC accordingly.

When a sensor node records an event, it sets the FH field of the RTS and DATA packets related to the transmission of the sensor record. All nodes overhearing the RTS with FH field set, determine that the transmission is related to a source functionality and perform E-MAC as will be explained in Section V-B. The recipient of this RTS packet sets the FH field of the CTS packet that is sent back to the source node. As a result, each neighbor of the sender and the receiver is informed about the type of the packet being transferred. Once a node receives the DATA packet, it clears the FH field, indicating that the packet is a route-thru packet. The node, then, simply forwards the packet to the next hop. Nodes accessing the medium for router functionality do not set the FH field in their RTS packets and perform N-MAC as will be explained in Section V-C.

B. Event MAC (E-MAC)

The Event MAC (E-MAC) protocol aims to filter out correlated event information by forming correlation regions based on the correlation radius, r_{corr}, obtained from the INS algorithm as shown in Section IV. In each correlation region, a single representative sensor node transmits data for a specific duration, while all other nodes stop their transmission attempts. After each transmission duration a new representative node is selected as a result of the contention protocol. All sensor nodes contend for the medium for the first time so that the representative nodes are selected by the help of the spatial-reuse property of the wireless channel. This initial phase is called as the *first contention phase* and is explained as follows.

- *First Contention Phase (FCP):* In the first contention phase, all nodes with event information contend for the medium for the first time using RTS/CTS/DATA/ACK structure [9]. Each of these nodes sets the FH field of the RTS packet and tries to capture the medium for transmission. At the end of this phase, some of the sensor nodes
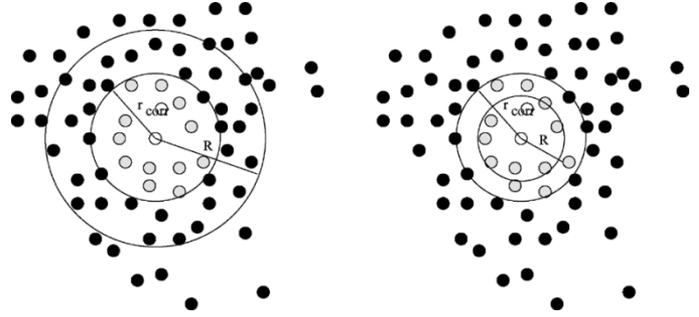
access the channel while others have to backoff. Note that more than one node can access the channel in this first phase due to the spatial reuse as shown in Fig. 7.

When a node n_i captures the channel after the first contention phase, it becomes the representative node of the area determined by the correlation radius r_{corr}. The node, n_i, continues to send information to the sink as a sole representative of its correlation region. Using the information in the ongoing transmission, other nodes hearing the transmission can determine whether they are correlation neighbors of node n_i. Every node, n_j, that listens to the RTS packet of the node n_i looks at the FH field and determines that the transmission is related to a source functionality. In addition, each node n_j determines $d_{(i,j)}$, its distance to node n_i. If $d_{(i,j)}$ is found to be less than the correlation radius, r_{corr}, then the node, n_j, determines that it is a correlation neighbor of the node n_i and stops its transmission. If the node is outside the correlation region of node n_i, then it contends for the medium if it has a packet to send.

The protocol procedure for the correlation neighbors depends on the relation between the transmission range, R, of the sensor nodes and the correlation radius, r_{corr}. For the case when $r_{\text{corr}} \leq R$, the transmission area of a node n_i contains the correlation area of the node as shown in Fig. 9. Hence, all the correlation neighbors of the node n_i can hear the transmission of node n_i. As a result, the redundancy due to correlation can be totally removed by the already ongoing transmission from the representative node. However, in the case when $r_{\text{corr}} > R$, some of the correlation neighbors of node n_i cannot hear the transmission of node n_i. Hence, the redundancy cannot be filtered out completely with respect to the total distortion constraint. Based on these observations, we first explain the procedure for the correlation neighbors for the first case and then point out the modifications for the second case.

Case 1: $r_{\text{corr}} \leq R$: In this case, when a node, n_j, determines that it is a correlation neighbor of a node, n_i, it suppresses its data transfer to the sink for a specific amount of time and performs necessary procedures for forwarding the packets. In addition, in order to conserve energy during the transmission of node n_i, each correlation neighbor enters *Suspicious Sleep State (SSS)* of duration T_{SSS}. As a result, during the SSS period, the representative node, n_i, continues sending information to the sink as a sole representative of its correlation region. Furthermore, the correlation neighbors defer contending for the medium for T_{SSS} regardless of new sensor records.

In SSS, the correlation neighbors switch to sleep state for the duration of the transmission, i.e., T_{tx}, which is extracted from the duration field of the RTS, CTS or DATA packets. However, in order to be able to forward the route-thru packets and to maintain network connectivity, correlation neighbors start to listen to the channel after a random sleep interval of t_{rs}, such that $T_{tx} < t_{rs} < T_{nextTx}$, where T_{nextTx} is the time when node n_i will begin the next transmission of the sensor record. T_{nextTx} is assumed to be set by higher layer protocols and is related to the transmission rate of the application. As a result, a correlation neighbor stays in sleep state during t_{rs}. The correlation neighbors, then, listen to the channel for EIFS sec and if there is a transmission destined to the neighbor during EIFS sec, the node performs N-MAC, otherwise it switches to sleep state again.

After T_{SSS} duration, node n_i releases its representative role by switching into SSS, leaving the medium to other nodes. The remaining nodes then go through the first contention phase again. As a result, an equal load-sharing among sensor nodes is provided.

Case 2: $r_{\text{corr}} > R$: In this case, some nodes correlated to the representative node, n_i, cannot be informed about their correlated data. This results in unnecessary contention for the medium from some of the correlation neighbors outside the communication range. However, trying to inform these nodes about node n_i's transmission requires additional transmission and contention for channel access, which increases the overhead of the protocol. Hence, there is a tradeoff between correlation filtering and protocol overhead in this case.

Another problem which needs to be addressed in this case is the routing of node n_i's packets out of the correlation region. Since all communication neighbors are also correlation neighbors, the packets generated by n_i should be routed through these nodes. However, if all correlation neighbors go into SSS as explained in Case 1, node n_i would not be able to send its packets outside the correlation region. In order to overcome this problem, we introduce the *directional sleeping technique*. The directional sleeping technique works as follows. After the first contention phase, all the correlation neighbors drop their in-queue packets. They do not try to send their packets for T_{SSS} seconds but continue to listen to the channel in order to route packets from node n_i. If, for a certain number κ of transmissions, a correlation neighbor does not receive an RTS packet destined for itself, it determines that the path from node n_i to the sink does not include itself and enters the SSS.

Other nodes that are in the route from node n_i to the sink continue to listen to the channel. This principle results in directional sleeping where nodes that are not in the path from node n_i to the sink can switch to sleep state. In addition, SSS helps those nodes not in the path to route packets coming from other sources to the sink as explained in Case 1. After the SSS, the remaining nodes enter the first contention phase and the whole process is repeated.

C. Network MAC (N-MAC)

As a node records an event and transmits its packets using E-MAC, these packets are forwarded through the network by intermediate nodes which perform the router functionality. In addition, node deployment over large sensor fields may have to deal with multiple concurrent events. Hence, when a packet is routed to the sink, it may traverse through nodes corresponding to other concurrent events. However, since the correlation has already been filtered out using E-MAC, the route-thru packet must be given priority over the packets generated by another concurrent event. This is the reason why we need a Network MAC (N-MAC) component.

When an intermediate node receives a DATA packet, it performs N-MAC to further forward that packet to the next hop. The route-thru packet is given precedence in two phases. When a correlation neighbor receives an RTS regarding a route-thru packet during the random listening period of the SSS, it switches from SSS to receive state and receives the packet. During the transmission, the representative node defers its transmission and the route-thru packet is received by the correlation neighbor.

In order to further exploit the higher priority of the route-thru packet, we use a priority scheme similar to that in IEEE 802.11 Point Coordinate Function (PCF) [9]. A node in a correlation region with a route-thru packet listens to the channel for *PCF Inter Frame Space* (PIFS) time units, which is smaller than the *DCF Inter Frame Space* (DIFS) used by the nodes performing E-MAC. The router node, then, sets its backoff window size to a random number which is between $[0, CW'_{\max} - 1]$, where $CW'_{\max}$ is a value smaller than the actual $CW_{\max}$ used by the representative node. Such a principle increases the probability that the router node captures the channel since the router node begins backoff before the representative node of the correlation region. As a result, the route-thru packet is given precedence. Since backoff procedure is still used, the collision between multiple route-thru packets that may be in the same correlation region is prevented. If, on the other hand, the representative node receives the route-thru packet, it simply gives precedence to the route-thru packet over its generated packet and forward the route-thru packet.

The two components of CC-MAC, i.e., E-MAC and N-MAC, provide a complete solution for exploiting spatial correlation in the MAC layer. E-MAC filters the highly correlated information by defining a correlation region for transmitting nodes and preventing correlation neighbors in that region from transmitting their data. N-MAC, then, provides precedence to the filtered out data while it traverses through the WSN. In Section VI, we investigate the performance of the CC-MAC protocol as well as the effects of protocol parameters.

VI. SIMULATION RESULTS

We use $ns-2$ [12] for our simulations. In order to gain more insight into the protocol operation, we first evaluate the effects of CC-MAC protocol parameters on the overall performance such as energy consumption and distortion achieved at the sink. We, then, present a comparative study of various MAC protocols. Along with our CC-MAC protocol, we simulate T-MAC [5], TRAMA [15], and S-MAC [21], which are energy aware MAC protocols designed specifically for wireless sensor networks, in addition to IEEE 802.11 [9] and a simple Carrier Sense Multiple Access (CSMA) protocol.

We present simulation results for sensor topology of 50 nodes randomly deployed in a $500 \times 500 \text{ m}^2$ sensor field. We assume

that one of the sensor nodes is a sink and all other nodes send their event information to that sink. The sensor nodes are modeled according to the $ns - 2$ wireless node module and energy model. The transmission range of each node is 100 m with average energy consumption of 24.75 mW, 13.5 mW, and 15 μW during transmitting, receiving, and sleeping, respectively. We assume that nodes consume the same energy for idle listening as receiving. The parameters $CW_{\max}$ and $CW'_{\max}$ are chosen as 1024 and 512. In each simulation, an event occurs with an event source located at the center of the sensor field. Each sensor node reports their event information to the sink. In order to investigate the effect of the traffic load, the simulation results are presented by varying the reporting period of the sensor nodes. The reporting period determines the period each node creates packets about the event information. Each simulation is performed for 600 s.

A. CC-MAC Parameters

Two parameters that are required for the CC-MAC operation are the *correlation radius*, r_{corr}, defined in Section IV and the *Suspicious Sleep State interval*, T_{SSS}, mentioned in Section V. We present the effects of these parameters on the CC-MAC performance according to the following performance metrics:

- *Average Energy Consumption (J):* is the average energy a node consumes during the simulation. The average energy consumed is the most important metric since the WSN rely on energy awareness as explained in Section I.
- *Distortion:* is the average distortion achieved at the sink according to the received sensor information. For each reporting interval, the distortion is calculated using (11), according to the information received at that interval. We use the *Power Exponential* model with $\theta_2 = 1$ and $\theta_1 = 5000$ as the covariance model. The distortion metric is presented in order to evaluate the reliability of the CC-MAC protocol in terms of distortion.

The first set of experiments are performed for the evaluation of the effect of correlation radius, r_{corr}, defined in Section IV. The correlation radius, r_{corr}, is required by the CC-MAC protocol in order to form the correlation regions for achieving the distortion constraint, $D_{\max}$, as explained in Section V. In order to observe the effect of the correlation radius, r_{corr}, on the performance of CC-MAC, we performed simulations by varying the correlation radius, r_{corr}, from 30 m to 100 m. We use Suspicious Sleep State (SSS) duration of $T_{SSS} = 50$ s for this set of experiments. This value is chosen such that the representative nodes could send multiple packets even for very low reporting rates, such that nodes do not enter the *First Contention Phase* (FCP), which is explained in Section V, for every packet. Note that this selection is made so that our CC-MAC protocol does not behave as a pure CSMA/CA with sleep periods. The results are shown in Figs. 10 and 11.

Fig. 10 shows the energy consumption for 4 different correlation radius values, i.e., $r_{\mathrm{corr}} = \{30\,\mathrm{m}, 50\,\mathrm{m}, 70\,\mathrm{m}, 100\,\mathrm{m}\}$. The average energy consumption decreases as the reporting period is increased, as expected, since less packets are generated during the simulation. It is clearly seen that significant energy conservation is possible by increasing the correlation radius, r_{corr}. Since less representative nodes transmit information to the
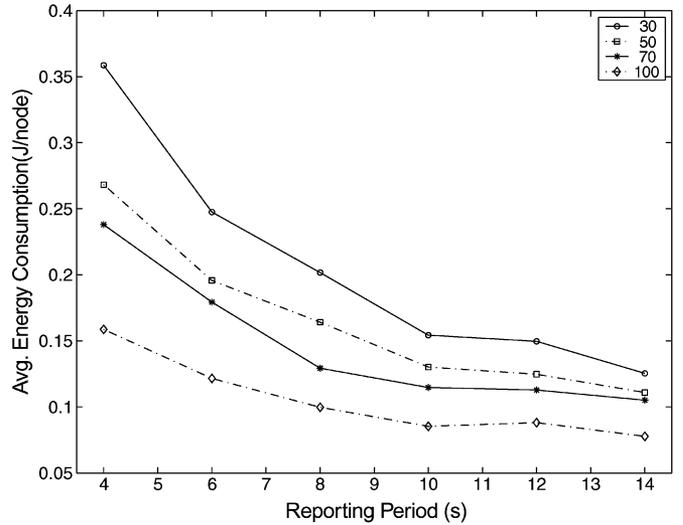

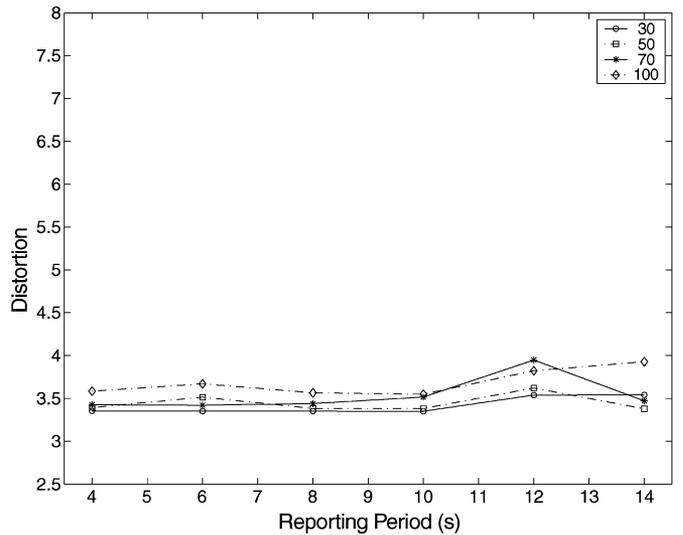Fig. 10. Energy consumption for different correlation radius values.

Fig. 11. Distortion achieved by different correlation radius values.

sink when correlation radius, r_{corr} is increased, less contention attempts are made in the overall network. The gain achieved by increasing the correlation radius is even more significant for high traffic load. As an example, increasing correlation radius from $r_{\mathrm{corr}} = 30$ m to $r_{\mathrm{corr}} = 100$ results in energy conservation of 56% for reporting period of 4 s, while the energy conservation of 38% is achieved for reporting period of 10 s. When the traffic load is increased, i.e., the reporting period is decreased, the collision probability during medium access also increases due to increased packet flow in the WSN. Hence, the effect of filtering correlated information has an increased impact on the energy conservation for high traffic load.

We show the effect of varying correlation radius, r_{corr} on the achieved distortion in Fig. 11. Fig. 11 shows that the achieved distortion is insensitive to the reporting interval. In addition, the distortion achieved stays relatively constant when the correlation radius, r_{corr}, is varied. These results are consistent with our theoretical results in Section III. In Fig. 2, the distortion stays relatively constant when the number of representative nodes is
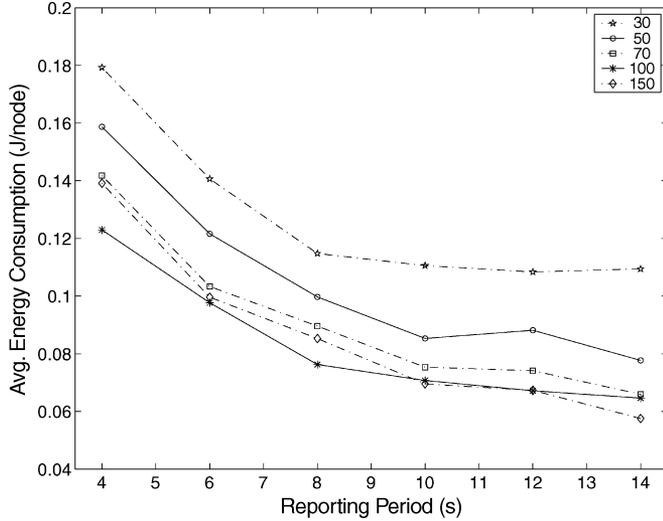
Fig. 12.   Average energy consumption for various SSS durations.

Fig. 13.   Distortion values for various SSS durations.

higher than 15. Since the correlation radius, r_{corr}, in effect, determines the number of representative nodes in CC-MAC, the same distortion behavior is also achieved in Fig. 11.

The second set of experiments considers the effect of the *Suspicious Sleep State* (SSS) duration, T_{SSS}, on the performance metrics. The SSS duration determines how long a node will stay as a representative node after capturing the channel during the *First Contention Phase* (FCP) as explained in Section V. Small T_{SSS} results in equal sharing of energy consumption by increasing the probability that every node becomes a representative node. Moreover, small T_{SSS} increases the probability that a node will be in the awake state for a route-thru packet. We perform simulations by changing the SSS duration, T_{SSS}, from 30 s to 150 s. We used $r_{\text{corr}} = 100$ m for this set of experiments.

The variation of energy consumption for various T_{SSS} is shown in Fig. 12. A decrease in the SSS duration from 150 s to 30 s, leads up to 60% increase in the energy consumption. As the SSS duration decreases, the selection of the representative nodes, i.e., *First Contention Phase* (FCP), is performed more frequently. Hence, nodes consume more energy due to the increase in number of contentions.

The distortion performance is shown in Fig. 13. The results show that the distortion is not affected by the choice of T_{SSS}. This follows from the fact that the same number of representative nodes sends information, regardless of the SSS duration. Hence, the choice of SSS duration, T_{SSS}, can be determined without any consideration about the distortion achieved at the sink.

B. Comparative Study

In this section we compare the performance results of our CC-MAC with TRAMA [15], S-MAC [21], T-MAC [5], IEEE 802.11 [9] and the basic CSMA protocol. Using the same sensor network setup as in Section VI-A, we evaluate the following performance metrics of the above protocols:

- *Medium Access Delay (s):* is the average time spent between the time a packet is handed to the MAC layer and the time it is received at the next hop. This delay accounts for
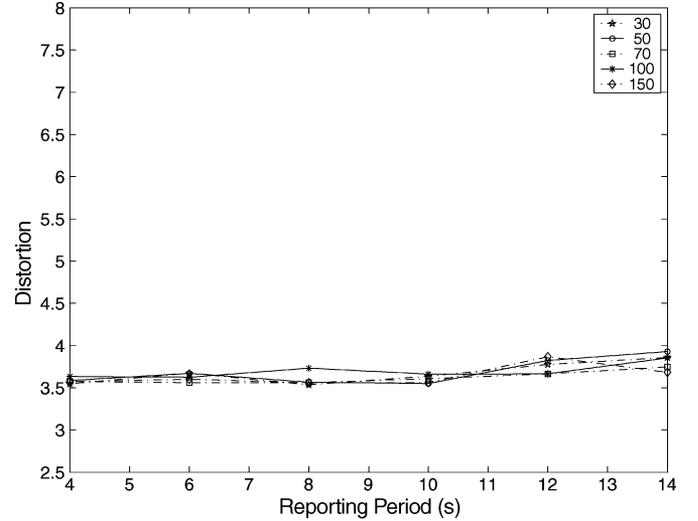
the contention delay in the case of contention-based protocols [5], [21] and scheduling delay in schedule-based protocols, [1], [15].

- *Packet Drop Rate:* is the fraction of packets that is dropped during the medium access. It is calculated as the percentage of dropped packets to the total packets sent from the MAC layer throughout the simulation. This metric shows the performance of the MAC protocol in terms of medium access overhead introduced in terms of wasted number of packets.

- *Goodput:* is the ratio between the total number of packets received at the sink and the total number of packets generated by all sensor nodes. As a result, the efficiency of the MAC protocol is investigated.

- *Average Energy Consumption (J):* is the average energy a sensor node consumes during the simulation.

TRAMA [15] is a schedule-based MAC protocol, designed specifically for WSN. Each node performs traffic-based scheduling using two-hop neighbor information in the network based on a schedule interval set according to the traffic rate. We use a schedule interval of 100 transmission slots of duration 6.82 ms for the TRAMA protocol. S-MAC [21] and T-MAC [5], on the other hand are contention-based MAC protocols that incorporate sleep cycles in order to conserve energy. Since T-MAC has been proposed to improve the energy efficiency of S-MAC by further trading off throughput and latency, we include T-MAC in our energy consumption analysis. Since the throughput and latency performance of T-MAC is designed to be inferior to S-MAC [5], we consider only S-MAC for the other performance metrics described above. We simulate the S-MAC protocol with frequency of schedule update of 10 s and 10% duty cycle. The T-MAC protocol is simulated with minimum idle listening period TA of 15 ms. The IEEE 802.11 is performed according to the $ns-2$ module existing in the $ns-2$ [12]. According to the INS protocol, we set the correlation radius $r_{\text{corr}} = 91.68$ m in order to achieve a distortion constraint of $D_{\max} = 3.16$ which accounts for 16 representative nodes in the INS solution. The SSS duration is set as $T_{SSS} = 100$ s.

Fig. 14 shows the medium access delay achieved by each MAC protocol. CSMA has the lowest medium access delay of
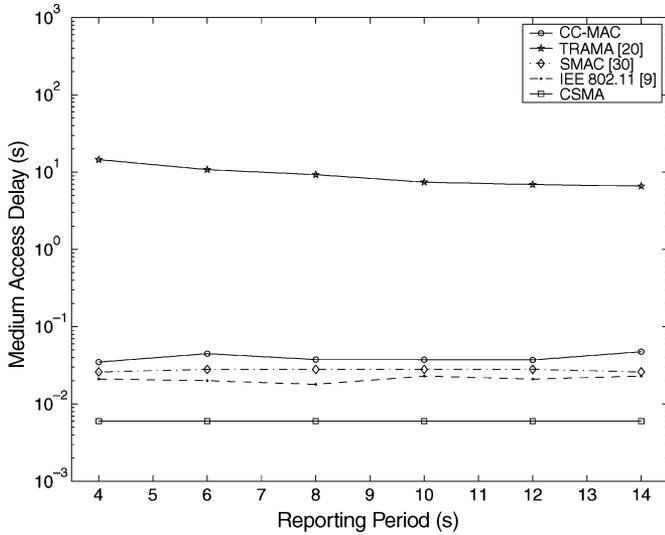
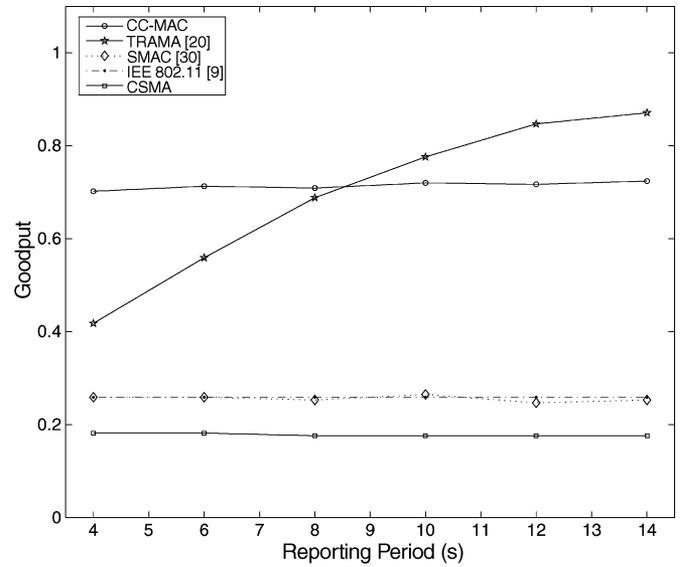Fig. 14. Medium access delay introduced by different MAC protocols.

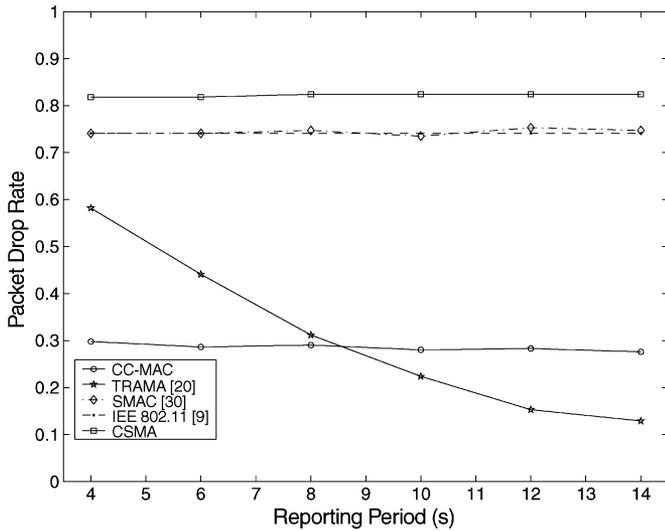Fig. 16. Goodput for different MAC protocols.

Fig. 15. Percentage of dropped packets for different MAC protocols.

6 ms, which is due to the lack of collision avoidance mechanism and energy-awareness. CC-MAC performs very close to the IEEE 802.11 and S-MAC performance with medium access delay below 50 ms. Note that the delay performance of three protocols is relatively constant for variable traffic load. TRAMA has a medium access delay of 10 s which is significantly higher than the medium access delay introduced by the contention-based protocols. This difference is due to the schedule-based medium access scheme.

The packet drop rate is shown for each of the MAC protocols in Fig. 15. S-MAC, IEEE 802.11 and CSMA achieve similar packet drop rates which is in the order of 80%, while TRAMA has a packet drop rate between 60% and 20% depending on the traffic load. CC-MAC protocol outperforms all the contention-based protocols, as well as TRAMA for high traffic load. Note that the packet drop rate is insensitive to traffic load in the case of S-MAC, IEEE 802.11, CSMA and CC-MAC. As a result, the packet drop rate depends on the number of

nodes contending for the medium. Since CC-MAC prevents correlation neighbors from contending for the medium, the packet drop rate is significantly lowered. The packet drop rate of the TRAMA protocol varies according to the traffic load. This is due to the scheduling approach of the protocol. Although collisions are prevented by scheduling transmissions, as the load increases, packet drop rate also increases since the packets cannot be accommodated in the transmission slots. For reporting period higher than 10 s, TRAMA has lower packet drop rate than CC-MAC.

The goodput of CC-MAC is shown in Fig. 16 along with CSMA, IEEE 802.11, S-MAC, and TRAMA protocols. Note that we do not compare the throughput since the efficiency of the protocols is more important than the total number of packets received at the sink. It is clear that CC-MAC would achieve lower throughput compared to other protocols, since it filters the redundant data injected into the network without compromising the distortion limits. However, sink is interested in uncorrelated data in order to construct the event features rather than the highly redundant data from each sensor node. Since the redundant contention attempts are eliminated, the goodput of CC-MAC is significantly higher than that of the other contention-based protocols achieving up to 180% increase compared to S-MAC [21] goodput. Moreover, CC-MAC outperforms TRAMA [15], which is a reservation-based protocol, when the nodes are transmitting at a high frequency. As the reporting period is increased, the frame structure of TRAMA can accommodate more nodes and goodput increases compared to CC-MAC. However, note that the number of nodes sending information to the sink is still much lower in CC-MAC than in TRAMA due to suppressed correlation neighbors. Hence, less packets are introduced into the network with less nodes transmitting data. As a result, energy consumption is decreased as shown in Fig. 17.

The energy consumption performance of CC-MAC with three energy-aware protocols, TRAMA, S-MAC and T-MAC is shown in Fig. 17. We also compare these four protocols
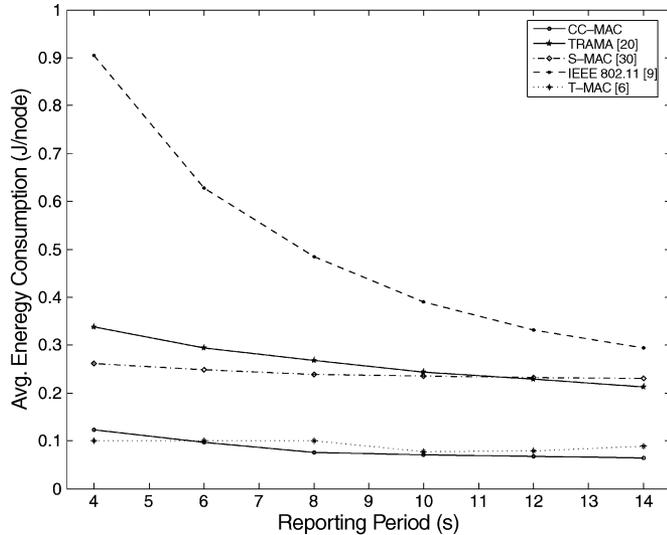
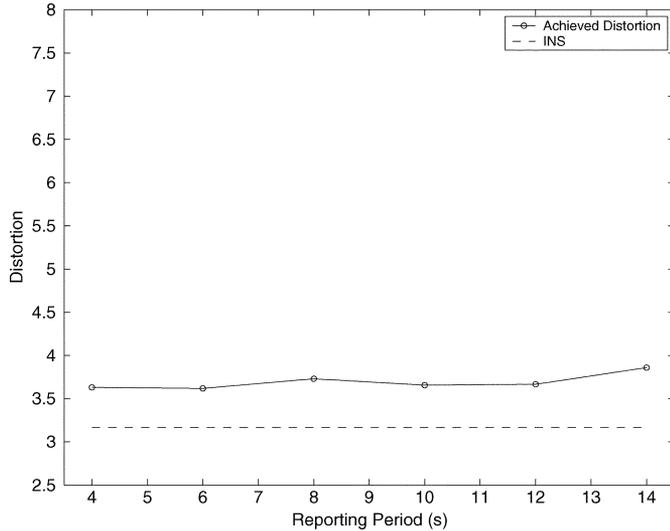Fig. 17.   Average energy consumption for different MAC protocols.

Fig. 18.   Achieved distortion and the distortion found by the INS algorithm.

against IEEE 802.11 for completeness. As shown in Fig. 17, CC-MAC has significant energy conservation compared to other energy-aware protocols. While S-MAC outperforms TRAMA for high reporting frequency, TRAMA achieves better energy consumption for increased reporting period. Since T-MAC provides variable sleep schedules based on the traffic, the energy consumption is significantly lower than S-MAC and TRAMA. However, with the help of spatial correlation-based approach of the CC-MAC protocol, the WSN consumes 25% less energy compared to T-MAC, 70% less energy compared to S-MAC and TRAMA, and 85% less energy compared to IEEE 802.11. As a result, CC-MAC protocol provides significant energy savings without compromising latency and throughput. This significant gain in energy consumption will increase the lifetime of the network.

CC-MAC also achieves distortion within 10% of the distortion constraint found theoretically by the INS algorithm as shown in Fig. 18. The increase in the distortion is due to the

lost packets during the transmission of event information to the sink. As shown in Fig. 15, 30% of the packets are dropped during transmission. The dropped packets result in an increase in the achieved distortion since less information about the event is received at the sink than anticipated by the INS algorithm. Hence, a reliability guarantee is required from higher layers for the exact achievement of required distortion. Moreover, the distributed nature of the CC-MAC protocol brings additional offset to the achieved distortion. Since the INS algorithm determines the correlation radius based on the statistical properties of the network, the realization of the network may be different due to changes in the statistical properties and different realizations. However, it is clear from our results in Section VI-A that the achieved distortion can further be decreased by decreasing the correlation radius further. Hence, by using less number of sensor nodes and filtering out the correlation between sensor observations, CC-MAC protocol achieves significant gains in terms of energy consumption in the MAC layer without compromising from latency, goodput, and distortion.

VII. CONCLUSION

Exploiting spatial correlation at the MAC layer is a powerful means of reducing the energy consumption in WSN under collective performance limits. This can be achieved by collaboratively regulating medium access so that redundant transmissions from correlation neighbors are suppressed. By allowing only a subset of sensor nodes to transmit their data to the sink, the proposed MAC protocol not only conserves energy, but also minimizes unnecessary channel access contention and thereby improves the packet drop rate without compromising the event detection latency. This is in contrast to the energy-latency tradeoffs that have been the main focus of many energy-efficient MAC proposals for WSN [5], [15], [21].

The spatial Correlation-based Collaborative MAC (CC-MAC) protocol proposed in this work is designed for distributed implementation and has two components: *Event MAC (E-MAC)* that filters out the correlation in source records and *Network MAC (N-MAC)* that prioritizes the transmission of route-thru packets over other packets. Route-thru packets are representative of an entire correlation region and hence given higher priority on their way to the sink. Using simulations, the performance of the CC-MAC protocol is investigated and significant performance gains in terms of energy consumption, latency and packet drop rate are shown.

Our work shows that by exploiting spatial correlation, the transmission of redundant nodes can be controlled. Moreover, controlling the transmission of sensor nodes has also been investigated in the application layer in terms of topology control [2]. Since these protocols focus on connectivity of the network and the traffic properties of the generated traffic, CC-MAC provides a localized control based on the spatial correlation in the physical phenomenon, when coupled with the existing topology management protocols, enables the realization of a complete solution for efficient communication in WSN.

In this paper, we consider only one type of phenomenon sensed by the sensor nodes. However, it is also possible to have multiple sensors in a node. Consequently, different phenomenon with various spatio-temporal properties can be

observed in the sensor field. Moreover, the information from these different types of sensors may have different priorities for delivery to the sink. For multiple types of traffic in WSN, CC-MAC protocol can be modified using multiple correlation radius values for each phenomenon. However, the quality of service requirements of various types of sensor information need to be considered as a future work.

REFERENCES

[1] K. A. Arisha, M. A. Youssef, and M. Y. Younis, "Energy-aware TDMA-based MAC for sensor networks," *Comput. Netw. J. (Elsevier)*, vol. 43, no. 5, pp. 539–694, Dec. 2003.

[2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Comput. Netw. J. (Elsevier)*, vol. 38, no. 4, pp. 393–422, Mar. 2002.

[3] S. Bandyopadhyay and E. J. Coyle, "Spatio-temporal sampling rates and energy efficiency in wireless sensor networks," in *Proc. IEEE INFOCOM*, Mar. 2004, vol. 3, pp. 1728–1739.

[4] J. O. Berger, V. de Oliviera, and B. Sanso, "Objective bayesian analysis of spatially correlated data," *J. Amer. Statistical Assoc.*, vol. 96, no. 456, pp. 1361–1374, 2001.

[5] T. van Dam and K. Langendoen, "An adaptive energy-efficient MAC protocol for wireless sensor networks," in *Proc. ACM SenSys 2003*, Los Angeles, CA, Nov. 2003, pp. 171–180.

[6] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," *Proc. 2nd Int. Workshop on Information Processing in Sensor Networks (IPSN'03)*, vol. 219, pp. 162–177, 2003.

[7] ——, "Power, spatio-temporal bandwidth, and distortion in large sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 745–754, Apr. 2005.

[8] T. J. Goblick, "Theoretical limitations on the transmission of data from analog sources," *IEEE Trans. Inf. Theory*, vol. IT-11, no. 4, pp. 558–567, Oct. 1965.

[9] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE 802.11, 1999.

[10] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, no. 1, pp. 84–95, Jan. 1980.

[11] S. Meguerdichian, F. Koushanfar, M. Potkonjak, and M. B. Srivastava, "Coverage problems in wireless ad hoc sensor networks," in *Proc. IEEE INFOCOM*, 2001, pp. 1380–1387.

[12] The Network Simulator—*ns*-2. [Online]. Available: http://www.isi.edu/nsnam/ns/index.html

[13] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 51–60, Mar. 2002.

[14] S. S. Pradhan, R. Puri, and K. Ramchandran, "n-channel symmetric multiple descriptions—part I: (n,k) source-channel erasure codes," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 47–61, Jan. 2004.

[15] V. Rajendran, K. Obraczka, and J. J. Garcia-Luna-Aceves, "Energy-efficient, collision-free medium access control for wireless sensor networks," in *Proc. ACM SenSys 2003*, Los Angeles, CA, Nov. 2003, pp. 181–192.

[16] A. Savvides, C. Han, and M. Srivastava, "Dynamic fine-grained localization in ad hoc networks of sensors," in *Proc. ACM MobiCom*, Jul. 2001, pp. 166–179.

[17] A. Scaglione and S. D. Servetto, "On the interdependence of routing and data compression in multi-hop sensor networks," in *Proc. ACM MobiCom*, Atlanta, GA, Sep. 2002, pp. 140–147.

[18] G. L. Stuber, *Principles of Mobile Communication*. Norwell, MA: Kluwer, 2001.

[19] Vector Quantization. [Online]. Available: http://www.dice.ucl.ac.be/~lee/software/vq/main.html

[20] M. C. Vuran, Ö. B. Akan, and I. F. Akyildiz, "Spatio-temporal correlation: theory and applications for wireless sensor networks," *Comput. Netw. J. (Elsevier)*, vol. 45, no. 3, pp. 245–259, Jun. 2004.

[21] W. Ye, J. Heidemann, and D. Estrin, "Medium access control with coordinated adaptive sleeping for wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 12, no. 3, pp. 493–506, Jun. 2004.

[22] L. Zhong, R. Shah, C. Guo, and J. Rabaey, "An ultra-low power and distributed access protocol for broadband wireless sensor networks," presented at the Networld + Interop: IEEE Broadband Wireless Summit, Las Vegas, NV, May 2001.

Mehmet C. Vuran (M'98) received the B.Sc. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2002, and the M.S. degree in electrical and computer engineering from the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, in 2004, where he is pursuing the Ph.D. degree.

He is currently a Research Assistant in the Broadband and Wireless Networking Laboratory, Georgia Institute of Technology. His research interests include adaptive medium access protocols for heterogeneous wireless architectures, wireless sensor networks, next-generation wireless networks and deep-space communication networks.

Ian F. Akyildiz (M'86–SM'89–F'96) received the B.S., M.S., and Ph.D. degrees in computer engineering from the University of Erlangen-Nuernberg, Germany, in 1978, 1981, and 1984, respectively.

Currently, he is the Ken Byers Distinguished Chair Professor with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, and Director of the Broadband and Wireless Networking Laboratory. He is an Editor-in-Chief of *Computer Networks* and the *Ad Hoc Networks Journal*. His current research interests are in sensor networks, IPN Internet, wireless networks, and satellite networks.

Dr. Akyildiz received the Don Federico Santa Maria Medal for his services to the Universidad de Federico Santa Maria in Chile in 1986. From 1989 to 1998, he served as a National Lecturer for ACM and received the ACM Outstanding Distinguished Lecturer Award in 1994. He received the 1997 IEEE Leonard G. Abraham Prize Award (IEEE Communications Society) for his paper entitled "Multimedia Group Synchronization Protocols for Integrated Services Architectures" published in the IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS (JSAC) in January 1996. He received the 2002 IEEE Harry M. Goode Memorial Award (IEEE Computer Society) with the citation "for significant and pioneering contributions to advanced architectures and protocols for wireless and satellite networking". He received the 2003 IEEE Best Tutorial Award (IEEE Communication Society) for his paper entitled "A Survey on Sensor Networks," published in *IEEE Communications Magazine* in August 2002. He also received the 2003 ACM Sigmobile Outstanding Contribution Award with the citation "for pioneering contributions in the area of mobility and resource management for wireless communication networks". He has been a Fellow of the Association for Computing Machinery (ACM) since 1996.