[Jean-François Chamberland and Venugopal V. Veeravalli]

Wireless Sensors in Distributed Detection Applications

[An alternative theoretical framework tailored to decentralized detection]

The ability to detect events of interest is essential to the success of emerging sensor network technology. Detection often serves as the initial goal of a sensing system. Indeed the presence of an object has to be ascertained before a sensor network can estimate attributes such as position and velocity. For systems observing infrequent events, detection may be the prevalent function of the network. Furthermore, in some applications such as surveillance, the detection of an intruder is the sole purpose of the sensor system. In the setting where local sensors preprocess observations before transmitting data to a fusion center, the corresponding decision making problem is termed decentralized detection.

1053-5888/07/$25.00©2007IEEE

Decentralized detection with fusion was an active research field during the 1980s and early 1990s, following the seminal work of Tenney and Sandell [1]. The application driver for this research was distributed radar. The high cost of data transfers at the time prompted system designers to quantize and compress data locally before information was relayed to the fusion center, hence the decentralized aspect of the problem. The goal was to design the sensor nodes and the fusion center to detect the event as accurately as possible, subject to an alphabet-size constraint on the messages transmitted by each sensor node [2]–[4].

More recently, decentralized detection has found applications in sensor networks. Wireless sensor nodes are typically subject to stringent resource constraints. To design an efficient system for detection in sensor networks, it is imperative to understand the interplay between data compression, resource allocation, and overall performance in distributed sensor systems. Classical results on inference problems and decentralized detection in particular can be leveraged and extended to gain insight into the efficient design of sensor networks. These results form a basis for much of the recent work on detection in sensor networks.

THE CLASSICAL DECENTRALIZED DETECTION FRAMEWORK

In the classical decentralized detection problem, a set of dispersed sensor nodes receives information about the state of nature. Based on its observation, sensor node ℓ selects one of D_ℓ possible messages and sends it to the fusion center via a dedicated channel. Upon reception of the data, the fusion center produces an estimate of the state of nature by selecting one of the possible hypotheses. Evidently, a distributed sensor system in which every sensor node transmits a partial summary of its own observation to the fusion center is suboptimal compared to a centralized system in which the fusion center has access to all the observations without distortion. Nevertheless, factors such as cost, spectral bandwidth limitations, and complexity may justify the use of compression algorithms at the nodes. A generic decentralized detection setting is illustrated in Figure 1. Resource constraints in the classical framework are captured by fixing the number of sensor nodes and further imposing a finite-alphabet constraint on the output of each sensor. This implicitly bounds the amount of data available at the fusion center, as both the number of nodes and the number of possible messages per node are finite. Perfect reception of the sensor outputs is typically assumed at the fusion center. It is important to recognize that once the structure of the information supplied by each sensor node is fixed, the fusion center faces a standard problem of statistical inference. As such, a likelihood-ratio test on the received data will minimize the probability of error at the fusion center for a binary hypothesis testing problem. The crux of a standard decentralized inference problem is to determine what type of information each sensor should send to the fusion center.

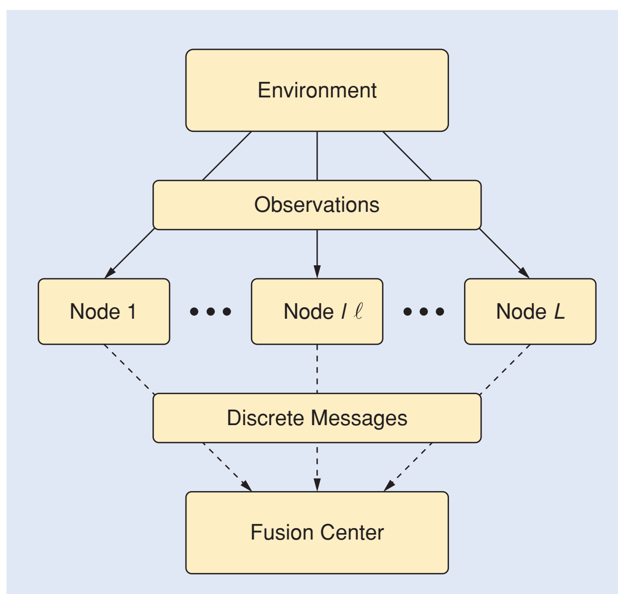
EXAMPLE 1

Consider a detection problem where the fusion center must distinguish between two hypotheses, H_0 and H_1, based on L observations. Each observation consists of one of two possible signals, $s_1 = -s_0 = m$, corrupted by additive noise

$$Y_\ell = s_j + N_\ell, \quad \ell = 1, \dots, L. \tag{1}$$

The observation noise is assumed to be a sequence of independent and identically distributed (i.i.d.) Gaussian components with zero-mean and variance σ^2. This implies that the observed process conditioned on the true hypothesis is a sequence of i.i.d. Gaussian random variables. Suppose that the observations are only available at some remote sensor locations and assume that each sensor must quantize its own observation to a single bit. One possible quantization rule for the sensors is a function that returns the sign of the observation. The information reaching the fusion center is of the form $U_\ell = \gamma(Y_\ell) = \text{sign}(Y_\ell)$. If we further assume that the two hypotheses are equally likely, then an optimal decision procedure at the fusion center for this special case is a majority rule on the received variables: H_0 is selected if more zeros are received and H_1 is picked otherwise [5]. We emphasize that while the optimality of the decision rule at the fusion center is evident, it is more difficult to qualify the suitability of the local quantization rules at the sensors.

A celebrated accomplishment in decentralized detection for binary hypothesis testing is the demonstration that, for the classical framework, likelihood-ratio tests at the sensor nodes are optimal whenever observations are conditionally independent given each hypothesis [2]. This property drastically reduces the search space for an optimal collection of local quantizers and, although the resulting problem is not necessarily easy, it is



[FIG1] Abstract representation of the classical decentralized detection framework.

amenable to analysis in many contexts. In general, it is reasonable to assume conditional independence across sensor nodes if the limited accuracy of the sensors is responsible for noisy observations. However, if the observed process is stochastic in nature or if the sensors are subject to external noise, this assumption may fail. Without the conditional independence assumption, the goal of finding the optimal solution to the decentralized detection problem rapidly becomes computationally intractable.

Even under a conditional independence assumption, finding optimal quantization levels at the sensor nodes remains, in most cases, a difficult task. This optimization problem is known to be tractable only under restrictive assumptions regarding the observation space and the topology of the underlying network. The solution does not scale well with the number of sensors except in some special cases, and it is not robust with respect to priors on the observation statistics.

[THE ABILITY TO DETECT EVENTS OF INTEREST IS ESSENTIAL TO THE SUCCESS OF EMERGING SENSOR NETWORK TECHNOLOGY.]

A popular heuristic method to design decentralized detection systems is to apply a person-by-person optimization (PBPO) technique. This technique consists in optimizing the decision rule of one sensor at a time while keeping the transmission maps of the remaining sensors fixed. The index of the sensor node being optimized changes with every step. The overall performance at the fusion center is guaranteed to improve (or, at least, to not worsen) with every iteration of the PBPO algorithm. Unfortunately, this algorithm does not necessarily lead to a globally optimal solution. Other notable heuristics applicable to the design of a decentralized detection system include the saddle-point approximation method [6] and techniques based on empirical risk minimization and marginalized kernels [7]. In contrast to a majority of the work on decentralized detection, the kernel method addresses system design for situations where only a collection of empirical samples is available; the joint distributions of the sensor observations conditioned on the possible hypotheses need not be known.

For wireless sensor networks with a small number of nodes, intuition regarding an optimal solution may be misleading. Consider a scenario where observations at the sensor nodes are conditionally i.i.d. The symmetry in the problem suggests that the decision rules at the sensors should be identical, and identical local decision rules are frequently assumed in many situations. However, counterexamples for which nonidentical decision rules are optimal have been identified [2]. Interestingly, identical decision rules are optimal in the asymptotic regime where the number of active sensors increases to infinity. For any reasonable collection of transmission strategies, the probability of error at the fusion center goes to zero exponentially as L grows unbounded. It is then adequate to compare collections of strategies based on their exponential rate of convergence to zero,

$$\lim_{L\to\infty} \frac{\log P_e(\mathcal{G}_L)}{L}. \tag{2}$$

We use $\mathcal{G}_L$ as a convenient notation for a system configuration that contains L sensors.

RESULT 1

Suppose that the observations are conditionally independent and identically distributed. Then, using identical local decision rules for all the sensor nodes is asymptotically optimal in terms of error exponent [8].

This result was originally proved by Tsitsiklis [8] through an application of the Shannon-Gallager-Berlekamp lower bound. An alternative derivation can be obtained using the encompassing framework of large deviations. Asymptotic regimes applied to decentralized detection are convenient because they capture the dominating behaviors of large systems. This leads to valuable insights into the problem structure and its solution.

DECENTRALIZED DETECTION IN WIRELESS SENSOR NETWORKS

The classical decentralized detection framework has limited application to modern wireless sensor networks, as it does not adequately take into account important features of sensor technology and of the wireless channel between the sensors and the fusion center. In particular, finite alphabet restrictions on the sensor outputs do not capture the resource constraints of cost, spectral bandwidth, and power adequately for efficient design. Furthermore, the assumption that sensor messages are received reliably at the fusion center ignores the link variability intrinsic to wireless communications.

Reevaluating the original assumptions of the classical decentralize detection framework is an instrumental step in deriving valuable guidelines for the efficient design of sensor networks. Many recent developments in the field have been obtained by studying the classical problem while incorporating more realistic system assumptions in the problem definition. The motivation underlying many of these new research initiatives is the envisioned success of future wireless sensor networks. An alternative theoretical framework tailored to decentralized detection over sensor networks is starting to emerge, as depicted in Figure 2.

Network architectures for distributed sensor systems come in many different flavors. A carefully deployed system usually forms a tree. In this configuration, the information propagates from the sensor nodes to the fusion center in a straightforward manner, following a unique deterministic path. The parallel architecture, a subclass of the tree category where each node communicates directly with the fusion center, has received much attention in the decentralized detection literature. A distributed sensor system can also assume the form of a self-configuring wireless sensor network. In such systems, nodes are positioned randomly in an environment and then cooperate with one another to produce a dynamic communication infrastructure for the resulting network. A reasonable abstraction for decentralized detection over sensor networks is one where the sensors local to an event of interest are used for sensing, and they transmit their information using a single hop or multiple hops to a fusion center. The other sensors in the system may be used as relays or routers. The fusion center is responsible for final decision making and further relaying of the information across the network if necessary.

Distributed sensing induces a natural tradeoff between performance, communication, and complexity. Combining information from neighboring nodes via in-network signal processing can improve reliability and reduce the amount of traffic on the network. On the other hand, the exchange of additional information could potentially yield better decisions. A simple technique to exchange information in the context of decentralized detection is proposed by Swaszek and Willett [9]. The authors explore the use of feedback, successive retesting, and rebroadcasting of the updated decisions as a means of reaching a consensus among sensors. Two modes of operation are discussed: a fast mode, where a decision is reached rapidly, and an optimum decision scheme that may require several rounds of information sharing before a consensus is reached. These two schemes illustrate well the natural tradeoff between resource consumption and system performance, as the more intricate scheme performs better.

Under a different setting, in-network signal processing is studied by D'Costa et al. [10]. In their work, observations are assumed to possess a local correlating structure that extends only to a limited area. As such, the sensor network can be partitioned into disjoint spatial coherence regions over which the signals remain strongly correlated. The resulting partition imposes a structure on the optimal decision rule that is naturally suited to the communication constraints of the network. Information is exchanged locally to improve the reliability of the measurements, while compressed data is exchanged among coherence regions. Under mild conditions, the probability of error of the proposed classification scheme is found to decay exponentially to zero as the number of measurements approaches infinity.

EMERGING FRAMEWORK

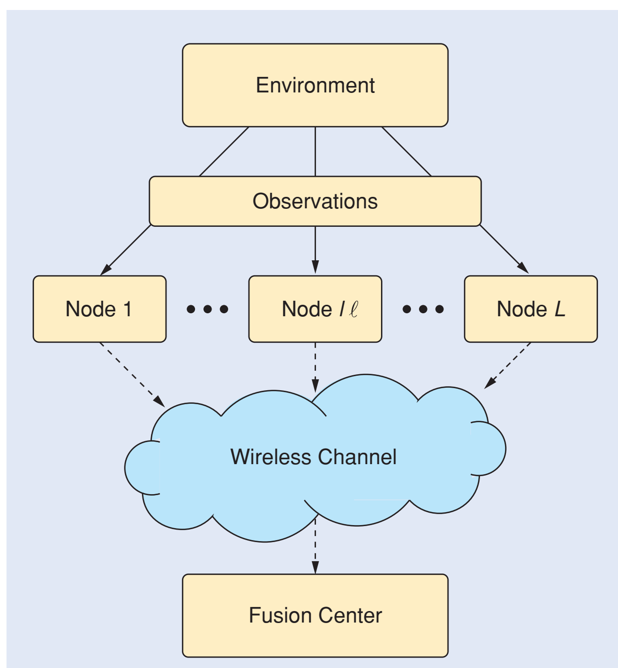
A significant departure from the classical decentralized detection framework comes from the realization that wireless sensors transmit data over a shared medium, the common wireless spectrum. A problem formulation that better accounts for the physical resource constraints imposed on the system is needed for accurate performance evaluation. As discussed above, sensor nodes are often subject to very stringent power requirements. A limited spectral bandwidth and a bound on the total cost of the system may further exacerbate the design process. A flexible and adequate solution to distributed sensing should account for these important factors. It is possible to extend the findings of Result 1 to the case where system resources rather than the number of sensors constitute the fundamental design limitation. For instance, this resource budget may represent a sum-rate constraint, a total power requirement, a bound on system cost, or a combination thereof.

RESULT 2

Suppose that the observations are conditionally independent and identically distributed. Then, under a global resource constraint, using identical transmission mappings for all the sensor nodes is asymptotically optimal [11].

A necessary condition for this result to hold is that the number of sensor nodes must tend to infinity as the actual resource budget grows without bound. This is usually the case, as the amount of information provided by a single observation is bounded and, consequently, the allocation of physical resources devoted to the corresponding sensor should also be finite. This result provides an extension to Result 1 and to the work by Tsitsiklis [8]. In the current setting, the resource budget rather than the number of sensor nodes forms the fundamental constraint on the sensor system. Moreover, the local decision rule γ need not be a finite-valued function, and the communication channels between the sensor nodes and the fusion center need not be noiseless. The optimality of wireless sensor networks with identical sensor nodes is encouraging. Such networks are easily implementable, amenable to analysis, and they provide robustness to the system through redundancy.

Asymptotic analyses based on error exponents also have the added benefit of decoupling the optimization across the sensors. The sensor mappings can be designed according to a local metric. For example, consider a Bayesian problem formulation where the probability of error at the fusion center is to be minimized. For wireless sensor networks with a large resource budget and conditionally i.i.d. observations, prospective sensor



[FIG2] Abstract representation of an alternative decentralized detection framework.

types should be compared according to their normalized Chernoff information

$$-\frac{1}{a(\gamma)}\min_{\lambda\in[0,1]}\left\{\log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}\left(\frac{d\mathcal{Q}_{1,\gamma}}{d\mathcal{Q}_{0,\gamma}}\right)^{\lambda}\right\}, \tag{3}$$

where $a(\gamma)$ is the expected amount of system resources consumed by a node of type γ, and $\mathcal{Q}_{j,\gamma}$ is the induced probability measure on the received information U_ℓ at the fusion center under hypothesis H_j. Intuitively, allocating a larger amount of resources per node implies receiving detailed information from each node at the fusion center. On the other hand, for a fixed budget, a reduction in resource consumption per node allows the system to operate with more active sensors. The normalized Chernoff information describes in mathematical terms how this tradeoff takes place: Chernoff information divided by consumed resources.

[THE CRUX OF A STANDARD DECENTRALIZED INFERENCE PROBLEM IS TO DETERMINE WHAT TYPE OF INFORMATION EACH SENSOR SHOULD SEND TO THE FUSION CENTER.]

We can extend the preceding result to the Neyman-Pearson variant of the detection problem with little effort. In this latter problem formulation, the normal operation of the system is H_0, while H_1 is considered a rare event. The prior probabilities on H_0 and H_1 are unknown. The function $a(\gamma_\ell)$ denotes the amount of resources consumed by sensor node ℓ under hypothesis H_0, and the global resource budget is a constraint on the behavior of the system under hypothesis H_0. In this case, the normalized relative entropy

$$\frac{1}{a(\gamma)}D\left(\mathcal{Q}_{0,\gamma}\|\mathcal{Q}_{1,\gamma}\right) \tag{4}$$

plays the role of the normalized Chernoff information. Indeed, in the Neyman-Pearson framework, prospective sensor types for a sensor network with a large resource constraint should be compared according to the normalized relative entropy.

When the observations are not conditionally i.i.d., the normalized Chernoff information (or relative entropy) can no longer be shown to be the right metric for optimizing the sensor mappings. However, even in this case, the findings described in Result 2 can be used as some justification to decouple the optimization across sensors. In this context, it is important to distinguish between the asymptotic results in the Bayesian and Neyman-Pearson formulations in that, in the latter formulation, the normalized relative entropy can be shown to be the right metric for optimizing the sensor mappings as long as the observations are conditionally independent and there are a large number of sensors of each type. The minimization over λ in (3) does not allow for a similar generalization in the Bayesian setting.

DETECTION UNDER CAPACITY CONSTRAINT

The admissible rate-region of a practical system with a simple encoding scheme may depend on the bandwidth, the signal power, the noise density, and the maximum tolerable bit-error rate at the output of the decoder. Specifying these quantities is equivalent to fixing the sum-rate of the corresponding multiple-access channel. A natural initial approach to the capacity-constrained problem is to overlook the specifics of these physical parameters and to constrain the sum-capacity of the multiple-access channel available to the sensors. Thus, the new design problem becomes selecting the number of sensors (L) and the number of admissible messages for each sensor (D_ℓ) to optimize system performance at the fusion center, subject to the capacity constraint

$$\sum_{\ell=1}^{L}\lceil\log_2(D_\ell)\rceil \le R. \tag{5}$$

For the time being, we neglect communication errors in the transmitted bits. Upon reception of the data, the fusion center makes a tentative decision about the state of nature.

We know that using identical transmission functions for all the sensor nodes is asymptotically optimal. In the framework of Result 2, a discrete transmission mapping γ^* is an optimal function if it maximizes the normalized Chernoff information,

$$\gamma^* = \arg\max_{\gamma} -\frac{1}{\lceil\log_2(D_\gamma)\rceil} \times \min_{\lambda\in[0,1]}\left\{\log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}\left(\frac{d\mathcal{Q}_{1,\gamma}}{d\mathcal{Q}_{0,\gamma}}\right)^{\lambda}\right\}. \tag{6}$$

As an immediate corollary to this result, it can be shown that binary sensors are optimal if there exists a binary quantization function γ_b whose Chernoff information exceeds half of the information contained in an unquantized observation [11], [12]. If the contribution of the first bit of quantized data to the Chernoff information exceeds half of the Chernoff information offered by an unquantized observation, then using binary sensors is optimal. This corollary is not too surprising in itself. However, the significance of this result is that the requirements of the corollary are fulfilled for important classes of observation models [12]. Binary sensor nodes are optimal for the problem of detecting deterministic signals in Gaussian noise and for the problem of detecting fluctuating signals in Gaussian noise using a square-law detector. That is, in these scenarios, the gain offered by having more sensor nodes outperforms the benefits of getting detailed information from each sensor.

This attribute can be generalized to a very important property that seems to be valid for a wide array of detection problems. In most detection settings, including the ones specified above, the number of bits necessary to capture most of the information contained in one observation appears to be very small. In other words, for detection purposes, the information contained in an observation is found in the first few bits of compressed data. The performance loss due to quantization decays very rapidly as the number of quantization levels increases. As such, message compression only plays a limited role in overall system performance. This property greatly simplifies quantizer design and system deployment. A second property worth mentioning at this point is, for conditionally i.i.d. observations, the diversity obtained by using multiple sensors more than offsets the performance degradation associated with receiving only coarse data from each sensor [11], [12].

WIRELESS CHANNEL CONSIDERATIONS

Most of the early results on decentralized detection assume that each sensor node produces a finite-valued function of its observation, which is conveyed reliably to the fusion center. In a wireless system, this latter assumption of reliable transmission may fail as information is transmitted over noisy channels [13]. This limitation is made worse by the fact that most detection problems are subject to very stringent delay constraints, thereby preventing the use of powerful error-correcting codes at the physical layer. Many recent research initiatives on decentralized detection consist in incorporating the effects of the wireless environment on the transmission of messages between the sensors and the fusion center.

EXAMPLE 2

Consider a distributed sensor network akin to the one introduced in Example 1, except that data must be transmitted over parallel wireless communication channels. The fusion center receives degraded information U_ℓ from sensor ℓ of the form

$$U_\ell = \gamma_\ell(Y_\ell) + W_\ell, \tag{7}$$

where W_ℓ is additive Gaussian noise with distribution $\mathcal{N}(0, \sigma_w^2)$. We study the simple situation where the additive noise is i.i.d. across sensor nodes. The hypothesis testing problem consists of deciding based on the received sequence $\{U_\ell\}$ whether the law generating $\{Y_\ell\}$ is $\mathcal{P}_0$ corresponding to hypothesis H_0, or $\mathcal{P}_1$ corresponding to hypothesis H_1. We focus on the specific detection problem where the wireless sensor network is subject to a total power constraint. The expected radiated power summed across all the sensor nodes may not exceed a given constraint A,

$$\sum_{\ell=1}^{L} a(\gamma_\ell) \leq A, \tag{8}$$

where $a(\gamma_\ell) > 0$ represents the expected power used by sensor node ℓ. This problem falls in the general framework of Result 2. Identical sensors are therefore optimal and system performance is maximized by using the normalized Chernoff information as a design criterion for individual sensors.

For the purpose of illustration, we study the class of nodes where each unit retransmits an amplified version of its own observation, $\gamma^{(s)}(y) = sy$. We can express the expected radiated power per sensor as $a(\gamma^{(s)}) = s^2(m^2 + \sigma^2)$ and the corresponding normalized Chernoff information as

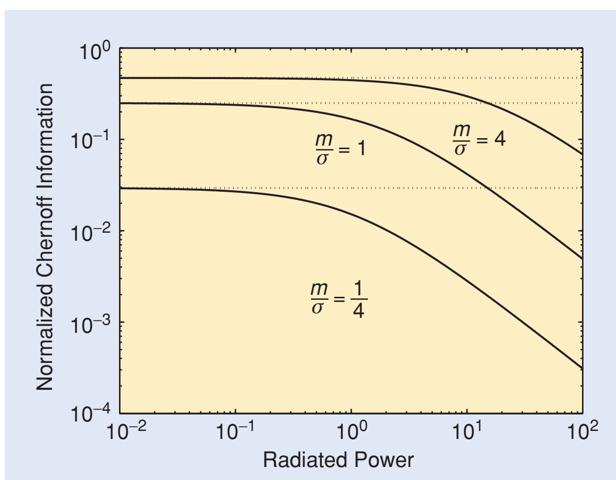
$$-\frac{1}{a\left(\gamma^{(s)}\right)} \log\left(\int_{-\infty}^{\infty} \sqrt{\mathcal{Q}_{0,\gamma^{(s)}}(u)\mathcal{Q}_{1,\gamma^{(s)}}(u)}du\right) = \frac{m^2}{2(m^2+\sigma^2)(s^2\sigma^2+\sigma_w^2)}, \tag{9}$$

where m is the signal amplitude and σ is the variance of the observation noise. Figure 3 plots the normalized Chernoff information, which is a monotone decreasing function of the radiated power in the present case.

Interestingly, although the problem definition of Example 2 constitutes a significant departure from the classical decentralized detection framework, a similar phenomenon is observed. Overall performance is optimized when the system uses as many independent sensors as possible, giving each sensor a minimum amount of system resources. The same conclusion can be reached for a class of sensor nodes where each node compresses its own observation to a one-bit summary message. Once again, the tradeoff between the number of sensors and the amount of resources allocated to each sensor seems to favor large networks composed of many nodes. The use of a restricted class of sensors in Example 2 underscores the difficulty of finding an optimal transmission mapping for the sensors. In general, identifying the best possible transmission map involves a non-convex optimization problem over a space of measurable functions. Such problems are typically very difficult to solve.

CORRELATED OBSERVATIONS

While the popular assumption that observations at the sensors are conditionally independent is convenient for analysis, it does not necessarily hold for arbitrary sensor systems. For instance, whenever sensor nodes lie in close proximity of one another, we expect their observations to become strongly correlated. Different approaches have been employed to study the latter problem, most of which focus on small sample sizes [14]. The theory of large deviations can be used to assess the performance of wireless sensor systems exposed to correlated observations



[FIG3] Normalized Chernoff information for analog relay amplifiers.

[15]. For differentiating between known signals in Gaussian noise, overall performance improves with sensor density.

EXAMPLE 3

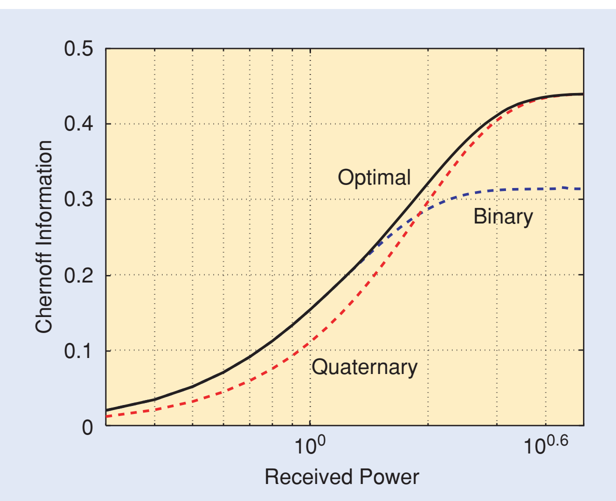
Consider the detection problem where observations become increasingly correlated as sensor nodes are placed in closer proximity. Mathematically, we adopt the same system as in Example 2. However, the observation noise sequence is equivalent to the sampling of a one-dimensional Gauss-Markov stochastic process. The covariance function of the observation noise is given by $\mathrm{E}[N_k N_\ell] = \sigma^2 \rho^{d(k,\ell)}$ where $d(k, \ell)$ is the distance between sensors k and ℓ. When the sensors are equally spaced at an interval of length d, the best possible error exponent becomes

$$\frac{m^2}{2(m^2+\sigma^2)} \frac{(1-\rho^d)}{\sigma_w^2(1-\rho^d)+s^2\sigma^2(1+\rho^d)}. \tag{10}$$

Correlation degrades overall performance. Still, it is interesting to note that performance improves with node density. Although correlation and observation signal-to-noise ratio affect the overall probability of error, they do not necessarily change the way the sensor network should be designed. Systems with many low-power nodes will perform well for the detection of deterministic signals in Gaussian noise. We stress that there exist situations where performance does not necessarily improve with node density [15].

ATTENUATION AND FADING

If sensor nodes are to be scattered around somewhat randomly, it is conceivable that their respective communication channels will feature different mean path gains, with certain nodes possibly having much better connections than others. Furthermore, changes in the environment, interference, and motion of the sensors can produce time-variations in the quality of the wireless channels. It is therefore of interest to quantify the impact of fading on the performance of distributed sensor systems.



[FIG4] Chernoff information for binary signaling, quaternary signaling, and for the adaptive scheme where power allocation is based on the state of the channel.

Chen et al. [16] modify the classical decentralized detection problem by incorporating a fading channel between each sensor and the fusion center. They derive a likelihood-ratio-based fusion rule for fixed local decision devices. This optimum fusion rule requires perfect knowledge of the local decision performance indices and the state of the communication channels over which messages are sent. A decision rule based on maximum-ratio combining and a two-stage approach inspired by the Chair-Varshney decision rule are also analyzed. These concepts are further researched by Niu et al. [17] for the scenario where instantaneous channel state information is not available at the fusion center. They propose a fusion rule that only requires knowledge of the channel statistics. In general, having channel state information at the sensors permits the usage of adaptive transmission mappings where a sensor decides what type of information to send based on the current quality of the channel.

EXAMPLE 4

We revisit the problem described in Example 2. This time, individual wireless connections are subject to fading. The sensors and the fusion center have perfect channel state information. Assume that the total power budget per transmission is fixed. We consider the specific case where the observation at each sensor node is quantized to two bits. These bits are sent directly to the fusion center over a Rayleigh fading channel. Depending on the specific realization of its channel gain, sensor node ℓ decides how much power should be allocated to the most significant bit and how much power should be given to the second bit.

The optimal power allocation is found to vary significantly with channel gain. At low signal-to-noise ratio, most of the power is given to the most significant bit; while at higher signal-to-noise ratio, power is split between the two bits. Figure 4 shows the Chernoff information as a function of received signal power for a binary signaling scheme, a quaternary signaling scheme with uniform bit-power, and the optimal allocation scheme. Clearly, channel state information at the sensor nodes increases overall performance by adapting the signaling schemes of individual sensor nodes based on the fade level of their respective communication channels. For encoded systems, this property entails using error-correcting codes with unequal bit protection.

NEW PARADIGMS

Recently, researchers have started to explore new paradigms for detection over wireless sensor networks. These alternate points of view offer vastly different solutions to the problem of distributed sensing. The disparity among the proposed solutions can be explained in part by the perceived operation of future wireless sensor networks. Some researchers envision sensor networks to be produced as application-specific systems, giving the designer much freedom on how to best use resources. Under this assumption, every component of the network can be engineered anew. Yet others believe that sensor networks will be subject to standard protocols and specifications, effectively imposing a rigid structure on the system. The exchange of information over future wireless networks may very well be governed by specifications similar to the Internet protocol suite (TCP/IP) or the Wi-Fi standard (IEEE 802.11). While more restrictive, the later philosophy insures the inter-operability of heterogeneous network components and it allows for mass production and cost reduction.

CONSTRUCTIVE INTERFERENCE

Adopting the application-specific viewpoint, Mergen and Tong [18] have proposed communication schemes for decentralized detection where nodes take advantage of the physical layer to transmit information efficiently and reliably. Their communication paradigms exploit the intrinsic broadcast nature of the wireless medium. When complete channel state information is available at the sensor nodes, it is possible for signals originating from various nodes to interfere constructively at the receiver through beam-forming. In a wireless environment, the superposition of multiple signals is equivalent to adding their amplitudes. This property can be employed to sum the local likelihood ratios produced by individual sensors through the wireless channel. The fusion center can then make a final decision by applying a threshold test on the amplitude of the received aggregate signal. In this model, the wireless medium is used both to communicate information to the fusion center and to add signals coherently. This greatly reduces the spectral bandwidth requirement for the system. Physical-layer schemes are found to be asymptotically optimal as the number of sensors increases, provided that the channels from the sensors to the fusion center are statistically identical.

In related research initiatives, Liu and Sayeed [19] and Mergen and Tong [20] propose communication schemes in which sensor nodes transmit according to the type of their observations. This strategy can be applied to decentralized parameter estimation and decentralized detection alike. The type-based multiple-access schemes lead to significant gains in performance when compared to the conventional architecture allocating orthogonal channels to the sensors. Based on the level of the received signal, the fusion center is able to make a decision. The broadcast nature of the wireless channel is exploited in a similar fashion by Hong and Scaglione [21]. In their work, the authors take advantage of the additive nature of a wireless multiple-access channel to address the problems of synchronization, cooperative broadcast, and decision making in sensor networks. Under suitable channel conditions, constructive interference techniques over multiple-access channels provide an interesting solution to the problems of decentralized detection and decentralized parameter estimation. However, certain technical issues such as symbol synchronization, phase synchronization, and security need to addressed before these techniques can be exploited effectively.

MESSAGE PASSING

A second paradigm that may reduce the need for spectral bandwidth is based on local message passing. In the message-passing approach, there is no designated fusion center and the goal is for all the sensors involved in the decision process to reach a consensus about the state of their environment. Every sensor possesses the same prior probability distribution about the true hypothesis and they share a common objective. They update their tentative decision whenever they make a new observation or when they receive additional information from a neighboring wireless sensor. Upon computing a new tentative decision, each sensor broadcasts its latest data to a randomly selected subset of neighbors. Sensors can exchange messages in a synchronous or asynchronous manner until consensus is reached. The design problem is to find communication protocols for communication between the sensors that results in an agreement in a reasonable time. This should be achieved while respecting the constraints imposed on the communication structure and on the system resources. Conditions for asymptotic convergence of the decision sequence made by each sensor and for asymptotic agreement among all the wireless nodes are of interest.

This line of work is heavily influenced by the pioneering work of Borkar and Varaiya [22] on distributed estimation. To facilitate analysis, the sensor network is modeled as a graph that represents the connectivity of various nodes. Data generally takes the form of a node's conditional marginal probability distribution over the possible hypotheses. Advantages of the message-passing paradigm include a simple communication infrastructure, scalability, robustness to sensor failures, and a possible efficient use of the limited system resources.

CROSS-LAYER CONSIDERATIONS

In the previous section, we presented local message-passing as a way to mitigate the effects of path loss and fading in wireless communications. Another way is for the nodes to exploit a multihop communication scheme where data packets are relayed from sensor to sensor until they reach their respective destinations. Although a multihop strategy necessitates more transmissions, the non-linear attenuation intrinsic to wireless channels insures overall savings.

If the data generated by the sensor nodes is to be conveyed over a multi-hop packet network, a few key observations are in order. In the context of decentralized detection, several studies point to the fact that most of the information provided by an observation can be compressed to a very few bits [11], [16]. Accordingly, the performance loss due to quantization decays rapidly as the number of information bits per transmission increases. Data packets carrying sensor information can then be assumed to contain only a few bits without much loss of generality. The exact number of bits per packet is unlikely to be a significant factor in energy consumption in view of the operations that take place at the onset of a wireless connection, and also taking into consideration the size of a typical packet header. Indeed, the payload of a packet in these situations is nearly of the same size or even smaller than its header. It is therefore safe to assume that once a communication link is established between two sensor nodes, the information content of an observation can be transferred essentially unaltered.

This characteristic leads to an all-or-nothing model for data transmission akin to the one put forth by Rago et al. [23].

ENERGY SAVINGS VIA CENSORING AND SLEEPING

A generic sensor node comprises four subsystems: a sensing unit, a microprocessor, a communication unit, and a power supply. Once the components of a sensor node are fixed, the only way to reduce the average power consumption at the node is to shut off some of its units periodically. Assuming that the sensing unit is coupled to the microprocessor and that the operation of the communication unit is contingent on the microprocessor being active, a wireless sensor node has three broad modes of operation. It can be active, with all of its units powered up. Alternatively, it can be in mute mode with its communication unit off, effectively isolating itself temporarily from the rest of the network. Finally, it can be sleeping with all of its units shut. In most sensor networks, substantial energy savings may be achieved by having nodes communicate with the rest of the network only when necessary [23], [24]. While censoring sensors is a straightforward scheme to save energy, a less intuitive one consists in shutting off the sensor node completely whenever the information content of its next few observations is likely to be small.

Sensor nodes can take advantage of past observations and a priori knowledge about the stochastic processes they are monitoring to save energy. A small hit in performance can result in considerable energy savings for a decentralized detection system. For example, a minimal increase in expected detection delay can more than double the expected lifetime of the sensor node. This result provides support for control policies in which wireless sensor nodes enter long sleep intervals whenever the information content of the next few observations is likely to be small. Conceptually, the sensor node uses a priori knowledge about the process it is monitoring together with its current and past observations to reduce energy consumption. When the event of interest becomes very unlikely, sensor nodes can afford to go to sleep for an extended period of time, thus saving energy. On the other hand, when in a critical situation, sensor nodes must stay awake.

EXTENSIONS AND GENERALIZATIONS

The detection problems described thus far are static problems in which the sensors receive either a single observation or a single block of observations and a binary decision needs to be made at the fusion center. Many extensions and generalizations of this formulation are possible. In the dynamic setting, each sensor receives a sequence of successive observations and the detection system has the option to stop at any time and make a final decision, or to continue taking observations. The simplest problem in this setting is that of decentralized binary sequential detection. A decentralized version of binary sequential detection, where sensors make final decisions at different stopping times is studied by Teneketzis and Ho [25]. A more general formulation of the fusion problem was introduced by Hashemi and Rhodes [26], and a complete solution to this problem was given by Veeravalli et al. [27].

A different binary sequential decision-making problem that first arose in quality control applications is the change detection problem. Here the distribution of the observations changes abruptly at some unknown time, and the goal is to detect the change as rapidly as possible after its occurrence, subject to constraints on the false alarm probability. A decentralized formulation of the change detection problem is considered by Crow and Schwartz [28] and Teneketzis and Varaiya [29] with the sensors implementing individual change detection procedures. A general formulation of decentralized change detection with a fusion center making the final decision about the change is given by Veeravalli [30].

The design of optimal decision rules for decentralized detection problems is based on the assumption that the probability distributions of the sensor observations are known. In many applications, however, the distributions of the sensor observations are only specified as belonging to classes which are referred to as *uncertainty classes*. The problem here is to design decision rules that are robust with respect to uncertainties in the distributions. A common approach for such a design is the minimax approach where the goal is to minimize the worst-case performance over the uncertainty classes. Extensions of the minimax robust detection problem to the decentralized setting are possible. Alternatives to robust detection when partial information is available about the distributions, include composite testing based on generalized likelihood ratios, locally optimal testing for weak signals, and nonparametric detection.

DISCUSSION AND CONCLUDING REMARKS

Detection problems provide a productive starting point for the study of more general statistical inference problems in sensor networks. In this article we reviewed the classical framework for decentralized detection and argued that, while this framework provides a useful basis for developing a theory for detection in sensor networks, it has serious limitations. The classical framework does not adequately take into account important features of sensor technology and of the communication link between the sensors and the fusion center. We discussed an alternative framework for detection in sensor networks that has emerged over the last few years. Several design and optimization strategies may be gleaned from this new framework.

The jointly optimum solution for the sensor mappings and fusion rule is difficult to obtain, complicated, and it does not scale well with the number of sensors. Attention should be focused on good (suboptimum) solutions that are robust and scalable. The asymptotic regime where the resource budget and the number of sensors become large leads to such scalable, tractable solutions. A performance metric based on error exponents results in a decoupling of the optimization problem across sensors, where transmission mappings are selected according to a local criterion. The number of sensors and the sensor density should be considered system design parameters that needs to be optimized before deployment. This is particularly important when the observations are conditionally dependent. Finally, the modes of operation of a sensor (censoring and sleeping) should be fully exploited to minimize resource consumption.

While much progress has been made towards the understanding of detection problem in sensor networks using the emerging framework described in this article, many interesting questions remain. How do we obtain observation statistics? How do we design adaptive and robust strategies that work even when such statistics are incomplete or partially known? What is the role of error-control coding applied to the sensor outputs? Is it better to use additional bits to protect sensor outputs or to transmit more information about the observations? What is the right architecture for the network in the context of detection applications, decentralized with fusion or distributed? How much do we gain by allowing the sensors to communicate with one another in the fusion configuration?

AUTHORS

Jean-François Chamberland (chmbrlnd@tamu.edu) received the Ph.D. degree in 2004 from the University of Illinois at Urbana-Champaign, the M.S. degree in 2000 from Cornell University, Ithaca, NY, and the B.Eng. degree in 1998 from McGill University, Montreal, Canada, all in electrical engineering. He joined Texas A&M University in 2004, where he is currently an Assistant Professor in the department of Electrical and Computer Engineering. His research interests include communication systems, detection and estimation theory, and statistical signal processing. He is a Member of the IEEE.

Venugopal V. Veeravalli (vvv@uiuc.edu) received the the Ph.D. degree in 1992 from the University of Illinois at Urbana-Champaign, the M.S. degree in 1987 from Carnegie-Mellon University, Pittsburgh, PA, and the B.Tech. degree in 1985 from the Indian Institute of Technology, Bombay, (Silver Medal Honors), all in electrical engineering. He joined the University of Illinois at Urbana-Champaign in 2000, where he is currently a professor in the department of Electrical and Computer Engineering and a research professor in the Coordinated Science Laboratory. He served as a program director for communications research at the U.S. National Science Foundation in Arlington, VA from 2003–2005. He was an assistant professor at Cornell University, Ithaca, NY from 1996 to 2000. His research interests include detection and estimation theory, information theory, wireless communications, and sensor networks. He is a Fellow of the IEEE and currently on the Board of Governors of the IEEE Information Theory Society. He was an associate editor for "Detection and Estimation" for *IEEE Transactions on Information Theory* from 2000–2003 and an associate editor for *IEEE Transactions on Wireless Communications* from 1999–2000. He received the IEEE Browder J. Thompson Best Paper Award in 1996, the National Science Foundation CAREER Award in 1998, and the Presidential Early Career Award for Scientists and Engineers (PECASE) in 1999.

REFERENCES

[1] R.R. Tenney and N.R. Sandell, "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 17, no. 4, pp. 501–510, 1981.

[2] J.N. Tsitsiklis, "Decentralized detection," *Advs. Statistical Signal Processing*, vol. 2, pp. 297–344, 1993.

[3] R. Viswanathan and P.K. Varshney, "Distributed detection with multiple sensors: Part I—fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.

[4] R.S. Blum, S.A. Kassam, and H.V. Poor, "Distributed detection with multiple sensors: Part II—advanced topics," *Proc. IEEE*, vol. 85, no. 1, pp. 64–79, Jan. 1997.

[5] Z. Chair and P.K. Varshney, "Distributed Bayesian hypothesis testing with distributed data fusion," *IEEE Trans. Syst., Man, Cybern.*, vol. 18, no. 5, pp. 695–699, Sept. 1988.

[6] S.A. Aldosari and J.M.F. Moura, "Saddlepoint approximation for sensor network optimization," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing*, Mar. 2005, vol. 4., pp. 741–744.

[7] X. Nguyen, M.J. Wainwright, and M.I. Jordan, "Nonparametric decentralized detection using kernel methods," *IEEE Trans. Signal Processing*, vol. 53, no. 11, pp. 4053–4066, Nov. 2005.

[8] J.N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Math. Control, Signals, Syst.*, vol. 1, no. 2, pp. 167–182, 1988.

[9] P. Swaszek and P. Willett, "Parley as an approach to distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 1, pp. 447–457, Jan. 1995.

[10] A. D'Costa, V. Ramachandran, and A.M. Sayeed, "Distributed classification of Gaussian space-time sources in wireless sensor networks," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1026–1036, Aug. 2004.

[11] J.-F. Chamberland and V.V. Veeravalli, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1007–1015, Aug. 2004.

[12] J.-F. Chamberland and V.V. Veeravalli,, "Decentralized detection in sensor networks," *IEEE Trans. Signal Processing*, vol. 51, no. 2, pp. 407–416, Feb. 2003.

[13] B. Chen and P.K. Willett, "On the optimality of the likelihood-ratio test for local sensor decision rules in the presence of nonideal channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 693–699, Feb. 2005.

[14] P. Willett, P.F. Swaszek, and R.S. Blum, "The good, bad and ugly: Distributed detection of a known signal in dependent Gaussian noise," *IEEE Trans. Signal Processing*, vol. 48, no. 12, pp. 3266–3279, Dec. 2000.

[15] J.-F. Chamberland and V.V. Veeravalli, "How dense should a sensor network be for detection with correlated observations?" *IEEE Trans. Inform. Theory*, vol. 52, no. 11, pp. 5099–5106, Nov. 2006.

[16] B. Chen, R. Jiang, T. Kasetkasem, and P.K. Varshney, "Channel aware decision fusion in wireless sensor networks," *IEEE Trans. Signal Processing*, vol. 52, no. 12, pp. 3454–3458, Dec. 2004.

[17] R. Niu, B. Chen, and P.K. Varshney, "Fusion of decisions transmitted over Rayleigh fading channels in wireless sensor networks," *IEEE Trans. Signal Processing*, vol. 54, no. 3, pp. 1018–1027, Mar. 2006.

[18] G. Mergen and L. Tong, "Sensor-fusion center communication over multiaccess fading channels," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing*, Mar. 2005, vol. 4., pp. 841–844.

[19] K. Liu and A.M. Sayeed, "Asymptotically optimal decentralized type-based detection in wireless sensor networks," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing*, May 2004, vol. 3, pp. 873–876.

[20] G. Mergen and L. Tong, "Type based estimation over multiaccess channels," *IEEE Trans. Signal Processing*, vol. 54, no. 2, pp. 613–626, Feb. 2006.

[21] Y.-W. Hong and A. Scaglione, "A scalable synchronization protocol for large scale sensor networks and its applications," *IEEE J. Select. Areas Commun.*, vol. 23, no. 5, pp. 1085–1099, May 2005.

[22] V. Borkar and P. Varaiya, "Asymptotic agreement in distributed estimation," *IEEE Trans. Automat. Contr.*, vol. 27, no. 3, pp. 650–655, June 1982.

[23] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: a low-communication-rate scheme for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 554–568, Apr. 1996.

[24] S. Appadwedula, V.V. Veeravalli, and D.L. Jones, "Energy-efficient detection in sensor networks," *IEEE J. Select. Areas Commun.*, vol. 23, no. 4, pp. 693–702, June 2005.

[25] D. Teneketzis and Y.C. Ho, "The decentralized Wald problem," *Inform. Computation*, vol. 73, no. 1, pp. 23–44, 1987.

[26] H.R. Hashemi and I.B. Rhodes, "Decentralized sequential detection," *IEEE Trans. Inform. Theory*, vol. 35, no. 3, pp. 509–520, May 1989.

[27] V.V. Veeravalli, T. Başar, and H.V. Poor, "Minimax robust decentralized detection," *IEEE Trans. Inform. Theory*, vol. 40, no. 1, pp. 35–40, Jan. 1994.

[28] R.W. Crow and S.C. Schwartz, "Quickest detection for sequential decentralized decision systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 1, pp. 267–283, Jan. 1996.

[29] D. Teneketzis and P. Varaiya, "The decentralized quickest detection problem," *IEEE Trans. Automat. Contr.*, vol. 29, no. 7, pp. 641–644, July 1984.

[30] V.V. Veeravalli, "Decentralized quickest change detection," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1657–1665, May 2001.