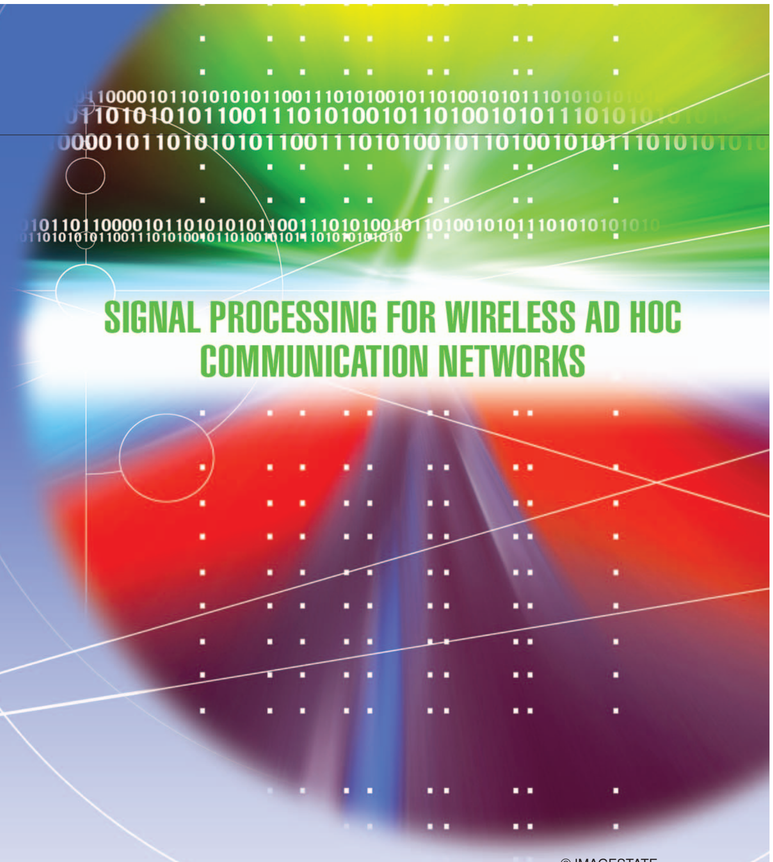
Youngchul Sung, Saswat Misra, Lang Tong, and Anthony Ephremides



Signal Processing for Application-Specific Ad Hoc Networks

The role of signal processing in protocol design

The layered architecture used in large data communication networks provides modularity at the expense of efficiency. This paradigm has been successful in wired and many wireless networks. Modularity simplifies the design of large heterogenous and complex networks via a "divide-and-conquer" approach, decoupling an enormously difficult problem into tasks that can be pursued independently. Because of this, new applications can be developed independently of the medium access and routing strategies specified at lower layers, and new physical layer techniques can be implemented without changing upper layer implementations. Perhaps the most remarkable feature of the layered architecture is that it makes the network design scalable. A prime example is the Internet, which has grown from a hand full of nodes in the ARPAnet [3] to hundreds of millions of nodes today.

But there is a different kind of network, one that is designed for specific applications. In contrast to the design of the Internet, serving individual nodes is not always the ultimate objective. Consider, for example, a sensor network deployed for target detection and tracking, environmental monitoring, or the detection of a specific chemical compound. In these applications, network performance should not be measured by general purpose metrics such as the data rate at the link level or by the throughput over the network. Conventional performance metrics such as throughput and delay do not necessarily translate to a performance measure suitable for signal detection and estimation applications. For application-specific networks, and sensor networks in particular, performance should be measured instead by application-defined metrics such as the miss detection and false alarm rates, the network lifetime for performing these tasks, and the energy efficiency of target detection, tracking, and estimation.

If an application-specific metric is to be optimized, signal processing may have a role in defining network architectures

1053-5888/06/$20.00©2006IEEE

and protocols. A signal processing perspective may change the way protocols are designed by shifting design from the user-centric view (as in the Internet) to an application- or data-centric viewpoint. Illustrated in Figure 1 is a schematic which shows that signal processing can be part of the protocol design. For example, to design a medium access control (MAC) scheme to collect data from distributed sensors that have correlated measurements, we may not want to allocate resources such as bandwidth and power among sensor nodes, as the classical layered architecture dictates. Instead, we may choose to allocate resources among the possible data types, minimizing the interference among transmissions of different measured phenomena, rather than among different nodes. Type-based multiple access (TBMA) [18] is an example of such a data-centric approach.

In this article, we examine the role of signal processing in protocol design for application-specific ad hoc networks. Our goal is to illustrate, using routing as an example, that a detection theoretic approach may suggest a different link metric, and that optimization of this link metric can provide superior application performance (for a fixed energy consumed) when compared to traditional approaches. It is our hope to demonstrate that time-honored signal processing concepts such as the innovations representation of the log-likelihood function have the potential to find their way into the design of application-specific network protocols.

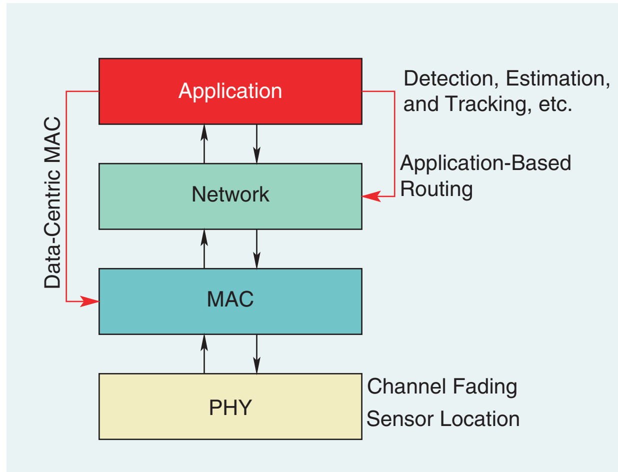
APPLICATION SPECIFIC ROUTING: AN ILLUSTRATIVE EXAMPLE

As an illustrative example, we show how to design application-specific routing to maximize the detection performance of an ad hoc network. Shown in Figure 2 is a large network with geographically distributed sensors, each taking measurements of a certain phenomenon. We assume that there is a fusion center (or gateway node) that is responsible for collecting data from sensors and drawing inferences from that data. We are interested in the detection of a spatially correlated random signal field. We assume that the field is a Gaussian field for which the two hypothesis are

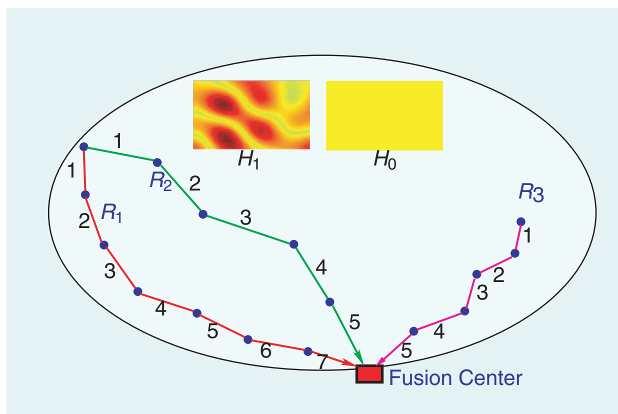
$$\begin{cases} \mathcal{H}_0 : & \text{IID Gaussian noise,} \\ \mathcal{H}_1 : & \text{correlated Gaussian random field} \\ & \quad \text{observed in IID Gaussian noise,} \end{cases} \tag{1}$$

where IID refers to independent and identically distributed. We assume that the sensors have already been placed in an arbitrary configuration (consistent with the notion of an ad hoc network) and that the correlation structure of the Gaussian signal field is known. Later we will impose a specific structure on the correlation to facilitate aspects of the analysis.

Sensors have limited transmission range, and they have to deliver their data to the fusion center cooperatively over certain routes. Obviously, the more data that is collected, the more accurate the inferences drawn about the phenomenon at the fusion center. What makes the problem interesting and practically relevant is that wireless sensors are subject to severe power and energy constraints. Each transmission will cost a certain amount of energy that will depend on the distance between the transmitter and the receiver and also on the amount of data that needs to be delivered. It is this tradeoff between performance and energy consumption that demands a different kind of design methodology.



[FIG1] Application-based cross-layer design.



[FIG2] Routing for a detection application.

There are several practical scenarios that present interesting design challenges. The first is related to the *event-driven* applications in which some sensor is alarmed by its observation and initiates a data collection by selecting the best possible route to the fusion center. The sensor may or may not know the location of the the fusion center, or, perhaps, there are several fusion centers to choose from. The other is related to the *clock-driven* applications in which the fusion center issues regularly scheduled (possibly random) data collections. The fusion center may or may not know the specific sensor locations.

We are interested in determining the *route* over which data collection should be performed. The problem would not be interesting if, under hypothesis $\mathcal{H}_1$, the sensor measurements of the signal field were independent, i.e., the conditionally IID model. In this case, each sensor provides equally valuable information conditioned on the observations of previously sampled sensors, and the "optimal" routing strategy would be to simply collect data from sensors closest to the fusion center to save transmission energy. However, routing becomes a nontrivial issue when sensor measurements are not IID under $\mathcal{H}_1$. In this case, sensors further away from previously sampled sensors may have more valuable (i.e., less correlated) observations of the signal field. However, this data can only be collected at the expense of using more transmission energy. Hence, a tradeoff emerges between detection performance and energy consumption. Furthermore, the measurements of sensor nodes are imperfect, and it is necessary to consider this measurement inaccuracy and aggregate the measurements of multiple sensors for the final decision.

[FOR APPLICATION-SPECIFIC NETWORKS, PERFORMANCE SHOULD BE MEASURED INSTEAD BY APPLICATION-DEFINED METRICS SUCH AS THE MISS DETECTION AND FALSE ALARM RATES, THE NETWORK LIFETIME FOR PERFORMING THESE TASKS, AND THE ENERGY EFFICIENCY OF TARGET DETECTION, TRACKING, AND ESTIMATION.]

Consider a clock-driven application in which the fusion center initiates data collection as depicted in in Figure 2. For any fixed route, we will assume that data is collected along *all* nodes along the route. (In the sequel we will propose a cooperative transmission scheme for routing data from a given source node to the fusion center. It will be seen that the transmission energy consumed in collecting data from all nodes en route is identical to that needed to simply relay a message from the source node to the fusion center. Therefore, for a fixed route, we assume that data is collected from all nodes along the route.) We highlight three potential routes $\mathcal{R}_1$, $\mathcal{R}_2$, and $\mathcal{R}_3$, and ask which route is preferred for detection. There are eight potential observations along $\mathcal{R}_1$ and six along $\mathcal{R}_2$. But measurements along $\mathcal{R}_1$ are more correlated than those along $\mathcal{R}_2$ because nodes are closer to each other. Thus the "information" content through $\mathcal{R}_1$ may not be as great as that through $\mathcal{R}_2$. Now $\mathcal{R}_3$ has the same number of nodes as $\mathcal{R}_2$, but the route length is shorter. The energy consumed in the collection through $\mathcal{R}_3$, conceivably, is lower than that through $\mathcal{R}_2$. But the limited coverage of $\mathcal{R}_3$ may result in a significant loss of performance. In an event-driven application, the argument above remains valid, but we should consider only those routes which originate from a common sensor, e.g., $\mathcal{R}_1$ and $\mathcal{R}_2$ above.

Intuitive concepts alone, e.g., "closely spaced nodes provide less information" and "collections over widely separated distances require more energy," will not carry us far in determining the optimal routes for the tradeoffs described above. To develop optimized strategies for application-specific routing, we will need an analytical characterization of performance, even if it is nonuniversal. It is only through such a characterization that we can hope to achieve the right tradeoff between detection performance and energy consumption.

CHERNOFF ROUTING, SCHWEPPE'S RECURSION, AND KALMAN AGGREGATION

Unfortunately, there is no general analytical form that describes detection performance along a given route. We can, of course, perform simulations. But simulations will not give us a way to develop a routing protocol that can be implemented in a distributed fashion at individual nodes. For this, we will need a link metric that accumulates the contributions of each link to detection performance in such a way that the total accumulation of a route is related to the overall detection performance of that route. If we can obtain such a measure, routing will be greatly simplified using the shortest path methodology [3]. To begin such an analysis, we will use performance bounds that are functions of network parameters such as distances between a pair of nodes and signal parameters such as signal-to-noise ratio (SNR) and correlation strength. The bound that we use to derive optimized routing is tight as the number of nodes in the route increases.

CHERNOFF ROUTING

To find a suitable bound, we digress briefly into the theory of large deviations [6]. In doing so, we will describe the tools that are needed to form a theoretical basis for Chernoff routing.

The Chernoff bound [25] is a well-known tool that can be used to provide an upper bound on the probability of detection error. Consider the simple binary hypotheses $\mathcal{H}_i : Y_k \sim p_i(y), i \in \{0, 1\}$, for $k = 1, 2, \ldots, n$. Define $Y \triangleq [Y_1, \ldots Y_n]$, and let $P_i(Y)$ denote the probability distribution of Y under $\mathcal{H}_i$, $i \in \{0, 1\}$. The optimal detector for this test (under either the Bayesian or Neyman-Pearson formulation) compares the log-likelihood ratio

$$l(Y) \triangleq \log \frac{P_1(Y)}{P_0(Y)} \underset{<_{\mathcal{H}_0}}{\geq^{\mathcal{H}_1}} \tau$$

to some choice of threshold τ. The false alarm probability can be upper bounded by

$$\Pr(l(Y) > \tau | \mathcal{H}_0) < \exp\{-\Lambda(\tau)\},$$

where the so-called error exponent $\Lambda(\tau)$ is the Fenchel-Legendre transform of the cumulant generating function $\mu(s) \triangleq \log \mathbb{E}\{e^{sl(Y)} | \mathcal{H}_0)\}$:

$$\Lambda(\tau) \triangleq \sup_{s>0} \{s\tau - \mu(s)\}.$$

There are two valuable features of the Chernoff bound when we have IID measurements. First, the error exponent $\Lambda(\tau)$ is additive. Specifically, when we have $Y_k \sim p_i(y), k = 1, \ldots, n$, the log-likelihood ratio is additive

$$l(Y_1, \ldots, Y_n) = \sum_k l(Y_k),$$

and the corresponding Chernoff bound on the false-alarm probability has the form

$$\Pr(l(Y) > \tau | \mathcal{H}_0) < \exp\{-n\Lambda_1(\tau)\},$$

where $\Lambda_1(\tau)$ is the Fenchel-Legendre transform of $\log \mathbb{E}\{e^{s\, l(Y_1)} | \mathcal{H}_0)\}$. It is this additivity that makes it possible to obtain an additive link metric. Second, the Chernoff bound is tight when n is large. In addition to this upper bound, we have a similar lower bound on the false alarm probability that states

$$\Pr(l(Y_1, \ldots, Y_n) > \tau | \mathcal{H}_0) = \exp\{-n\Lambda_1(\tau) + o(n)\},$$

where $o(n)$ is such that $\lim_{n\to\infty} o(n)/n = 0$. Here, we interpret $\Lambda_1(\tau)$ as the decay rate of the false alarm probability. In fact, $\Lambda_1(\tau)$ can be shown to be the largest possible decay rate

$$\Lambda_1(\tau) = \lim_{n\to\infty} \frac{1}{n} \log \Pr(l(Y_1, \ldots, Y_n) > \tau | \mathcal{H}_0).$$

Under the Bayesian setup, the two types of detection error probabilities, false alarm and miss detection, are balanced by the priors of the two hypotheses. However, the largest decay rate for the average error probability, $P_e = \Pr(\mathcal{H}_0) \Pr(\text{Error} \,|\, \mathcal{H}_0) + \Pr(\mathcal{H}_1) \Pr(\text{Error} \,|\, \mathcal{H}_1)$, does not depend on prior probabilities, and is given by the Chernoff information defined as

$$C \triangleq \Lambda_1(0) = \sup_{s>0}\{-\mu(s)\}\,. \tag{2}$$

This concludes our digression.

By Chernoff routing we mean routing where the Chernoff information is used as a route metric. For a fixed route, say $\mathcal{R}_1$ in Figure 2, we have a set of measurements $\{y_i\}$. Note, however, that y_is are not IID, the Chernoff information for such a case is a function of the distribution of $l(y_1, \cdots, y_n)$, which, in turn, is a function of the route. Denoting the Chernoff information associated with a specific route $\mathcal{R}$ as $C(\mathcal{R})$, Chernoff routing aims to select a route that maximizes $C(\mathcal{R})$. Denoting $E(\mathcal{R})$ as the energy consumed when data are routed through route $\mathcal{R}$, we obtain an energy constrained form of Chernoff routing

$$\max_{\mathcal{R}} C(\mathcal{R}) \text{ subject to } E(\mathcal{R}) \le \varepsilon\,. \tag{3}$$

LINK METRIC VIA INNOVATIONS REPRESENTATION

Although (3) captures the essence of optimal routing subject to an energy constraint, it does not provide a practical scalable protocol that can be implemented in a distributed fashion. To be able to use well-established techniques such as shortest path routing (i.e., Bellman-Ford optimization), we need to have an additive *link* metric such that the accumulated value of the link costs on $\mathcal{R}$ is proportional to the value of $C(\mathcal{R})$. Unfortunately, the standard expression of the Chernoff information for the Gaussian hypotheses is given in terms of the eigenvalues of the covariance matrix of signal samples [25] and does not allow the decomposition of the overall performance into a sum of the incremental performance gains at each link.

The key to obtaining an additive link metric, as proposed in [29] and [30] by Sung et al., is the use of the innovations representation of the log-likelihood function [26]. To understand this crucial step, we note that the Chernoff information associated with y_i is not additive because the log-likelihood function under $\mathcal{H}_1$ is not additive. It is thus natural to seek independent innovations. In the context of signal processing, this can often be achieved using recursive techniques. The idea of using the innovations representation to obtain the likelihood function recursively was first proposed by Schweppe [26], and the Schweppe's recursion leads to the decomposition of Chernoff information into an additive link metric.

For a fixed route $\mathcal{R}$, assuming a Gaussian signal along the route, Sung et al. show in [29] and [30] that the Chernoff information $C(\mathcal{R})$ is approximately equal to the sum of the logarithm of the innovations variance $R_{e,i}$ (normalized by the measurement noise variance) at each link, i.e.,

$$C(\mathcal{R}) \approx \sum_i C_i, \quad C_i = \frac{1}{2} \log \frac{R_{e,i}}{\sigma_w^2}\,, \tag{4}$$

at high SNR, where SNR is defined as the observational SNR at each sensor, σ_w^2 is the variance of measurement noise at each sensor, $\hat{y}_{i|i-1} \triangleq \mathbb{E}\{y_i | y_0, \ldots, y_{i-1}\}$ is the minimum mean square error (MMSE) estimate of y_i given all upstream measurements, and $R_{e,i} \triangleq \mathbb{E}\{|y_i - \hat{y}_{i|i-1}|^2\}$ is the MMSE of the estimation process. When the random process is Markovian, the link metric is almost memoryless. This crucial property makes it possible to apply Bellman-Ford and similar routing algorithms.

A comment on the form of the link metric C_i is in order. A simple derivation shows that

$$C_i = \frac{1}{2} \log \left(1 + \frac{P_{i|i-1}}{\sigma_w^2}\right)\,,$$

where $P_{i|i-1}$ is the variance of signal innovation at node i with respect to all its upstream nodes. This provides an intuitively satisfying interpretation: an appropriate link metric is the mutual information between a node and its neighbor. Hence, the optimal route is the one that has the maximal accumulated innovations entropy.

Next, we need to connect the innovation variance $P_{i|i-1}$ to physical parameters such as the distance and signal field correlation between two nodes and the SNR of each sensor observation. For the Gauss-Markov random field, this connection is easily obtained as

$$C_i \approx \frac{1}{2}\log\left\{\mathrm{SNR} + 1 - (\mathrm{SNR} - 1)e^{-2A\Delta_i}\right\}, \quad (5)$$

where $\Delta_i > 0$ is the link length, and $A > 0$ describes the correlation strength and is the diffusion constant of the first order stochastic differential equation of the Gauss-Markov model. Note that as $A \to \infty$ the sensor observations approach statistical independence and that $A \to 0$ corresponds to the fully correlated case.

A numerical evaluation of C_i as a function of link length Δ_i provides useful insights. Figure 3(a) shows the link metric as a function of link length Δ_i. For SNR ≥ 1, the metric is strictly increasing, strictly concave, bounded from above, and achieves a maximum value of $(1/2)\log(1 + \mathrm{SNR})$. Thus, this value represents the maximum information that a link can provide, and it is attained if the two sensors at each end of the link have independent observations of the signal field (such may be the case if the sensor are spaced far enough apart, or if the field is sufficiently weak in correlation).

We can now bring energy consumption into the framework. The energy used by a particular node en route can represented by the sum of the processing energy $E_p \geq 0$ and transmission energy to the next node, i.e.,

$$E_i = E_p + E_{t,0}\Delta_i^{\nu}, \qquad \nu \geq 2, \quad (6)$$

where $E_{t,0} \geq 0$ is a constant. Thus, the link *efficiency* can be defined as

$$\eta \triangleq \frac{C_i}{E_i}. \quad (7)$$

Now, the tradeoff between having large C_i and low E_i emerges. Figure 3(b) shows the detection efficiency for several values of processing energy E_p at each sensor. The transmission energy at each link increases without bound as the link length is increased. However, note that the link efficiency peaks before decreasing with increasing link length. Hence, we conclude that there is an optimal link length for optimal detection efficiency.

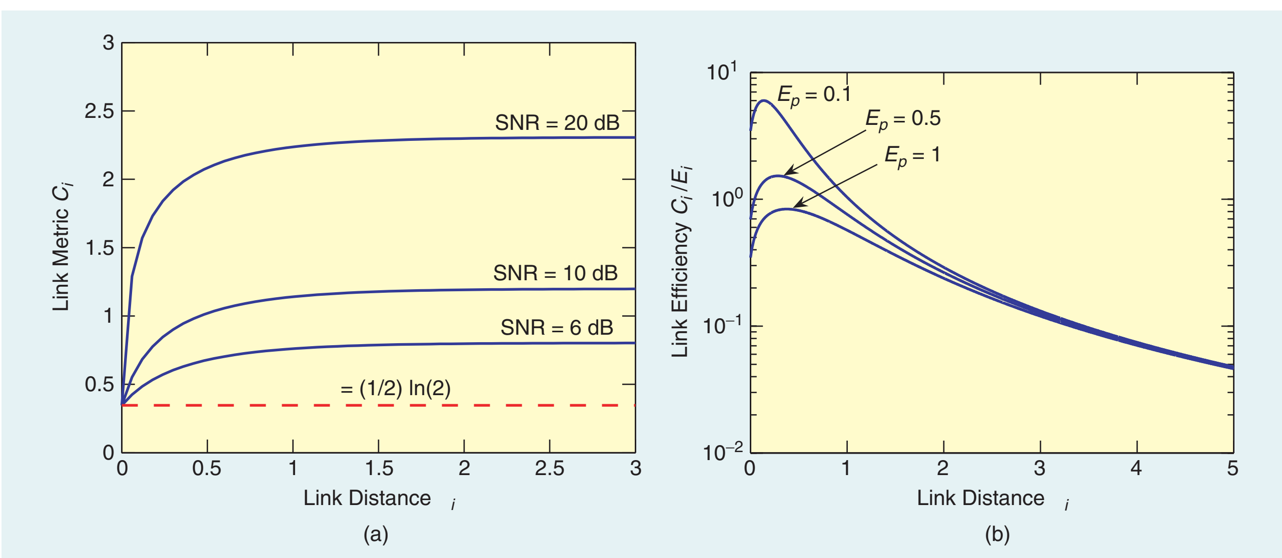
KALMAN AGGREGATION

We wish to perform cumulative data aggregation along each route, so that all sensor observations are delivered to the fusion center. Many such data aggregation methods are available based on, e.g., taking a minimum number of observations, a maximum number of observations, and the most recent observations. However, for the detection of a *Gaussian* random field, the optimal data aggregation can be implemented easily using a recursion based on the *innovations representation of the log-likelihood function*, as discussed previously. Suppose that we have observations $\{y_1, y_2, \ldots, y_n\}$ at n sensor nodes along a fusion route. Consider the observations under $\mathcal{H}_1$. (The case when $\mathcal{H}_0$ is given by a simplified version of the following derivation.) In this case, the log-likelihood can be rewritten in a recursive fashion:

$$\log p_1(y_1, y_2, \ldots, y_i) = \log p_1(y_1, y_2, \ldots, y_{i-1}) + \log p_1(y_i|y_1, \ldots, y_{i-1}). \quad (8)$$

The conditional distribution $p_1(y_i|y_1, \ldots, y_{i-1})$ is Gaussian with mean $\mathbb{E}\{y_i|y_1, \ldots, y_{i-1}\}$ and variance $R_{e,i}$ (this recurrence of notation will be resolved shortly). Thus, the log-likelihood l_i up to the ith observation along the route can be expressed

$$l_i = l_{i-1} - \frac{1}{2}\log(2\pi R_{e,i}) - \frac{1}{2}\frac{e_i^2}{R_{e,i}}, \quad (9)$$



[FIG3] Detection-based link metric: (a) link metric C_i as a function of link length ($A = 1$) and (b) link efficiency as a function of length link ($A = 1$, SNR $= 10$ dB, $\nu = 2$, and $E_{t,0} = 1$).

where the *innovation* is given by $e_i \triangleq y_i - \hat{y}_{i|i-1}$, where $\hat{y}_{i|i-1} = \mathbb{E}\{y_i | y_1, \ldots, y_{i-1}\}$ is the MMSE prediction of y_i given the previous observations, as defined previously. Now, note that $R_{e,i} = \mathbb{E}\{e_i^2\}$ which is consistent with its previous definition.

The update terms in the recursion can be easily calculated using Kalman filter theory under the Makovian assumption [16]. Suppose that each sensor knows its own location. Then, the only other necessary location information for sensor N_i is that of the next sensor for the prediction step in the Kalman recursion. The information flow between nodes along the route is illustrated in Figure 4. Node N_i needs to receive the log-likelihood l_{i-1}, the prediction for its observation $\hat{y}_{i|i-1}$, and the error variance $P_{i|i-1}$ from node N_{i-1}, and the location information x_{i+1} from N_{i+1} so that it may predict the observation to be made at node N_{i+1}. Note that the aggregation via Kalman recursion eliminates the need to deliver all the observations and sensor locations to the fusion center. In case of non-Markovian signals, a similar calculation is possible but requires more computation [16].

BELLMAN-FORD IMPLEMENTATION

We now consider a routing protocol that incorporates the proposed link metric C_i. The optimal fusion algorithm can be easily derived using the Kalman filter with the knowledge of the field correlation and sensor locations. Each sensor is then required to transmit the aggregated sufficient statistic (accumulated log-likelihood) to its neighboring sensor on the route [29]. When a certain sensor in the field observes the signal to be above its local threshold, it initiates the fusion process along the predetermined route to the fusion center. The route from each sensor to the fusion center is typically determined using a shortest path algorithm based on some link metric. Here, we simply use a suitably modified version of C_i as a new link metric for shortest path routing.

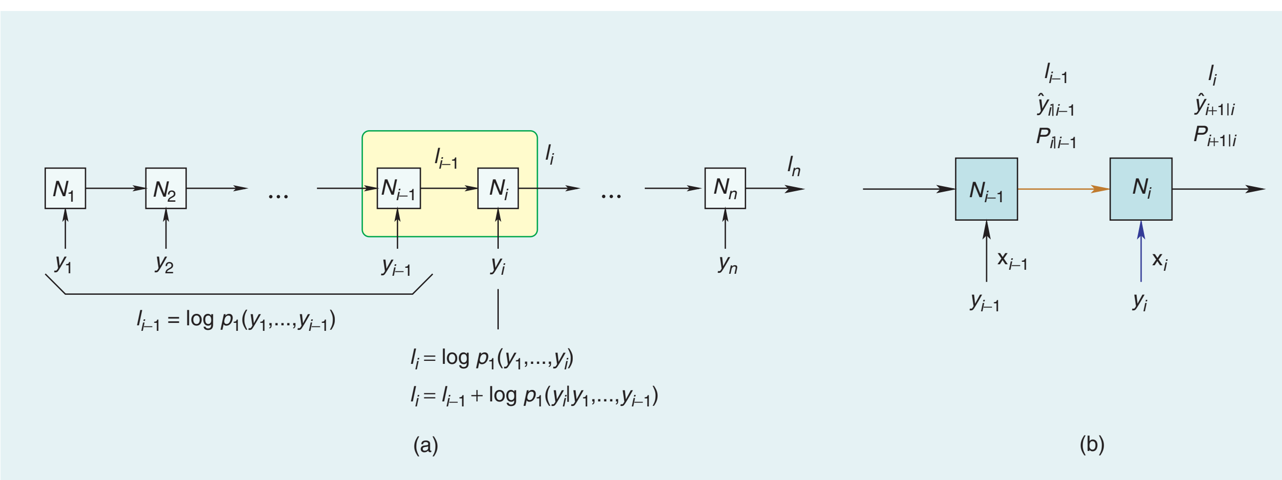
Typically in shortest path routing algorithms, each link in the network, say the one connecting node i to node j, is assigned a nonnegative link cost $\gamma_{i,j}$. The shortest path algorithm then finds an optimal route (path) from each source node to the destination (in our case, the fusion center), where an optimal route is defined such that the sum of all link costs along the optimal route is no larger than that of any other route from the same source to the destination. To apply the shortest path routing algorithm a suitable modification to the metric (5) is necessary. To see this, suppose that we use C_i in a shortest path routing algorithm without any additional criteria (i.e., actually, a function such as $1/C_i$ would be used since using C_i directly would *minimize* the accumulated Chernoff information, which the exact opposite effect we seek). Then, the algorithm would try to accumulate as much innovation entropy as possible. Since this quantity is a strictly increasing function of the number of links, the algorithm would only converge after visiting every node in the network, a paradigm which is unsuitable for very large networks. Furthermore, detection performance saturates after a certain number of observations even in moderately sized networks, and this makes visiting every node extremely detection inefficient. That is, after a certain number of links, each additional link will consume additional energy while providing little detection benefit.

We now propose a solution that addresses the energy constrained version of Chernoff routing as described in (3). Although the solution to (3) is desirable, it is not readily available in a closed form that is also amenable to implementation. As an alternative, we propose to balance the detection performance with the energy consumed through the following simple link metric

$$\gamma_{i,j} = \begin{cases} (E_i - \lambda C_i)^+_\epsilon & \text{if nodes } i \text{ and } j \text{ are connected,} \\ \infty & \text{otherwise,} \end{cases} \quad (10)$$

where

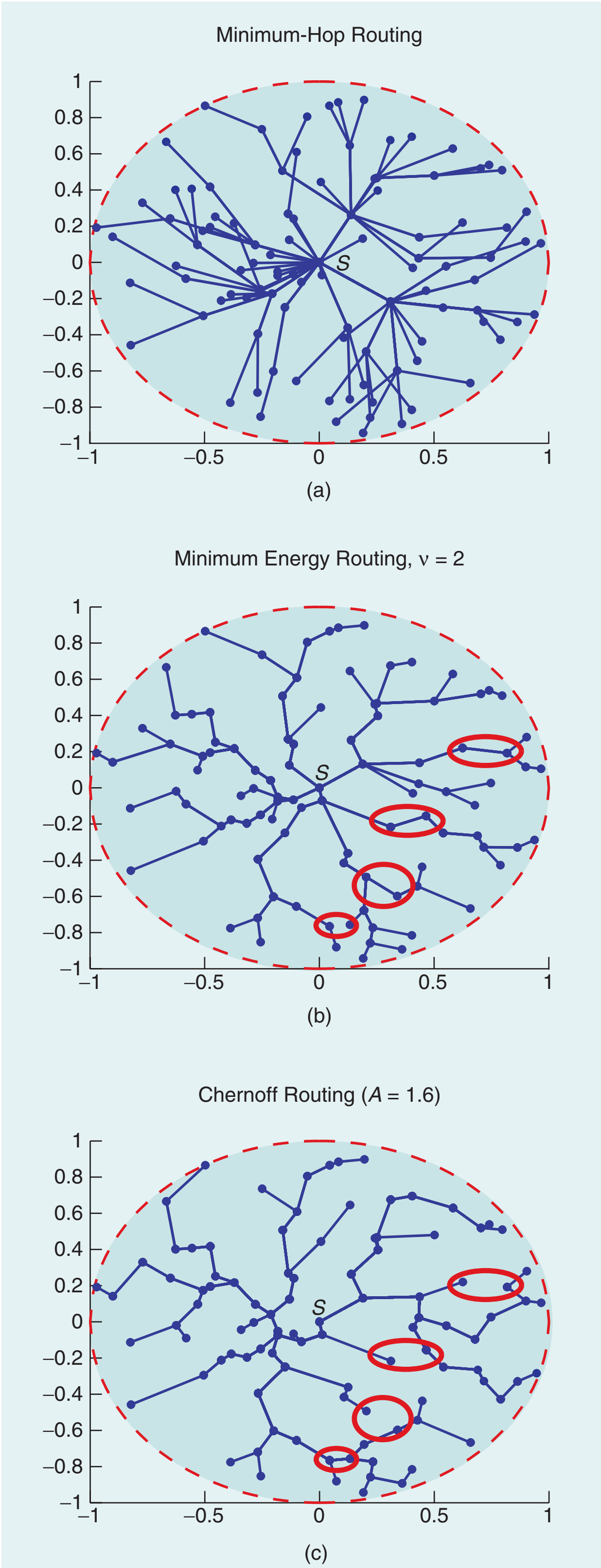
$$(x)^+_\epsilon \triangleq \begin{cases} x & x > 0, \\ \epsilon & x \leq 0, \end{cases}$$



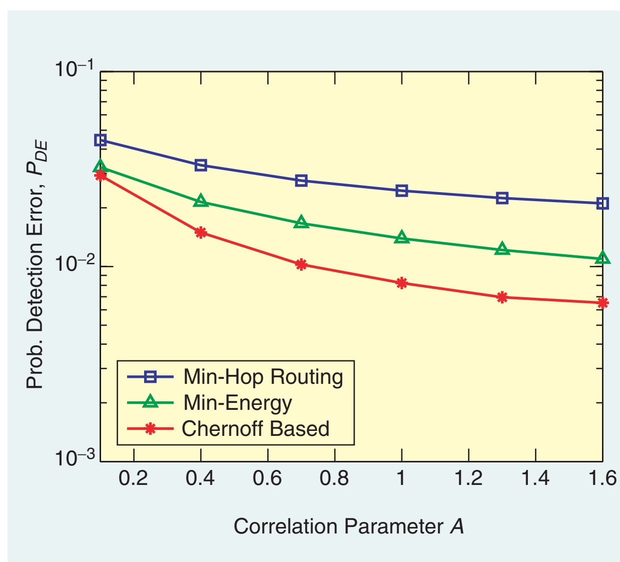
[FIG4] Data aggregation using Kalman recursion.

and $\epsilon > 0$ is a constant. Note that the cases $\lambda \to \infty$ and $\lambda \to 0$ correspond to the well-known minimum-hop and minimum-energy routing strategies, respectively.

We are now ready to provide some numerical insights by considering a sensor network with 100 sensors placed on a circular field with radius one. Figure 5 shows the shortest-path route from each node to the fusion center, which is located at the center of the field and denoted S, for each of the routing strategies: minimum-hop routing, minimum-energy routing, and Chernoff routing (i.e., shortest path routing based on the metric (10) with a nontrivial value of λ, i.e., $0 < \lambda < \infty$). (For simplicity, the Gauss-Markov model with diffusion constant A is used to describe the signal evolution along the route.) The differences in the route topology for these schemes is evident. While minimum-hop routing results in a few, large, well-directed hops to the fusion center, minimum energy and Chernoff routing take smaller and more scattered steps. This is because the transmission energy is a convex function of link length. The nodes that lead to major topological differences between the two latter strategies are circled in the figure. As expected, Chernoff routing produces routes that deviate from those of minimum-energy routing and for which the detection performance is presumably better. Figure 6 is a plot of the probability of detection error $P_{\rm DE}$ for the topology shown in the previous figure averaged over all N routes (i.e., $P_{\rm DE} = (1/N)\sum_{n=1}^{N} P_e(\mathcal{R}_n)$ where $\mathcal{R}_n$ is the optimal route from node n to the fusion center, determined separately for each routing scheme). The probability that the signal field is absent is $\Pr(\mathcal{H}_0) = 0.75$. Here, we used a more realistic signal correlation for the detection, i.e., the actual correlation between two sensors is a function of their Euclidean distance. That is, while the Chernoff routes are assigned assuming the



[FIG5] The shortest-path route from each node to the fusion center S for three different routing strategies: (a) minimum-hop routing, (b) Minimum-energy routing, and (c) Chernoff-routing ($A = 1.6$, $\lambda = 0.01$, SNR = 15 dB, $\nu = 2$, $N = 100$, $E_{t,0} = 1$, and $E_p = 0$).

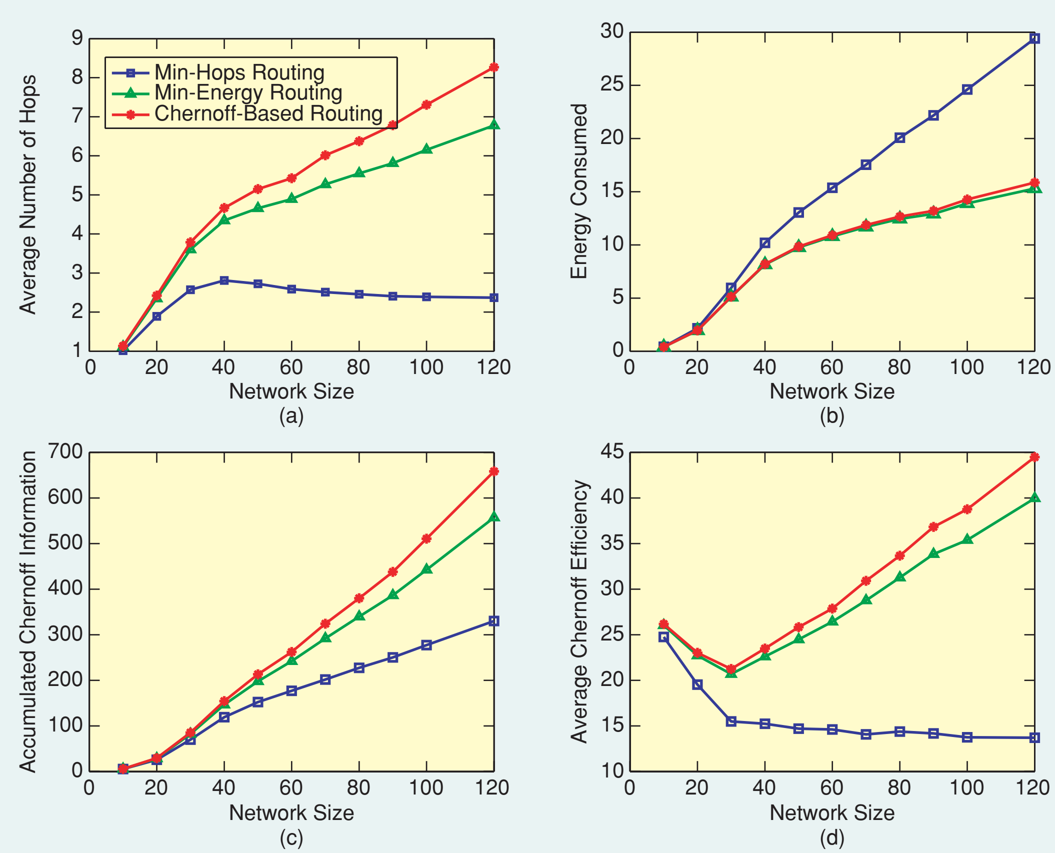


[FIG6] Probability of detection error $P_{\rm DE}$ averaged over all routes for the network topology of Figure 5 under three different routing strategies ($\Pr(\mathcal{H}_0) = 0.75$ and all other parameters are the same as in Figure 5).

Gauss-Markov model, the P_{DE} is determined assuming the more realistic model. Note that Chernoff-routing provides about a 40% reduction in the P_{DE} compared to minimum energy routing.

Figure 7 shows the average routing characteristics when SNR = 15 dB. Here, for each value of the network size, the performance is averaged over realizations of the network topology, (i.e., the sensor locations) to extract the fundamental network behavior. Figure 7(a) shows the average number of hops from all potential sources to the fusion center. As expected, the minimum-hop routing gives the smallest number of hops while Chernoff routing provides the largest. Figure 7(b) shows the average energy required by each scheme. It is seen that Chernoff routing requires almost the same as the minimum-energy routing, providing the largest accumulated Chernoff information as shown in Figure 7(c). In Figure 7(d) it is seen that Chernoff routing results in the maximum detection efficiency as expected. Finally, Figure 8 shows the average detection error probability as a function of the network size. Note that the network size can be reduced significantly in the same area for the same error rate when we use Chernoff routing over the conventional routing methods.

Finally, we describe a modification to Chernoff routing that can provide an additional several orders of magnitude reduction in the P_{DE} compared to the conventional schemes. The idea is to tune the value of the weighing parameter λ for each pairwise link in the network as a function of parameters that are known at each sensor: the SNR, the correlation parameter, and the link distance of the next hop. Figure 9 provides some insight into the potential gains of such an approach. Specifically, we have plotted the P_{DE} for a fixed route (the one originating at coordinates $\approx (4.25, -4.50)$ in the topology graph of Figure 5) versus λ for several values of the correlation parameter A. It is seen that performance is highly sensitive to λ and that choosing λ suboptimally can result in a loss of several orders of magnitude in P_{DE} performance, particularly when the field is weakly correlated. This effect has been observed to be more severe for routes containing more hops and less severe for routes with fewer hops. Therefore, it is expected that analytic approaches to choosing λ optimally, either on a per-link basis or as a single value chosen for all network links, can further boost the performance of Chernoff routing beyond the results we have detailed in this work. Analytic approaches to this problem are an interesting avenue of future research.

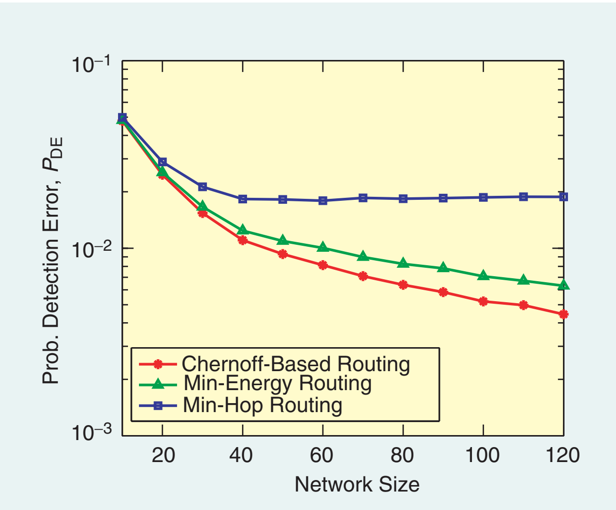


[FIG7] Performance analysis of the three routing strategies in terms of the: (a) average number of hops, (b) total energy consumed, (c) accumulated Chernoff information, and (d) average Chernoff (detection) efficiency.

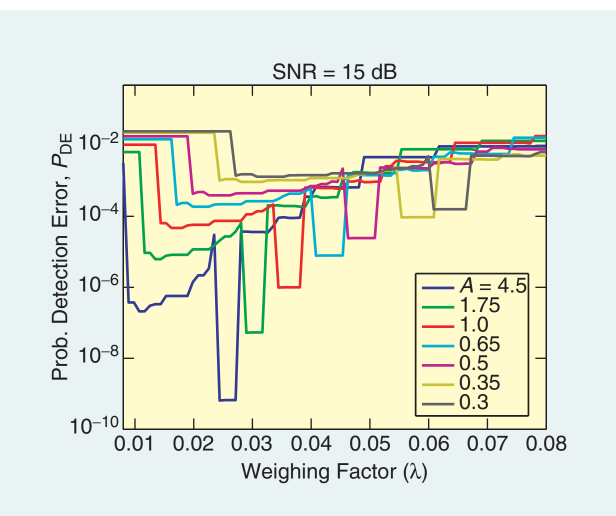
CONCLUSIONS AND RELATED WORK

In this article, we have applied classical signal processing tools to illustrate how application-specific design can provide significant performance gains in wireless ad hoc sensor networks. As an illustrative example, we have presented a new approach to routing that maximizes detection performance in such a network. By considering a well-specified application (rather than general characteristics of a set of applications), we were able to develop new protocols that depend explicitly on the application parameters and hence outperform conventional approaches. Although we considered one specific example of application-dependent design, we believe that it exemplifies many of the defining characteristics of the application-specific approach to protocol design.

The literature on routing for ad hoc network is vast. Common route metrics include the hop count [14], [20], [23], traffic level and corresponding delay [1], [2], [35], packet success rate [12], and interference level [11], [33], [34]. There is a growing body of work on routing for sensor networks. The idea of cross-layer design has been explored by many in recent years; see, e.g., [8], [13], [17], and [22]. Energy is a particularly well-studied metric, and approaches based on required communications energy, processing energy, system lifetime, and/or residual battery energy have been given in [4], [5], [9], [10], [19], [21], [27], [28], and [32]. Each of these metrics is justifiably popular for an application or set of applications in traditional wireless networks, and it would be easy to assume that similar conclusions carry over to the ad hoc wireless sensor network setting.



[FIG8] Probability of detection error P_{DE} averaged over all routes and network topology under three different routing strategies (SNR = 15 dB, $\nu = 2$, $\lambda = 0.01$, $A = 4$, $\Pr(\mathcal{H}_0) = 0.75$).



[FIG9] Probability of detection error P_{DE} for Chernoff routing for a fixed route versus the weighing factor λ, for several values of the correlation parameter A. Here, SNR = 15 dB, $\nu = 2$, and $\Pr(\mathcal{H}_0) = 0.75$.

Application-based protocol design for wireless ad hoc sensor networks is a promising and largely unexplored research area. We hope that we have impressed upon the reader that there exist many critical applications for which the strategies developed for conventional wireless networks can be improved upon. The authors hope to stimulate interest of the readers through the discussion in this article.

Finally, we note that we have not addressed several issues, such as the effect of transmission interference, the potential for frequency reuse, and the design of a carefully constructed MAC and PHY layer, that would be fundamental to a real-world implementation of the proposed design. Also of practical and perhaps theoretical interest is the effect of quantization both the real-valued observations at each sensor and the scalar quantities used for prediction and propagation in the Kalman filter implementation of the proposed routing algorithm. Although some partial results exist for general models, they have not yet been applied to the specific model used in the current work. Such issues present an open avenue of further research.

ACKNOWLEDGMENTS

This work was supported in part by the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011 and the National Science Foundation under Contract CNS-0435190.

AUTHORS

Youngchul Sung (ysung@qualcomm.com) received the B.S. and M.S. degrees from Seoul National University, Korea, in electronics engineering in 1993 and 1995, respectively, and the Ph.D. degree in electrical engineering from Cornell University, Ithaca, New York, in 2005. He is currently a senior engineer at Qualcomm, Inc. His areas of interest include mobile communication network analysis and design, communication theory, large-scale sensor networks, and statistical signal processing.

Saswat Misra (sm353@cornell.edu) received the B.S. in electrical engineering from the University of Maryland,

College Park, in 2000 and the M.S. in electrical engineering from the University of Illinois, Urbana-Champaign, in 2002. Since 2002, he has been a research scientist at the Army Research Laboratory, Adelphi, Maryland, in the Communications and Network Systems division. He has previously worked on optimal training design for wireless communication systems; an area in which he has published several papers and holds two patents (pending). Since 2005, he has been a Ph.D. candidate at Cornell University. He is currently studying routing and security issues in wireless military networks.

Lang Tong (ltong@ece.cornell.edu) is a professor in the School of Electrical and Computer Engineering, Cornell University, Ithaca, New York. He received the B.E. degree from Tsinghua University, Beijing, China, in 1985 and the M.S. and Ph.D. degrees in electrical engineering in 1987 and 1991, respectively, from the University of Notre Dame, Indiana. He received the Young Investigator Award from the Office of Naval Research in 1996, the Outstanding Young Author Award from the IEEE Circuits and Systems Society, the 2004 Best Paper Award (with M. Dong) from the IEEE Signal Processing Society, and the 2005 Leonard G. Abraham Prize Paper Award (with P. Venkitasubramaniam and S. Adireddy) from the IEEE Communications Society.

Anthony Ephremides (tony@eng.umd.edu) received his B.S. degree from the National Technical University of Athens (1967) and M.S. (1969) and Ph.D. (1971) degrees from Princeton University, all in electrical engineering. He has been at the University of Maryland since 1971 and currently holds a joint appointment as professor in the Electrical Engineering Department and the Institute of Systems Research (ISR). He is cofounder of the NASA Center for Commercial Development of Space on Hybrid and Satellite Communications Networks. He received the 1992 IEEE Donald E. Fink Prize Paper Award. He has been the president of the IEEE Information Theory Society (1987) and served on the IEEE Board of Directors (1989 and 1990).

REFERENCES

[1] R. Agarwal and A. Goldsmith, "Joint rate allocation and routing for multihop wireless networks with delay-constrained data," Tech. Rep., Wireless Systems Lab, Stanford University, CA, 2004.

[2] M. Andrews, A. Fernandez, M. Harchol-Balter, T. Leighton, and L. Zhang, "General dynamic routing with per-packet delay guarantees of O(distance +1/session rate)," *SIAM J. Computing*, vol. 30, no. 5, pp. 1594–1623, 2002.

[3] D.P. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1992.

[4] H. Chang and L. Tassiulas, "Routing for maximum system lifetime in wireless ad-hoc networks," in *Proc. 37th Allerton Conf. Commun., Control, Computing*, Monticello, IL, Sep. 1999, pp. 1–10.

[5] H. Chang and L. Tassiulas, "Energy conserving routing in wireless ad-hoc networks," in *Proc. IEEE INFOCOM 2000*, Tel-Aviv, Israel, Mar. 2000, pp. 22–31.

[6] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*. Boston, MA: Jones and Bartlett, 1993.

[7] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[8] T. Girici and A. Ephremides, "Joint routing and scheduling metrics for ad hoc wireless networks," in *Proc. 36th Asilomar Conf. Signals, Systems and Computers*, Nov. 2002, pp. 1–5.

[9] T. Girici and A. Ephremides, "Optimal power control for wireless queueing networks," in *Proc. CISS 2004*, Princeton, NJ, Mar. 2004.

[10] T. Girici and A. Ephremides, "Energy-efficient power control in wireless queueing networks," in *Proc. 37th Annual Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep. 2004.

[11] R. Gupta, Z. Jia, T. Tung, and J. Walrand, "Interference-aware QoS Routing (IQRouting) for ad-hoc networks," in *Proc. ICC 2005, Seoul*, Korea, May 2005, pp. 1–6.

[12] L. Iannone, R. Khalili, K. Salamatian, and S. Fdida, "Cross-layer routing in wireless mesh networks," in *Proc. ISWCS2004*, Mauritius, Sept. 2004, pp. 1–5.

[13] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: A scalable and robust paradigm for sensor networks," in *Proc. 6th ACM/Mobicom Conf.*, Boston, MA, 2000, pp. 1–13.

[14] D.B. Johnson and D.A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*, T. Imielinski and H.F. Korth, Ed. Norwell, MA: Kluwer Academic, Mar. 1996, pp. 1–18.

[15] T. Kailath, "The innovations approach to detection and estimation theory," *Proc. IEEE*, vol. 58, no. 5, pp. 680–695, May 1970.

[16] T. Kailath, A.H. Sayed, and B. Hassibi, *Linear Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 2000.

[17] B. Krishnamachari, D. Estrin, and S. Wicker, "Modeling data centric routing in wireless sensor networks," in *Proc. IEEE INFOCOM*, New York, June 2002, pp. 1–11.

[18] G. Mergen and L. Tong, "Estimation over deterministic multiaccess channels," in *Proc. 42nd Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Oct. 2004, pp. 1–10.

[19] A. Michail and A. Epremides, "Energy efficient routing for connection-oriented traffic in wireless ad hoc networks," *Mobile Networks Applicat.*, vol. 8, no. 1, pp. 517–533, 2003.

[20] V.D. Park and S. Corson, "A highly adaptive distributed routing algorithm for mobile wireless networks," in *Proc. IEEE INFOCOM'97*, 1997, vol. 3, pp. 1405–1413.

[21] S. Park and M. Srivastava, "Power aware routing in sensor networks using dynamic source routing," in *ACM MONET Special Issue on Energy Conserving Protocols in Wireless Networks*, 1999.

[22] S. Pattem, B. Krishnamachari, and R. Govindan, "The impact of spatial correlation on routing with compression in wireless sensor networks," in *Proc. IPSN 2004*, Berkeley, CA, Apr. 2004, pp. 1–8.

[23] C.E. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers," in *Proc. ACM SIGCOMM*, 1994, pp. 1–11.

[24] C.E. Perkins and E.M. Royer, "Ad hoc on-demand distance vector routing," in *Proc. 2nd IEEE Workshop on Mobile Computing Systems and Applications*, Feb. 1999, pp. 90–100.

[25] H.V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer-Verlag, 1994.

[26] F.C. Schweppe, "Evaluation of likelihood functions for Gaussian signals," *IEEE Trans. Inform. Theory*, vol. IT-1, no. 11, pp. 61–70, 1965.

[27] R. Shah and J. Rabaey, "Energy aware routing for low energy ad hoc sensor networks," in *Proc. 1st Int. Workshop Sensor Network Protocol and Applications*, May 2003, pp. 1–6.

[28] S. Singh, M. Woo, and C. Raghavendra, "Power-aware routing in mobile ad hoc networks," in *Proc. 4th ACM/IEEE Conf. Mobile Computing and Networking*, Oct. 1998, pp. 1–10.

[29] Y. Sung, S. Misra, L. Tong, and A. Ephremides, "A new link metric for the detection of correlated signal fields in multi-hop wireless sensor networks," Adaptive Commun. Signal Processing Group, School of Elect. Comput. Eng., Cornell Univer., Tech. rep. ACSP-01-06-01, Jan. 2006.

[30] Y. Sung, L. Tong, and A. Ephremides, "A new metric for routing in multihop wireless sensor networks for detection of correlated random fields," in *Proc. MILCOM 2005*, Atlantic City, NJ, Oct. 2005, pp. 1–6.

[31] Y. Sung, L. Tong, and H.V. Poor, "A large deviations approach to sensor scheduling for detection of correlated random fields," in *Proc. ICASSP 2005*, Philadelphia, PA, Mar. 2005, pp. 1–4.

[32] J.E. Wieselthier, G.D. Nguyen, and A. Ephremides, "On the construction of energy-efficient broadcast and multicast trees in wireless networks," in *Proc. INFOCOM 2000*, Tel-Aviv, Israel, Mar. 2000, pp. 1–10.

[33] Y. Yang and R. Kravets, "Contention-aware admission control for ad hoc networks," *IEEE Trans. Mobile Computing*, vol. 4 , no. 4 , pp. 363–377, July 2005.

[34] C. Zhu and M.S. Corson, "QoS routing for mobile ad hoc networks," in *Proc. INFOCOM 2002*, New York, NY, pp. 1–10.

[35] T. Znati and R. Melhem, "Node delay assignment strategies to support end-to-end delay requirements in heterogeneous networks," *IEEE/ACM Trans. Netw.*, vol. 12, no. 5, pp. 879–892, 2004.