



# Spatio-temporal correlation: theory and applications for wireless sensor networks

Mehmet C. Vuran<sup>\*</sup>, Özgür B. Akan, Ian F. Akyildiz

*Broadband and Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology,  
Atlanta, GA 30332, USA*

Available online 31 March 2004

## Abstract

Wireless Sensor Networks (WSN) are characterized by the dense deployment of sensor nodes that continuously observe physical phenomenon. Due to high density in the network topology, sensor observations are highly correlated in the space domain. Furthermore, the nature of the physical phenomenon constitutes the temporal correlation between each consecutive observation of a sensor node. These spatial and temporal correlations along with the collaborative nature of the WSN bring significant potential advantages for the development of efficient communication protocols well-suited for the WSN paradigm. In this paper, several key elements are investigated to capture and exploit the correlation in the WSN for the realization of advanced efficient communication protocols. A theoretical framework is developed to model the spatial and temporal correlations in WSN. The objective of this framework is to enable the development of efficient communication protocols which exploit these advantageous intrinsic features of the WSN paradigm. Based on this framework, possible approaches are discussed to exploit spatial and temporal correlation for efficient medium access and reliable event transport in WSN, respectively.

© 2004 Elsevier B.V. All rights reserved.

*Keywords:* Spatial correlation; Temporal correlation; MAC protocol; Transport protocol; Wireless sensor networks

## 1. Introduction

The technological advances in the micro-electro-mechanical systems (MEMS) and the wireless communications have enabled the deployment of wireless sensor networks which may provide sensing capabilities in space and time that surpass the traditional sensing approaches. Wireless sensor

networks (WSN) are event-based systems that rely on the collective effort of densely deployed several microsensor nodes which continuously observe physical phenomenon.

The main objective of the WSN is to reliably detect/estimate event features from the collective information provided by sensor nodes. Therefore, the energy and hence processing constraints of small wireless sensor nodes are overcome by this collective sensing notion which is realized via their networked deployment. While the collaborative nature of the WSN brings significant advantages over traditional sensing including greater accuracy, larger coverage area, and extraction of

<sup>\*</sup> Corresponding author. Tel.: +1-404-894-5141; fax: +1-404-894-7883.

*E-mail addresses:* [mcvuran@ece.gatech.edu](mailto:mcvuran@ece.gatech.edu) (M.C. Vuran), [akan@ece.gatech.edu](mailto:akan@ece.gatech.edu) (Ö.B. Akan), [ian@ece.gatech.edu](mailto:ian@ece.gatech.edu) (I.F. Akyildiz).

localized features; the spatio-temporal correlation among the sensor observations is another significant and unique characteristic of the WSN which can be exploited to drastically enhance the overall network performance. The characteristics of the correlation in the WSN can be summarized as follows:

- *Spatial correlation.* Typical WSN applications require spatially dense sensor deployment in order to achieve satisfactory coverage [4,11]. As a result, multiple sensors record information about a single event in the sensor field. Due to high density in the network topology, spatially proximal sensor observations are highly correlated with the degree of correlation increasing with decreasing internode separation.
- *Temporal correlation.* Some of the WSN applications such as event tracking may require sensor nodes to periodically perform observation and transmission of the sensed event features. The nature of the energy-radiating physical phenomenon constitutes the temporal correlation between each consecutive observation of a sensor node [10]. The degree of correlation between consecutive sensor measurements may vary according to the temporal variation characteristics of the phenomenon.

In addition to the collaborative nature of the WSN, the existence of above mentioned spatial and temporal correlations bring significant potential advantages for the development of efficient communication protocols well-suited for the WSN paradigm. For example, intuitively, due to the spatial correlation, data from spatially separated sensors is more useful to the sink than highly correlated data from nodes in proximity. Therefore, it may not be necessary for every sensor node to transmit its data to the sink; instead, a smaller number of sensor measurements might be adequate to communicate the event features to the sink within a certain reliability/fidelity level. Similarly, for a certain event tracking application, the measurement reporting frequency, at which the sensor nodes transmit their observations, can be adjusted such that temporal-correlated phenomenon signal is captured at the sink within a certain

distortion level and with minimum energy-expenditure.

There has been some research effort to study the correlation in WSN [6,13,14,18]. However, most of these existing studies investigate the information theoretical aspects of the correlation, and they do not provide efficient networking protocols which exploit the correlation in the WSN. For example, in [6], a joint source-channel coding paradigm is proposed for optimal performance in a class of sensor networks. On the other hand, there exist some proposals which attempt to exploit spatial correlation in WSN [13,14]. However, these schemes aim to find the optimum rate to compress redundant information in the sensor observations and they also do not propose to exploit correlation for developing efficient communication protocols for the WSN. In a recent effort, the joint routing and source coding is introduced in [18] to reduce the amount of traffic generated in dense sensor networks with spatially correlated records. While joint routing and source coding reduces the number of transmitted bits; from the network point of view, the number of transmitted packets remains unchanged, which can be further minimized by regulating the network access based on the spatial correlation between the sensor nodes.

On the other hand, there already exists significant amount of research on the communication protocols for sensor networks in the literature [1]. For example, there exist some proposals to address the medium access control (MAC) problems in wireless sensor networks. However, these solutions mostly focus on energy-latency tradeoffs. S-MAC [24] aims to decrease the energy consumption by using sleep schedules with virtual clustering. A variant of S-MAC, T-MAC, incorporates variable sleep schedules to further increase the energy consumption [5]. However, in both protocols, since spatial correlation is not exploited, sensor nodes continue to send redundant data with increased latency due to sleep durations. In addition to contention-based protocols, TDMA-based protocols have also been proposed. In [2], an energy-aware TDMA-based MAC protocol is presented where cluster-based centralized slot assignment is used. IEEE 802.15.4 is also being proposed for low data rate wireless networks [9]. Although these

protocols aim energy efficiency; both of the protocols assume a cluster-based topology, which requires significant additional processing complexity in the overall sensor network. In [16], energy efficient collision-free MAC protocol is presented. The protocol is based on a time-slotted structure and uses a distributed election scheme based on traffic requirements of each node to determine the time slot that a node should use. 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, which, in the case of high density sensor networks, may result in either incomplete neighbor information due to collisions or high energy consumption due to signaling cost. In summary, none of these MAC protocols take advantage of the correlation in the WSN in order to improve energy-efficiency without compromising on the access latency.

Similarly, there also exist transport layer proposals for wireless sensor networks in the current literature. In [23], the PSFQ (Pump Slowly, Fetch Quickly) mechanism is proposed for reliable re-tasking/reprogramming in WSN. However, PSFQ is applicable only for strict sensor-to-sensor reliability and for purposes of control and management in the reverse direction from the sink to sensor nodes. Individual data flows are correlated and loss tolerant to the extent that desired event features are collectively and reliably informed to the sink. Hence, the use of PSFQ for the forward direction can lead to a waste of valuable resources. In [19], the RMST (Reliable Multi-Segment Transport) protocol is proposed to utilize in-network caching and provide guaranteed delivery of the data packets generated by the event flows. However, such guaranteed reliability via in-network caching may bring significant overhead for the sensor networks with power and processing limitations. The congestion detection and avoidance (CODA) protocol for sensor networks is presented in [22]. However, the CODA does not address the reliable event transport to the sink. Similarly, what is common in all of these works is that none of these solutions exploit the correlation to achieve energy-efficient communication in WSN.

Consequently, despite the existing proposed works for medium access control and transport

layer solutions in sensor networks discussed above, these protocols do not exploit the correlation in the sensed phenomenon. In this paper, several key elements are investigated to capture and exploit the correlation in the WSN for the realization of advanced efficient communication protocols. We first develop a theoretical framework to model the spatial and temporal correlations in sensor networks. The objective of this framework is to enable the development of efficient communication protocols which exploit these advantageous intrinsic features of the WSN paradigm. Based on this framework, possible approaches are discussed to exploit spatial and temporal correlation for efficient medium access and reliable event transport in WSN, respectively.

The remainder of this paper is organized as follows. In Section 2, the theoretical framework is developed to model the spatial and temporal correlations in wireless sensor networks. In Section 3, applications of the results of this framework are explored for energy-efficient communication in WSN. More specifically, in Section 3.1, we discuss an efficient medium access control approach in WSN which aims to reduce the energy consumption of the network by exploiting spatial correlation in the WSN without compromising the access latency. In Section 3.2, we explore a reliable event transport mechanism exploiting temporal correlation with an objective of reliable event detection with minimum energy expenditure. Finally, the concluding remarks are discussed in Section 4.

## 2. Spatio-temporal correlation in wireless sensor networks

In this section, we develop the theoretical framework for the spatio-temporal correlation in wireless sensor networks. The results of this analysis are then further explored in the context of correlation-based medium access control and reliable event transport approaches for WSN.

### 2.1. Architecture and correlation model for WSN

In a sensor field, each sensor observes the noisy version of a physical phenomenon. The sink is

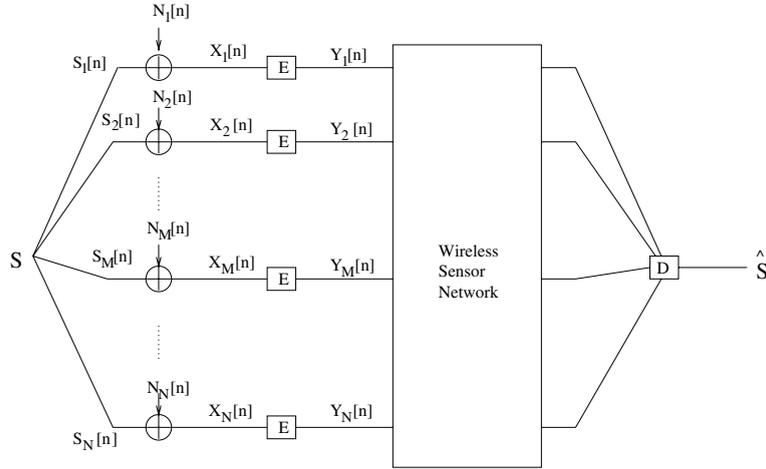


Fig. 1. Correlation model and architecture.

interested in observing the physical phenomenon using the observations from sensor nodes with the highest accuracy. The physical phenomenon in interest can be modeled as a spatio-temporal process  $s(t, x, y)$  as a function of time  $t$  and spatial coordinates  $(x, y)$ .

Depending on the specific sensor application, the physical phenomenon may be a spatio-temporal process generated by a point source in case of applications such as object tracking. In this case, the sink is interested in reconstructing the source signal at a specific location  $(x_0, y_0)$  based on sensor observations. In other applications, the spatio-temporal process may be a combination of multiple point sources where the sink is interested in reconstructing the signal in multiple locations or over an event area. Although the reconstruction is application specific, the properties of the observations can be modeled based on the spatio-temporal process  $s(t, x, y)$ .

The model for the information gathered by  $N$  sensors in an event area is illustrated in Fig. 1. The sink is interested in estimating the event source,  $S$ , according to the observations of the sensor nodes,  $n_i$ , in the event area. 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 get the estimate,  $\hat{S}$ , of the event source  $S$ . The encoders and the decoders are labelled as  $E$  and  $D$  in Fig. 1, respectively. Using this model, we will exploit various aspects of correlation among sensor readings both in terms of time and space.

Each observed sample,  $X_i[n]$ , of sensor  $n_i$  at time  $n$  is represented as

$$X_i[n] = S_i[n] + N_i[n], \quad (1)$$

where the subscript  $i$  denotes the spatial location of the 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 time  $t = t_n$ <sup>1</sup> and  $(x, y) = (x_i, y_i)$ , 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  $\sigma_N^2$ . 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  $\forall n$ .

As it is shown in Fig. 1, each observation  $X_i[n]$  is then encoded into  $Y_i[n]$  by the source-coding at the sensor node as

$$Y_i[n] = f_i(X_i[n]) \quad (2)$$

<sup>1</sup> Note that, we use a discrete-time model since each node is assumed to sample the physical phenomenon synchronously after the initial wake-up.

and then sent through the network to the sink. The sink decodes the received data to reconstruct an estimation  $\hat{S}$  of the source  $S$ :

$$\hat{S} = g(Y_1[n_1], \dots, Y_1[n_\tau]; \dots; Y_N[n_1], \dots, Y_N[n_\tau]) \quad (3)$$

based on the data received from  $N$  nodes in the event area over a time period  $\tau = t_{n_\tau} - t_{n_1}$ . The sink is interested in reconstructing the source  $S$  according to a distortion constraint

$$D = E[d(S, \hat{S})]. \quad (4)$$

In the next sections, the general distortion function in (4) will be used to independently obtain the distortion functions for spatial and temporal correlation in the WSN, which can then be explored in discussing the correlation-based medium access control and reliable event transport approaches for WSN in Sections 3.1 and 3.2, respectively.

## 2.2. Spatial correlation in WSN

In this section, we model the spatial correlation between observations of each sensor node. The information gathered by  $N$  sensors in an event area can be modeled as shown in Fig. 1. The sink is assumed to be interested in a point source  $S$ . Since we only consider the spatial correlation between nodes, in this analysis, we assume that the samples are temporally independent. Hence, by dropping the time index  $n$ , (1) can be restated as

$$X_i = S_i + N_i, \quad i = 1, \dots, N. \quad (5)$$

The sink is interested in reconstructing the source  $S$  according to observations of nodes  $n_i$  which observe the spatially correlated version of  $S$  at  $(x_i, y_i)$ , i.e.,  $S_i$ . The physical phenomenon is modeled as joint gaussian random variables (JGRVs) at each observation point as

$$E\{S_i\} = 0, \quad i = 1, \dots, N,$$

$$\text{var}\{S_i\} = \sigma_S^2, \quad i = 1, \dots, N,$$

$$\text{cov}\{S_i, S_j\} = \sigma_S^2 \text{corr}\{S_i, S_j\},$$

$$\text{corr}\{S_i, S_j\} = \rho_{i,j} = K_\theta(d_{i,j}) = \frac{E[S_i S_j]}{\sigma_S^2},$$

where  $d_{i,j} = \|\mathbf{s}_i - \mathbf{s}_j\|$  denotes the distance between nodes  $n_i$  and  $n_j$  located at coordinates  $\mathbf{s}_i$  and  $\mathbf{s}_j$ , respectively, and  $K_\theta(\cdot)$  is the correlation model. The covariance function is assumed to be non-negative 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 groups [3]:

- *Spherical*:

$$K_\theta^S(d) = \begin{cases} 1 - \frac{3}{2} \frac{d}{\theta_1} + \frac{1}{2} \left(\frac{d}{\theta_2}\right)^3 & \text{if } 0 \leq d \leq \theta_1; \\ 0 & \text{if } d > \theta_1 \end{cases}$$

$$\theta_1 > 0.$$

In this model, two observations taken more than  $\theta_1$  distance apart are uncorrelated.

- *Power exponential*:

$$K_\theta^{\text{PE}}(d) = e^{(-d/\theta_1)^{\theta_2}}; \quad \theta_1 > 0, \theta_2 \in (0, 2].$$

For  $\theta_2 = 1$ , the model becomes exponential, while for  $\theta_2 = 2$  squared exponential.

- *Rational quadratic*:

$$K_\theta^{\text{RQ}}(d) = \left(1 + \left(\frac{d}{\theta_1}\right)^2\right)^{-\theta_2}; \quad \theta_1 > 0, \theta_2 > 0.$$

- *Matérn*:

$$K_\theta^{\text{M}}(d) = \frac{1}{2^{\theta_2-1} \Gamma(\theta_2)} \left(\frac{d}{\theta_1}\right)^{\theta_2} \mathcal{K}_{\theta_2}\left(\frac{d}{\theta_1}\right);$$

$$\theta_1 > 0, \theta_2 > 0,$$

where  $\mathcal{K}_{\theta_2}(\cdot)$  is the modified Bessel function of second kind and order  $\theta_2$ .

The correlation model can be chosen according to the properties of the physical phenomenon the sink is interested in. Since we are interested in  $S$ , which is also a JGRV, we use a special notation with

$$\text{var}\{S\} = \sigma_S^2,$$

$$\text{corr}\{S, S_i\} = \rho_{s,i} = K_\theta(d_{s,i}) = \frac{E[SS_i]}{\sigma_S^2},$$

where  $d_{s,i}$  denotes the distance between the source  $S$  and the node  $n_i$ . The observation noise  $N_i$  of each

node  $n_i$  is modeled as i.i.d. Gaussian random variable with zero mean and variance  $\sigma_N^2$ , i.e.,  $N_i \sim \mathcal{N}(0, \sigma_N^2)$ .

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 joint source-channel coding outperforms separate coding [6,15]. In addition, for Gaussian sources, if the source is Gaussian and the cost on the channel is the encoding power, then uncoded transmission is optimal for point to point transmission [7]. Furthermore, for sensor networks with finite number of nodes, uncoded transmission outperforms any approach based on the separation paradigm leading to the optimal solution for infinite number of nodes [6]. Hence, we adopt uncoded transmission for the sensor observations in this work. Each node  $n_i$  sends to the sink, a scaled version,  $Y_i$ , of the observed sample  $X_i$  according to encoding power constraint  $P_E$ :

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

where  $\sigma_S^2$  and  $\sigma_N^2$  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 source  $S$ . Since uncoded transmission is used, it is well known that minimum mean square error (MMSE) estimation is the optimum decoding technique [12]. Hence, the estimation,  $Z_i$ , of the event information  $S_i$  is simply the MMSE estimation of  $Y_i$ , which is given by

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

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 source,  $S$ , with a distortion constraint, intuitively, this constraint can still be met by using a smaller number of sensor nodes rather than all the nodes in the event area. In order to investigate the distortion achieved when smaller number of nodes sending information, we assume that only  $M$  out of  $N$  packets are received by 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

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

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

$$D(M) = E[(S - \hat{S}(M))^2], \quad (9)$$

where we use the mean-squared error as the distortion metric. Using (5) and (6) in (7), the estimate  $Z_i$  of each event information  $S_i$  can be written as

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

denoting  $\alpha = \sqrt{P/(\sigma_S^2 + \sigma_N^2)}$ ,

$$E[S_i Y_i] = \alpha \sigma_S^2,$$

$$E[Y_i^2] = \alpha^2 (\sigma_S^2 + \sigma_N^2),$$

then, (10) is restated as

$$Z_i = \frac{\sigma_S^2}{\sigma_S^2 + \sigma_N^2} (S_i + N_i). \quad (11)$$

Using (11) and (8) in (9), the distortion function  $D(M)$  is found to be

$$D(M) = \sigma_S^2 - \frac{\sigma_S^4}{M(\sigma_S^2 + \sigma_N^2)} \left( 2 \sum_{i=1}^M \rho_{(s,i)} - 1 \right) + \frac{\sigma_S^6}{M^2(\sigma_S^2 + \sigma_N^2)^2} \sum_{i=1}^M \sum_{j \neq i}^M \rho_{(i,j)}. \quad (12)$$

$D(M)$  shows the distortion achieved at the sink as a function of number of 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 source  $S$  and node  $n_i$ , respectively. Based on the distortion function, we discuss possible approaches that can be used in the Medium Access Control (MAC) protocols for WSN in Section 3.1.

### 2.3. Temporal correlation in WSN

As mentioned in Section 1, the energy-radiating physical phenomenon constitutes the temporal correlation between each consecutive observation of a sensor node [10]. For the periodic sensing applications such as event tracking, each consecutively taken sensor observations are temporally correlated to a certain degree. In this section, we establish the theoretical analysis for this temporal correlation, which will be further elaborated in the context of correlation-based reliable event transport approach discussed in Section 3.2.

Here, we consider the temporal correlation between the sensor observations and hence we omit the spatial variation in this analysis. We are interested estimating the signal  $s(t)$  in a decision interval of  $\tau$ . In our theoretical analysis, we model an event-to-sink distortion metric, where all the information coming from the sensor nodes in the event area is considered as if it is generated by a single source node during the decision interval  $\tau$ .

Assume that the sensed information from the sensors are sent to the sink using a reporting frequency of  $f$ . In this case, we seek to control the reporting frequency  $f$  such that a desired distortion level is not exceeded in the estimation of the event features at the sink. The event signal  $s(t)$  is assumed to be a Gaussian random process with  $\mathcal{N}(0, \sigma_s^2)$ . The sink is interested in finding the expectation of the signal  $s(t)$  over the decision interval  $\tau$ , i.e.,  $S(\tau)$ . Assuming the observed signal  $s(t)$  is wide-sense stationary (WSS), the expectation of the signal over the decision interval  $\tau$  can be calculated by the time average of the observed signal [8], i.e.

$$S(\tau) = \frac{1}{\tau} \int_{t_0}^{t_0+\tau} s(t) dt, \quad (13)$$

where  $t_0$  is the time the sensor node wakes up for the sampling of the signal. With a change of variables,  $S(\tau)$  can be shown as

$$S(\tau) = \frac{1}{\tau} \int_0^\tau s(t_0 + \Gamma) d\Gamma. \quad (14)$$

We define the value of the signal at each sampling interval as

$$S[n] = s\left(t_0 + \frac{n}{f}\right), \quad (15)$$

where  $f$  is the sampling frequency and  $S[n]$  are JGRV with  $\mathcal{N}(0, \sigma_s^2)$ .<sup>2</sup> For the derivation of the distortion function, the following definitions are needed:

$$E\{S[n]\} = 0,$$

$$E\{(S[n])^2\} = \sigma_s^2,$$

$$E\{S[n]S[m]\} = \sigma_s^2 \hat{\rho}_s(n, m),$$

$$E\{s(t)s(t + \delta)\} = \sigma_s^2 \rho_s(\delta),$$

where  $\hat{\rho}_s(n, m) = \rho_s(|m - n|/f)$  is the covariance function that depends on the time difference between signal samples. Although our results about the distortion function apply to all the covariance models introduced in Section 2.2, we use the power exponential model in the derivation since the physical event information such as, electromagnetic waves, is modeled to have an exponential autocorrelation function [20]. Hence, the covariance function becomes

$$\rho_s(\delta) = e^{-|\delta|/\theta_1}. \quad (16)$$

Each sensor node observes the noisy version of the signal given as

$$X[n] = S[n] + N[n] \quad (17)$$

and the transmitted signal is expressed by

$$Y[n] = \sqrt{\frac{P_E}{\sigma_s^2 + \sigma_N^2}} X[n] \quad (18)$$

<sup>2</sup> Note that, the samples of a Gaussian random process are jointly Gaussian [8].

based on the discussion in Section 2.2. Using the MMSE estimator at the sink, each sample is estimated as

$$Z[n] = \frac{E[S[n]Y[n]]}{E[Y^2[n]]} Y[n]. \quad (19)$$

Hence, each estimated sample from the sensor nodes can be represented as

$$Z[n] = \frac{\sigma_S^2}{\sigma_S^2 + \sigma_N^2} (S[n] + N[n]). \quad (20)$$

After collecting all the samples of the signal in the decision interval  $\tau$ , the sink estimates the expectation of the signal over the last decision interval by

$$\hat{S}(\tau) = \frac{1}{\tau f} \sum_{k=1}^{\tau f} Z[k], \quad (21)$$

where  $\tau f$  is the total number of sensor samples taken within a decision interval with duration of  $\tau$ . As a result, the distortion achieved by using  $\tau f$  samples to estimate the event is given as

$$D = E[(S(\tau) - \hat{S}(\tau))^2]. \quad (22)$$

Using the definitions above and substituting (14), (20), and (21) into (22); the distortion function can easily shown to be

$$\begin{aligned} D(f) = & \sigma_S^2 + \frac{\sigma_S^4}{\tau f (\sigma_S^2 + \sigma_N^2)} \\ & + \frac{\sigma_S^6}{\tau^2 f^2 (\sigma_S^2 + \sigma_N^2)^2} \sum_{k=1}^{\tau f} \sum_{l \neq k} e^{-(|k-l|/f)/\theta_1} \\ & - \frac{2\sigma_S^4 \theta_1}{\tau^2 f (\sigma_S^2 + \sigma_N^2)} \sum_{k=1}^{\tau f} (2 - e^{-k/(f\theta_1)} \\ & \quad - e^{-(\tau-k)/\theta_1}). \quad (23) \end{aligned}$$

It is observed from (23) that the distortion in the estimation decreases with increasing  $f$ . Note that a distortion level  $D$  for the estimation of event features from the sensor observations means the reliability level of the event-to-sink communication in the WSN. In Section 3.2, this distortion function will be further explored in the context of reliable event transport in WSN.

### 3. Exploiting correlation in WSN

Spatial and temporal correlation, in addition to the collaborative nature of the WSN, bring significant potential advantages for the development of efficient communication protocols well-suited for the WSN paradigm. In this section, we discuss possible approaches exploiting spatial and temporal correlation to achieve energy-efficient medium access and reliable event transport in WSN, respectively.

#### 3.1. Correlation-based medium access control

The shared wireless channel between sensor nodes and energy considerations of the WSN make the Medium Access Control (MAC) a crucial part of the wireless sensor phenomenon. The distributed architecture and the application-oriented traffic properties of the WSN pose interesting challenges for the development of MAC protocols. Furthermore, the scarce energy sources of sensor nodes necessitate energy aware MAC protocols. Hence, MAC protocols for WSN should be developed tailored to the physical properties of the sensed phenomenon and the specific network properties so that the access to the channel is coordinated with minimum collisions without effecting the connectivity throughout the network.

In WSN, many individual nodes deployed in large areas sense events and send corresponding information about these events to the sink. When an event occurs in the sensor field, all the nodes in an event area collect information about the event taking place and try to send this information to the sink. Due to the physical properties of the event, this information may be highly correlated in nature according to the spatial correlation between sensor nodes. Intuitively, data from spatially separated sensors is more useful to the sink than highly correlated data from closely located sensors. Hence, it may not be necessary for every sensor node to transmit its data to the sink; instead, a smaller number of sensor measurements might be adequate to communicate the event features to the sink within a certain distortion constraint. As a result, the MAC protocol can *reduce the energy consumption of the network by exploiting spatial cor-*

relation in the WSN without compromising on the access latency as well as the distortion achieved.

In accordance with our intuition, in Section 2.2, the theoretical framework of spatial correlation in WSN is developed and the distortion resulting from sending a subset  $M$  of total  $N$  nodes in the event area is derived (12). The relations between the positions of the sensor nodes in the event area and the event estimation reliability is also important for exploiting spatial correlation.

In order to gain more insight to our intuitions, we performed a case study using the distortion function (12). In a 500 by 500 grid, we deployed 50 sensor nodes randomly. We use the *Power Exponential* model with  $\theta_2 = 1$  and  $\theta_1 = \{10, 50, 100, 500, 1000, 5000, 10000\}$  as the covariance model for the covariance function,  $K_\vartheta(\cdot)$  in (1). The parameter,  $\theta_1$ , controls the relation between the distance of the nodes and the correlation coefficient. For each value of  $\theta_1$  we calculate the distortion function (12) varying the number of sensor nodes sending information. Starting from 50 nodes, we decrease the number of nodes that send event information to the sink. We refer to these nodes as the *representative nodes*.

The simulations were done for 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 these nodes. The average distortion calculated from these simulations and the distribution of the distortion for each number of representative nodes 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. In addition, with increasing  $\theta_1$ , the observed event distortion decreases since close nodes become less correlated with increasing  $\theta_1$ . Based on the results shown in Fig. 2 and the distortion function (12), the following discussions about the observed distortion at the sink can be made.

**Remark 1.** The minimum distortion is achieved when all the 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 saving is possible by allowing less number of nodes send information to the sink about an event.

**Remark 2.** Based on (12), there are two factors effecting the distortion other than the number of representative nodes:

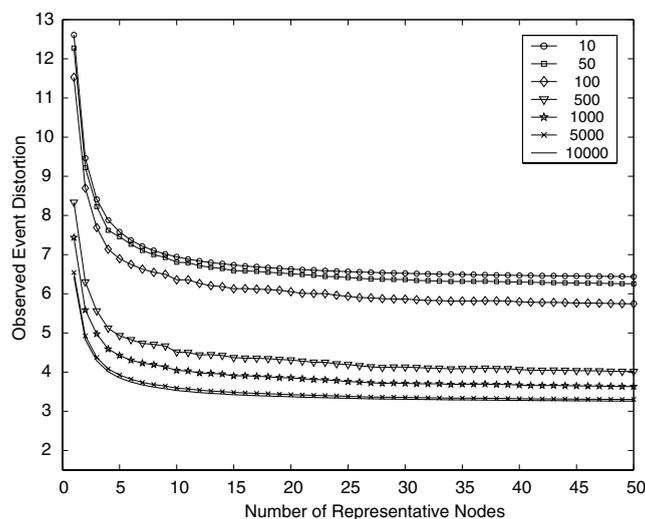


Fig. 2. Observed event distortion for different  $\theta_1$  values according to changing number of representative nodes.

1. The correlation coefficient,  $\rho_{(s,i)}$ , between a node  $n_i$  sending 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 node is chosen apart from the source, it observes relatively inaccurate data resulting in higher distortion at the sink.
2. The correlation coefficient,  $\rho_{(i,j)}$ , between each representative node  $n_i$  and  $n_j$  affects the distortion positively. As the distance between nodes increases, distortion decreases. Since further apart 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 saving can be achieved by choosing representative nodes among the nodes in the event area without degrading the achieved distortion at the sink. It is clear that reduced number of nodes transmitting information decreases contention in the wireless medium resulting in decreased energy consumption. Energy consumed from both transmission of packets and collision penalties can be reduced drastically if the spatial correlation is exploited. As a result, it is important to find the minimum number of representative nodes that achieve the distortion constraint given by the sensor application. This minimum number can be given as

$$M^* = \underset{M}{\operatorname{argmin}} \{D(M) < D_{\max}\},$$

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 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

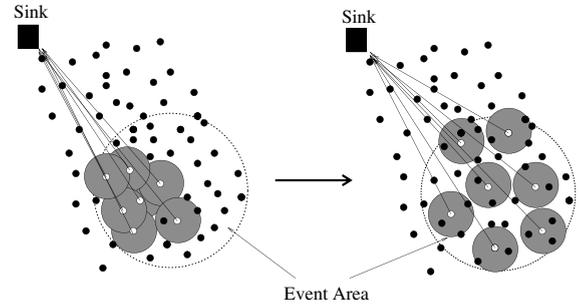


Fig. 3. Spatial re-usage in sensor networks.

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 performance 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 a recent work [21], the authors proposed a MAC protocol that exploits the spatial correlation between closely located sensor nodes that regulates medium access and prevents redundant transmissions from closely located sensors. Based on the spatial correlation among sensor nodes, the MAC protocol collaboratively regulates medium access so that redundant transmissions from correlation neighbors are suppressed. In addition, necessary mechanisms for the efficient transmission of the information from the sensor nodes to the sink has been proposed. The experimental results in [21] reveal that significant performance gains are obtained from exploiting spatial correlation in the MAC layer.

### 3.2. Correlation-based reliable event transport

In order to realize the potential gains of the WSN, it is imperative that desired event features are reliably communicated to the sink. To accomplish this, a reliable transport mechanism is required in addition to an efficient media access scheme as discussed in Section 3.1. The main

objective of the transport layer mechanism in WSN is to achieve reliable collective transport of event features from the sensors within the coverage of the phenomenon, i.e., event area, to the sink. In order to provide reliable event detection at the sink, possible congestion in the forward path should also be addressed by the transport layer. Once the event is sensed by a number of sensor nodes within the event area, significant amount of traffic is triggered by these sensor nodes, which may easily lead to congestion in the forward path. Furthermore, the error and congestion control objectives must be achieved with minimum possible energy expenditure. Energy efficiency must be also considered in transport mechanism design by shifting the burden to the high-powered sink in the WSN in order to conserve limited sensor resources.

Unlike traditional communication networks, the sensor network paradigm necessitates that the event features are estimated within a certain distortion bound, i.e., required reliability level, at the sink as discussed in Section 2. Reliable event detection at the sink is based on collective information provided by source nodes and not on any individual report. Hence, conventional end-to-end reliability definitions and solutions are inapplicable in the WSN regime and would only lead to over-utilization of scarce sensor resources. On the other hand, the absence of reliable transport altogether can seriously impair event detection which is the main objective of WSN deployment. Hence, the WSN paradigm necessitates a collective *event-to-sink* reliability notion rather than the traditional end-to-end notion [17]. Such *event-to-sink reliable transport* notion based on collective identification of data flows from the event to the sink is illustrated in Fig. 4 and depends on following definitions:

**Definition 1.** The *observed event distortion*  $D_i$  is the distortion achieved, i.e., as in (23), when the sink performs estimation of the signal  $S$  being tracked in decision interval  $i$ .

**Definition 2.** The *desired event distortion*  $D^*$  is the maximum distortion allowed to assure reliable event detection in the estimation performed by the

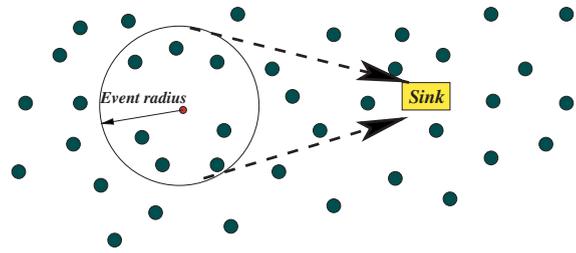


Fig. 4. Typical sensor network topology with event and sink. The sink is only interested in collective information of sensor nodes within the event radius and not in their individual data.

sink. This upper bound for the distortion level is determined by the application and based on the physical characteristics of the signal  $S$  being tracked.

Based on the packets generated by the sensor nodes in the event area, the sink estimates the event features to determine the necessary action and observes  $D_i$  at each decision interval  $i$ . Note that a distortion level  $D$  for the estimation of event features from the sensor observations corresponds to the reliability level of the event-to-sink communication in the WSN. If observed event distortion is less than the distortion bound, i.e.,  $D_i < D^*$ , then the event is deemed to be reliably detected. Else, appropriate action needs to be taken to assure the desired reliability level in the event-to-sink communication.

The main rationale behind such *event-to-sink reliability* notion is that the data generated by the sensors are temporally correlated which tolerates individual packets to be lost to the extent where the desired event distortion  $D^*$  is not exceeded. Let  $f$  be the reporting frequency of a sensor node defined as the number of samples taken and hence packets sent out per unit time by that node for a sensed phenomenon. This reporting frequency can be attributed to increase in sampling rate, increase in number of quantization levels, number of sensing modalities etc. Hence, the reporting frequency  $f$  controls the amount of traffic injected to the sensor field while regulating the number of temporally-correlated samples taken from the phenomenon. This, in turn, affects the observed event distortion, i.e., event detection reliability. Thus, the reliable event transport problem in WSN

is to determine the reporting rate ( $f$ ) of source nodes so that the maximum event estimation distortion bound  $D^*$  is not exceeded, i.e., required event detection reliability is achieved at the sink, with minimum resource utilization.

The determination of an appropriate reporting frequency  $f$  in order to assure the desired event distortion with minimum energy expenditure and without causing congestion is a challenging issue. As derived in Section 2.3, the distortion  $D_i$  observed in the estimation of the signal  $S$  being tracked depends on the reporting frequency  $f$  used by the sensor nodes sending their readings to the sink in the decision interval  $i$ . A case study with the same network configuration and parameters in Section 3.1 is also performed to observe the variation of the observed event distortion at the sink for varying reporting frequency  $f$ , i.e., distortion function  $D(f)$  in (23). It is observed from (23) and Fig. 5 that the observed event distortion at the sink decreases with increasing  $f$ . This is because the number of samples received in a decision interval  $i$  increases with increasing  $f$  conveying more information to the sink from the event area. Note that after a certain reporting frequency  $f$ , the observed event distortion cannot be further reduced. Therefore, a significant energy saving can be achieved by selecting small enough  $f$  which

achieves desired event distortion  $D^*$  and does not lead to an overutilization of the scarce sensor resources.

On the other hand, any  $f$  chosen arbitrarily small to achieve a certain distortion bound  $D^*$  using (23) may not necessarily achieve the desired distortion level and hence assure the event transport reliability. This is mainly because all of the sensor samples generated with this chosen reporting frequency may not be received because of packet losses in the sensor network due to link errors and network disconnectivity. Similarly, as very high values of  $f$  do not bring any additional gain in terms of observed event distortion as shown in Fig. 5; on the contrary, it may endanger the event transport reliability by leading to congestion in the sensor network. Let  $f_{\max}$  be the maximum reporting frequency which the network capacity can accommodate. Thus,  $f > f_{\max}$  leads to congestion and hence packet losses resulting in an increase in the observed event distortion.

This has been also observed in the preliminary simulation experiments in [17]. In these experiments, the normalized event transport reliability, i.e.,  $\eta = M_i/M^*$  where  $M_i$  and  $M^*$  are respectively the received and desired number of sensor samples in a decision interval, is observed to first increase with reporting frequency until  $f = f_{\max}$  is reached.

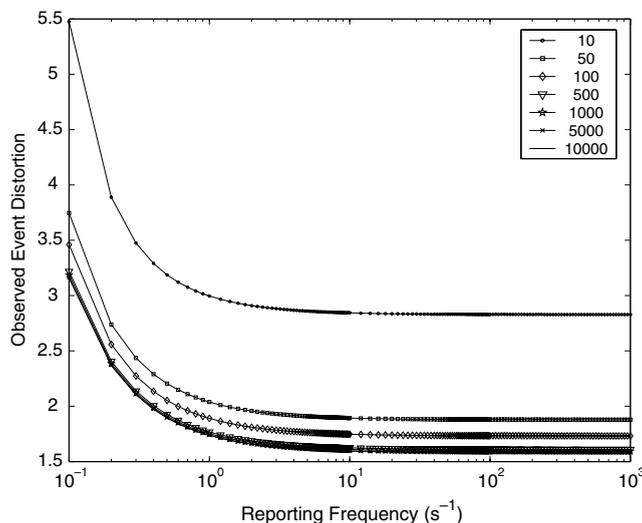


Fig. 5. Observed event distortion for varying normalized reporting frequency.

After this point, increasing reporting frequency is observed to impair the number of samples  $M_i$  received at the sink. This is because excessive packet transmissions result in network congestion and hence packets are discarded at the routing sensor nodes. Moreover, in the worst case, the event transport reliability is not achieved at all because of the high distortion in the estimation of the tracked signal  $S$  due to reception of inadequate number of sensor observations at the sink.

According to the observed event reliability and the current congestion condition in the network, the network operation regions are determined as shown in Table 1. Here, *OOR* is defined as the optimal operating point around  $\eta = 1$  within a tolerance zone of width  $2\epsilon$ . Note that low and high reliability levels correspond to high and low observed event distortion at the estimation of the signal being tracked at the sink, respectively. Hence, the objective is to operate the network at OOR, i.e., achieve required event reliability level (desired distortion level in the estimation) with minimum energy-expenditure. To achieve this objective and address the reliable event transport problem, an event-to-sink reliable transport (ESRT) protocol is also proposed in [17] based on the *event-to-sink reliability* notion for WSN. The objective of this scheme is to achieve reliable event transport with minimum energy expenditure and congestion control by exploiting the correlation and the collaborative nature of the WSN. To help accomplish this, the protocol uses a congestion control mechanism that serves the dual purpose of reliable detection and energy conservation. For example, in the states with the observed reliability is greater than that required, i.e., very low observed event distortion, and there is no congestion, the protocol conservatively reduces the reporting

frequency  $f$  to conserve energy, while not compromising on the event estimation distortion. On the other hand, the protocol pursues more aggressive update policies in the states with congestion and low event reliability, i.e., high observed event distortion [17]. Simulation experiments and analytical study in [17] show that ESRT protocol indeed achieves event-to-sink reliability with minimum energy consumption with the help of the correlation and collaborative nature of the WSN. As a result, temporal correlation conveyed in the physical characteristics of the phenomenon is exploited in addressing reliable event transport problem in WSN.

#### 4. Conclusions

In addition to the collaborative nature of the WSN, the existence spatial and temporal correlations among the sensor observations are significant and unique characteristics of the WSN. The correlation in WSN can be considered in developing new energy-efficient networking protocols specifically tailored for WSN paradigm. These protocols utilizing the correlation to conserve energy resources may drastically enhance the overall network performance. In this paper, we introduced a theoretical framework to capture the spatial and temporal correlations in wireless sensor networks. Our theoretical framework constitutes a basis for the development of such energy-efficient communication protocols for WSN. Moreover, based on our framework, we discussed possible efficient medium access and reliable event transport approaches taking advantage of the spatial and temporal correlations in WSN, respectively. We showed via mathematical analysis, their results,

Table 1  
Network operation regions based on congestion and reliability levels

State	Decision boundaries	Definition
(NC, LR)	$f < f_{\max}$ and $\eta > 1 + \epsilon$	(No congestion, Low reliability)
(NC, HR)	$f \leq f_{\max}$ and $\eta < 1 - \epsilon$	(No congestion, High reliability)
(C, HR)	$f > f_{\max}$ and $\eta < 1$	(Congestion, High reliability)
(C, LR)	$f > f_{\max}$ and $\eta \geq 1$	(Congestion, Low reliability)
OOR	$f < f_{\max}$ and $1 - \epsilon \leq \eta \leq 1 + \epsilon$	(Optimal operating region)

case studies and discussions that correlation in WSN can be exploited to significantly improve the energy-efficiency in WSN.

## References

- [1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, Wireless sensor networks: a survey, *Comput. Networks* 38 (4) (2002) 393–422.
- [2] K.A. Arisha, M.A. Youssef, M.Y. Younis, Energy-aware management in cluster-based sensor networks, *Comput. Networks* 43 (5) (2003) 649–668.
- [3] J.O. Berger, V. de Oliveira, B. Sanso, Objective bayesian analysis of spatially correlated data, *J. Am. Statist. Assoc.* 96 (2001) 1361–1374.
- [4] T. Clouqueur, V. Phipatanasuphorn, P. Ramanathan, K. Saluja, Sensor deployment strategy for target detection, in: *Proceedings of the ACM WSN 2002*, Atlanta, USA, September 2002.
- [5] T. van Dam, K. Langendoen, An adaptive energy-efficient MAC protocol for wireless sensor networks, in: *Proceedings of the ACM SenSys 2003*, Los Angeles, CA, November 2003.
- [6] M. Gastpar, M. Vetterli, Source-channel communication in sensor networks, in: *Proceedings of the 2nd International Workshop on Information Processing in Sensor Networks (IPSN'03)*, 2003.
- [7] T.J. Goblick, Theoretical limitations on the transmission of data from analog sources, *IEEE Trans. Inform. Theory* IT-11 (4) (1965) 558–567.
- [8] S. Haykin, *Communication Systems*, third ed., Wiley, New York, 1994.
- [9] IEEE 802.15.4/D17, Draft Standard Low-Rate Personal Area Networks, October, 2002.
- [10] J. Kusuma, L. Doherty, K. Ramchandran, Distributed compression for sensor networks, in: *Proceedings of the IEEE Image Processing 2001*, vol. 1, October 2001, pp. 82–85.
- [11] S. Meguerdichian, F. Koushanfar, M. Potkonjak, M.B. Srivastava, Coverage problems in wireless ad-hoc sensor networks, in: *Proceedings of the IEEE INFOCOM 2001*, Anchorage, AK, April 2001.
- [12] V. Poor, *An Introduction to Signal Detection and Estimation*, second ed., Springer, Berlin, 1994.
- [13] S.S. Pradhan, K. Ramchandran, Distributed source coding: symmetric rates and applications to sensor networks, in: *Proceedings of the Data Compression Conference 2000*, 2000, pp. 363–372.
- [14] S.S. Pradhan, J. Kusuma, K. Ramchandran, Distributed compression in a dense microsensor network, *IEEE Signal Process. Mag.* 19 (2) (2002) 51–60.
- [15] S.S. Pradhan, R. Puri, K. Ramchandran, n-Channel symmetric multiple descriptions—Part I:  $(n, k)$  source-channel erasure codes, *IEEE Trans. Inform. Theory* 50 (1) (2004) 47–61.
- [16] V. Rajendran, K. Obraczka, J.J. Garcia-Luna-Aceves, Energy-efficient, collision-free medium access control for wireless sensor networks, in: *Proceedings of the ACM SenSys 2003*, Los Angeles, CA, November 2003.
- [17] Y. Sankarasubramaniam, O.B. Akan, I.F. Akyildiz, ESRT: event-to-sink reliable transport for wireless sensor networks, in: *Proceedings of the ACM MOBIHOC 2003*, June 2003, pp. 177–188.
- [18] A. Scaglione, S.D. Servetto, On the interdependence of routing and data compression in multi-hop sensor networks, in: *Proceedings of the ACM MOBICOM 2002*, Atlanta, GA, USA, September 2002.
- [19] F. Stann, J. Heidemann, RMST: Reliable data transport in sensor networks, in: *Proceedings of the IEEE SNPA 2003*, Anchorage, AK, May 2003, pp. 102–112.
- [20] G.L. Stuber, *Principles of Mobile Communication*, Kluwer Academic Publisher, Dordrecht, 2001.
- [21] M.C. Vuran, I.F. Akyildiz, Spatial correlation-based collaborative medium access control in wireless sensor networks, submitted, December 2003.
- [22] C.-Y. Wan, S.B. Eisenman, A.T. Campbell, CODA: congestion detection and avoidance in sensor networks, in: *Proceedings of the ACM SenSys 2003*, November 2003.
- [23] C.Y. Wan, A.T. Campbell, L. Krishnamurthy, PSFQ: a reliable transport protocol for wireless sensor networks, in: *Proceedings of the WSN 2002*, Atlanta, GA, USA, September 2002.
- [24] W. Ye, J. Heidemann, D. Estrin, An energy-efficient MAC protocol for wireless sensor networks, in: *Proceedings of the INFOCOM 2002* vol. 3, 2002, pp. 1567–1576.



**Mehmet C. Vuran** received his B.Sc. degree in Electrical and Electronics Engineering from Bilkent University, Ankara, Turkey, in 2002. He is currently a Research Assistant in the Broadband and Wireless Networking Laboratory and pursuing his Ph.D. degree at the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA. His current research interests include adaptive medium access protocols for heterogeneous wireless architectures, next generation wireless networks, and wireless sensor networks.



**Özgür B. Akan** received his B.Sc. and M.Sc. degrees in Electrical and Electronics Engineering from Bilkent University and Middle East Technical University, Ankara, Turkey, in 1999 and 2001, respectively. He is currently a Research Assistant in the Broadband and Wireless Networking Laboratory and pursuing his Ph.D. degree at the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA. His current research interests include adaptive transport protocols for heterogeneous wireless architectures, next generation wireless networks, wireless sensor networks, and deep space communication networks.



**Ian F. Akyildiz** 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 Broadband and Wireless Networking Laboratory. He is an Editor-in-Chief of *Computer Networks* and for the newly launched journal *Ad Hoc Net-*

*works* and an Editor for *ACM Journal of Wireless Networks*. His current research interests are in sensor networks, IPN Internet, wireless networks, and satellite networks.

He received the “Don Federico Santa Maria Medal” for his services to the Universidad of Federico Santa Maria, 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.