OPTIMAL DISTRIBUTED DETECTION AND ESTIMATION IN STATIC AND MOBILE WIRELESS SENSOR NETWORKS

A Dissertation
Presented to
The Academic Faculty

by

Xusheng Sun

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
August 2012

OPTIMAL DISTRIBUTED DETECTION AND ESTIMATION IN STATIC AND MOBILE WIRELESS SENSOR NETWORKS

Approved by:

Professor Edward J. Coyle, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Douglas M. Blough
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Mary Ann Ingram
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Fumin Zhang
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Yajun Mei
School of Industrial and Systems Engineering
Georgia Institute of Technology

Date Approved: 20 June 2012

To my parents

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Prof. Edward J. Coyle, for his invaluable advice and persistent support during my Ph.D. study. Without his guidance, I would not have been able to successfully complete my dissertation. Moreover, his guidance in both my research and life has pretty much reshaped my way of thinking and working in the past several years and certainly for the future.

I would also like to thank: Prof. Douglas M. Blough and Prof. Mary Ann Ingram for serving on my reading committee; Prof. Saibal Mukhopadhyay for serving on my proposal committee; and, Prof. Fumin Zhang and Prof. Yajun Mei for serving on my defense committee. Their advice and comments have helped me to improve my dissertation.

I want to thank the many other faculty, staff, and students at Georgia Tech who have given me a lot of inspiration and help during my study here, especially my lab mates: Ghaith M. Matalkah, Seskan Laitrakun and Ximeng Mao. I would also like to extend special thanks to Dr. Junlin Li, who was the earliest one to introduce to me the topic of distributed signal processing in wireless sensor networks and some interesting research by Prof. Coyle and his previous students.

Finally I would like to express my deepest gratitude to my parents for their unconditional love and support. It is to them that I dedicate my dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This dissertation develops optimal algorithms for distributed detection and estimation in static and mobile ad hoc networks. In distributed detection or estimation scenarios in clustered wireless sensor networks, sensor motes observe their local environment, make decisions or quantize these observations into local estimates of finite length, and send/relay them to a Cluster-Head (CH). For event detection tasks that are subject to both measurement errors and communication errors, we develop an algorithm that combines a Maximum a Posteriori (MAP) approach for local and global decisions with low-complexity channel codes and processing algorithms. For event estimation tasks that are subject to measurement errors, quantization errors and communication errors, we develop an algorithm that uses dithered quantization and channel compensation to ensure that each mote's local estimate received by the CH is unbiased and then lets the CH fuse these estimates into a global one using a Best Linear Unbiased Estimator (BLUE). We then determine both the minimum energy required for the network to produce an estimate with a prescribed error variance and show how this energy must be allocated amongst the motes in the network.

In mobile wireless sensor networks, the mobility model governing each node will affect the detection accuracy at the CH and the energy consumption to achieve this level of accuracy. Correlated Random Walks (CRWs) have been proposed as mobility models that accounts for time dependency, geographical restrictions and nonzero drift. Hence, the solution to the continuous-time, 1-D, finite state space CRW is provided and its statistical behavior is studied both analytically and numerically. The impact of the motion of sensor on the network's performance and the strategies to improve the performance are also studied.

CHAPTER I

INTRODUCTION

The market for wireless sensor networks has yet to reach the multi-billion dollar levels predicted ten years ago. Sales of ZigBee chip sets, for example, are at most a few million per year [73]. One reason for this smaller than expected market is that most sensor networks considered so far have been dedicated to a single task, such as environmental monitoring or energy management, and the resulting benefits do not justify the cost of installing and maintaining the network. This will eventually change as the cost of sensors and motes decline, batteries and energy scavenging techniques improve, communication capacity and efficiency increase, and software tools improve.

One important possibility missing from the above scenario is the use of a sensor network for multiple tasks. The combined benefits from performing these multiple tasks might then be sufficient to justify the network. A situation like this has arisen in the eStadium project at Purdue and Georgia Tech [109, 110, 10]. Its goal is to gather game-, venue- and fan-generated multimedia content and deliver it to fans and stadium personnel via their wireless devices. Some of this content is gathered via a wireless sensor network deployed in the stadium – Fig. 1 and Fig. 2 show a Cluster-Head (CH) and mote installed in the stadium at Purdue. The tasks this network performs fall into two categories: safety/security and enhancing fans' enjoyment of the game. The first is essential but does not generate revenue; the second is non-essential but is a potentially significant source of revenue. Their support by a dual-purpose network should achieve economic sustainability for both.

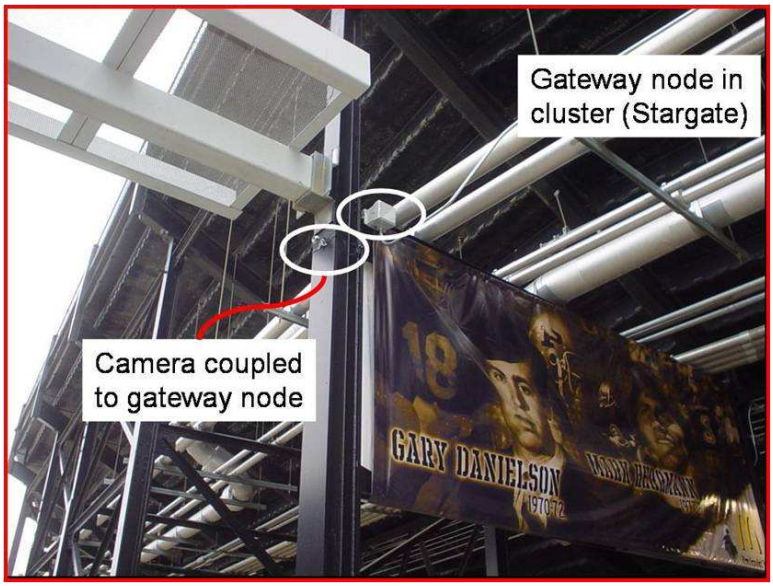


Figure 1: A Cluster-Head (CH) of the eStadium sensor network in Ross Ade stadium at Purdue. It is a Crossbow Stargate powered via a Power-over-Ethernet (PoE) enabled switch. Its USB web-cam captures images of the concourse area of the stadium and its attached sensor mote collects data and supports communication with other, untethered motes.

Figure 2: A Mica2 wireless sensor mote inside an enclosure that is strapped to a steel girder of the stadium.

This dissertation focuses on the event detection and estimation tasks the eStadium sensor network will perform. These include detecting or estimating: the occurrence/loudness of cheering or booing, which is used to tag the video clips of plays of greatest interest to fans; the presence/coverage of WiFi, Bluetooth and ZigBee networks that interfere with the stadium's wireless networks; and, in the future, the presence/concentration of chemicals, such as smoke, in or near the stadium. If all of these detection and estimation tasks are to be supported by the same network, it is essential that each one is very lightweight.

The support of two or more tasks or applications by a single sensor network means that the requirements of each one must be very well understood to ensure they can coexist in such a resource constrained environment. In particular, the minimum processing and communication resources required for the application to perform well must be determined. The patterns of use of these resources by each optimized application must then be determined so that they can be optimally scheduled/managed. We have thus begun an effort to characterize and optimize the use of resources by classes of applications. We began by characterizing applications based on distributed detection [90] and then move to applications based on distributed estimation, which is generally more complicated. See, for example, [94, 18, 93] for excellent tutorials on distributed detection, and [95, 33, 49, 68] for fundamental results on distributed estimation.

The systems-level work reported in this dissertation for distributed detection/estimation in mobile wireless sensor networks is also motivated by a security scenario that has arisen in the context of the eStadium project, which develops multi-media web applications, wireless communication networks, and wireless sensor networks for football games and other large-scale events. One security scenario of great interest occurs in the hour immediately before or immediately after a large event. At these times, large numbers of fans, stadium personnel, and concession staff are typically walking

toward or away from the stadium. This is a time when they are very vulnerable to surreptitious exposure to hazardous chemical or biological agents.

Suppose that the smartphones that the fans are carrying, or possibly the tickets that they have bought, have sensors embedded in them that can detect and/or estimate the concentration of these agents. These sensors would necessarily be small and inexpensive and might thus have a potentially high false-alarm rate and/or low rate of correct detection. It is thus important that as many as possible of their individual decisions and estimates be fused with each other to ensure they are correct and accurate. They must therefore be able to communicate with each other or with a designated Cluster Head (CH) to accomplish this goal. These mobile, battery-powered, wireless sensors can thus be modeled as a mobile ad-hoc network that is supporting applications that perform distributed detection or estimation. Hence it is interesting and necessary in this scenario to model and analyze the motion of sensors and its effects on the performance and cost during their execution of detection/estimation tasks.

1.1 Overview of Wireless Sensor Networks

Advances in MEMS and wireless communications technologies have enabled the deployment of wireless sensors for applications involving environmental sensing, battlefield surveillance, health monitoring, process control, and so on. Each wireless sensor node is equipped with an array of transducers that can convert a variety of physical phenomenon, such as temperature, humidity, light and sound levels, into electric signals. In a typical scenario, these motes are either deterministically or randomly deployed in a geographical area and they automatically self-organize into a network. Then they perform their sensing tasks and transmit the data gathered to the cluster head, which fuses the data collected from all motes in the network. The most important constraints for the design of wireless sensor networks are the limits on

their power consumption and transmission rate. Sensor motes are usually powered by batteries and their transmission rate is about 250kbps. Hence, there have been many research efforts whose goals are to save energy and bandwidth in the design of wireless sensor networks. For example, see [103, 29, 27, 74, 69, 28] for MAC layer designs and [35, 39, 1, 5] for routing algorithms for wireless sensor networks.

The proliferation of mobile computing and communication devices such as smartphones, laptops and handheld digital devices has been driving the evolution of mobile ad hoc sensor networks. Mobile ad hoc networks represent distributed systems comprised of wireless mobile nodes that can freely and dynamically self-organize into arbitrary and temporary ad hoc network topologies without assuming any pre-existing communication infrastructure. See [31, 26] for general tutorials on mobile ad hoc networks, [78] for a survey that focuses on routing protocol design for the organization of mobile ad hoc networks, and [105] for a typical application that generates a signal processing task in mobile sensor networks.

This dissertation will focus on signal processing aspects in both static and mobile wireless sensor networks.

1.2 Overview of Distributed Signal Processing

The idea of distributed signal processing is to get an acceptable signal processing result with a number of devices with limited capabilities to acquire, process, and transmit signals. To achieve the objective, these devices have to execute the task in a distributed and collaborative manner. Distributed signal processing in wireless sensor networks is a typical example because the sensing accuracy, power consumption, processing ability, and communication bandwidth of each mote is typically very limited. Hence the target of the system is to obtain a much better result based on a number of coarse results from each mote through data fusion.

As to sensor motes, communication is the most energy consuming operation; when

compared with sensing and computation tasks [6]. Thus processing at each sensor and collaboration within the entire network are the most important factors that affect the performance of the final result. Concretely speaking, each mote makes observations of the event the network is interested in, such as light, fire and sound, when it is scheduled to be active. Then each mote will perform local signal processing individually such as filtering, quantization, and encoding. In the next step, sensors transmit their locally processed result to each other or a fusion center. Finally the fusion center or one designated sensor will extract the useful information from the data it received and fuse them together to generate the final result. Distributed signal processing includes several different kinds of tasks, such as distributed data compression [70, 101], distributed estimation [9, 12, 13, 11, 36, 46, 51, 52, 53, 63, 96, 97, 100, 95, 33, 60, 99, 62, 76, 75, 98], distributed detection [94, 18, 93, 92, 91, 22, 59, 61, 24, 23, 66, 55, 40, 30, 57, 8, 64, 67, 102, 56], localization [25, 43, 49], and tracking [34, 50, 58, 107, 108].

Distributed detection and estimation are the most fundamental tasks among distributed signal processing tasks. They are duals of each other, but an estimation problem is typically more complicated than a detection problem. Therefore, this dissertation will focus on distributed estimation and detection, with detection problem being treated first.

The rest of the dissertation is organized as follows. In Chapter II, an algorithm of low complexity for distributed detection in wireless sensor networks is developed to achieve the objective of energy/bandwidth efficiency. To ensure that a multi-hop cluster of battery-powered, wireless sensor motes can complete all of its tasks, each task must minimize its use of communication and processing resources. For event detection tasks that are subject to both measurement errors by sensors and communication errors in the wireless channel, this implies that: (i) the Cluster-Head (CH) must optimally fuse the decisions received from its cluster in order to reduce

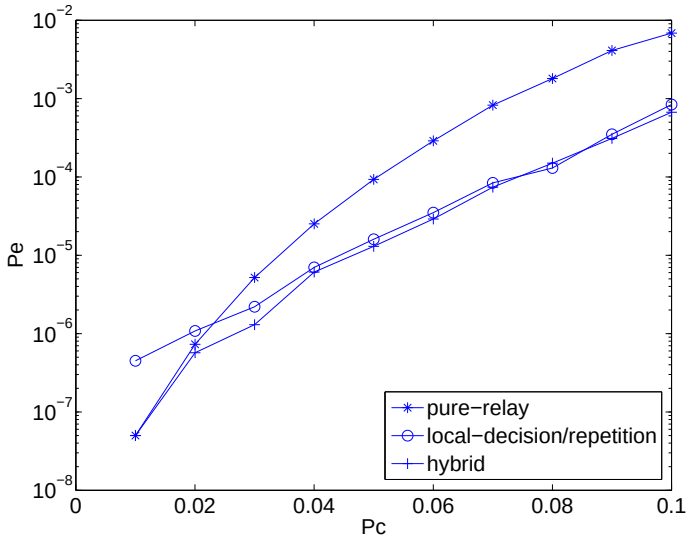


Figure 3: Comparisons via simulation of the pure-relay, local-decision/repetition, and hybrid strategies. The hybrid decision strategy lowers the probability of error by using local decisions when needed.

the effect of measurement errors; (ii) the CH and all motes that relay other motes' decisions must adopt low-complexity processing and coding algorithms that minimize the effects of communication errors.

This chapter combines a Maximum a Posteriori (MAP) approach for local and global decisions in multi-hop sensor networks with low-complexity repetition codes and processing algorithms. It is shown by analysis and confirmed by simulation that there exists an an odd integer M and an integer K_M such the decision error probability at the CH is reduced when: (1) nodes in rings $k \leq K_M$ hops from the CH directly relay their decisions to the CH; (2) nodes in rings $k > K_M$ locally fuse groups of M decisions and then use a repetition code to forward these fused decisions to the CH; and (3) K_M is a nondecreasing function of M. This algorithm – and hybrid, hierarchical, and compression approaches based on it – enable tradeoffs amongst the probability of error, energy usage, compression ratio, complexity, and time to decision. In Fig. 3, we can see that hybrid scheme outperforms both the pure relay scheme and local decision/repetition schemes.

In Chapter III, a combination of dithered quantization, channel compensation and energy allocation is proposed for distributed estimation in wireless sensor networks is proposed. In clustered networks of wireless sensors, each sensor collects noisy observations of the environment, quantizes these observations into a local estimate of finite length, and forwards them through one or more noisy wireless channels to the Cluster Head (CH). The measurement noise is assumed to be zero-mean and have finite variance and each wireless hop is modeled as a Binary Symmetric Channel (BSC) with a known crossover probability. A novel scheme is proposed that uses dithered quantization and channel compensation to ensure that each sensor's local estimate received by the CH is unbiased. The CH fuses these unbiased local estimates into a global one using a Best Linear Unbiased Estimator (BLUE). Analytical and simulation results show that the proposed scheme can achieve much smaller mean square error (MSE) than two other common schemes while using the same amount of energy. The sensitivity of the proposed scheme to errors in estimates of the crossover probability of the BSC channel is studied by both analysis and simulation. We then determine both the minimum energy required for the network to produce an estimate with a prescribed error variance and how this energy must be allocated amongst the sensors in the multi-hop network. In Tab. 1, we can see the optimal energy allocation scheme among the entire network to achieve the threshold on MSE of the final estimate with the minimum amount of energy.

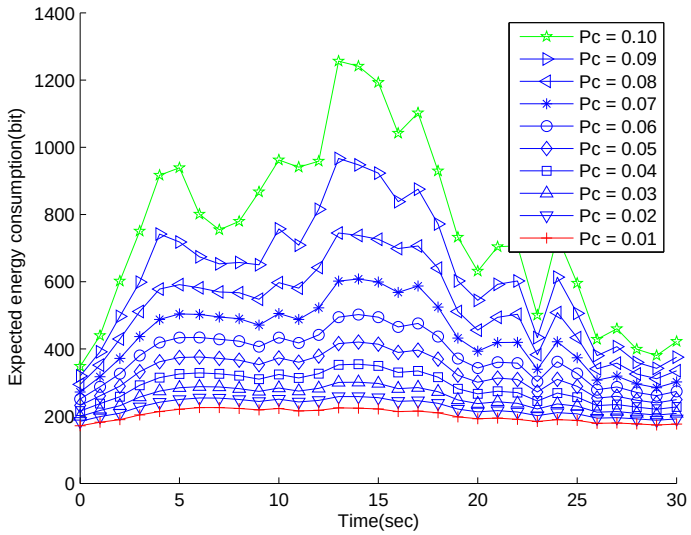
The effects of motion on distributed detection/estimation are studied in Chapter IV. A large set of mobile wireless sensors observe their environment as they move about. We consider the subset of these sensors that each made observations about a brief, localized event at the time when they were near that location. As they continue to move, one of them eventually finishes processing its observations and makes a decision or an estimate and determines that must be reported if other sensors confirm its results. This sensor then assumes the role of a Cluster-Head (CH) and requests

Table 1: Comparison of the energy allocation schemes under different crossover probabilities of the BSC channel. Bits are allocated first to the sensors closest to the CH, with sensors in the third ring contributing more bits as the channel crossover probability p_c increases. Recall that p_{abc} is the probability that node b in ring a transmits $2^c - 1$ bits.

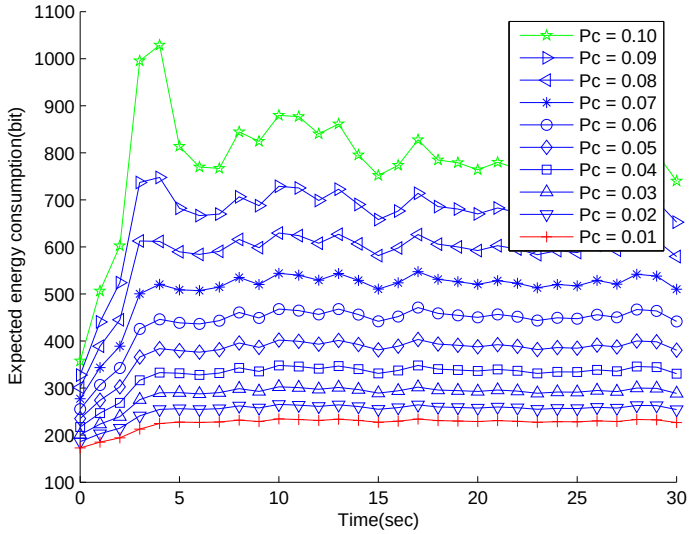
p_c	ring 1	ring 2	ring 3
0.02	$p_{1i3} = 1$	$p_{2i3} = 1$	$p_{3i2} = 0.2433$
0.04	$p_{1i4} = 1$	$p_{2i3} = 1$	$p_{3i2} = 0.5721$
0.06	$p_{1i4} = 1$	$p_{2i3} = 1$	$p_{3i2} = 0.6477, p_{3i3} = 0.3523$
0.08	$p_{1i4} = 1$	$p_{2i4} = 1$	$p_{3i2} = 0.6838, p_{3i3} = 0.3162$
0.10	$p_{1i5} = 1$	$p_{2i4} = 1$	$p_{3i3} = 0.9789, p_{3i4} = 0.0211$
0.12	$p_{1i5} = 1$	$p_{2i5} = 1$	$p_{3i2} = 0.1291, p_{3i3} = 0.8709$
0.14	$p_{1i5} = 1$	$p_{2i5} = 1$	$p_{3i4} = 0.9346, p_{3i5} = 0.0654$

that all other sensors that collected observations at that time/location reply to it with their decisions/estimates. The motion of the sensors since the observation time determines how many wireless hops their decision/estimates must cross to reach the CH. We analyze the effect of this motion in the 1D case by modeling each sensor's motion as a Correlated Random Walk (CRW), which can account for realistic transient behavior, geographical restrictions, and nonzero drift. We also account for observation errors and errors in each hop in the wireless channel. Quantities, such as the error probability of the final decision at the CH and the minimum energy required to collect the estimates from all relevant sensors to achieve a prescribed estimation accuracy, can then be determined as functions of time and the parameters of the CRW, the measurement noise and the channel noise. These results thus allow rapid calculation of the time-dependence of detection and estimation algorithms that are being executed in realistic mobile sensor networks. In Fig. 4, we can see clearly the tradeoff between delay and energy of estimation in mobile sensor networks.

Chapter V concludes the dissertation and discusses future research topics. The results in this dissertation have been published in several journal papers [90, 87, 85] and a number of conference papers [80, 88, 83, 89, 82, 81, 84, 86].



(a) All sensors move according to asymmetric CRWs with p_1=0.8, p_2=0.2



(b) All sensors move according to symmetric CRWs with p_1=0.9, p_2=0.9

Figure 4: Minimum energy, as a function of time, required in a mobile sensor network to achieve a specified MSE for different values of the crossover probabilities of the BSC channel. The MSE requirement is 0.015. The The network has 100 nodes that each start at position zero and move independently according to asymmetric CRWs on [0,10] with a mean rate of $\lambda = 1$ unit/sec. The transmission radius of each sensor is $r = 1$; the signal and measurement noise are both uniformly distributed on $[-1, 1]$. This scenario is a reasonable model of a group of people/vehicles who start moving after a traffic light turns and proceed to another traffic light where they again bunch up and wait. Some of them stop for some time, or turn-around occasionally.

CHAPTER II

DISTRIBUTED DETECTION IN WIRELESS SENSOR NETWORKS

2.1 Research Motivation

This chapter will focus on distributed detection in wireless sensor networks. For each event of interest, a sensor mote periodically produces a single bit to indicate whether the event has occurred or not. Each decision bit is based on noise-corrupted measurements of the environment in the stadium. These individual decisions are then transmitted/relayed, either individually or after being gathered into packets, over noisy wireless channels to the CH. When these decisions are relayed, they may be processed by the relaying motes before being forwarded. The CH gathers the decisions it receives from all of its motes and fuses them into a cluster-wide decision about the event.

Fig. 5 is a schematic of the 1-D sensor net deployed along the edge of the stadium [109]. The motes are organized into multi-hop clusters, either directly or by a fast, energy-minimizing algorithm, such as the one in [15]. The CH's make cluster-wide decisions that they forward to a server that produces a stadium-wide status map for each type of event being monitored. The diameter of each cluster is less than the spatial-correlation length of the observed phenomenon. The motes in a cluster are thus observing and making decisions about a phenomenon that we can assume to be spatially-invariant.

The use of many inexpensive sensor motes in these clusters provides resilience against failures, increases the spatial sampling rate, and yields decision error probabilities at the CH that decrease exponentially as the number of motes increases. Each

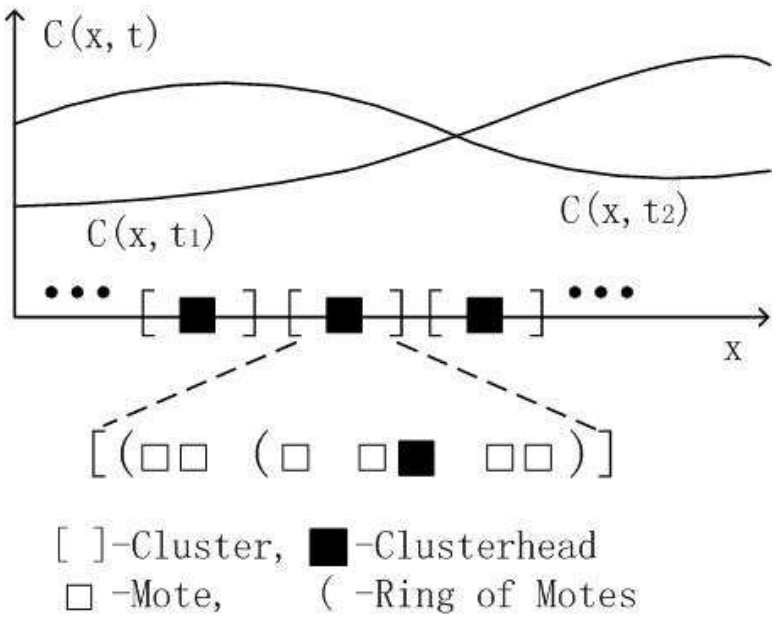


Figure 5: Layout of a 1-D sensor net that may be monitoring $C(x,t)$, the space- and time-varying concentration of an airborne contaminant, such as smoke, along the edge of the stadium. The expanded view of a two-hop cluster shows its CH and motes.

mote makes decisions/estimates about a wide variety of phenomena [110] and can thus combine the bits from all of its tasks into a single packet to transmit. The relaying motes can likewise add relayed decisions and estimates to packets they generate. This minimizes overhead, thus decreasing the energy expended.

A more general two-hop cluster is shown in Fig. 28. The 7 motes in the first ring are one hop from the CH; the 15 motes in the second ring are two hops away. Three motes in a *sector* of ring 2 are shown forwarding their decisions to a *relay* mote in ring 1. Every mote's observations are affected by measurement noise and all communications throughout the network are affected by channel noise, fading, and transceiver errors. Because the cluster is multi-hop, decisions forwarded by motes in the outer rings suffer repeated exposure to these sources of communication error. Understanding the combined effects of the communication and measurement errors and developing lightweight algorithms to mitigate these effects are the goals of this chapter.

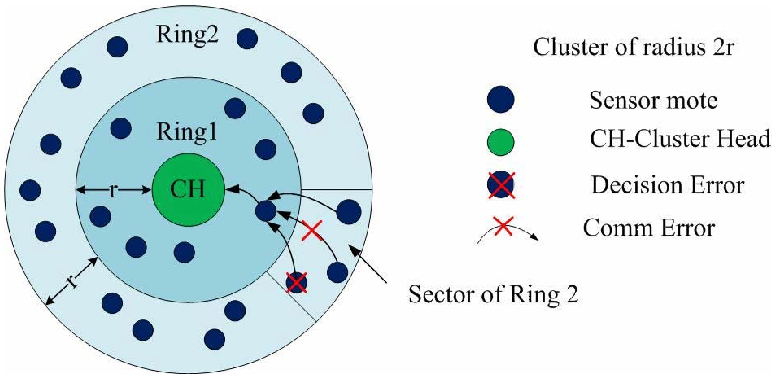


Figure 6: A two-hop, 2-D cluster. Each mote's binary decision may be incorrect because of measurement errors. Decisions transmitted over wireless channels may be received incorrectly because of communication errors.

The approach we use is a combination of large deviation techniques, direct calculation, and simulation. Large deviation techniques, which were used to analyze errors in a single-hop cluster in [22], are used to obtain the error exponent and upper bound for the decision error probability at the CH for complex, multi-hop clusters. This technique accurately characterizes the rate of decay of the tail of the decision error probability for large numbers of motes but is often not precise for small or moderate numbers of motes. For small numbers of motes, we thus use direct calculation of the error probability. For moderate numbers of motes, the computational complexity is too great, so simulation is used. The combination of these techniques enables us to completely characterize the performance of the new strategies we propose.

This chapter is organized as follows: Section 2.2 reviews prior work on the distributed detection problem in single- and multi-hop sensor networks. Section 2.3 describes the system model and summarizes results for the Maximum a Posteriori (MAP) detector in [91], which is adopted in this chapter. In Section 2.4, large deviation techniques are used to bound the decision error probability of the MAP detector for the multi-hop case. Section 2.5 defines new strategies that combine decision fusion in sectors of the cluster with a repetition code for forwarding these decisions. They are shown to reduce the probability of a decision error at the CH. In Section

2.6, the bounds from Section 2.2 are used to determine when a cluster's detection performance can be improved by combining local decision fusion and forwarding in the outer rings with simple forwarding of decisions in the inner rings. In Section 2.7, simulations are used to compare the performance of a pure-relay strategy, a local-decision/repetition strategy, a hybrid strategy, and some compression strategies that sacrifice detection performance to save energy. In Section 2.8, implementation issues are discussed. Finally, Summary is provided in Section 2.9.

2.2 Most Relevant Prior Results

Many papers consider the task of distributed detection in single-hop sensor networks. In [59], an optimal distributed detection strategy in a single-hop network was studied. In [61], a decentralized detection problem under bandwidth constrained communication was investigated. The noise at different sensors was assumed to be independent, but the statistics of the noise were assumed to be unknown to the CH, so it treats all received detection results equally. It is shown in [24] how the performance of detection algorithms is improved by knowledge of the channel. The limits of detection performance in a one-hop sensor cluster with non-ideal channels were determined in [23].

In [66], the performance of distributed detection in a random sensor field is analyzed. [55] and [40] discuss the distributed detection problem for Gaussian signals under communication constraints. In [30], the authors consider the detection and localization problem of material releases with sparse sensor configurations. In [57], the sensors adopt robust binary quantizers for distributed detection. Censoring and sequential tests have also been adopted in distributed detection to achieve the same detection accuracy with less energy [8], [64], and they have been combined to further save energy [67], [102].

There are many papers on cooperative communication in sensor networks; see, for

example, [45]. They assume that relay nodes will use a collaborative amplify-and-forward strategy to help ensure correct reception of a node's packet by the destination. This approach is thus a blend of one- and two-hop networks.

The paper most closely related to this one is [56] because it considers multi-hop fusion schemes. The schemes differ in the form of the transmitted messages, the fusion rules and the communication structure. The Multi-hop Forwarding (MF) algorithm is similar to the pure-relay scheme that we define and use as a baseline for comparison. The Multi-hop Forwarding algorithm (MF) and Multihop Histogram Fusion (HF) algorithm use encoding schemes at the relay motes to reduce the number of bits that must be forwarded to the CH, thus saving energy. The effect of channel errors on these encoding algorithms is not considered; an error in the most-significant bit can have a major impact on the probability of error. One solution to this problem is dithered quantization and channel compensation [88]. A lower complexity approach is provided in this paper.

In [91], a MAP approach was developed for the distributed detection problem in a multi-hop cluster in a sensor network. Based on the results of the proposed MAP detector, this chapter adopts the same system model and focuses on the optimal relay strategy to improve the system performance without increasing of energy consumption. The system assumptions and preliminary results are reviewed in detail in the next section.

2.3 System Assumptions

This chapter assumes a multi-hop sensor network in which each wireless hop is modeled as a Binary Symmetric Channel (BSC). The BSC cross-over probability captures the effect of channel errors on each individual decision bit and is easier to estimate than the full characteristics of the channel. The assumptions of symmetry and the same cross-over probability for each mote-to-mote channel are easily relaxed – they

are used only to simplify the analysis so that important trends can be discerned. This chapter also assumes that the error probabilities of the individual received detection results are learned over time by the CH. Hence, the CH is assumed to know the optimal weights for the weighted median in the MAP detector [91] it uses to fuse the received detection results.

The system assumptions used in this chapter are summarized in the following:

1. The *deployment* assumption: The cluster has K rings and N_k motes in the kth ring.

2. The *task* assumption: Each mote makes a decision between two hypotheses, $s_0 = 0$ and $s_1 = 1$, where "1" denotes that an event has occurred and "0" that it has not.

3. The *measurement* assumption: The detection results by different motes are assumed to be i.i.d. Bernoulli random variables, each with a detection error probability $p_m < 1/2$. The noise processes in different wireless channels are assumed to be independent and white.

4. The *channel* assumption: Each hop in the network is modeled as a BSC with cross-over probability $p_c < 1/2$.

5. The *relay* assumption: The decisions made by motes in the outer rings are relayed to the CH by the motes in inner rings.

Let the error probability of the detection results received by the CH from the kth ring be $p_{e,k}$. Then, for example, $p_{e,1} = p_c(1-p_m) + p_m(1-p_c)$. Denote the detection result received by the CH from the ith mote in the kth ring by $r_{k,i}$ and arrange the detection results in the same ring in a vector $\bar{r}_k = (r_{k,1}, r_{k,2}, \ldots, r_{k,N_k}), k = 1, 2, \ldots, K$. Let E denote the event that a decision error happens at the CH.

In a *one*-hop cluster, suppose the correct decision should be s. Assume the prior probability $p(s = s_0) = p < 1/2$ and find a real number χ such that $\ln((1-p)/p) = \chi \ln((1-p_{e,1})/p_{e,1})$. Define $W = \lfloor N_1/2 + \chi/2 \rfloor$. The MAP-based decision bit at the CH is $\hat{r} = (\bar{r}_1)_W = W$'th order statistic of $(r_1, r_2, \ldots, r_{N_1})$. The decision error probability at the CH is [91]:

$$\begin{aligned} P(E) = (1-p) & \left(\sum_{i=W}^{N_1} \binom{N_1}{i} (p_{e,1})^i (1-p_{e,1})^{N_1-i} \right) \\ & + p \left(\sum_{i=N_1-W}^{N_1} \binom{N_1}{i} (p_{e,1})^i (1-p_{e,1})^{N_1-i} \right) \end{aligned} \tag{1}$$

It is shown in [91] that, as an estimate of the true decision bit, this weighted order statistic is biased but asymptotically unbiased. When the prior probability is $p = 1/2$, it simplifies into a binary median filter.

Calculating the decision error probability in (1) is computationally complex when N_1 is large. The asymptotic behavior of the median filter in one-hop clusters has thus been studied using large deviation techniques in [22]:

$$\begin{aligned} \lim_{N_1 \to \infty} & -\frac{1}{N_1} P(E) = C(p_{e,1}) \\ & C(p_{e,1}) = -\ln(2) - \frac{1}{2} \ln(p_{e,1}(1-p_{e,1})) \end{aligned} \tag{2}$$

Now we review results in [91] for the *multi*-hop case. For a mote that is k hops away from the CH, the detection result received by the CH after being relayed over these hops has error probability:

$$p_{e,k} = \frac{1}{2} - \frac{1}{2}(1-2p_m)(1-2p_c)^k, k \geq 1 \tag{3}$$

We use the notation $W \diamondsuit x$, which means x should be duplicated W times. For simplicity, suppose the prior probabilities are $p(s = s_0) = p(s = s_1) = 1/2$.

Theorem 1: Define $\chi_k = \ln((1-p_{e,k})/p_{e,k})$, and assume these χ_k's can be scaled so that $\chi_1 : \chi_2 : \ldots : \chi_K = W_1 : W_2 : \ldots : W_K$, where the W_k's are positive integers with $gcd(W_1, W_2, \ldots, W_K) = 1$. The MAP-based decision bit is then given

by $\hat{r} = Median(W_1 \diamondsuit \bar{r}_1, W_2 \diamondsuit \bar{r}_2, \ldots, W_K \diamondsuit \bar{r}_K)$. The decision error probability at the CH is given by:

$$P(E) = \sum_{\sum W_k(2c_k - N_k) > 0} \prod_{k=1}^{K} \binom{N_k}{c_k} (p_{e,k})^{c_k} (1 - p_{e,k})^{N_k - c_k} \tag{4}$$

where c_k is the number of occurrences of $1 - s$ in the vector $\bar{r}_k, k = 1, 2, \ldots, K$.

■

Because the weighted median is the MAP detector, we assume – even for sector-level fusion algorithms – that the CH calculates this weighted median after receiving inputs from all motes in the cluster. To understand what happens in these more complex cases, the asymptotic behavior of the weighted median must be determined for the single- *and* multi-hop cases; otherwise, it is too difficult to determine the probability of a cluster-wide decision error at the CH.

2.4 Error Exponent and Error Bounds for Multi-Hop Networks

We now find the error exponent for the decision error probability for the multi-hop case.

Theorem 2: In multi-hop sensor networks in which $p_{e,k}$ is the error probability of the individual detection results received by the CH from nodes in ring k, the error probability of the MAP detector is upper bounded by:

$$\ln(P(E)) \leq \sum_{k=1}^{K} N_k [\ln(2) + \frac{1}{2} \ln(p_{e,k}(1 - p_{e,k}))] \tag{5}$$

Proof: In the context of the weighted median filter, the decision error probability at the CH is:

$$\begin{aligned} P(E) &= P(\sum_{k=1}^{K} W_k \sum_{i=1}^{N_k} r_{k,i} \geq \sum_{k=1}^{K} W_k N_k / 2) \\ &= P(\sum_{k=1}^{K} \sum_{i=1}^{N_k} W_k (r_{k,i} - 1/2) \geq 0) \end{aligned} \tag{6}$$

Since $P(Z \geq 0) \leq E[e^{Zt}]$, let $r'_{k,i} = r_{k,i} - 1/2$, and $E[r'_{k,i}] = p_{e,k} - 1/2 < 0$. We have:

$$\begin{aligned} P(E) &\leq E[e^{t\sum_{k=1}^{K}\sum_{i=1}^{N_k} W_k r'_{k,i}}] \\ &= E[\prod_{k=1}^{K}\prod_{i=1}^{N_k} e^{tW_k r'_{k,i}}] \\ &= \prod_{k=1}^{K}\prod_{i=1}^{N_k} E[e^{tW_k r'_{k,i}}] \end{aligned} \tag{7}$$

For each term in the product, $E[e^{tW_k r'_{k,i}}] = p_{e,k}e^{\frac{1}{2}tW_k} + (1 - p_{e,k})e^{-\frac{1}{2}tW_k}$. Setting $\frac{\partial E[e^{tW_k r'_{k,i}}]}{\partial t} = 0$, we find:

$$\begin{aligned} \frac{1}{2}W_k p_{e,k} e^{\frac{1}{2}tW_k} &= \frac{1}{2}W_k(1 - p_{e,k})e^{-\frac{1}{2}tW_k} \\ tW_k &= \ln((1 - p_{e,k})/p_{e,k}) \\ t &= c \end{aligned} \tag{8}$$

This last equality follows because the results for the MAP detector in (4) state that $W_k = c\ln((1-p_{e,k})/p_{e,k})$, where c is constant for all k. The optimal t's are thus same for every ring. When $t = c$, $E[e^{tW_k r'_{k,i}}] = 2\sqrt{p_{e,k}(1 - p_{e,k})}$. Thus:

$$\begin{aligned} P(E) &\leq \prod_{k=1}^{K}\prod_{i=1}^{N_k} E[e^{W_k r'_{k,i}}] \\ &= \prod_{k=1}^{K}\prod_{i=1}^{N_k} 2\sqrt{p_{e,k}(1 - p_{e,k})} \\ &= \prod_{k=1}^{K}\left[2\sqrt{p_{e,k}(1 - p_{e,k})}\right]^{N_k} \end{aligned} \tag{9}$$

Taking the logarithm on both sides,

$$\ln(P(E)) \leq \sum_{k=1}^{K} N_k[\ln(2) + \frac{1}{2}\ln(p_{e,k}(1 - p_{e,k}))] \tag{10}$$

■

This error exponent is accurate and the bound is tight when the motes are densely deployed; i.e., the number of sensors is large. Minimizing $P(E)$ when the number of motes is finite is a very difficult combinatorial problem, so we can minimize this

upper bound instead. Also note that the effect of each ring of motes on the decision error probability at the CH is apparent in this bound. It may thus be used to simplify many optimization problems in distributed detection in multi-hop scenarios. In the next section, it is used to show when and to what degree our proposed local-fusion/repetition strategy provides a smaller upper bound on the decision error probability than traditional pure-relay strategies. In a later section, direct calculation and simulations are used to confirm that these results hold for small to moderate numbers of motes.

Now consider the simplifying case that a spatial Poisson process governs the locations of the motes. Assume that the spatial density of the motes is λ_R and the width of each ring is r, so $\lambda_k = E[N_k] = \lambda_R[\pi k^2 r^2 - \pi(k-1)^2 r^2]$. Then the effect on the decision error probability of varying the density of the motes is given by:

$$\begin{aligned} P(E) &\leq \prod_{k=1}^{K} \sum_{N_k=0}^{\infty} \left[2\sqrt{p_{e,k}(1-p_{e,k})}\right]^{N_k} \lambda_k^{N_k} e^{-\lambda_k}/N_k! \\ &= \prod_{k=1}^{K} e^{-\lambda_k\left[1-2\sqrt{p_{e,k}(1-p_{e,k})}\right]} \end{aligned} \tag{11}$$

The decision error probability, as expected, decreases exponentially as the density of the motes is increased.

Fig.7 compares the large deviation bound on the error probability with the error probability from simulations for a 3-ring cluster. It shows the desired result that the large deviation error exponent is accurate – the bound is parallel with the simulation result over the entire range of spatial densities. Of course, the bound is high by a multiplicative factor, which is often the case with large deviation techniques. The conclusion is that it is reasonable to optimize the error exponent because it will optimize the rate of decay of the actual error probability itself. Simulations are used later to confirm that this is the case.

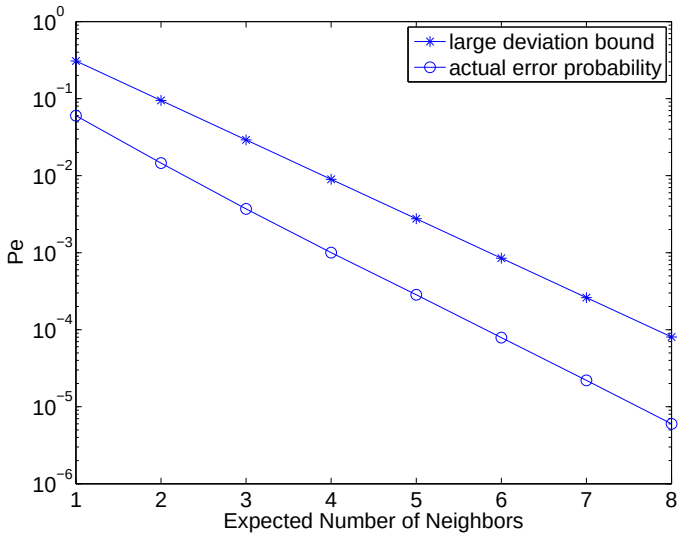


Figure 7: Comparison for a 3-ring cluster of the error probability obtained via simulation and its large deviations bound. The error exponent is clearly correct and the difference approaches zero asymptotically. Comparing the performance of complex strategies via their error exponents will thus correctly show which strategy is best when motes have large numbers of 1-hop neighbors. Direct calculations and simulations are used to confirm those results for small to moderate numbers of neighbors.

2.5 Local-Decision/Repetition Strategy

We now analyze a decision and relay strategy that reduces the decision error probability at the CH for decisions forwarded from motes in a single sector of a ring. Consider the case that one mote in ring i receives a set of individual detection results from M motes in a sector of ring $i+1$. Assume w.l.o.g. that each individual detection result has an equal probability to be "0" or "1".

The relay mote has two options. It can relay every received bit toward the CH – this is called the *pure-relay* decision strategy and is the most commonly assumed strategy. The alternative is for the relay mote to make a local decision by computing the locally optimal weighted median of the M decision bits it receives from the motes in a sector of the next furthest ring. The relay mote then transmits this decision M times toward the CH. This will be called the *local-decision/repetition* strategy.

Note that the number of bits transmitted by the relay mote is the same in both

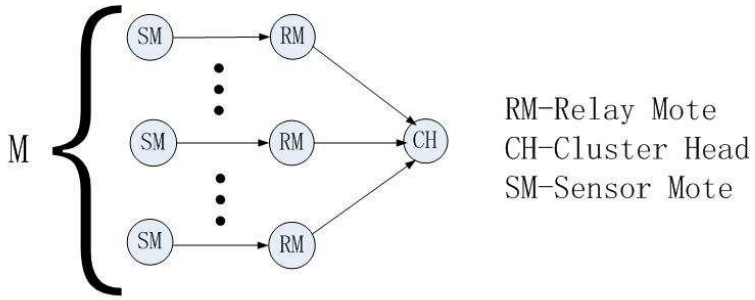


Figure 8: The *pure-relay* strategy in a 2-hop network.

strategies. Both thus use the same communication energy. The local-decision/repetition strategy uses slightly more processing energy because the relay mote computes the median of the bits it receives.

Figs. 8 and 9 show block diagrams of the pure-relay and local-decision/repetition strategies for handling bits from one sector of a ring. Assume w.l.o.g. that the CH makes a decision based on relayed detection results using the median filter. Then, since both strategies use the same energy, we want to know which one results in a lower probability of error when the CH makes a decision based solely on the bits it eventually receives from this sector. Clearly, the local-decision/repetition strategy should be better – in fact, the following theorem and corollary show that it is significantly better for sectors in the outer rings because it helps counteract the effects of multiple exposures to communication errors as the decisions are hopped to the CH.

Theorem 3: Consider $M > 1$ motes in a sector of ring $k > 1$ that transmit their decisions to a single relay mote in ring $k-1$. Assume the individual detection errors for these motes are i.i.d. with probability of error $0 < p_d < .5$ and the communication errors experienced by bits forwarded by the relay mote to the CH are i.i.d. with $0 < p_c < .5$. Then, the probability of error for the decision the CH makes with the M bits it receives under the local-decision/repetition strategy is always smaller than

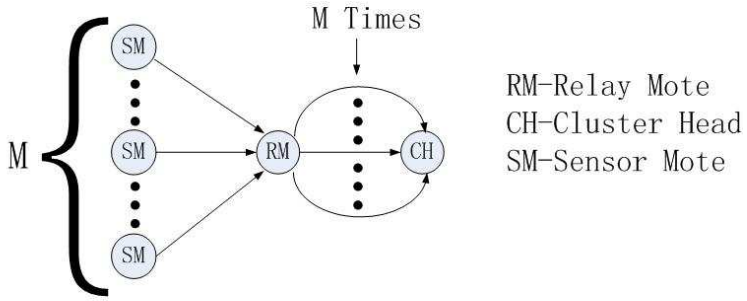


Figure 9: The *local-decision/repetition* strategy in a 2-hop network.

when its decision is based on the M bits received under the pure-relay strategy.

Proof: Please see Appendix A.

Thus, given that we must either locally fuse and repeat all M decisions or relay all of them, it is always better to fuse them for all rings $k \geq 2$. The local-decision/repetition strategy thus outperforms the pure-relay strategy for any given ring sector with M motes. The benefit increases with the number of hops between the mote performing the local fusion and the CH. In other words, it is best to do the local fusion as soon as possible – in ring $k-1$ for decisions originating from a sector of ring k:

Corollary 1: Consider $M > 1$ sensor motes in a sector of ring $K > 1$ of a multi-hop cluster. Define G_0 as the strategy that relays every bit from these motes to the CH and then makes a global decision. Define $G_k, 1 \leq k \leq K-1$ as strategies that first relay every bit to a mote that is k hops away from the CH and then uses the local-decision/relay strategy to forward the M bits for each remaining hop. The decision error probabilities, $P_e(G_k)$, of these strategies satisfy:

$$P_e(G_0) > P_e(G_1) > \ldots > P_e(G_{K-1}) \tag{12}$$

■

These results lead to three conclusions: (a) Both strategies use exactly the same

amount of communication energy, so, for any global decision based solely on nodes from one sector, the local-decision/repetition strategy is better than the pure-relay strategy; (b) It is best to use the local-decision/repetition strategy at the first hop along the path from the sector to the CH; (b) The performance degradation is cumulative from right to left in (12), so the benefits of the local-decision/repetition strategy are greater for sectors farther from the CH; and, (c) The more motes there are in a given sector, the greater the advantage of using the local-decision/repetition strategy.

2.6 Hybrid Decision Strategies

The previous section considered decisions from a single sector of some ring in the cluster. We now consider an entire ring of motes. If the shortest paths are used for moving the decisions from ring k to the CH, then they can not all be fused in ring $k-1$ as if the ring were a single sector. Theorem 3 and its Corollary can thus not be used. Instead, the ring must be divided into sectors and the local-decision/repetition strategy should be evaluated for each sector.

We must thus determine the tradeoff between the number of nodes in a sector and the number of sectors in a ring. The number of motes in a ring is fixed, so increasing the number of motes in each sector decreases the number of sectors. The number of sectors reporting via the local-decision/repetition scheme (that fuses *all* decision from a given sector) is the number of independent sets of decisions arriving at the CH. On the other hand, local-decision/repetition reduces the impact of communication errors on decisions from each sector – and this effect increases as the number of motes in a sector increases. To evaluate this tradeoff, we will determine the effect on the global decision at the CH of each possible sector size.

There is thus a tradeoff between the quantity of independent data, which is the *number* of independent sectors, and the *quality* of the data, which is the number of

motes per sector when each sector uses local-decision/repetition. We now evaluate this tradeoff under the assumptions that: (i) all sectors have exactly M motes; and (ii) the total number of motes in the ring is large enough for the large deviations results to be accurate. The choice is then which rings should just relay decisions of all motes and which rings should be divided into sectors, each of which uses the local-decision/repetition strategy to fuse the M decisions from that sector. We will see that for any fixed M, the inner rings of a cluster should use the pure-relay strategy and the outer rings should use the local-decision/repetition strategy.

Theorem 4: Consider a sensor network with a node density that is large enough for the large deviation bounds derived earlier to be tight. Given a local-decision/repetition strategy that fuses M local decisions, there exists a finite $\overline{K}$ such that the strategy that minimizes the decision error probability at the CH is for motes in rings $k < \overline{K}$ to use the pure-relay strategy and those in rings $k \geq \overline{K}$ to use the local-decision/repetition strategy.

Proof: For simplicity, we assume that the number of nodes in ring k, N_k, is divisible by M for every k. Since M is usually a small integer and N_k is a large integer, the number of sectors, N_k/M, is large. For the pure-relay and local-decision/repetition strategies, we have

$$\begin{aligned} &\lim_{N_k \to \infty} -\frac{1}{N_k} P(E_G) = C(p_{e,k}) \\ &C(p_{e,k}) = -\ln(2) - \frac{1}{2}\ln(p_{e,k}(1-p_{e,k})) \\ &p_{e,k} = \frac{1}{2} - \frac{1}{2}(1-2p_m)(1-2p_c)^k \end{aligned} \tag{13}$$

and

$$\begin{aligned} &\lim_{N_k \to \infty} -\frac{1}{N_k} P(E_L) = C^{'}(M, p^{'}_{e,k}) \\ &C^{'}(M, p^{'}_{e,k}) = \frac{1}{M}[-\ln(2) - \frac{1}{2}\ln(p^{'}_{e,k}(1-p^{'}_{e,k})] \\ &p^{'}_{e,k} = \frac{1}{2} - \frac{1}{2}(1-2h(M,p_{e,1}))(1-2h(M,p_c))^{k-1} \end{aligned} \tag{14}$$

respectively, where $p_{e,k}^{'}$ can be obtained by solving the first-order difference equation $p_{e,k}^{'} = p_{e,k-1}^{'}(1 - h(M, p_c)) + (1 - p_{e,k-1}^{'})h(M, p_c), k \geq 2$ with the initial condition $p_{e,1}^{'} = h(M, p_{e,1})$. In each ring, the motes can compare these two large deviation rates, and adopt the relay strategy corresponding to the larger rate. We consider two extreme cases here. First, when $k = 1$, there is no relay step during the transmission between the motes and the CH. Also, the pure-relay strategy is already known to minimize the decision error probability for this case. Hence motes in the first ring should always adopt the pure-relay strategy. For the completeness of the analysis, we give a concise proof that $C(p_{e,1}) > C^{'}(M, p_{e,1}^{'})$. For any positive odd number M, based on the Cauchy-Schwarz inequality,

$$\begin{aligned}
&h(M, p_{e,1})(1 - h(M, p_{e,1})) \\
&\quad = \left(\sum_{i=\lceil M/2 \rceil}^{M} \binom{M}{i} p_{e,1}^{i} (1 - p_{e,1})^{M-i} \right) \\
&\qquad \left(\sum_{i=\lceil M/2 \rceil}^{M} \binom{M}{M-i} p_{e,1}^{M-i} (1 - p_{e,1})^{i} \right) \\
&\quad \geq \left(\sum_{i=\lceil M/2 \rceil}^{M} \sqrt{\binom{M}{i}\binom{M}{M-i} p_{e,1}^{M} (1 - p_{e,1})^{M}} \right)^2 \\
&\quad = (2^M/2)^2 p_{e,1}^{M} (1 - p_{e,1})^{M} \qquad (15)
\end{aligned}$$

Note that the equality can not hold unless the two vectors are proportional to each other, i.e., $p_{e,1} = 0$, $1/2$ or 1. We can further get

$$h(M, p_{e,1})(1 - h(M, p_{e,1})) > (2^M/2)^2 p_{e,1}^{M} (1 - p_{e,1})^{M} \qquad (16)$$

$$\begin{aligned}
&\frac{1}{2} \ln(h(M, p_{e,1}) (1 - h(M, p_{e,1}))) \qquad (17) \\
&\qquad > (M - 1) \ln(2) + \frac{M}{2} \ln(p_{e,1}(1 - p_{e,1}))
\end{aligned}$$

$$\begin{aligned}
&\frac{1}{M}[- \ln(2) - \frac{1}{2} \ln(h(M, p_{e,1})(1 - h(M, p_{e,1})))] \qquad (18) \\
&\qquad < - \ln(2) - \frac{1}{2} \ln(p_{e,1}(1 - p_{e,1}))
\end{aligned}$$

$$C^{'}(M, p^{'}_{e,1}) < C(p_{e,1}) \tag{19}$$

For any positive even number M, $h(M, p_{e,1}) = h(M-1, p_{e,1})$, $C^{'}(M, p^{'}_{e,1}) < C^{'}(M-1, p^{'}_{e,1})$, and the inequality still holds.

Now, assuming K is sufficiently large, we consider the case for $k = K$. Since $\ln(1+x) \cong x$ when x is sufficiently small, we have $C(p_{e,K}) \cong \frac{1}{2}(1-2p_m)^2(1-2p_c)^{2K}$ and $C^{'}(M, p^{'}_{e,K}) \cong \frac{1}{2M}(1-2h(M, p_{e,1}))^2(1-2h(M, p_c))^{2(K-1)}$. Comparing $C(p_{e,K})$ and $C^{'}(M, p^{'}_{e,K})$, we can get $\ln(C(p_{e,K})) - \ln(C^{'}(M, p^{'}_{e,K})) \cong 2K[\ln(1-2p_c) - \ln(1-2h(M, p_c))] + 2\ln(1-2p_m) + \ln(M) - 2\ln(1-2h(M, p_{e,1})) + 2\ln(1-2h(M, p_c))$. Note that $p_c > h(M, p_c)$, so $\ln(1-2p_c) - \ln(1-2h(M, p_c)) < 0$. Since all the other terms in the difference are independent of K, when K is sufficiently large, $C(p_{e,K}) < C^{'}(M, p^{'}_{e,K})$. Hence the motes many hops away from the CH should always adopt the local-decision/relay strategy.

■

Now define $\overline{K}_M$ as the ring index beyond which the local-decision/repetition strategy of fusing M decisions is better than the pure-relay strategy. Then, under the same assumptions used in Theorem 4, it is straightforward to show that:

Corollary: $\overline{K}_M$ is a nondecreasing function of M.

Fig. 11 compares the large deviations rates between the pure-relay strategy and local-decision/repetition strategies with $M = 3$ and $M = 5$. As expected, the largest rate for the innermost rings, rings 1 and 2, is that of the pure-relay strategy. The communication error that nodes in these rings suffer is smaller than for all other rings, so they should just forward their decisions to the CH. For rings 3 and 4, the rate for the local-decision/repetition strategy with $M = 3$ is largest, so nodes in those rings should adopt that strategy. For rings 5 and 6, the rate for the local-decision/repetition strategy with $M = 5$ is largest, so these rings should adopt that strategy. For rings sufficiently further out, the local-decision/repetition strategy with $M = 7$ will then best, and so forth.

As the probability of communication error increases, the points at which the curves in Fig. 11 cross will move to the left. In other words, the best value of M for a given ring will increase when the probability of communication error increases. Of course, nodes in ring 1 should always send their decisions directly to the CH.

Now consider the simplest case in which only one value of M, typically $M = 3$ is feasible because of network architecture, routing constraints, or other considerations. To minimize the upper bound on the CH's overall decision error probability, the large deviation rate can be used to determine which rings should fuse local decisions. The motes will switch from the pure-relay strategy to the local decision strategy when k is large enough to make $C^{'}(M, p^{'}_{e,k})$ greater than $C(p_{e,k})$. We call this the *hybrid-decision* strategy. Denote the event that the decision error happens at the CH using the hybrid decision strategy by E_H. Given M, the upper bound of the decision error probability at the CH is:

$$\ln(P(E_H)) \leq -\sum_{k=1}^{K} N_k \max\{C(p_{e,k}), C^{'}(M, p^{'}_{e,k})\} \tag{20}$$

which can be rewritten as

$$\begin{aligned} P(E_H) &\leq e^{-\sum_{k=1}^{K} N_k \max\{C(p_{e,k}), C^{'}(M, p^{'}_{e,k})\}} \\ &= \prod_{k=1}^{K} e^{-N_k \max\{C(p_{e,k}), C^{'}(M, p^{'}_{e,k})\}} \\ &\leq \min\{e^{-\sum_{k=1}^{K} N_k C(p_{e,k})}, e^{-\sum_{k=1}^{K} N_k C^{'}(M, p^{'}_{e,k})}\} \\ &= \min\{P(E_G), P(E_L)\} \end{aligned} \tag{21}$$

These three decision and relay strategies – pure-relay, local-decision/repetition, and the hybrid just defined – are now compared via simulation. When the motes are uniformly deployed in a plane, the average number of motes in ring k is $N_k = d(2k-1)$, where d is the node density. Fig. 12 compares the three relay strategies when the measurement error probability at each node is $P_m = 0.2$, there are $K = 5$ rings, the node density is $d = 3$, and there are $M = 3$ motes in each sector. The hybrid strategy

is, of course, the best for all values of the communication error probability p_c. For high values of p_c, the hybrid decision strategy is close to the local-decision/repetition strategy; for low values of p_c, the hybrid decision strategy is close to the pure-relay strategy.

2.7 Hybrid, Hierarchical and Compression Strategies

In this section, six strategies for distributed detection in multi-hop sensor networks are compared via simulation. The first three were analyzed above; the others are variations of them that either save energy or improve performance.

1. The *pure-relay* strategy: The decision reached by each sensor mote is relayed to the CH by other motes over as many hops as necessary. No fusion of decisions takes place anywhere along the route

2. The *local-decision/repetition* strategy: For each i, a local decision (median) is made by a relay mote in ring i based on the M independent decisions transmitted to it by the M motes in a sector of ring $i+1$. This local decision is transmitted M times by the relay mote in ring i to a relay mote in ring $i-1$. The relay mote in ring $i-1$ computes the median of these M bits and then transmits its decision M times to a relay mote in ring $i-2$. This *fuse/repeat* process continues until M bits reach the CH. It is also performed for every set of M motes in every ring except the first ring.

3. The *hybrid* strategy: Sensor motes in different rings adopt the *pure-relay* (1) or *local-decision/repetition* (2) strategy by choosing the one with the largest large-deviation exponent. When all motes have the same measurement error probability and every transmission has the same communication error probability, the innermost rings might adopt the *pure-relay* strategy while all others adopt the *local-decision/repetition* strategy.

4. The *hierarchical decision* strategy: The local decisions are made in a hierarchical manner. Each mote in a set of M motes in ring i transmits its decisions to a relay mote in ring $i-1$. The relay mote fuses these received bits by computing the median and then relays this bit M times to a relay mote in ring $i-2$. The relay mote in ring $i-2$ fuses the M bits repeated to it by each of M relay motes in ring $i-1$ and then transmits its decision M^2 times to the relay mote in ring $i-3$.

5. The *local compression* strategy: This is the *local-decision/repetition* strategy in (2) but without repetition. Thus, M sensor motes in ring i send their independent decisions to a relay mote in ring i-1. The relay mote fuses (median) these M bits and transmits the single bit that results to a relay mote in ring i-1. This bit is forwarded by each relay mote after that until it reaches the CH – no further fusion takes place.

6. The *hierarchical compression* strategy: It is the same as the *hierarchical decision* strategy, but local fusions at each hop are transmitted only once to the next hop. This significantly reduces the number of bits transmitted and thus the energy used.

We first compare the number of transmissions, and thus the energy used, by these strategies for a sensor network that has a tree topology with the CH as the root. Suppose there are Q motes in the inner-most ring around the CH and that the k'th ring has $N_k = L^{k-1}Q,$, $1 \leq k \leq K$ motes, where L is an integer. There are thus $(L^K-1)Q/(L-1)$ motes in the network. All motes except those in the outermost ring are responsible for relaying the decision results transmitted from the L motes in the corresponding sector in the next closest ring. In this case, strategies (1) through (4) each require a total of $(KL^{K+1}-(K+1)L^K+1)Q/(L-1)^2$ transmissions. Strategy (5) requires $(L^K-1)Q/(L-1)+((K-1)L^K-KL^{K-1}+1)Q/(L-1)^2$ transmissions

and strategy (6) requires $(L^{K+1} - KL + K - 1)Q/(L-1)^2$.

Fig. 13 shows the results of simulations to determine the detection error performance of the 6 strategies when the measurement error probability is $P_m = 0.2$, there are $K = 5$ rings, there are $Q = 3$ motes in ring 1, the factor by which the number of motes increases from ring to ring is $L = 3$, and there are $M = 3$ nodes in each sector.

Fig. 14 shows a comparison when $P_m = 0.2$, $K = 3$, $Q = 3$, $L = 3$, $M = 3$. When the communication noise is severe, strategies 2 and 4 perform very well and the detection error probabilities of 1, 5, and 6 are similar. When the communication noise is negligible, strategy 1 is always the best, 5 and 6 are the worst, and 2 and 4 are in the middle.

Another important measure of performance is the time required to collect one decision from each node in the multi-hop network. The compression strategies require less time because fewer bits or bits/packets are transmitted. To quantify this advantage, the techniques in [17] can be used to obtain lower bounds on the time required. We thus assume: the nodes in a 7 hop cluster are distributed according to a spatial Poisson process with density λ; the width of each ring is r; and the time required to transmit a packet between two motes is t. With these assumptions, the minimum time to collect all decisions under the pure-relay strategy is $106\lambda_R \pi r^2 t$. Under a local compression strategy with $M = 3$ motes in each sector, the minimum time required is $38.67\lambda_R \pi r^2 t$, which is 36.48% of that using the pure-relay strategy. The structure of the network prevents this figure from being the 33% that one would expect, but the improvement is still substantial.

2.8 Implementation Issues

When a relay mote sends multiple copies of its local decision result to the next hop, those replicas can be encapsulated in a packet with decisions/estimates related to other tasks. This minimizes the addressing overhead and results in fewer and longer

packets, which help improve the throughput if a CSMA/CA-based protocol is used. The number of hops that the bits from each detection result have traveled to reach the current ring can be included in this packet.

The BSC assumption implies that the errors in successive bits in a packet are independent. By interleaving the relay node's replicated decision bits with other bits in the packet, the independence of these errors for each decision application can be guaranteed. This makes the BSC-model of the channel very realistic. The BSC model can also account for processing errors by nodes, including the injection of incorrect decisions by compromised nodes, when the probability that this occurs can be measured by temporary forwarding of preset bit sequences.

To simplify the analysis, each sector in a ring was assumed to have the same number of nodes, M. If nodes and the routes are deployed deterministically, this can usually be arranged. If the motes are deployed randomly and algorithms like the one in [15] are used for creating multi-hop clusters and routes, then the number of nodes in a sector will also be random. New algorithms that attempt to optimize the number of motes per sector are thus needed. This must be balanced with algorithms that rotate the responsibility for relaying among a set of motes in order to prevent rapid exhaustion or any one mote.

Finally, the local-decision/repetition, hybrid, and compression algorithms introduced in this chapter will be tested in the eStadium sensor net currently under development at Georgia Tech.

2.9 Summary

This chapter developed tools to determine when local fusion of decisions improves the performance of distributed detection algorithms operating in multi-hop clusters of sensor motes subject to both measurement and communication errors. Our current efforts include development of optimization tools that enable the best tradeoffs

between detection-performance, delay, energy-usage, etc. to be found for any sensor network application.

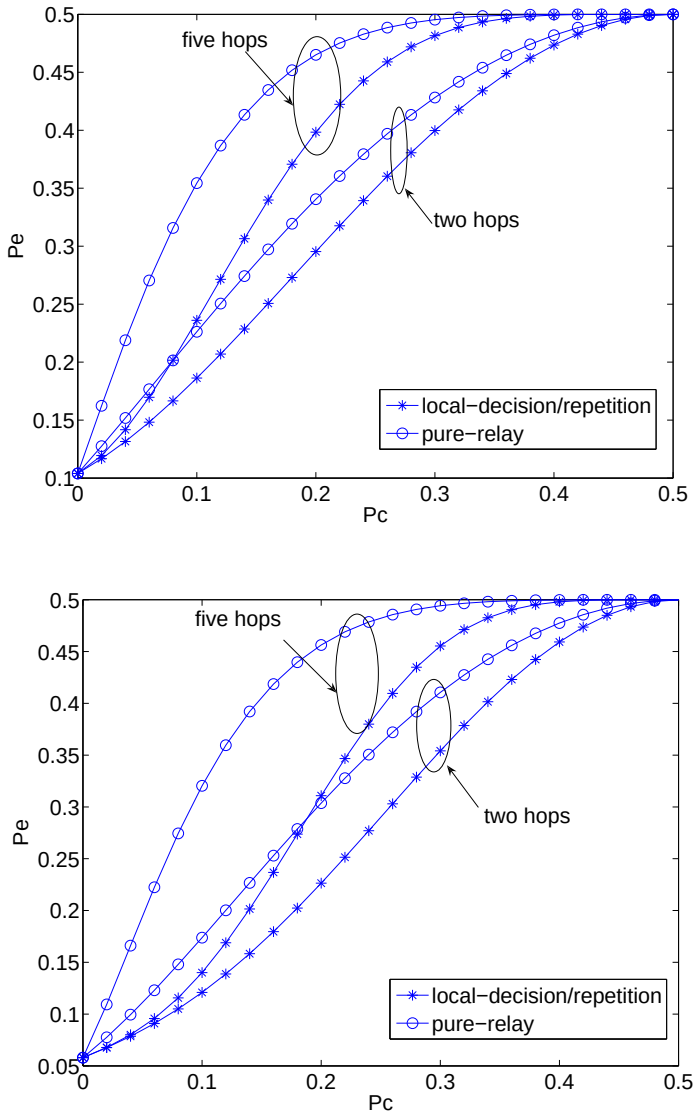


Figure 10: Calculations of the probability of error for the local-decision/repetition strategy for small numbers of motes show how much it improves the detection performance for a range of communication error probabilities. In the upper figure, decisions from M=3 motes are fused and the measurement error probability is p_m=0.2; in the lower figure, M=5, p_m=0.2. The improvement increases with both the number of wireless hops and the number of motes, M, whose decisions are fused. The cases shown are for decisions made at the CH based on the M decisions it receives from M motes in one sector of one ring in a multi-hop cluster – the full clusters' performances are shown in Figure 12–13.

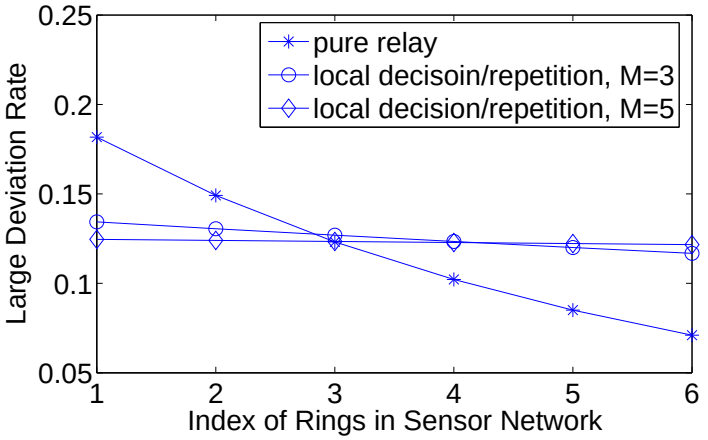


Figure 11: Large deviation rates as a function of the number of rings are shown for the pure-relay and the M=3 and M=5 local-decision/repetition strategies. The probability of measurement error was $p_m = 0.2$ and the probability of communication error was $p_c = 0.04$. The crossing of the rates implies that the error probabilities for a finite numbers of motes should cross as well. Motes in rings 1 and 2 should clearly use the pure-relay strategy. Nodes in rings 3 and 4 should fuse decisions from sectors of size $M = 3$. Nodes in rings 5 and 6 can use either $M = 3$ or $M = 5$ because they provide almost the same performance. If a node in ring 2 receives decisions from six nodes in ring 3, it should break them into two sets of 3 decisions and use the local-decision/repetition strategy on each set.

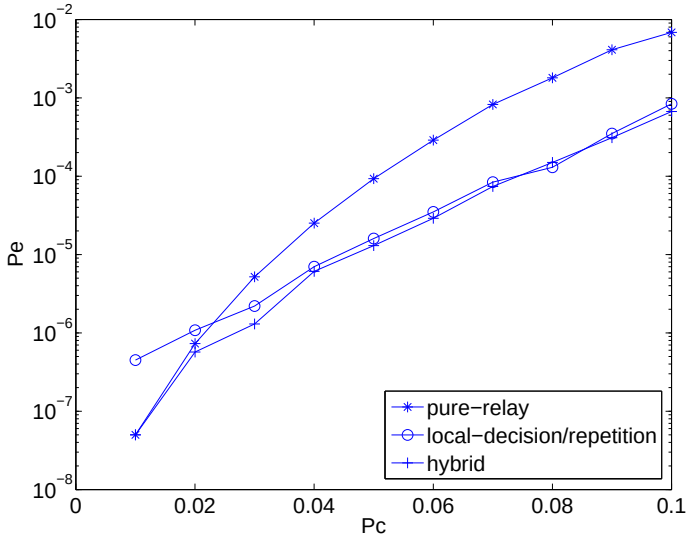


Figure 12: Comparisons via simulation of the pure-relay, local-decision/repetition, and hybrid strategies. The hybrid decision strategy lowers the probability of error by using local decisions when needed.

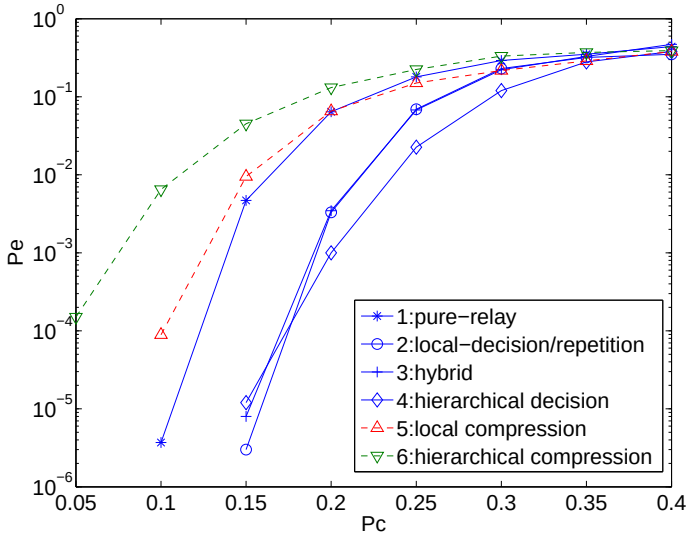


Figure 13: Simulations showing error probabilities achieved by six strategies in a 5 ring network. The ones in solid use the same amount of communication energy. The compression strategies (dashed) use less energy because they make fewer transmissions.

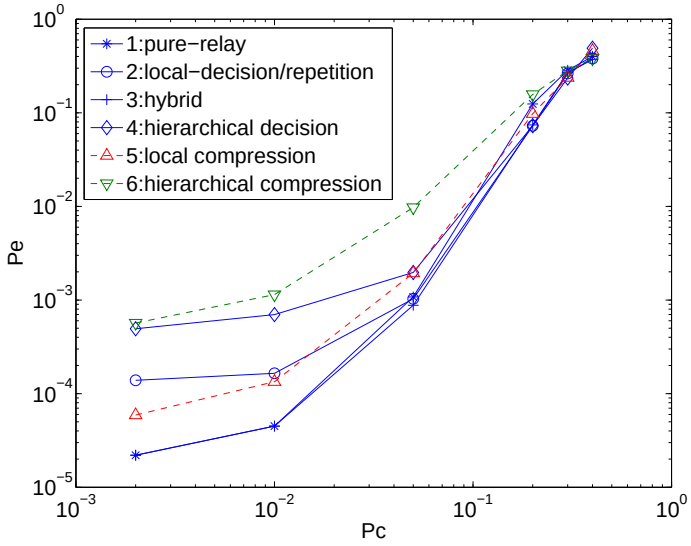


Figure 14: Simulations Showing error probabilities achieved by six strategies in a 3 ring network. The ones in solid use the same amount of communication energy. The compression strategies (dashed) use less energy because they make fewer transmissions.

CHAPTER III

DISTRIBUTED ESTIMATION IN WIRELESS SENSOR NETWORKS

3.1 Research Motivation

This chapter will focus on distributed estimation in wireless sensor networks. In this scenario for estimation in wireless sensor networks, we assume that many small, battery-powered sensors with limited sensing, processing and transmission capabilities are spread over a region [6]. They each make local estimates of the parameters of some phenomenon that affects that region. These local estimates are then transmitted over noisy wireless channels to a processing center that fuses them into global estimates of the phenomenon.

To reduce the energy used for communication, the sensors self-organize into a multi-hop clustered architecture. This architecture may have one level of multi-hop clusters [7] or be a hierarchy of two or more levels of clusters [15]. In such cases, each cluster's Cluster-Head (CH) will fuse the decisions it receives from nodes on its level into a cluster-wide decision. This cluster-wide decision is then forwarded to a processing center, either directly or via a higher level of clusterheads.

The clusters at the lowest level of the hierarchy typically have diameters that are less than the spatial correlation length of the phenomenon being observed. The sensors in these low-level clusters are thus making local estimates about a spatially-invariant phenomenon. Each of these local estimates may be affected by measurement noise and their transmission through the network may be affected by noise, fading, transceiver errors, errors injected by compromised nodes, etc. Analysis of the effects of these measurement and communication errors, typically individually but sometimes

together, has been the goal of many studies in distributed detection and estimation.

The focus of this chapter is the distributed estimation problem that CHs in the first level of the hierarchy must solve. [60] proposes an isotropic decentralized estimation scheme in a homogeneous environment. Each sensor transmits just one bit, with probability 1/2 to transmit the most significant bit, probability 1/4 to transmit the second-most significant bit, etc. [99] extends this result to the distributed estimation problem in an inhomogeneous environment by letting each sensor transmit a number of bits proportional to the logarithm of its local signal to noise ratio (SNR). [55] analyzed the optimal energy allocation problem in this case by giving priority to the transmissions with high energy efficiency.

[62] discusses optimal local estimation and final fusion schemes under the constraint that the communication from each sensor to the fusion center must be a one-bit message. In a similar case with a noisy communication channel, a MAP estimator and its variations are given in [12]. Each sensor is assumed to transmit a single bit that results from comparison of its local measurement with a pre-set threshold. The information these bits provide decreases significantly when the preset threshold is not equal to the true but typically unknown signal value. To reduce the required number of sensors and the sensitivity to the choice of threshold, a modified scheme is provided in [79] in which the sensors use nonidentical thresholds that are uniformly distributed in an interval that must contain the true signal value. [76] and [75] propose to let the sensors transmit one bit when the noise is relatively high, and multiple bits when the noise is relatively low, both with nonidentical thresholds obtained via convex optimization techniques.

[98] discusses the power scheduling problem for universal decentralized estimation in sensor networks. There are, however, two issues not addressed by this chapter: power scheduling among different bits of each quantized local estimate has not been considered, and the final estimate is generally biased. [92] derives a MAP estimator

for distributed estimation in multi-hop clustered wireless sensor networks, but its realization by line search is of high computational complexity.

There are many papers that discuss the consensus problem in sensor networks. In this case, all sensors smooth their local estimates by exchanging them with their neighbors. This topic is only partially related to this chapter; please consult [47, 37, 38, 13] for information about it.

The results in this chapter improve upon those discussed above in five ways: (1) The system model is extended to the more general case that each sensor can transmit up to a fixed number of bits; (2) The more realistic scenario in which the individual estimates received by the CH are corrupted by both measurement noise and communication noise is addressed; (3) The computational complexity of our algorithm is fairly low and its performance is guaranteed; (4) The overhead required for implementation of our approach is low – the CH only needs a rough estimate of the crossover probability of each one-hop channel and the variance of the measurement noise; and (5) An algorithm is developed that both minimizes the energy required for the network to produce an estimate with a prescribed error variance and determines how this energy should be allocated amongst the sensors in the multihop network.

This rest of the paper is organized as follows: Section 3.2 describes the system model for the distributed estimation problem in clustered sensor networks and reviews the use of the BLUE estimator in the noiseless channel case. In Section 3.3, an unbiased estimator for distributed estimation in the noisy-channel case is proposed and compared with two other commonly used schemes. In section 3.4, the sensitivity of the proposed schemes to the estimation accuracy of the crossover probability of the BSC channel is analyzed. Section 3.5 investigates the tradeoffs in energy allocation involving measurement noise, communication noise, and quantization noise. Section 3.6 discusses the difference between the proposed scheme and other schemes such as those using thresholds. Section 3.7 provides some summarizing remarks.

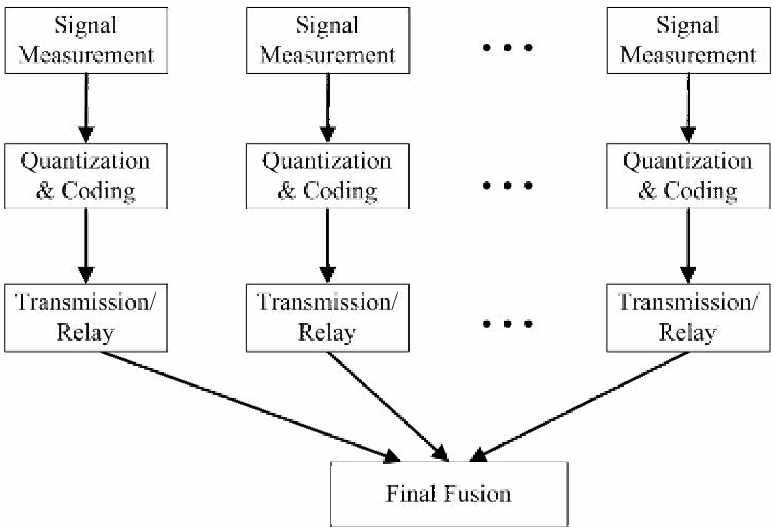


Figure 15: Block diagram illustrating distributed estimation in clustered wireless sensor networks.

3.2 System Model

We consider the problem of distributed estimation in a clustered sensor network. Suppose that the number of rings in a cluster is K and the number of sensors in the k'th ring is N_k. The objective is to provide an accurate estimate of a random signal based on a set of coarse estimates corrupted by both measurement noise and communication noise. The process consists of four steps: (1) noisy signal measurements at individual sensors, (2) quantization and coding of these measurements, (3) transmission of the quantized and coded message along noisy multi-hop channels, (4) final data fusion at the cluster head. A flowchart of the entire process can be seen in Fig. 15. The details of each step are described below.

3.2.1 Signal Measurement

Denote the random signal being observed by x and the measurement result at the ith sensor of the kth ring by $x_{k,i}$. The measurement process at each sensor can be described as

$$x_{k,i} = x + n_{k,i} \tag{22}$$

where $n_{k,i}$ is the measurement noise at the ith sensor of the kth ring. Assume that the random signal is distributed in the interval $[-U, U]$ and the measurement noise is zero-mean and bounded by the interval $[-V, V]$. Letting $W = U + V$, the corrupted signal lies in the interval $[-W, W]$. Suppose the measurement noise is independent of the random signal and that the measurement noise processes at different sensors are mutually independent. Assume the variance of the measurement noise is known by both the local sensor and the CH, but the probability density functions of both the random signal and the measurement noise are unknown to either of them. For simplicity only, assume the measurement noises at all the sensor are identically distributed. Let the variance of the measurement noise be σ_m^2. Then each $x_{k,i}$ is an unbiased estimate of x with $Var(x_{k,i}) = \sigma_m^2$.

3.2.2 Quantization & Coding

Since a sensor can only transmit a finite number of bits, it has to quantize each noisy measurement it makes. The measurement result $x_{k,i}$ is quantized to $x'_{k,i}$ according to some predetermined rule. Two different quantization schemes will be proposed and compared in Section 3. Also, to increase the accuracy of the final estimate, a coding technique should be considered to combat channel errors. Since energy efficiency is a main concern in the design of algorithms implemented by sensor networks, many powerful but complex coding techniques are not appropriate here. We thus adopt a simple repetition code because it is easy to implement and, as will be seen, it can also be completely analyzed and optimized. The intuition is to transmit more replicas of the bits of more significance, as discussed in Section 5.

3.2.3 Transmission / Relay

After quantization and coding, the coded message is transmitted or relayed to the CH. Suppose the sensors use a BPSK modulation scheme here, so that the transmission and relay is bit by bit. For simplicity, assume all the one-hop communication channels

between neighboring sensors are binary symmetric channels with crossover probability p_c. Then the equivalent crossover probability of the k-hop communication channel $p_{c,k}$ satisfies the recursive equation:

$$p_{c,k} = p_{c,k-1}(1 - p_c) + (1 - p_{c,k-1})p_c, k \geq 2 \tag{23}$$

The solution to this first-order difference equation with initial condition $p_{c,1} = p_c$ is:

$$p_{c,k} = \frac{1}{2} - \frac{1}{2}(1 - 2p_c)^k \tag{24}$$

3.2.4 Final Fusion

Since only the variance of the measurement noise and the crossover probabilities of the BSC channels are known, the BLUE estimator is a reasonable and computationally tractable one for the CH to adopt. It achieves the lowest variance among all linear estimators and only requires first and second order statistics of each individual local estimate. Denote the local estimate received at the CH from the ith sensor in the kth ring by $x''_{k,i}$ and the final estimate made by the CH by $\hat{x}$. If we can ensure that the individual estimates from the sensors in the kth ring have zero mean and can show that σ_k^2 is an upper bound on its variance, then the global estimate at the fusion center can be represented as:

$$\hat{x} = \frac{\sum_{k=1}^{K}\sum_{i=1}^{N_k} x''_{k,i}/\sigma_k^2}{\sum_{k=1}^{K}\sum_{i=1}^{N_k} 1/\sigma_k^2} \tag{25}$$

The variance of the global estimate is then:

$$Var(\hat{x}) \leq \left(\sum_{k=1}^{K}\sum_{i=1}^{N_k} 1/\sigma_k^2\right)^{-1} = \left(\sum_{k=1}^{K} N_k/\sigma_k^2\right)^{-1} \tag{26}$$

3.3 Distributed Estimation in Clustered Sensor Networks

In this section, three schemes for distributed estimation in clustered sensor networks are proposed and compared.

3.3.1 BLUE Estimator

Assume the measurement result is quantized into L bits. Then uniformly divide the interval $[-W, W]$ into $2^L - 1$ subintervals, each with length $\Delta = (2W)/(2^L - 1)$. While doing quantization, the each sensor rounds its measurement result to the closest border of the subinterval in which it falls. If $d\Delta - W \leq x_{k,i} \leq (d+1)\Delta - W$, set $r = (x + W - d\Delta)/\Delta$, we have:

$$\begin{aligned} x'_{k,i} &= d\Delta - W, \quad if \quad r <= 1/2 \\ x'_{k,i} &= (d+1)\Delta - W, \quad if \quad r > 1/2 \end{aligned} \tag{27}$$

Let $b'_{k,i} = E[x'_{k,i} - x_{k,i}]$, and then $|b'_{k,i}| = |E[x'_{k,i} - x_{k,i}]| \leq E[|x'_{k,i} - x_{k,i}|] = E[\min\{r, 1-r\}\Delta] \leq \frac{1}{2}\Delta$. Obviously, $x'_{k,i}$ is a biased estimate of $x_{k,i}$.

Then the index d or $d+1$ is coded to a binary message to transmit to the CH. Avoiding any coding specifics for now, suppose the lth most significant bit of the message is m_l. Clearly, we have:

$$x'_{k,i} = \sum_{l=1}^{L} m_l 2^{L-l}\Delta - W \tag{28}$$

All the m_l's are sent to the next hop in order. Denote the received bit at the CH based on noise-corrupted m_l by m'_l and let the value represented by the possibly noise corrupted received message be $x''_{k,i}$. They satisfy the following relationship:

$$x''_{k,i} = \sum_{l=1}^{L} m'_l 2^{L-l}\Delta - W \tag{29}$$

Obviously, $E[x''_{k,i}]$ may deviate from $E[x'_{k,i}]$ since the BSC channel is naturally biased.

Denote $b''_{k,i} = E[x''_{k,i} - x'_{k,i}]$, and we have:

$$\begin{aligned}
&E[x''_{k,i} - x'_{k,i}|x'_{k,i}] \\
&\quad = E\left[\sum_{l=1}^{L}(m'_l - m_l)2^{L-l}\Delta|x'_{k,i}\right] \\
&\quad = \sum_{l=1}^{L} E[m'_l - m_l|m_l]2^{L-l}\Delta \\
&\quad = \sum_{l=1}^{L}(1-2m_l)p_{c,k}2^{L-l}\Delta \\
&\quad = \sum_{l=1}^{L} p_{c,k}2^{L-l}\Delta - 2\sum_{l=1}^{L} m_l p_{c,k}2^{L-l}\Delta \\
&\quad = p_{c,k}(2^L - 1)\Delta - 2p_{c,k}\sum_{l=1}^{L} m_l 2^{L-l}\Delta \\
&\quad = 2p_{c,k}W - 2p_{c,k}(x'_{k,i} + W) \\
&\quad = -2p_{c,k}x'_{k,i}
\end{aligned} \tag{30}$$

so that,

$$b''_{k,i} = E[E[x''_{k,i} - x'_{k,i}|x'_{k,i}]] = -2p_{c,k}E[x'_{k,i}] \tag{31}$$

Note that $E[m'_l - m_l|m_l] = (1-2m_l)p_{c,k}$; i.e.,

$$\begin{aligned}
E[m'_l - m_l|m_l = 0] &= p_{c,k}, \\
E[m'_l - m_l|m_l = 1] &= -p_{c,k},
\end{aligned} \tag{32}$$

which means that the BSC channel is linear in the statistical sense. Based on this result, the received individual estimates at the CH can be made unbiased via a linear mapping at either the transmitter or the receiver – this will be discussed in detail in the next subsection. The bias depends on the transmitted message, and its absolute value is bounded by:

$$2p_{c,k}\Delta \leq |b''_{k,i}| \leq 2p_{c,k}W \tag{33}$$

The total bias $b_{k,i} = b'_{k,i} + b''_{k,i}$, and

$$|b_{k,i}| \leq |b'_{k,i}| + |b''_{k,i}| \leq \frac{1}{2}\Delta + 2p_{c,k}W \tag{34}$$

Now we calculate the mean square error (MSE):

$$
\begin{aligned}
&E[(\ x''_{k,i} - x'_{k,i})^2] \\
&\quad = E\left[\left(\sum_{l=1}^{L} m'_l 2^{L-l}\Delta - \sum_{l=1}^{L} m_l 2^{L-l}\Delta\right)^2\right] \\
&\quad = E\left[\left(\sum_{l=1}^{L} (m'_l - m_l) 2^{L-l}\Delta\right)^2\right] \\
&\quad \le E\left[\left(\sum_{l=1}^{L} |m'_l - m_l| 2^{L-l}\Delta\right)^2\right] \\
&\quad = \sum_{l=1}^{L} E[|m'_l - m_l|^2] 2^{2(L-l)}\Delta^2 \\
&\qquad + \sum_{l=1}^{L}\sum_{s=1,s\neq l}^{L} E[|m'_l - m_l||m'_s - m_s|] 2^{2L-l-s}\Delta^2 \\
&\quad = \left(\sum_{i=1}^{L} 2^{2(L-i)}\Delta^2\right) p_{c,k} \\
&\qquad + \left(\left(\sum_{i=1}^{L} 2^{(L-i)}\Delta\right)^2 - \sum_{i=1}^{L} 2^{2(L-i)}\Delta^2\right) p_{c,k}^2 \\
&\quad = \frac{1}{3}(2^{2L}-1)\Delta^2(p_{c,k} - p_{c,k}^2) + (2^L - 1)^2\Delta^2 p_{c,k}^2
\end{aligned} \tag{35}
$$

Also, $E[(x_{k,i} - x)^2] = \sigma_m^2$, and $E[(x'_{k,i} - x_{k,i})^2] = E[\min\{r^2, (1-r)^2\}\Delta^2] \le \frac{1}{4}\Delta^2$. Let

$$
\begin{aligned}
\sigma_{c,k}^2 &= \frac{1}{3}(2^{2L}-1)\Delta^2(p_{c,k} - p_{c,k}^2) + (2^L - 1)^2\Delta^2 p_{c,k}^2 \\
\sigma_q^2 &= \frac{1}{4}\Delta^2
\end{aligned} \tag{36}
$$

We have

$$
E[(x''_{k,i} - x)^2] = E[((x''_{k,i} - x'_{k,i}) + (x'_{k,i} - x_{k,i}) + (x_{k,i} - x))^2] \tag{37}
$$

All the differences depend on the measurement noise $n_{k,i}$, so in general they are not mutually independent. Nevertheless, based on the Cauchy-Schwarz inequality, we have

$$
E[(x''_{k,i} - x)^2] \le (\sigma_m + \sigma_q + \sigma_{c,k})^2 \tag{38}
$$

Set $\sigma_k^2 = (\sigma_m + \sigma_q + \sigma_{c,k})^2$ and apply the BLUE estimator. Then using the Cauchy-Schwarz inequality again, we can get an upper bound on the mean square error (MSE):

$$\begin{aligned} E[(\hat{x} - x)^2] &= E\left[\left(\frac{\sum_{k=1}^{K}\sum_{i=1}^{N_k}(x''_{k,i} - x)/\sigma_k^2}{\sum_{k=1}^{K}\sum_{i=1}^{N_k} 1/\sigma_k^2}\right)^2\right] \\ &\leq \frac{\left(\sum_{k=1}^{K}\sum_{i=1}^{N_k}\sqrt{E[(x''_{k,i} - x)^2]}/\sigma_k^2\right)^2}{\left(\sum_{k=1}^{K}\sum_{i=1}^{N_k} 1/\sigma_k^2\right)^2} \\ &= \frac{(\sum_{k=1}^{K} N_k/\sigma_k)^2}{(\sum_{k=1}^{K} N_k/\sigma_k^2)^2} \end{aligned} \tag{39}$$

The square root of the total mean square error is thus a convex combination of the square roots of the individual mean square errors. Letting $N = \sum_{k=1}^{K} N_k$, when the number of sensors is large, $\lim_{N\to\infty} E[(\hat{x} - x)^2] \neq 0$. While it is not zero in the limit, it is bounded by the result in the preceding equation.

3.3.2 BLUE Estimation with Dithered Quantization

As proposed in [99], the quantization can be done in a probabilistic manner to remove bias; i.e., the corrupted signal is rounded to either of the two borders of the subinterval it falls in according to different probabilities. If $x_{k,i} \in [d\Delta - W, (d+1)\Delta - W]$ and r is as defined at the beginning of Section 3.1, we have

$$\begin{aligned} P(x'_{k,i} = d\Delta - W) &= 1 - r \\ P(x'_{k,i} = (d+1)\Delta - W) &= r \end{aligned} \tag{40}$$

It is easy to verify that $E[x'_{k,i}] = E[x_{k,i}]$. Thus

$$\begin{aligned} b_{k,i} = b''_{k,i} &= -2p_{c,k}E[x'_{k,i}] = -2p_{c,k}x \\ |b_{k,i}| &= |b''_{k,i}| \leq 2p_{c,k}U \end{aligned} \tag{41}$$

Since $E[(x'_{k,i} - x_{k,i})^2] = E[(r - r^2)\Delta^2] \leq \Delta^2/4 = \sigma_q^2$ and

$$\begin{aligned} &E[(x'_{k,i} - x_{k,i})(x_{k,i} - x)] \\ &= E[E[(x'_{k,i} - x_{k,i})(x_{k,i} - x)|x_{k,i}]] = 0 \end{aligned} \tag{42}$$

we have

$$\begin{aligned} E[(x'_{k,i} - x)^2] &= E[(x'_{k,i} - x_{k,i})^2] + E[(x_{k,i} - x)^2] \\ &\leq \sigma_m^2 + \sigma_q^2 \end{aligned} \tag{43}$$

Hence we find

$$E[(x''_{k,i} - x)^2] \leq \left(\sqrt{\sigma_m^2 + \sigma_q^2} + \sigma_{c,k}\right)^2 \tag{44}$$

For the same reason as in the first scheme, when the number of sensors is large the bias does not go to zero: $\lim_{N\to\infty} E[(\hat{x} - x)^2] \neq 0$.

3.3.3 BLUE Estimator with Dithered Quantization and Channel Compensation

Clearly, to make the final estimate unbiased, we also need to make the BSC channel unbiased. This can be accomplished by modifying either the channel's input or output. We have

$$E[x''_{k,i}|x'_{k,i}] = x'_{k,i} - 2p_{c,k}x'_{k,i}. \tag{45}$$

Let

$$\tilde{x}''_{k,i} = \frac{x''_{k,i}}{1 - 2p_{c,k}}. \tag{46}$$

Then we can guarantee that

$$\begin{aligned} E[\tilde{x}''_{k,i}|x'_{k,i}] &= x'_{k,i} \\ E[\tilde{x}''_{k,i}] &= E[x'_{k,i}] \end{aligned} \tag{47}$$

This can also be interpreted as a mapping of bits. Suppose the output "0" is

mapped to a and the output "1" is mapped to b. Then a and b should satisfy:

$$\begin{aligned} a(1-p_{c,k}) + bp_{c,k} &= 0 \\ b(1-p_{c,k}) + ap_{c,k} &= 1 \end{aligned} \tag{48}$$

Solving these linear equations, we find

$$\begin{aligned} a &= \frac{p_{c,k}}{2p_{c,k}-1} \\ b &= \frac{p_{c,k}-1}{2p_{c,k}-1} \end{aligned} \tag{49}$$

Note that if the channel is asymmetric, we can solve a similar set of linear equations to make it unbiased. Suppose the crossover probability from "0" to "1" is p_1 and the crossover probability from "1" to "0" is p_2. Then the solution in the general case is:

$$\begin{aligned} a &= \frac{p_1}{p_1+p_2-1} \\ b &= \frac{p_1-1}{p_1+p_2-1} \end{aligned} \tag{50}$$

A unique solution exists if and only if $p_1 + p_2 \neq 1$, which implies $p_{c,k} \neq 1/2$ for the binary symmetric channel. Note that with this mapping scheme the received message will always be unbiased no matter what coding technique is used. A comparison between the BSC channel and the modified BSC channel can be seen in Fig. 16 and Fig. 17.

The modification could also be done at the input of the BSC channel. The sensors can amplify $x'_{k,i}$ by the factor $\alpha = 1/(1-2p)$ before transmission to make the final estimate unbiased. Then the interval of the corrupted signal is $[-\alpha W, \alpha W]$. If the sensors still transmit L bits, the real number to quantize is $\frac{\alpha x'_{k,i}+\alpha W}{\alpha\Delta} = \frac{x'_{k,i}+W}{\Delta}$, which is the same as before. So this scheme is mathematically the same as doing the modification at the output of the BSC channel. A comparison between the quantization step and the modified quantization step can be seen in Fig. 18 and Fig. 19.

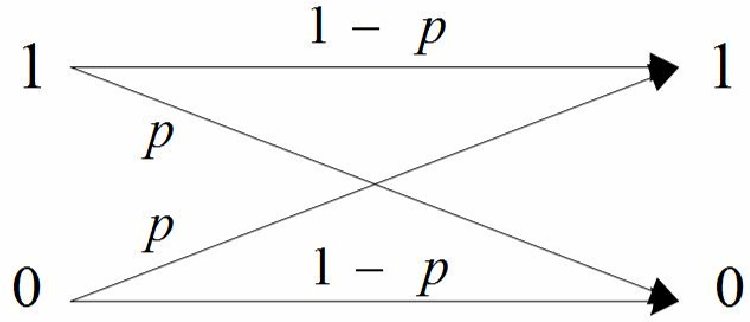


Figure 16: Binary symmetric channel model with crossover probability of p. We use this model because it allows us to account for many types of errors in the relaying of bits/packets from sensors to the CH. In addition to actual channel/decoding errors, it can include, for example, processing errors by each relay mote and bit errors maliciously injected by compromised relay nodes. More complex error models can be substituted; this one is sufficient to demonstrate the effects of errors that accumulate over multiple hops. The value of ϵ must be estimated – the effects of errors in this estimate are considered later in this chapter.

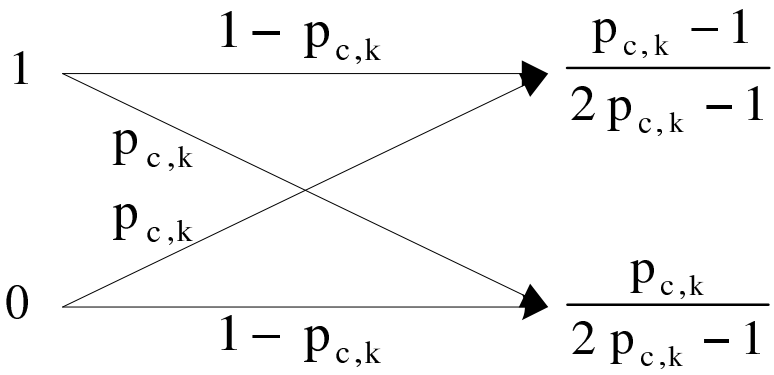


Figure 17: Binary symmetric channel model with modification to channel output. This may include multiple hops, starting from a sensor node to the CH, so the overall crossover probability is $p_{c,k}$.

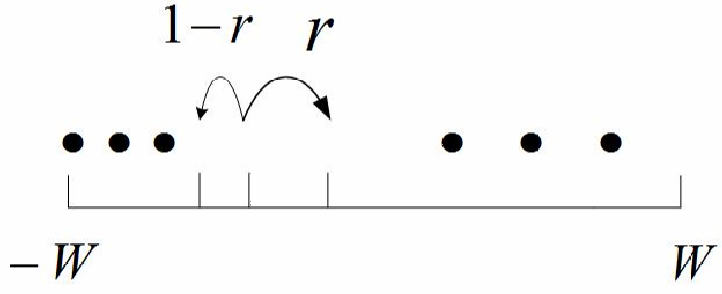


Figure 18: Dithered quantization method.

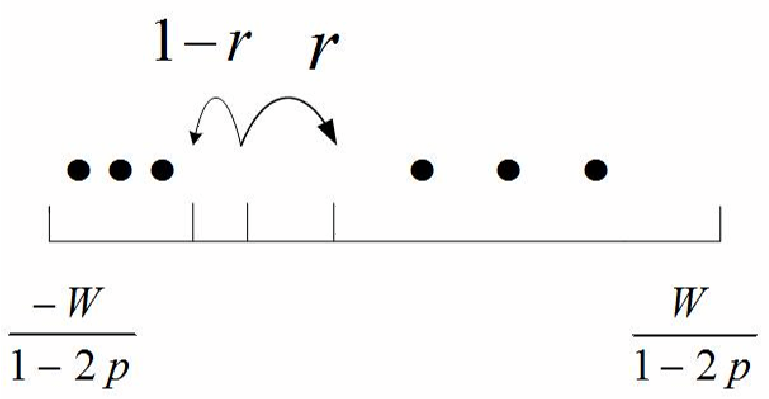


Figure 19: Dithered quantization method with modification to channel input.

Now we calculate the variance. Denote the mapping result of m'_l by $\tilde{m}'_l$. Thus

$$\begin{aligned}
E[(\tilde{x}''_{k,i} - x'_{k,i})^2] &= E\left[\left(\sum_{l=1}^{L} \tilde{m}'_l 2^{L-l}\Delta - \sum_{l=1}^{L} m_l 2^{L-l}\Delta\right)^2\right] \\
&= \sum_{l=1}^{L} E[(\tilde{m}'_l - m_l)^2] 2^{2(L-l)}\Delta^2 \\
&= \sum_{l=1}^{L} E[E[(\tilde{m}'_l - m_l)^2 | m_l]] 2^{2(L-l)}\Delta^2 \\
&= \sum_{l=1}^{L} \frac{p_{c,k}(1-p_{c,k})}{(1-2p_{c,k})^2} 2^{2(L-l)}\Delta^2 \\
&= \frac{1}{3}(2^{2L}-1)\frac{p_{c,k}(1-p_{c,k})}{(1-2p_{c,k})^2}\Delta^2 \qquad (51)
\end{aligned}$$

Let $\sigma'^2_{c,k} = \frac{1}{3}(2^{2L}-1)\frac{p_{c,k}(1-p_{c,k})}{(1-2p_{c,k})^2}\Delta^2$. Since

$$\begin{aligned}
&E[(\tilde{x}''_{k,i} - x'_{k,i})(x'_{k,i} - x)] \\
&= E[E[(\tilde{x}''_{k,i} - x'_{k,i})(x'_{k,i} - x)|x'_{k,i}]] = 0, \qquad (52)
\end{aligned}$$

we have:

$$\begin{aligned}
&E[(\tilde{x}''_{k,i} - x)^2] = E[(\tilde{x}''_{k,i} - x'_{k,i})^2] + E[(x'_{k,i} - x)^2] \\
&\leq \sigma_m^2 + \sigma_q^2 + \sigma'^2_{c,k} \qquad (53)
\end{aligned}$$

Set $\sigma_k^2 = \sigma_m^2 + \sigma_q^2 + \sigma'^2_{c,k}$ and apply the BLUE estimator. Then we can get an upper

bound on the mean square error (MSE):

$$E[(\hat{x}-x)^2] \leq \left(\sum_{k=1}^{K} N_k/(\sigma_m^2 + \sigma_q^2 + \sigma_{c,k}'^2) \right)^{-1} \tag{54}$$

When the number of sensors is large, we now have asymptotic efficiency: $\lim_{N\to\infty} E[(\hat{x}-x)^2] = 0$.

Note that $\tilde{x}''_{k,i} = \frac{x''_{k,i}}{1-2p_{c,k}}$. We can thus find the variance of the final estimate in the second scheme:

$$Var(x''_{k,i}) = (1-2p_{c,k})^2 Var(\tilde{x}''_{k,i}) \tag{55}$$

In the case when $L=1$, $Var(x'_{k,i}) = \sigma_m^2 + \sigma_q^2$ and

$$\begin{aligned} Var(x''_{k,i} - x'_{k,i}) &= E[Var(x''_{k,i} - x'_{k,i}|x'_{k,i})] \\ &\quad + Var(E[x''_{k,i} - x'_{k,i}|x'_{k,i}]) \\ &\geq E[Var(x''_{k,i} - x'_{k,i}|x'_{k,i})] \\ &= (p_{c,k} - p_{c,k}^2)\Delta^2 \end{aligned} \tag{56}$$

Considering the bounds instead of the true values, we can roughly get the following relationship in the second scheme,

$$Var(x''_{k,i}) \leq Var(x''_{k,i} - x'_{k,i}) + Var(x'_{k,i}) \tag{57}$$

which means that $x''_{k,i} - x'_{k,i}$ and $x'_{k,i}$ are negatively correlated. This coincides with the observation that $x''_{k,i} - x'_{k,i}$ decreases when $x'_{k,i}$ increases.

Finally, it is worthwhile to mention that in the first two schemes, the range of the final estimate is $[-W, W]$, and in the third scheme, the range of the final estimate is $[-\alpha W, \alpha W]$. Hence, there is a possibility that the final estimate falls outside of the interval $[-U, U]$. It can be projected back onto the interval $[-U, U]$ without increasing the total mean square error, namely

$$x^* = \max\{-U, \min\{U, \bar{x}\}\} \tag{58}$$

Fig. 20 shows simulations that compare the MSE if the BLUE is used at the CH when neither, either, or both of the bias-removal schemes are used. In all cases, the signal and noise are uniformly distributed with $U = 1$, $V = 0.25$, respectively; and there are $N_1 = 50$ nodes. Fig. 20(a), shows what happens when $L = 1$, in which case the quantization of the sensor output is very coarse because there are only $2^L = 2$ quantization levels; the output of the quantizer is basically the sign of the input. When neither dithered quantization nor channel compensation are used, the MSE is unexpectedly low for some values of p_c because the bias is very large. When dithered quantization is used by itself, most of the bias is removed but the residual bias due to lack of channel compensation grows as p_c increases. When channel compensation is used by itself, the MSE increases significantly because the compensation basically amplifies the errors caused by the coarse quantizer. *When coarse quantization is used, it is thus essential that dithered quantization and channel compensation be used together.*

In 20(b), $L = 2$ and $2^L = 4$ quantization levels are used. This somewhat finer quantization eliminates much of the unusual, bias-induced behavior of the MSE in 20(a), but not all of it. The primary improvement occurs when channel compensation is used. The improvement is not as great as what will be seen in the next figure because the error variance due to quantization is still high.

Fig. 20(c) shows the effect of the bias removal schemes when $L = 5$, so there are $2^L = 32$ quantization levels. In this case the quantization error and its bias are both negligible, so the use of dithered quantization makes very little difference. Channel compensation does significantly reduce the MSE, both with and without dithered quantization. Thus, dithered quantization can be skipped in order to simplify processing at the nodes when the input is finely quantized. Note that the MSE in 20(c) is significantly lower than any of the MSEs in 20(a) or 20(b), which shows the advantage of fine over coarse quantization.

3.4 Sensitivity Analysis

In section 3.3, we proposed an unbiased estimator that uses channel compensation when the channel is noisy. When the number of sensors is sufficiently large, it can definitely outperform the other two estimators. However, if the single-bit error rate of the channel is not perfectly estimated, the expected value of the estimate still deviates from the true value. Our focus in this section is thus to evaluate the sensitivity of the performance of channel compensation to mis-estimation of the crossover probability.

For fairness, we only compare the second scheme and the third scheme. Denote the CH's estimate of the single-bit error rate $p_{c,k}$ by $p'_{c,k}$. In the second scheme, the bias and variance stay the same as before since $p'_{c,k}$ has nothing to do with the estimator. We have

$$\begin{aligned} b_{k,i} &= -2p_{c,k}x \\ Var(x''_{k,i}) &= (1-2p_{c,k})^2(\sigma_m^2+\sigma_q^2+\sigma_{c,k}'^2) \end{aligned} \tag{59}$$

In the third scheme, the one using channel compensation, since it is already known that

$$\tilde{x}''_{k,i} = \frac{x''_{k,i}}{1-2p_{c,k}} \tag{60}$$

is an unbiased estimator of x, we have

$$\bar{x}''_{k,i} = \frac{x''_{k,i}}{1-2p'_{c,k}} = \frac{1-2p_{c,k}}{1-2p'_{c,k}}\tilde{x}''_{k,i} \tag{61}$$

Hence, we can further get

$$\begin{aligned} b_{k,i} &= \frac{2(p'_{c,k}-p_{c,k})}{1-2p'_{c,k}}x \\ Var(\bar{x}''_{k,i}) &= \frac{(1-2p_{c,k})^2}{(1-2p'_{c,k})^2}(\sigma_m^2+\sigma_q^2+\sigma_{c,k}'^2) \end{aligned} \tag{62}$$

Note that the bias will dominate the performance of distributed estimation when

the number of sensors is large. The absolute value of the bias in the dithered quantization scheme is given by

$$|b| = \frac{\sum_{k=1}^{K}\sum_{i=1}^{N_k} 2p_{c,k}|x|/\sigma_k^2}{\sum_{k=1}^{K}\sum_{i=1}^{N_k} 1/\sigma_k^2} \tag{63}$$

The absolute value of the bias in the channel compensation scheme is given by

$$|b| \leq \frac{\sum_{k=1}^{K}\sum_{i=1}^{N_k} \frac{2|p'_{c,k}-p_{c,k}|}{1-2p'_{c,k}}|x|/\sigma_k^2}{\sum_{k=1}^{K}\sum_{i=1}^{N_k} 1/\sigma_k^2} \tag{64}$$

Hence a sufficient condition for the scheme with channel compensation to have a smaller upper bound is

$$\begin{aligned}
&\frac{2|p'_{c,k}-p_{c,k}|}{1-2p'_{c,k}} < 2p_{c,k} \\
&-p_{c,k}(1-2p'_{c,k}) < p'_{c,k}-p_{c,k} < p_{c,k}(1-2p'_{c,k}) \\
&0 < p'_{c,k} < \frac{2p_{c,k}}{1+2p_{c,k}} \\
&-p_{c,k} < p'_{c,k}-p_{c,k} < \frac{1-2p_{c,k}}{1+2p_{c,k}}p_{c,k}
\end{aligned} \tag{65}$$

Clearly, the bias will vanish as the sensors gain a perfect knowledge of $p_{c,k}$. Also, if $E[p_{c,k}-p'_{c,k}]=0$ and $E[|p_{c,k}-p'_{c,k}|] << p_{c,k}$, the estimator approaches unbiasedness. When an accurate estimate of $p_{c,k}$ is not achievable, with the same distance to the true value, an underestimate of $p_{c,k}$ is slightly better than an overestimate of $p_{c,k}$ because an overestimate will amplify the bias.

The absolute values of the residual coefficient under different $p_{c,k}$'s and $p'_{c,k}$'s are shown in Fig. 21. Also, Fig. 22 shows the comparison based on the mean square error of four different cases in a one-hop network when both the signal and noise are uniformly distributed with $U = 1$, $V = 0.25$, $L = 5$, and $N_1 = 50$. In the first case, $p_{c,k}$ is perfectly estimated; in the second case, $p'_{c,k} = 0.8p_{c,k}$; in the third case, $p'_{c,k} = 1.2p_{c,k}$; in the last case, $p'_{c,k}$ is uniformly distributed in the interval $[0.8p_{c,k}, 1.2p_{c,k}]$.

3.5 Energy Allocation

3.5.1 Energy Allocation by Each Sensor

We now consider the problem of allocating energy among different bits of the quantized message at each sensor for the channel compensation scheme. Assume the quantized message is of length L and the overall energy available for transmission is limited in order to maximize the lifetime of the network. Suppose the lth most significant bit m_l is repeated b_l times. The mean of the received replicas of the lth most significant bit is $u_l = \left(\sum_{n=1}^{b_l} \tilde{m}'_n\right)/b_l$. Define

$$\dot{x}''_{k,i} = \sum_{l=1}^{L} 2^{L-l} u_l \Delta - W \tag{66}$$

It is easy to verify that $\dot{x}''_{k,i}$ is an unbiased estimate of x. Let

$$\beta_k = \frac{p_{c,k}(1-p_{c,k})}{(1-2p_{c,k})^2} \tag{67}$$

and we have

$$\begin{aligned}
E[(\dot{x}''_{k,i} - x_{k,i})^2] &= E\left[\left(\sum_{l=1}^{L} 2^{2(L-l)} u_l \Delta^2 - \sum_{l=1}^{L} 2^{2(L-l)} m_l \Delta^2\right)^2\right] \\
&= \sum_{l=1}^{L} 2^{2(L-l)} E[(u_l - m_l)^2] \Delta^2 \\
&= \sum_{l=1}^{L} 2^{2(L-l)} \frac{\beta_k}{b_l} \Delta^2 \\
&= \sum_{l=1}^{L} \sum_{m=1}^{2^{L-l}} 2^{L-l} \frac{\beta_k}{b_l} \Delta^2 \\
&\geq \frac{(\sum_{l=1}^{L} 2^{L-l})^2 \Delta^2}{\sum_{l=1}^{L} \sum_{m=1}^{2^{L-l}} 2^{-(L-l)} \frac{b_l}{\beta_k}} \\
&= \frac{(\sum_{l=1}^{L} 2^{L-l})^2 \Delta^2}{\sum_{l=1}^{L} \frac{b_l}{\beta_k}} \\
&= \frac{(2^L - 1)^2 \Delta^2}{M/\beta_k}
\end{aligned} \tag{68}$$

The above inequality uses the result that the arithmetic mean is always no smaller than the harmonic mean. Furthermore, the equality holds if and only if

$$\frac{2^{L-l}}{b_l} = \frac{2^{L-s}}{b_s}, \forall 1 \leq l, s \leq L, l \neq s \tag{69}$$

This result is similar to the solutions proposed in [60] and [41], but is actually an extension of them because there are no assumptions about the channel noise or code distribution. However, the problem itself is a discrete optimization problem rather than a continuous optimization problem. Hence the above optimal solution is usually not achievable. To facilitate the analysis, we round the optimal solution to the nearest integer that is greater than the real value. Then, by using at most $L-1$ extra bits, the performance can be guaranteed. We can further get the total variance

$$\begin{aligned} E[(\dot{x}''_{k,i} - x)^2] &= E[(\dot{x}''_{k,i} - x_{k,i})^2] + E[(\dot{x}''_{k,i} - x)^2] \\ &\leq \frac{(2^L - 1)^2 \Delta^2}{M/\beta_k} + \frac{1}{4}\Delta^2 + \sigma_m^2 \\ &= 4\beta_k W^2/M + \frac{W^2}{(2^L - 1)^2} + \sigma_m^2 \end{aligned} \tag{70}$$

Note that the first term and the last term do not change with respect to L, and the middle term decreases as L increases. Also the bound on the possible total energy consumption increases as L increases. So the sensors have to trade quantization error for communication error. We look at two special cases here. If the single bit error rate is negligible, then $\beta_k = 0$ and

$$E[(\dot{x}''_{k,i} - x)^2] = W^2/(2^L - 1)^2 + \sigma_m^2 \tag{71}$$

The optimal energy allocation strategy is then: $b_l = 1, 1 \leq l \leq M$, so that $L = M$ and $E[(\dot{x}''_{k,i} - x)^2] = W^2/(2^M - 1)^2 + \sigma_m^2$.

The energy allocation problem among different sensors has been studied in [99] and [55]. The intuition here is when β_k is large, the bits are repeated multiple times in proportion to their weights; when β_k is small, each bit is transmitted only once. If

$2^L - 1 = M$,

$$E[(\dot{x}''_{k,i} - x)^2] = 4\beta_k W^2/M + W^2/M^2 + \sigma_m^2 \tag{72}$$

In this case, there are actually no differences among different bits and the fusion center sums up all the received bits to get a maximum likelihood estimate of the quantized message. To simplify the analysis, let $L = \lceil \log(M+1) \rceil$ here. The performance is bounded as before, and it may be improved by adding a few more bits.

Fig. 23 shows the mean square error of the different energy allocation schemes in the one-hop case when both the signal and the measurement noise are uniformly distributed with $U = 1$ and $V = 0.25$, respectively. Each sensor transmits $M = 7$ bits and there are $N_1 = 50$ sensors.

3.5.2 Energy Allocation in the One-Hop Case

We now consider the energy allocation problem among different sensors in a one-hop network with N_1 sensors. The goal is to minimize the MSE subject to a limit on the energy that can be used. The energy used is generally proportional to the total number of bits, B, that are transmitted.

Suppose there are a total of S sensors actually participating in the distributed estimation task and that the i'th one uses a total of $M_{1,i}$ bits to transmit its local estimate to the CH. Minimizing the MSE subject to an energy bound can then be formulated, following (51), as finding the $M_{1,i}$'s that achieve:

$$\begin{aligned} &\max \sum_{i=1}^{S} \frac{1}{4\beta_1 W^2/M_{1,i} + W^2/M_{1,i}^2 + \sigma_m^2} \\ &\text{subject to:} \quad \sum_{i=1}^{S} M_{1,i} \le B \\ &S \le N_1, \quad M_{1,i} > 0, \quad 1 \le i \le S. \end{aligned} \tag{73}$$

Since S is fixed, we can introduce a Lagrange multiplier and define:

$$G(M_{1,i},\lambda) = \sum_{i=1}^{S} \frac{1}{4\beta_1 W^2/M_{1,i} + W^2/M_{1,i}^2 + \sigma_m^2} + \lambda(\sum_{i=1}^{S} M_{1,i} - B). \tag{74}$$

Following standard techniques leads to the following optimality conditions:

$$\partial(\frac{1}{4\beta_1 W^2/M_{1,i} + W^2/M_{1,i}^2 + \sigma_m^2})/\partial M_{1,i} = 0 \tag{75}$$

$$\sum_{i=1}^{S} M_{1,i} = B \tag{76}$$

Obviously, each sensor participating in distributed estimation should transmit the same number of bits. Hence, the goal is to choose $M_{1,i}$'s that achieve

$$\max \frac{B}{(4\beta_1 W^2/M_{1,i} + W^2/M_{1,i}^2 + \sigma_m^2)M_{1,i}} \tag{77}$$

This objective is maximized when

$$4\beta_1 W^2 + W^2/M_{1,i} + \sigma_m^2 M_{1,i} \tag{78}$$

is minimized. When $B < (W/\sigma_m)N_1$, this minimum is achieved when $B(\sigma_m/W)$ sensors are randomly selected and each one transmits $M_{1,i} = W/\sigma_m$ bits. The overall error variance is then:

$$\frac{1}{B}(4\beta_1 W^2 + 2W\sigma_m) \tag{79}$$

To guarantee performance, each sensor can transmit $M_{1,i} = \lceil W/\sigma_m \rceil$ bits. If $B < W/\sigma_m$, the network can randomly select one sensor to send $M_{1,i} = B$ bits and let all other sensors keep silent. If $B > (W/\sigma_m)N_1$, the sensor can typically transmit $M_{1,i} = \lceil B/N_1 \rceil$ bits.

Fig. 24 compares the mean square error of different energy allocation schemes in the one-hop case when both the signal and noise are uniformly distributed with $U = 1$ and $V = 0.25$, respectively. In the cases shown, the energy constraint was $B = 100$ bits and there were $N_1 = 50$ nodes.

3.5.3 Energy Allocation in Multi-Hop Cases

We now consider the multi-hop case. In this case, the goal is allocate bits amongst rings and sensors within rings to minimize the energy required to achieve a MSE that is below a specified level.

We must first define a function $f(k)$ that is the energy expended to move a bit from ring k of the network to the CH. In this dissertation, we assume it takes exactly one transmission per hop, in which case $f(k) = \alpha k$ and we can let $\alpha = 1$. Depending on the approach used for communication, such as collaborative communication or a standard random access protocol, $f(k)$ may not be linear in k.

Now let $M_{k,i}$ be the number of bits used by the ith sensor in ring k to transmit its local estimate. The optimization problem can then be formulated as choosing the $M_{k,i}$'s that achieve:

$$\begin{aligned} &\min \sum_{k=1}^{K} \sum_{i=1}^{N_k} f(k) M_{k,i} \\ &s.t. \sum_{k=1}^{K} \sum_{i=1}^{N_k} \frac{1}{4\beta_k W^2 / M_{k,i} + W^2 / M_{k,i}^2 + \sigma_m^2} \geq T \\ &M_{k,i} \geq 0, \quad 1 \leq k \leq K, \quad 1 \leq i \leq N_k \end{aligned} \tag{80}$$

where T is fixed. Note that the total error variance is then upper bounded by $1/T$. To facilitate the implementation, suppose each sensor can only transmit $2^j - 1, 1 \leq j \leq Q$ bits, where Q is restricted by the residual energy of the sensor(s). Further suppose that each sensor can transmit different numbers of bits with different probabilities. To this end, define $p_{k,i,j}$ to be the probability that the ith sensor in the kth ring transmits $2^j - 1$ bits. Then, introducing this randomization into the above optimization problem

leads to the following one.

$$\begin{aligned}
&\min \sum_{k=1}^{K}\sum_{i=1}^{N_k}\sum_{j=1}^{Q} p_{k,i,j} f(k)(2^j-1) \\
&s.t. \sum_{k=1}^{K}\sum_{i=1}^{N_k}\sum_{j=1}^{Q} p_{k,i,j} \frac{1}{4\beta_k W^2/(2^j-1) + W^2/(2^j-1)^2 + \sigma_m^2} \geq T \\
&0 \leq p_{k,i,j} \leq 1, 1 \leq k \leq K, 1 \leq i \leq N_k, 1 \leq j \leq Q \\
&\sum_{j=1}^{Q} p_{k,i,j} \leq 1, \quad 1 \leq k \leq K, \quad 1 \leq i \leq N_k
\end{aligned} \tag{81}$$

This optimization problem can be solved by linear programming. The result can be converted back to a deterministic policy by letting $p_{k,i,j}N_k$ sensors in the kth ring transmit $2^j - 1$ bits.

If there is no limit on the maximum number of bits that each sensor can transmit, we can increase Q to track the solution of the linear program for each Q until all $p_{k,i,Q}$'s in the solution equal zero.

Fig. 37 shows the optimal energy consumption under different crossover probabilities of the BSC channel for a three-hop network as a function of the bound on the variance of the estimate. The signal for all cases was uniformly distributed on $[-U, U] = [-1, 1]$ and the noise was independent but also uniformly distributed on $[-V, V] = [-1, 1]$. There were 60 nodes distributed among the three rings as: $N_1 = 10$, $N_2 = 20$, and $N_3 = 30$. Table 2 shows the corresponding energy allocation schemes amongst rings and nodes when $1/T = 0.015$.

Fig. 26 compares the optimal energy consumption in a three-hop network as the limit V of the measurement noise (which is uniformly distributed on $[-V, V]$) is varied. The signal is again uniformly distributed on $[-1, 1]$, the channel crossover probability is $p_c = 0.05$, and the 60 sensors are again spread across the rings as $N_1 = 10$, $N_2 = 20$, and $N_3 = 30$. Table 3 shows the corresponding energy allocation schemes when $1/T = 0.015$.

There is considerable freedom in the above optimization problem to account for

Table 2: Comparison of the energy allocation schemes under different crossover probabilities of the BSC channel. Bits are allocated first to the sensors closest to the CH, with sensors in the third ring contributing more bits as the channel crossover probability p_c increases. Recall that p_{abc} is the probability that node b in ring a transmits $2^c - 1$ bits.

p_c	ring 1	ring 2	ring 3
0.02	$p_{1i3} = 1$	$p_{2i3} = 1$	$p_{3i2} = 0.2433$
0.04	$p_{1i4} = 1$	$p_{2i3} = 1$	$p_{3i2} = 0.5721$
0.06	$p_{1i4} = 1$	$p_{2i3} = 1$	$p_{3i2} = 0.6477, p_{3i3} = 0.3523$
0.08	$p_{1i4} = 1$	$p_{2i4} = 1$	$p_{3i2} = 0.6838, p_{3i3} = 0.3162$
0.10	$p_{1i5} = 1$	$p_{2i4} = 1$	$p_{3i3} = 0.9789, p_{3i4} = 0.0211$
0.12	$p_{1i5} = 1$	$p_{2i5} = 1$	$p_{3i2} = 0.1291, p_{3i3} = 0.8709$
0.14	$p_{1i5} = 1$	$p_{2i5} = 1$	$p_{3i4} = 0.9346, p_{3i5} = 0.0654$

Table 3: Energy allocation across rings and sensors as the measurement noise is varied. As the measurement noise becomes more severe (as V increases), sensors in the outer rings must contribute – but the energy cost of transmitting over multiple hops ensures that the sensors in the inner rings use the most bits. $p_c = .05$ and the MSE is held at 0.015.

V	ring 1	ring 2	ring 3
0.2	$p_{1i4} = 0.2817$		
0.4	$p_{1i3} = 0.9160$ $p_{1i4} = 0.0840$		
0.6	$p_{1i4} = 1$	$p_{2i3} = 0.1515$	
0.8	$p_{1i3} = 0.7452$ $p_{1i4} = 0.2548$	$p_{2i3} = 1$	
1.0	$p_{1i4} = 1$	$p_{2i3} = 1$	$p_{3i2} = 0.9441$
1.2	$p_{1i4} = 1$	$p_{2i3} = 0.6678$ $p_{2i4} = 0.3322$	$p_{3i3} = 1$
1.4	$p_{1i5} = 1$	$p_{2i4} = 0.9369$ $p_{2i5} = 0.0631$	$p_{3i4} = 1$

additional constraints. For example, energy tracking or scavenging algorithms might impose constraints on the maximum amount of energy that can be used by each ring of motes. An upper bound on the energy that can be used by ring k could take the form:

$$\sum_{i=1}^{N_k} \sum_{j=1}^{M} p_{k,i,j} f(k) 2^{j-1} < E_k \tag{82}$$

These additional constraints are also linear; hence, an optimal allocation that satisfies these constraints can still be found via a linear program.

3.6 Summary

We now discuss the impact of various modeling assumptions made in this chapter and briefly mention possible future work.

3.6.1 The Binary Symmetric/Asymmetric Channel Models

The channel models used in this chapter were the Binary Symmetric and Asymmetric Channels. They are reasonable choices because they cover some practical situations, even when the estimates the sensors make are included in packets with other data. The reason is that wireless sensors typically use very simple error control codes in order to minimize both the computation involved in preparing a packet and the number of bits transmitted. There will thus be a nontrivial probability of error for each bit of the packet.

The assumption that errors affecting successive bits are independent is also reasonable. Possible correlations between bits in the estimate that is being transmitted can be avoided by interleaving these bits with those of other data being carried in packets.

3.6.2 What Sensors and Clusterheads Must Know or Estimate

This chapter first addressed the issue of bias of the local estimates received by the CH. This bias is introduced when the sensors quantize their local estimates and when these

quantized estimates are transmitted over noisy wireless channels. We thus proposed novel but simple quantization and channel coding schemes to ensure that each local estimate received by the CH is unbiased.

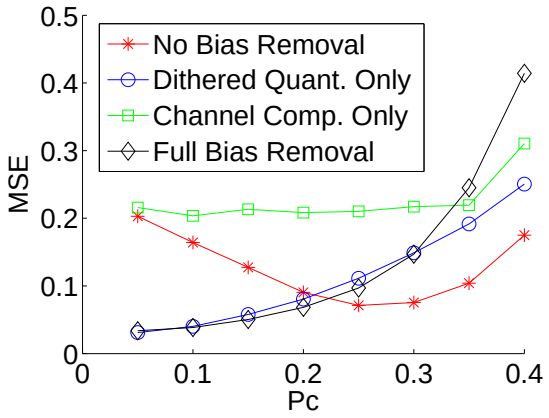
To implement these schemes, the clusterhead and the sensors must know the variance of the measurement errors made by the sensors. This is generally available either from models of the hardware used by the sensors or via estimation of it by the sensors. We also assumed that the CH knows the channel crossover probabilities for the wireless channels but then demonstrated that reasonably accurate estimates of them are sufficient for good performance. The sensors should thus be expected to estimate the bit-error rates on each link that they use, which is something they would usually be expected to do as part of any routing protocol for a wireless network.

With knowledge of the noise variance and estimates of the channel crossover probability, the bias in the received local estimates can thus be eliminated. The BLUE estimator can then be used to fuse these local estimates. The analysis of the variance of this estimator then led to techniques for optimal allocation amongst the sensors of the bits to be transmitted. The minimum energy required to achieve an error variance below a certain threshold could thus be found. Furthermore, an allocation of bits across rings and amongst the sensors in each ring that achieves this minimum energy was determined.

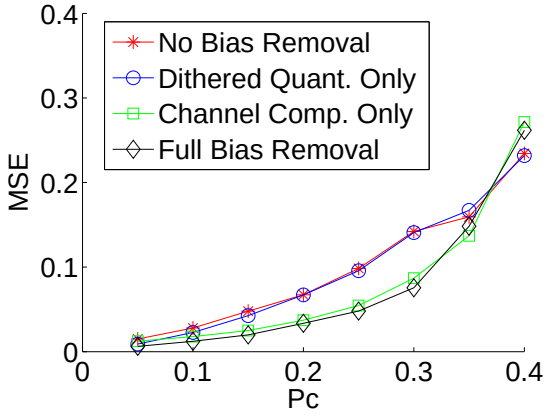
3.6.3 Future Work

One of the promising next steps for the work in this chapter is its use in mobile ad hoc networks. The tractability of the coding, quantization, estimation, and energy allocation schemes in this paper means that they can be used to rapidly study, via analysis and optimization, the effect of mobility of sensors on an application like estimation. For example, sensors that were near the same time and place may try to fuse their estimates to increase their accuracy. Because of the time spent gathering

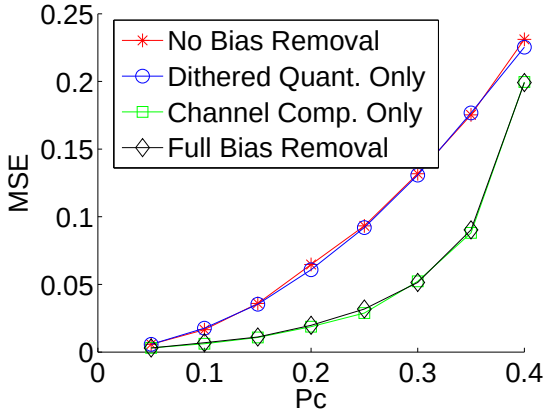
a sample, making measurements of the sample, and computing estimates based on these measurements, the sensors may have moved a substantial distance before their estimates are ready for transmission. The number of hops they have moved in this time will have a significant effect on the accuracy of the fusion algorithm. Our goal will be to quantify this effect as a function of time. Please see [89] for preliminary work on the effect of mobility on distributed detection applications.



(a) $L = 1$



(b) $L = 2$



(c) $L = 5$

Figure 20: Simulations comparing the mean square error of four different estimation schemes with different L's, where 2^L is the number of quantization levels used. *No Bias Removal* is the BLUE estimator without dithered quantization or channel compensation; *Dithered Quant. Only* is the BLUE estimator with dithered quantization; *Channel Comp. Only* is the BLUE estimator with channel compensation; *Full Bias Removal* is the BLUE estimator with dithered quantization and channel compensation.

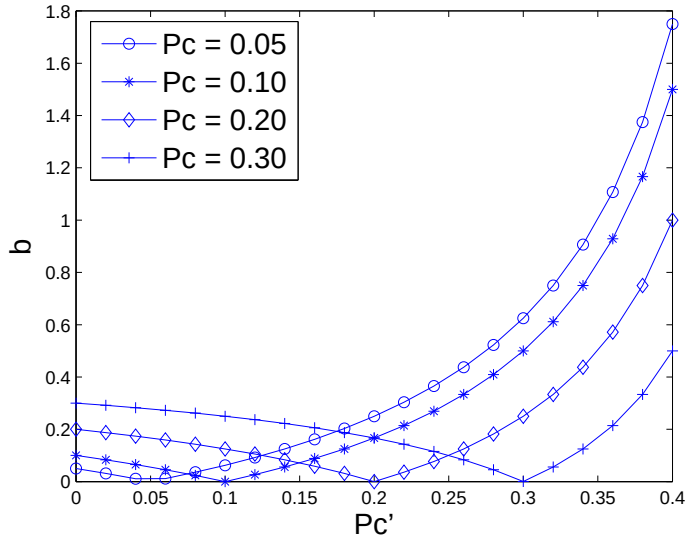


Figure 21: The bias after channel compensation when the estimate of the crossover probability of the BSC channel is not perfect.

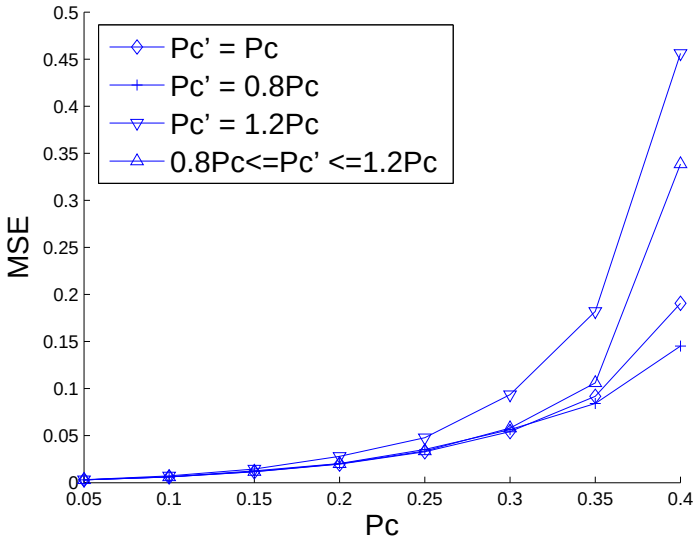


Figure 22: Comparison of the mean square errors under different estimates of the crossover probability of the BSC channel.

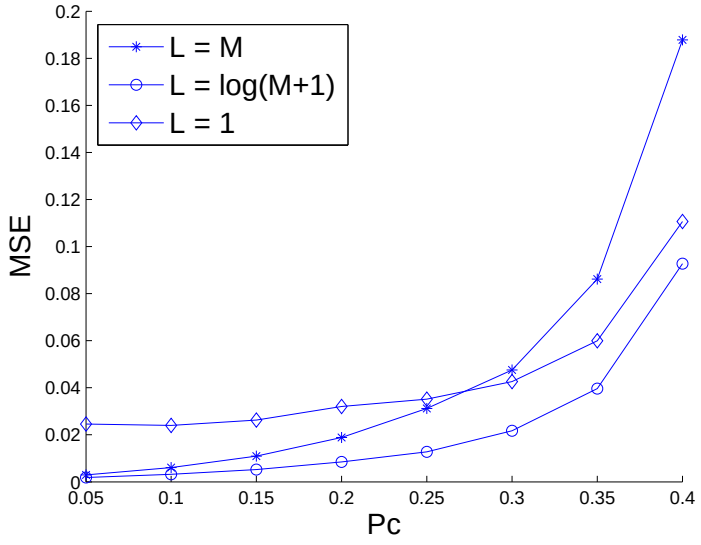


Figure 23: The mean square error as a function of the channel crossover probability for three energy allocation schemes. Each sensor transmits $M = 7$ bits. In the case $L = 1$, each sensor quantizes its local estimate to 1 bit and transmits it $M = 7$ times; for $L = M$, each sensor quantizes its local estimate to $M = 7$ bits and transmits each bit once. For the optimal case of $L = log(M + 1) = 3$, each sensor quantizes its estimate to 3 bits, repeats the most significant bit (msb) 4 times, the next msb 2 times, and the least significant bit once.

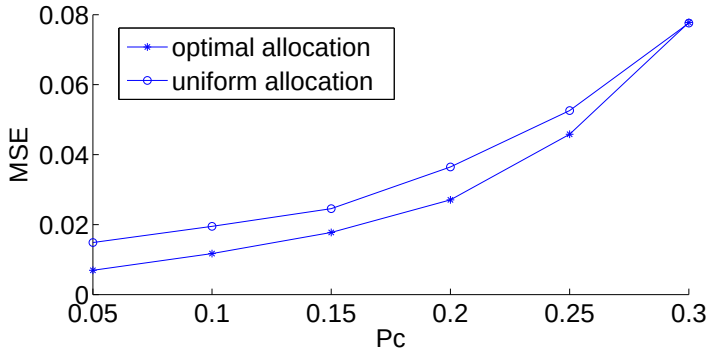


Figure 24: The mean square error of different energy allocation schemes for a one-hop network. The difference in performance between the optimal energy allocation scheme – each sensor quantizes its estimate to one bit and transmits it twice – and the uniform energy allocation scheme – each sensor quantizes its estimate to two bits and transmits each bit once – decreases as the channel becomes noisier.

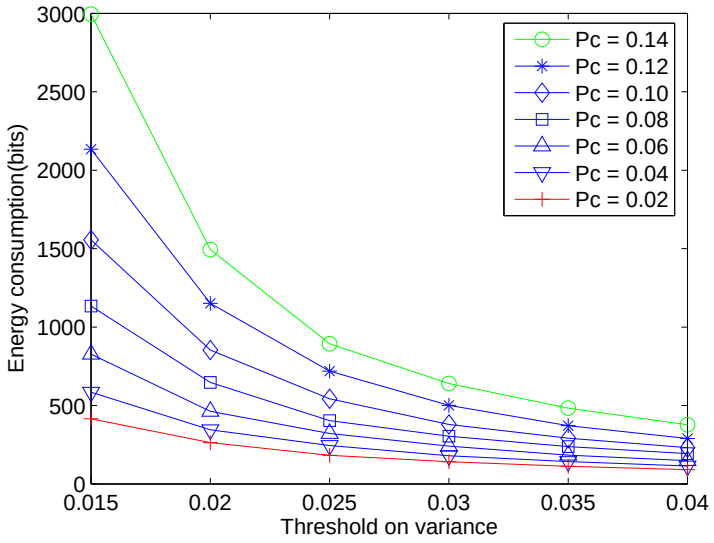


Figure 25: Minimum energy required in a three-hop network to achieve a specified MSE for different values of the crossover probabilities of the BSC channel. The network has 60 sensors: $N_1 = 10$, $N_2 = 20$, and $N_3 = 30$; the signal and measurement noise are both uniformly distributed on $[-1, 1]$.

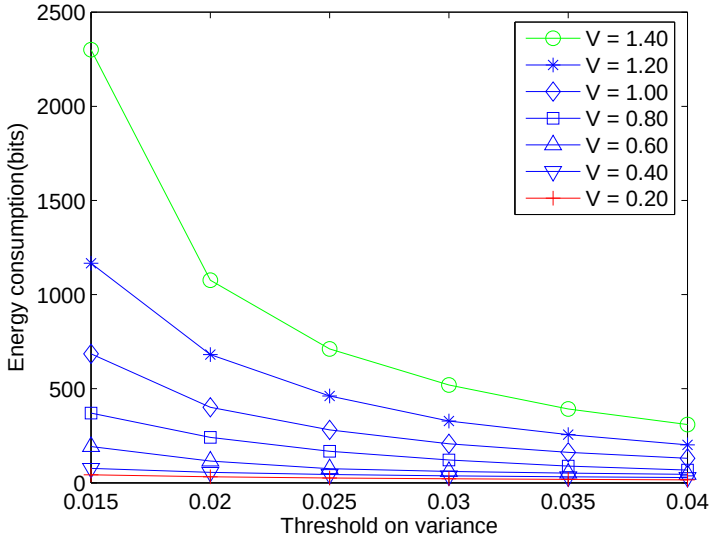


Figure 26: Minimum energy required in a three-hop network to achieve a specified MSE as the limit V on the measurement noise, which is uniformly distributed on $[-V, V]$, is varied. The network has 60 sensors: $N_1 = 10$, $N_2 = 20$, and $N_3 = 30$; the signal is uniformly distributed on $[-1, 1]$; and the channel crossover probability is $p_c = 0.05$.

CHAPTER IV

EFFECTS OF MOTION ON DISTRIBUTED DETECTION AND ESTIMATION IN MOBILE AD HOC SENSOR NETWORKS

4.1 Research Motivation

This chapter will focus on the effects of motion on distributed detection and estimation in mobile ad hoc networks. Our first goal is to develop algorithms to efficiently calculate the performance of fusion algorithms in this very dynamic and complex scenario. These numerical algorithms should account for as many factors as possible that could affect the system's performance, including: the motions that are typical of people in crowds, measurement errors in the sensors, errors made during wireless communication, and fusion algorithms with low enough complexity that they can be executed on very low-power processors or on processors that are shared by many different applications.

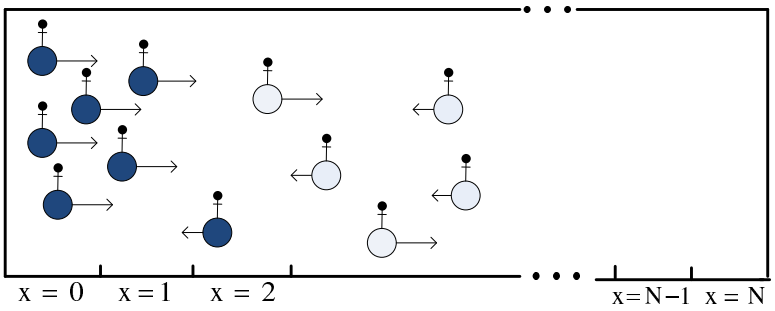
We will illustrate how to approach the analysis of such systems via a particular example. For this example, we assume a situation like that shown in Figure 1. The wireless nodes shown in the figure are carried by people who are moving. They have been moving for some time, generally toward the the stadium entrance, which is at position $x = N$ at the right end of each subfigure. Their motion to the right is thus strongly dominant over motion to the left; which is shown by arrows to the right that are larger than arrows pointing to the left. Each node may occasionally switch directions of motion and briefly move against the general flow, just as people do in a real crowd.

At time t_0 we assume without loss of generality that the sensors are in a configuration like that shown shown in Fig. 1(a). We assume that the sensors that are colored blue have each collected observations/samples that have been affected by a chemical or biological agent. It takes time for the sensors to process these observations/samples and the processors managing the sensors may have to finish other tasks before processing them. During this processing delay, the people carrying the sensors continue to move toward the stadium.

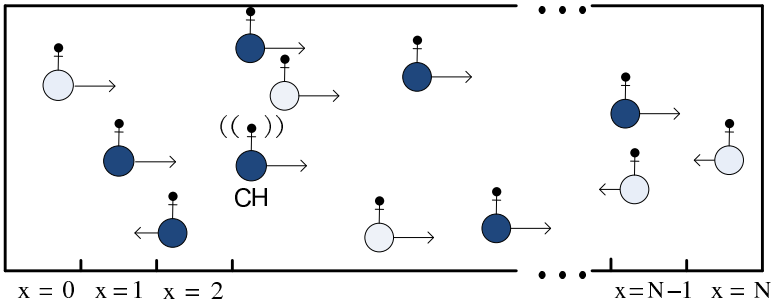
At time t_1 one of the sensors shown in Fig. 1(b) has finished processing its observations. The results satisfy some criterion – a preliminary positive detection or a preliminary estimate that exceeds a threshold – that requires that sensor to seek other sensor's results so they can be fused with its own to make a highly reliable decision/estimate. This sensor declares itself to be a clustered (CH) and floods the network of sensors – shown in Fig. 1(b) as one sensor transmitting at time t_2 – with a request for other nodes to send it their results if they were at approximately the same location and time that the CH acquired its samples. This flooding could be accomplished in many ways, including, for example, by the A-OLA cooperative, multihop broadcasting scheme proposed in [42]. This first request for data will suppress or supersede any other requests by any other sensors for data generated for this location and time. Thus, all other sensors with data about this event will forward it to the CH. For our analysis we will consider two situations: (i) the CH remains stationary; (ii) the CH continues to move with crowd.

At time t_3, the CH's request for reports have reached all other sensors, including those with nothing to report, as shown in Fig. 1(c). Sensors with data to report, shown in blue, begin transmitting their results to the CH. In some cases, the CH may be multiple wireless hops away, so intermediate nodes will relay these transmissions. We assume that no aggregation takes place as the nodes' data is relayed to the CH.

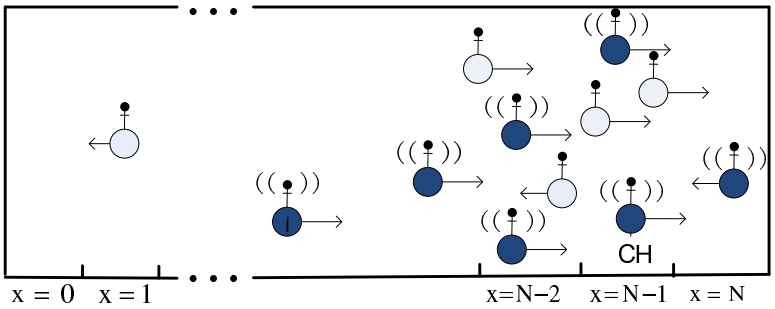
It is during this data collection phase that the behavior becomes quite interesting.



(a) The nodes at the time t_0 at which an event occurs. The nodes that make observations affected by the event are colored blue. All nodes are moving, generally toward the stadium entrance at the right, $x = N$, where they form a queue to enter the stadium.



(b) The nodes at time t_1, the time when the first node to process its observations decides that decisions and estimates from other nodes should be processed, collected and fused as soon as possible. It floods the group of sensors with a message with which it declares itself to be the CH, organizes the nodes into a multi-hop cluster to collect the data, and requests all nodes with relevant data – the ones colored blue – to report.



(c) The nodes at time t_3. All nodes that were affected by the event have, on request, completed processing their observations and are reporting their results back through the cluster to the CH.

Figure 27: A security scenario that arises when large crowds are moving toward a stadium or other venue to participate in a event, such as a football game, soccer match, or concert.

From the moment it starts we want to determine the best decision or estimate that the CH can make given the multi-hop network that exists at time t_3. If t_3 is close to t_1, the sensors would not have spread out very much, so each one may be within one hop of the CH. If t_3 is significantly greater than t_1, then the nodes may be either very spread out or, if they have all gotten close to the stadium entrance, have begun to bunch up again. In any of these cases, there will be communication errors, but more errors will occur when the nodes are spread out because additional transmissions are needed to relay the data. The algorithms we use for distributed detection and estimation account for these situations and can even determine how much energy each node should dedicate to sending its data – where the energy is measured in terms of the number of bits the node sends. For instance, it would determine how many bits each node that is n hops from the CH should use to encode each bit in its quantized estimate.

We develop analytical models of this scenario and *numerical methods* to calculate the best possible performance that could be expected at any time t. We are thus interested in the behavior of the network as it changes over time, where these changes are due to the motion of the crowd. The transient behavior of stochastic models of this motion are critical to understanding this scenario and eventually determining the optimal time at which to make a decision.

This chapter is organized as follows. Section 4.2 reviews the most relevant prior results in mobile ad hoc networks. In section 4.3, the transient behavior of semi-infinite correlated random walks is reviewed. The solution to the finite case, which is a good model for the motion of people the scenario of interest to us, is provided in section 4.4. Its statistical properties that have the greatest impact on detection and estimation applications are analyzed in section 4.5. In section 4.6 and 4.7, the performance of detection and estimation tasks in mobile wireless sensor networks are determined. Numerical results are provided in section 4.8. They demonstrate: (i)

the use of the mobility model in calculating the minimum energy required to collect data from each sensor participating in the data fusion process; (ii) the decision error probability at the cluster head once the decisions from all participating sensors have been gathered; and (iii) the minimum energy required to collect the estimates from all relevant sensors to achieve a prescribed estimation accuracy. These three quantities are functions of both time and the parameters of the mobility model and the ability to rapidly and accurately calculate them is valuable in developing optimal strategies form managing mobile sensor networks. Summary and a discussion of future work are provided in Section 4.9.

4.2 Most Relevant Prior Results

We must first model how sensors organize into a cluster, which must be done very rapidly in mobile, event-driven scenarios. We thus assume the existence of a very fast, low-complexity algorithm for clustering that is triggered by an event – in this case the request by one node for other nodes' data. The algorithm we choose is the one defined and analyzed in [15, 17]. We will only consider the the single-cluster case that is created in the scenario in this chapter; the more general case of hierarchical clustering is relevant if we need to collect results from *all* sensors participating in an event. In this more general case, many thousands of sensors may be involved; one for each person attending the event.

In the single-cluster scenario, the communication architecture is some variation of the 2-hop cluster shown in Fig. 28. The sensor in the center of the cluster, called the CH, is the one that requested the decisions/estimates from other sensors. The sensors shown in the first ring are one wireless hop from the CH; the ones in the second ring are two hops away, etc. Three sensors in a sector of ring 2 are shown forwarding their decisions to one sensor in ring 1. Each sensor's observation is affected by measurement noise and all communications throughout the network will

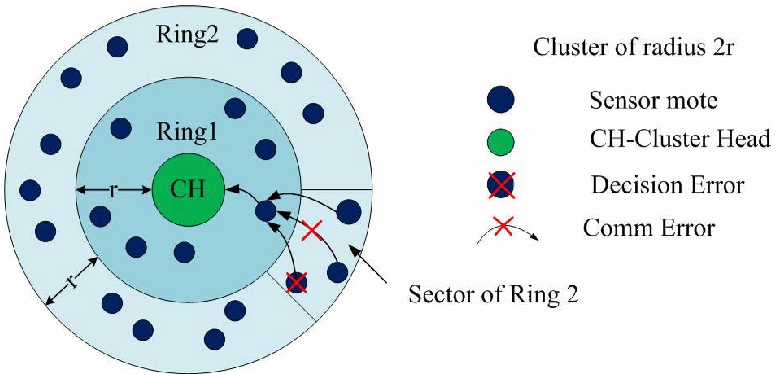


Figure 28: A 2-hop, 2-D cluster. The mobility of the sensors produces this temporary, multi-hop cluster for applications such as detection/estimation. As they move, the cluster changes size and more hops may be required for all relevant sensors to participate. Each sensor's decision/estimate may be incorrect because of measurement noise; transmitted packets may suffer bit errors because of noisy communication channels. If the number of hops in the cluster increases with time, the energy required for communication will increase and detection/estimation performance will decrease.

be affected by channel noise, fading, and transceiver errors. Because the cluster is multi-hop, decisions forwarded from the outer rings will suffer repeated exposure to the sources of communication error and more energy must expended to get them to the CH.

The mobility model governing each sensor and the time since the event of interest will clearly affect the number of hops, and thus both the detection/estimation accuracy at the CH and the energy consumed to achieve this level of accuracy. It will also affect the network's performance in terms of coverage, maximum throughput, and throughput-delay trade-offs.

Many mobility models have been proposed for analyzing the behavior of mobile ad hoc networks. In [14, 21], those models are categorized into four classes: random models, models with temporal dependency, models with spatial dependency, and models with geographic restrictions. For example, random walks and random waypoint models are random models; Gauss-Markov mobility models and Smooth Random models are models with temporal dependency; group mobility models [14] are models with spatial dependency; and, Pathway mobility models and Obstacle

mobility models are models with geographic restrictions. These models are useful as analytical or simulation tools but none of them can account for all of the constraints or features of real systems.

We thus use Correlated Random Walks (CRWs) as the model for the motion of the sensors in our security scenario. They can account for time dependency, geographical restrictions, and nonzero drift. They are more general than random walks but are still amenable to analysis, sometimes in closed-form. The limiting distributions for a discrete-time 1-D completely infinite CRW was derived in [32]. The probabilities of being at any lattice point at the nth step for a discrete-time 1-D CRW in different cases was found in [72, 19, 71, 48]. The absorbing probability and expected duration of a discrete-time CRW was found in [106]. The *transient behavior* of the continuous-time CRW on a semi-infinte, 1-D state space, is solved, in closed-form in some cases, in [16]. The behavior of multiple sensors whose motion is modeled by discrete-time processes related to CRWs has been studied in [20]. In this chapter, we use continuous-time, 1-D CRWs on finite state-spaces to calculate the behavior of mobile sensors in the security scenario described above.

Ultimately, our goal is to determine – by numerical analysis instead of simulation – the effect of the mobility of the nodes on the detection and estimation problem that exists at the application layer of the mobile, ad-hoc network. The performance of these applications will vary with time because of the nodes' motions. Understanding the significance of these transient effects, and being able to precisely determine such quantities as a lower bound on the probability of detection or MSE of an estimate, is critical to developing a strategy for reaching a decision meeting certain criteria in the minimum amount of time.

In [59], an optimal distributed detection strategy in a single-hop network was studied. In [61], a decentralized detection problem under bandwidth constrained communication was investigated. The noise at different sensors was assumed to be

independent, but the statistics of the noise were assumed to be unknown to the CH, so it treats all received detection results equally. It is shown in [24] how the performance of detection algorithms is improved by knowledge of the channel. The limits of detection performance in a one-hop sensor cluster with non-ideal channels were determined in [23].

In [66], the performance of distributed detection in a random sensor field is analyzed. [55] and [40] discuss the distributed detection problem for Gaussian signals under communication constraints. In [30], the authors consider the detection and localization problem of material releases with sparse sensor configurations. In [57], the sensors adopt robust binary quantizers for distributed detection. Censoring and sequential tests have also been adopted in distributed detection to achieve the same detection accuracy with less energy [8], [64], and they have been combined to further save energy [67], [102].

There are also many papers that consider the task of distributed estimation in sensor networks. [60] proposes an isotropic decentralized estimation scheme in a homogeneous environment. Each sensor transmits just one bit, with probability 1/2 to transmit the most significant bit, probability 1/4 to transmit the second-most significant bit, etc. [99] extends this result to the distributed estimation problem in an inhomogeneous environment by letting each sensor transmit a number of bits proportional to the logarithm of its local signal to noise ratio (SNR). [51] analyzed the optimal energy allocation problem in this case by giving priority to the transmissions with high energy efficiency.

[62] discusses optimal local estimation and final fusion schemes under the constraint that the communication from each sensor to the fusion center must be a one-bit message. In a similar case with a noisy communication channel, a MAP estimator and its variations are given in [12]. Each sensor is assumed to transmit a single bit that results from comparison of its local measurement with a pre-set threshold.

The information these bits provide decreases significantly when the preset threshold is not equal to the true but typically unknown signal value. To reduce the required number of sensors and the sensitivity to the choice of threshold, a modified scheme is provided in [79] in which the sensors use nonidentical thresholds that are uniformly distributed in an interval that must contain the true signal value. [76] and [75] propose to let the sensors transmit one bit when the noise is relatively high, and multiple bits when the noise is relatively low, both with nonidentical thresholds obtained via convex optimization techniques.

4.3 Basic Results on Transient Analysis of a Correlated Random Walk on $\{0, 1, \ldots, \infty\}$

In this section, we briefly review results on the transient probability distributions of continuous-time, 1D CRWs on $\{0, 1, 2, \ldots, \infty\}$[16]. In this case, the sensor moves according to the following rules:

It takes a step in the same direction as its previous step with probability p_1 or p_2 depending on whether its previous step was in the positive or negative direction, respectively. It takes a step in the opposite direction of its previous step with probability $q_1 = 1 - p_1$ or $q_2 = 1 - p_2$ depending on whether its previous step was in the positive or negative direction, respectively. On reaching a (reflecting) boundary, it takes a step in the opposite direction with probability one.

The time at which the sensor takes its next step is governed by a Poisson process of intensity λ.

A CRW on $\{0, 1, \ldots, \infty\}$ with a reflecting boundary at 0 can be modeled as a quasi-birth-death (QBD) process [65] with the state transition diagram shown in Fig. 29. The state $n-$ at any level n is entered when the sensor moves to location n from location $n+1$. The state $n+$ any level n is entered when the sensor moves to location n from location $n-1$.

Letting Q denote the generator for this QBD process, the Laplace transform $\Pi(s)$

of the vector $\pi(t)$ of transient probabilities for all states in the process satisfies the equations [104]:

$$\Pi(s)(Q - sI) = -\pi(0), \pi(0) = \pi(t)|_{t=0} \tag{83}$$

For this QBD process,

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ q_1\lambda & -\lambda & 0 & p_1\lambda & 0 & 0 & 0 & 0 & \dots \\ p_2\lambda & 0 & -\lambda & q_2\lambda & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & q_1\lambda & -\lambda & 0 & p_1\lambda & 0 & 0 & \dots \\ 0 & 0 & p_2\lambda & 0 & -\lambda & q_2\lambda & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & q_1\lambda & -\lambda & 0 & p_1\lambda & \dots \\ 0 & 0 & 0 & 0 & p_2\lambda & 0 & -\lambda & q_2\lambda & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & & \end{bmatrix} \tag{84}$$

For simplicity only, assume the initial position of the sensor is 0; i.e., $\pi(0) = [1, 0, 0, \cdots]$. Define $\Pi_n(s)$ to be the transform of $\pi_n(t) = [\pi_{n+}(t), \pi_{n-}(t)]$, the row-vector of transient probabilities for states on level n of the process. Then,

$$-(\lambda + s)\Pi_0(s) + \Pi_1(s) \begin{bmatrix} q_1\lambda \\ p_2\lambda \end{bmatrix} = -1, \tag{85}$$

$$\lambda\Pi_0(s) + \Pi_1(s) \begin{bmatrix} -(\lambda + s) \\ 0 \end{bmatrix} = 0, \tag{86}$$

$$\Pi_n(s)B(s) + \Pi_{n+1}(s)C(s) = 0, n \geq 1, \tag{87}$$

where

$$B(s) = \begin{bmatrix} 0 & p_1\lambda \\ -(\lambda + s) & q_2\lambda \end{bmatrix} \tag{88}$$

$$C(s) = \begin{bmatrix} q_1\lambda & -(\lambda + s) \\ p_2\lambda & 0 \end{bmatrix} \tag{89}$$

Thus we get

$$\Pi_{n+1}(s) = \Pi_n(s)W(s), \tag{90}$$

where

$$W(s) = -B(s)[C(s)]^{-1} = \begin{bmatrix} \frac{p_1\lambda}{\lambda+s} & -\frac{p_1q_1\lambda}{p_2(\lambda+s)} \\ \frac{q_2\lambda}{\lambda+s}\,\frac{\lambda+s}{p_2\lambda} & -\frac{q_1q_2\lambda}{p_2(\lambda+s)} \end{bmatrix}. \tag{91}$$

The boundary variables must also satisfy the following equation:

$$\Pi_1(s)v_1(s) = 0. \tag{92}$$

where $v_1(s)$ denotes the right eigenvector corresponding to the eigenvalue of $W(s)$ whose magnitude is greater than or equal to 1 for all possible values of s. For this CRW, $W(s)$ is diagonalizable and its eigenvalues are given by:

$$\gamma_k(s) = \frac{f(s) - (-1)^k\sqrt{(f(s))^2 - 4p_1p_2\lambda^2(\lambda+s)^2}}{2p_2\lambda(\lambda+s)} \tag{93}$$

where $f(s) = (p_1+p_2)\lambda^2 + 2s\lambda + s^2$, for $k = 1, 2$. The right eigenvectors $v_k(s)$, $k = 1, 2$, corresponding to the eigenvalues $\gamma_1(s)$ and $\gamma_2(s)$ are

$$v_k(s) = \begin{bmatrix} \frac{f(s)-2p_1p_2\lambda^2+(-1)^k\sqrt{(f(s))^2-4p_1p_2\lambda^2(\lambda+s)^2}}{2p_2(p_2-1)\lambda^2} \\ 1 \end{bmatrix} \tag{94}$$

We find numerically that, for all possible values of p_1, p_2, λ, and s, the magnitude of $\gamma_1(s)$ is greater than or equal to 1. Hence,

$$\Pi_{1+}(s) = \frac{2\lambda(1-p_2)}{A\lambda^2 + B(p_2)(s+\lambda)^2 + g(s)} \tag{95}$$

and

$$\Pi_{1-}(s) = \frac{f(s) - C\lambda^2 - \sqrt{(f(s))^2 - 2C\lambda^2(\lambda+s)^2}}{p_2\lambda\left(A\lambda^2 + B(p_2)(s+\lambda)^2 + g(s)\right)} \tag{96}$$

where $g(s) = \sqrt{(s+\lambda)^4 - [B(p_1)B(p_2)+1]\lambda^2(\lambda+s)^2 + A^2\lambda^4}$ and $A = p_1 + p_2 - 1$, $B(x) = 1 - 2x$, $C = 2p_1p_2$.

Thus by recursion,

$$\Pi_n(s) = \Pi_1(s)\tilde{W}^{n-1}(s), \tag{97}$$

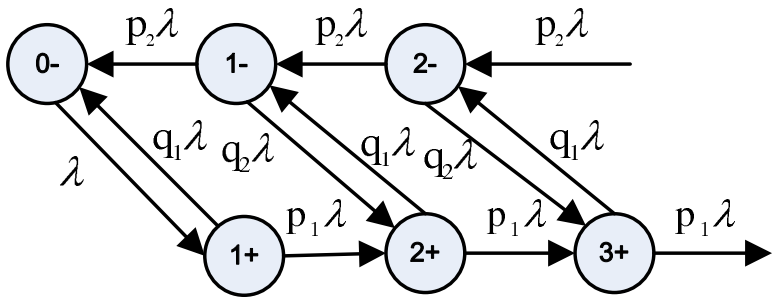


Figure 29: A Correlated Random Walk (CRW) on $\{0, 1, \ldots, \infty\}$.

where $\tilde{W}(s)$ is $W(s)$ after the mode whose eigenvalue is outside of the unit circle has been removed [104]. This ensures the stability of the recursion even when numerical errors occur in enforcing the orthogonality condition in Eq. 92.

4.4 New Results on Transient Analysis of Correlated Random Walks on $\{0, 1, \ldots, N\}$

A CRW on $\{0, 1, \ldots, N\}$ with reflecting boundaries at 0 and N can be modeled as a QBD process with the state transition diagram shown in Fig. 30 and the following generator:

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & 0 & 0 & \ldots & 0 & 0 \\ q_1\lambda & -\lambda & 0 & p_1\lambda & 0 & 0 & \ldots & 0 & 0 \\ p_2\lambda & 0 & -\lambda & q_2\lambda & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & q_1\lambda & -\lambda & 0 & p_1\lambda & \ldots & 0 & 0 \\ 0 & 0 & p_2\lambda & 0 & -\lambda & q_2\lambda & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & q_1\lambda & -\lambda & 0 & p_1\lambda \\ 0 & 0 & 0 & 0 & \ldots & p_2\lambda & 0 & -\lambda & q_2\lambda \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 & \lambda & -\lambda \end{bmatrix} \tag{98}$$

We assume w.l.o.g. that the initial position of the sensor is 0; i.e., $\pi(0) = [1, 0, \ldots, 0]$. Then we get

$$-(\lambda + s)\Pi_0(s) + \Pi_1(s) \begin{bmatrix} q_1\lambda \\ p_2\lambda \end{bmatrix} = -1, \tag{99}$$

$$\lambda\Pi_0(s) + \Pi_1(s) \begin{bmatrix} -(\lambda + s) \\ 0 \end{bmatrix} = 0, \tag{100}$$

$$\Pi_n(s)B(s) + \Pi_{n+1}(s)C(s) = 0, \quad 1 \leq n \leq N - 2. \tag{101}$$

$$\lambda\Pi_N(s) + \Pi_{N-1}(s) \begin{bmatrix} 0 \\ -(\lambda + s) \end{bmatrix} = 0, \tag{102}$$

$$-(\lambda + s)\Pi_N(s) + \Pi_{N-1}(s) \begin{bmatrix} p_1\lambda \\ q_2\lambda \end{bmatrix} = 0, \tag{103}$$

Thus, solving for $\Pi(s)$ in $\Pi(s)(Q - sI) = -\pi(0)$ is reduced to solving

$$[\Pi_0(s), \Pi_1(s), \Pi_{N-1}(s), \Pi_N(s)]\tilde{Q}(s) = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{104}$$

where, with I as 2x2 identity matrix,

$$\tilde{Q}(s) = \begin{bmatrix} -(\lambda + s) & \lambda & & & & \\ q_1\lambda & -(\lambda + s) & I & & & \\ p_2\lambda & 0 & & & & \\ & & W^{N-2}(s) & 0 & p_1\lambda \\ & & & -(\lambda + s) & q_2\lambda \\ & & & \lambda & -(\lambda + s) \end{bmatrix}$$

and the 2x2 matrix $W^{N-2}(s)$ can be computed, via the Cayley-Hamilton theorem, as a linear combination of $W(s)$ and I. It is then straightforward to calculate the

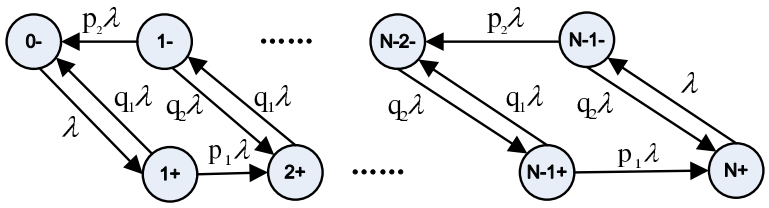


Figure 30: A Correlated Random Walk on $\{0, 1, \ldots, N\}$.

transient distributions of the process via stable algorithms for transform inversion [2, 4, 3].

For large values of t, the transient distribution for the finite CRW always converges to a unique limiting distribution. The limiting probability of the CWR on $\{0, 1, \ldots, N\}$ being at any location n is [48]:

$$\begin{aligned} \bar{\pi}_0 &= \frac{1}{2}\frac{(p_1/p_2) - 1}{(p_1/p_2)^N - 1}, \\ \bar{\pi}_n &= \bar{\pi}_0(1 + p_1/p_2)(p_1/p_2)^{N-1}, n = 1, 2, \ldots, (N-1). \\ \bar{\pi}_N &= \bar{\pi}_0(p_1/p_2)^{N-1}, \end{aligned} \tag{105}$$

For the special case of $p_1 = p_2 = p$, $\bar{\pi}_0 = \bar{\pi}_N = 1/(2N)$ and $\bar{\pi}_n = 1/N, 1, 2, \ldots, (N-1)$.

4.5 Statistical Properties of Distances of Nodes From the CH

Suppose the CH is motionless and there are a total of K sensors in the cluster. As described in the introduction, they start at the the same location as the CH ($n = 0$) and then move independently based on identical but independent CRW models. Define $x_i(t)$ as the location of the ith sensor at time t. Further define $d_{\max}(t)$ and $d_{\min}(t)$ as the maximum distance and the minimum distance between any sensor in the cluster and the CH, respectively. Then,

$$\begin{aligned} P(d_{\max}(t) = n) &= P(d_{\max}(t) \leq n) - P(d_{\max}(t) \leq n-1) \\ &= P(x_1(t) \leq n)^K - P(x_1(t) \leq n-1)^K \\ &= \left(\sum_{l=0}^{n} \pi_l(t)\right)^K - \left(\sum_{l=0}^{n-1} \pi_l(t)\right)^K \end{aligned} \tag{106}$$

and

$$\begin{aligned} P(d_{\min}(t)=n) &= P(d_{\min}(t)\geq n) - P(d_{\min}(t)\geq n+1) \\ &= P(x_1(t)\geq n)^K - P(x_1(t)\geq n+1)^K \\ &= \left(\sum_{l=n}^{N}\pi_l(t)\right)^K - \left(\sum_{l=n+1}^{N}\pi_l(t)\right)^K \end{aligned} \tag{107}$$

Now assume the CH is also mobile, that its position is governed by the same CRW as any sensor, and, as before, their motions are all independent. Then $d_{\max}(t)$ and $d_{\min}(t)$ become the maximum distance and the minimum distance between a specified sensor and any other sensor in the cluster. The probabilities remain the same whether the sensor is designated as the CH before or after motion of the sensors begins. Denote the position of the CH by $x_0(t)$. We have

$$\begin{aligned} &P(\ d_{\ \max}(t)=n) \qquad (108)\\ &= \sum_{m=0}^{N} P(x_0(t)=m)P(d_{\max}(t)=n|x_0=m) \\ &= \sum_{m=0}^{N} P(x_0(t)=m)[P(d_{\max}(t)\leq n|x_0=m) \\ &\quad -P(d_{\max}(t)\leq n-1|x_0=m)] \\ &= \sum_{m=0}^{N} P(x_0(t)=m)[P(m-n\leq x_i(t)\leq m+n, 1\leq i\leq K) \\ &\quad -P(m-n+1\leq x_i(t)\leq m+n-1, 1\leq i\leq K)] \\ &= \sum_{m=0}^{N}\pi_m(t)\left[\left(\sum_{l=m-n}^{m+n}\pi_l(t)\right)^K - \left(\sum_{l=m-n+1}^{m+n-1}\pi_l(t)\right)^K\right] \\ &P(\ d_{\ \min}(t)=n) \qquad (109)\\ &= \sum_{m=0}^{N} P(x_0(t)=m)[P(d_{\min}(t)\geq n|x_0=m) \\ &\quad -P(d_{\min}(t)\geq n+1|x_0=m)] \\ &= \sum_{m=0}^{N}\pi_m(t)\left(\sum_{l=0}^{m-n}\pi_l(t)+\sum_{l=m+n}^{N}\pi_l(t)\right)^K \\ &\quad -\sum_{m=0}^{N}\pi_m(t)\left(\sum_{l=0}^{m-n-1}\pi_l(t)+\sum_{l=m+n+1}^{N}\pi_l(t)\right)^K \end{aligned}$$

The probability distribution of the distance between any two sensors $d(t)$ is, for

$n = 0$:

$$P(d(t) = 0) = \sum_{m=0}^{N} \pi_m(t)^2 \tag{110}$$

and, for $n \neq 0$:

$$P(d(t) = n) = \sum_{m=0}^{N} \pi_m(t)(\pi_{m+n}(t) + \pi_{m-n}(t)) \tag{111}$$

Define $D_{\max}(t)$ and $D_{\min}(t)$ as the maximum distance and the minimum distance between any two sensors in the cluster. Then,

$$\begin{aligned} &P\left(D_{\max}(t) = 0\right) \\ &= \sum_{m=0}^{N} P(x_i(t) = m, 1 \leq i \leq K) = \sum_{m=0}^{N} \pi_m(t)^K \end{aligned} \tag{112}$$

and for $n \neq 0$,

$$\begin{aligned} &P(\, D_{\max}(t) = n) \\ &= \sum_{m=0}^{N-n} P(\min\{x_i(t)\} = m, \\ &\quad \max\{x_i(t)\} = m + n, 1 \leq i \leq K) \\ &= \sum_{m=0}^{N-n} \sum_{l=1}^{K-1} \sum_{r=1}^{K-l} \binom{K}{l} \binom{K-l}{r} \pi_m(t)^l \pi_{m+n}(t)^r \\ &\quad \left(\sum_{s=m+1}^{m+n-1} \pi_s(t)\right)^{K-l-r} \end{aligned} \tag{113}$$

where l sensors are at the location m and r sensors are at the location $m+n$, while all the others are in between. The calculation of the probability distribution of $D_{\min}(t)$ is much more complicated and, when $K > N$, it is always zero.

Thus, for both the motionless- and mobile-CH cases, computing $D_{\max}(t)$, $D_{\min}(t)$, $d_{\max}(t)$, and $d_{\min}(t)$ have been reduced to computing the transient distributions of CRWs. The computational techniques in Section II and III thus enable fast and stable computation of statistics of interest in the structure of a cluster of mobile nodes as it evolves over time.

4.6 Detection and Estimation in mobile ad hoc networks

The main results of distributed detection and estimation in wireless sensor networks are included in chapter II and III, respectively. Based on that, some results for the mobile case are derived here.

If the CH is static, then we have

$$\begin{aligned} P(E) &\leq E\left[\prod_{i=1}^{K} 2\sqrt{p_{e,i}(1-p_{e,i})}\right] \\ &= \prod_{i=1}^{K} E\left[2\sqrt{p_{e,i}(1-p_{e,i})}\right] \\ &= \left[E\left[2\sqrt{p_{e,1}(1-p_{e,1})}\right]\right]^{K} \end{aligned} \tag{114}$$

where

$$p_{e,1} = \frac{1}{2} - \frac{1}{2}(1-2p_m)(1-2p_c)^{\lceil x_1(t)/r \rceil} \tag{115}$$

The expected energy consumption for collecting the data is then

$$C = KE\left[\lceil x_1(t)/r \rceil\right] \tag{116}$$

If the CH is also mobile and its position is governed by a CRW, then we have

$$\begin{aligned} P(E) &\leq \sum_{m=0}^{N} P(x_0(t)=m) P(E|x_0(t)=m) \\ &= \sum_{m=0}^{N} P(x_0(t)=m) \prod_{i=1}^{K} E\left[2\sqrt{p_{e,i}(1-p_{e,i})}|x_0(t)=m\right] \\ &= \sum_{m=0}^{N} P(x_0(t)=m) \left[E\left[2\sqrt{p_{e,1}(1-p_{e,1})}|x_0(t)=m\right]\right]^{K} \end{aligned} \tag{117}$$

where

$$p_{e,1} = \frac{1}{2} - \frac{1}{2}(1-2p_m)(1-2p_c)^{\lceil d(t)/r \rceil} \tag{118}$$

The expected energy consumption for collection of the data is

$$C = KE\left[\lceil d(t)/r \rceil\right] \tag{119}$$

4.7 Numerical Results

Fig. 31 shows the transient probability distribution of two CRWs on [0,10] with different parameters. One CRW is asymmetric, which means the walker tends, in the case shown, to prefer motion to the right. The other CRW is symmetric, which means that the walker has no preference for one direction or the other – but its motion is still correlated, unlike in a standard random walk.

Fig. 32 shows the transient probability distribution of the distance from a static clusterhead to the furthest of five mobile sensors for each point in time. The static clusterhead is assumed to be at location 0. This distance d is important for calculating the number of wireless hops between each sensor and the CH that is trying to gather data. If the transmission radius of each sensor is r, then the number of hops to the CH is the ceiling of d/r.

Fig.33 shows the transient probability distribution of the maximum distance between any of the five sensors and the CH when the mobile CH moves according to the same mobility model as the sensors. This is probably the most realistic case for mobile networks. It leads to an interesting phenomenon when the motion is on a finite grid and the CRWs are all asymmetric in the same direction. The sensors start at 0 when they gathered their observations. With time, they spread out but they all tend to move to the right. As they get close to the other barrier, they then tend to bunch up again. This effect is similar to a crowd moving in a given direction and then gathering around that destination. The distance here, when divided by the transmission radius of the sensors tells the maximum size, in wireless hops, of the cluster.

Fig. 34 shows the expected energy consumed – as a function of the time at which data collection starts – for a static or mobile CH to collect one packet of data from five mobile sensors. The six sensors all collected measurements at the same time/place but then continued to move. By the time a request to send in all data has been received,

they are at locations in the state space with probabilities determined by their CRWs. The data collected and decisions made by the sensors that have wandered the furthest require the most energy to collect – that data must travel over the largest number of hops to reach the CH.

Also note that the mobility of the CH has a significant impact on the energy consumed to collect data. If it is mobile and all sensors are moving according to a CRW that is asymmetric, then the energy required for collection first increases and then decreases. In this case, it is better to collect data either very quickly after the event, or, if that is not possible because of the time to process measurements, to wait until the CH and the other sensors bunch up again at their destination.

Fig. 35 shows the large deviation bound of the error probability as a function of time for a static or mobile CH to collect one packet of data from five mobile sensors. It shows good agreement with the results in Fig.34, which shows the expected energy consumption as a function of time for a static or mobile CH to collect one packet of data from five mobile sensors.

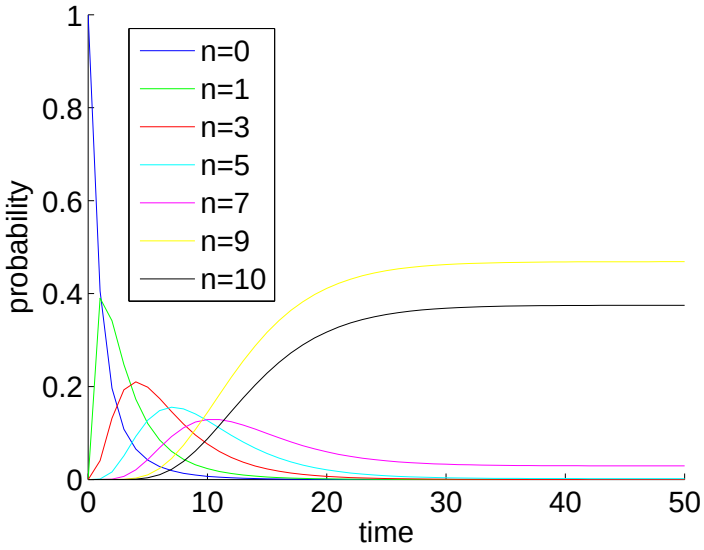
Fig. 36 shows the expected error probability as a function of time for a static or mobile CH to collect one packet of data from five mobile sensors. The same comments made previously about the effects of the mobility of the CH, and the effects of symmetric and asymmetric CRWs, apply here.

Fig. 37 shows the optimal energy consumption under different crossover probabilities of the BSC channel for a mobile ad hoc network as a function of time. The signal for all cases was uniformly distributed on $[-U, U] = [-1, 1]$ and the noise was independent but also uniformly distributed on $[-V, V] = [-1, 1]$. There were 100 nodes, and the MSE requirement is 0.015.

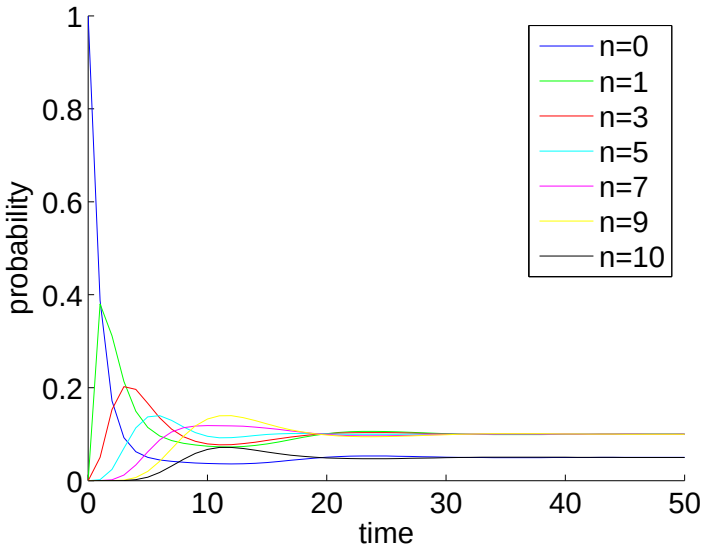
The peaks in Fig. 37 follow from the motion of the sensors. They all start at location 0 at time 0 and proceed, moving at most one spatial unit at a time, in the general direction of location 10. Each location might, for example, represent a

storefront on a city block and location 10 is the next traffic light or is a stadium entrance. They spend a random amount of time, whose average is 1 second, at each location before moving on. They may occasionally turn around but their preference is to move from location 0 to location 10. At between 10 and 15 seconds, they are maximally spread between locations 0 and 10, so data must travel more hops than at other times. After that, they begin bunching up again, but at location 10 instead of location 0. The energy used thus starts at a minimum at time zero, grows to a peak, and then declines toward the equilibrium values that would be expected when enough time has passed for the nodes' locations to be distributed according to the limiting distributions of their CRWs.

The behavior in Figure 37 is clearly quite complex. For low channel error probability, the energy required to produce the desired quality of estimate stays reasonably low and varies slowly with time. This is because, for most times, there are enough sensors within a few hops of the (moving) CH that their data, received essentially error-free, is sufficient for the CH to form its estimate – so the remaining sensors do not need to transmit at all. When the channel error probability is large, many more sensors must participate and they may be several hops away from the CH. It is then that the motion of the entire group of sensors comes into play, and then that the strange-at-first-glance dips and peaks appear in the energy required. These dips and peaks have to do with the times that the nodes that move fastest to the right or left start bouncing off the boundaries at 1 and 10. The important point to note in these cases is that up 300% more energy may be required in some cases to produce an estimate of a desired quality at a given time. It is thus very clear that the transient behavior of nodes in mobile networks can have a very significant impact on overall network performance.

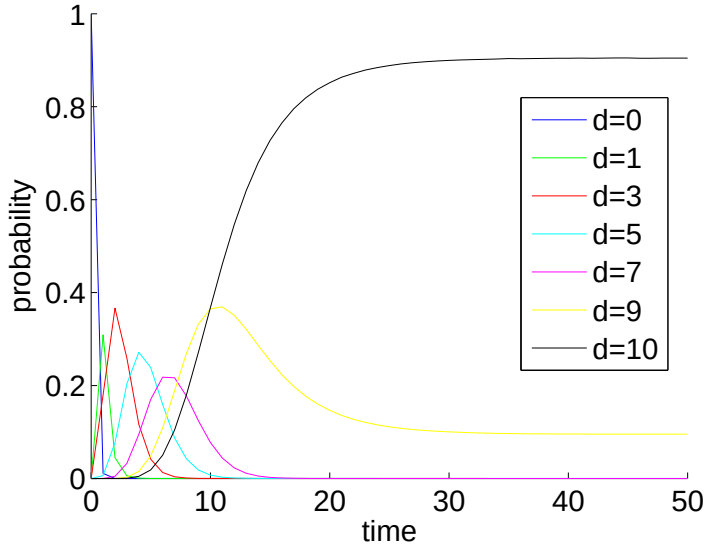


(a) An asymmetric CRW with p_1=0.8, p_2=0.2. The walker starts at 0 and then drifts toward and tends to stay near 10.

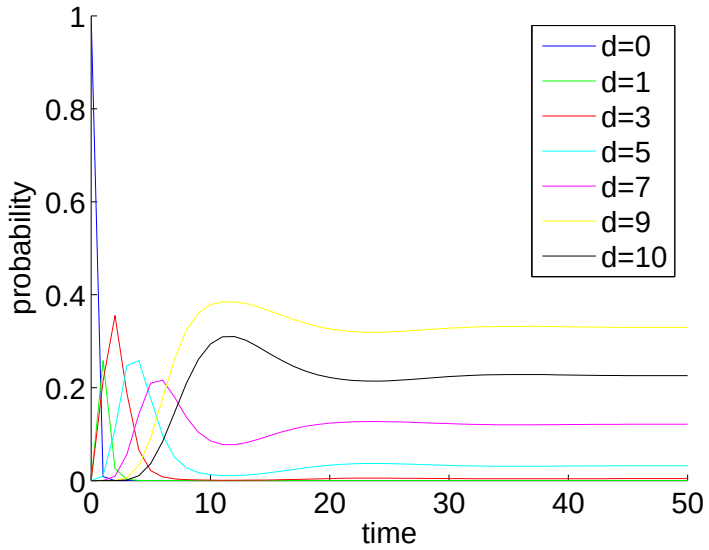


(b) A symmetric CRW with p_1=0.9, p_2=0.9. The walker starts at 0 and then it is eventually equally likely to be in each position except a boundary state.

Figure 31: Transient probability distribution of two CRWs on [0,10]. The curve for a given n shows the probability that the walker is at position n at time t.

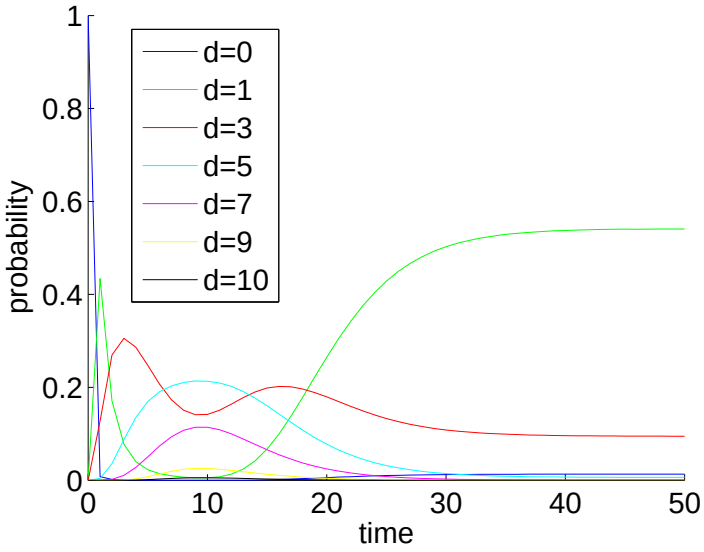


(a) All sensors move according to asymmetric CRWs with p_1=0.8, p_2=0.2

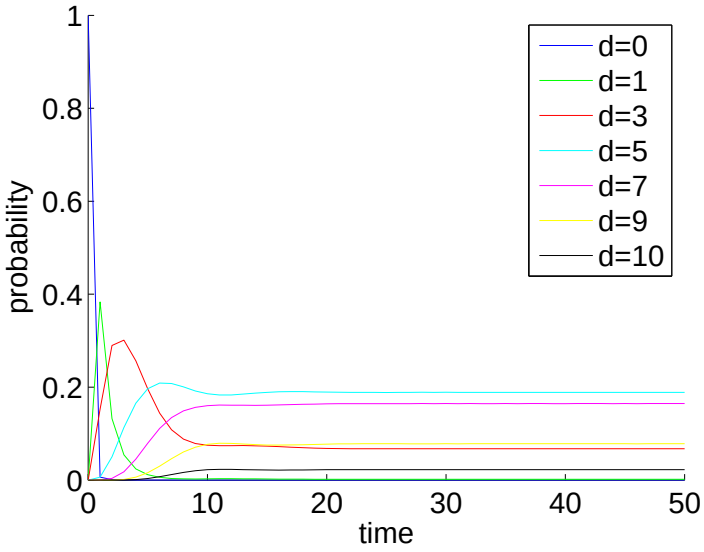


(b) All sensors move according to symmetric CRWs with p_1=0.9, p_2=0.9

Figure 32: Transient probability distribution of the distance from a static CH to the furthest of five sensors that are moving according to independent CRWs on [0,10].

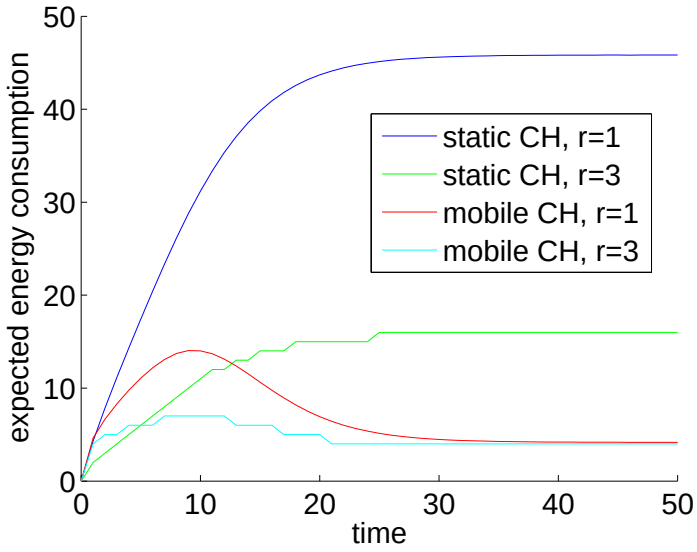


(a) All sensors move according to asymmetric CRWs with p_1=0.8, p_2=0.2

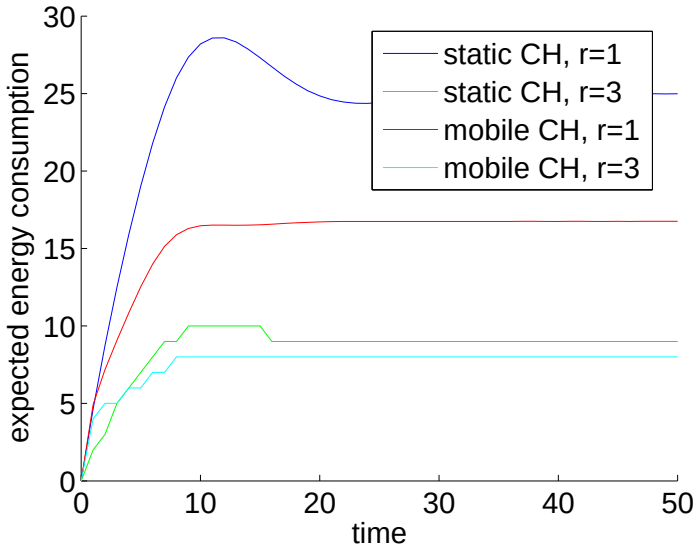


(b) All sensors move according to symmetric CRWs with p_1=0.9, p_2=0.9

Figure 33: Transient probability distribution of the distance from a mobile CH to the furthest of five sensors. The CH and the sensors are all moving independently and according to CRWs on [0,10].

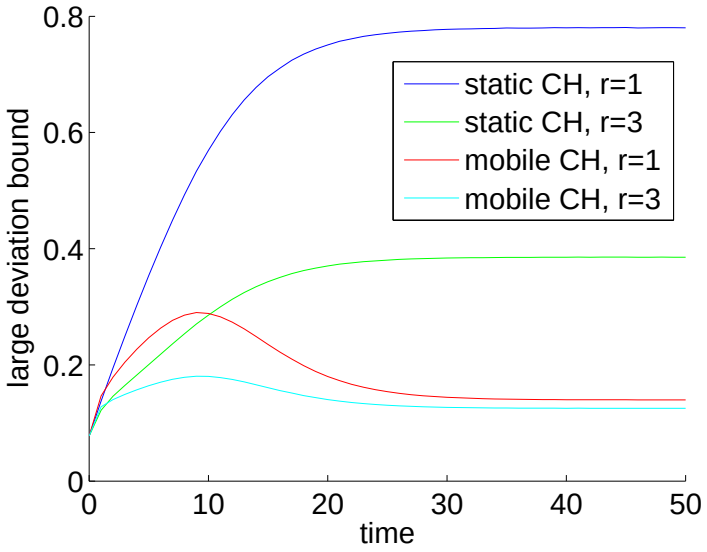


(a) All sensors move according to asymmetric CRWs with p_1=0.8, p_2=0.2

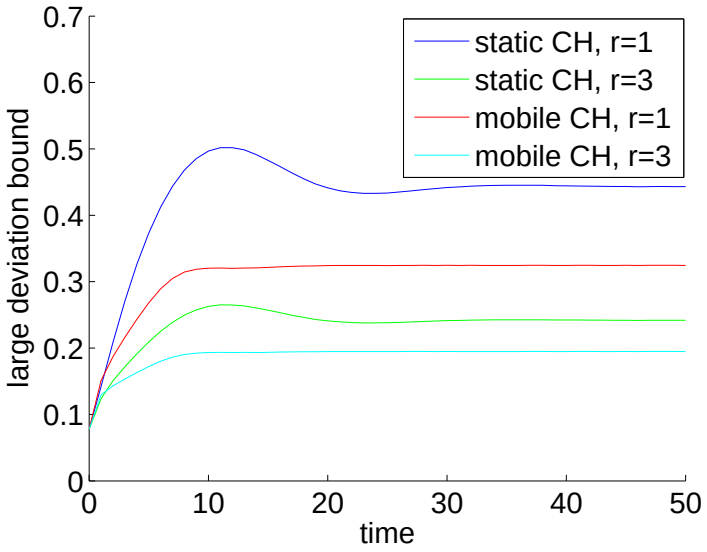


(b) All sensors move according to symmetric CRWs with p_1=0.9, p_2=0.9

Figure 34: Expected energy consumed by the network when the collection of data from the five mobile sensors begins at time t. The cases considered are when the transmission radius of each sensor is *r=1* or *r=3* and all sensors move according to symmetric or asymmetric CRWs.

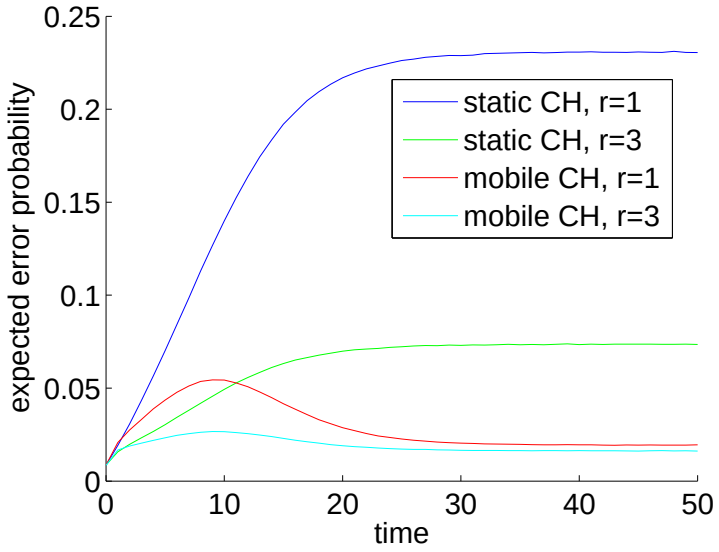


(a) All sensors move according to asymmetric CRWs with p_1=0.8, p_2=0.2

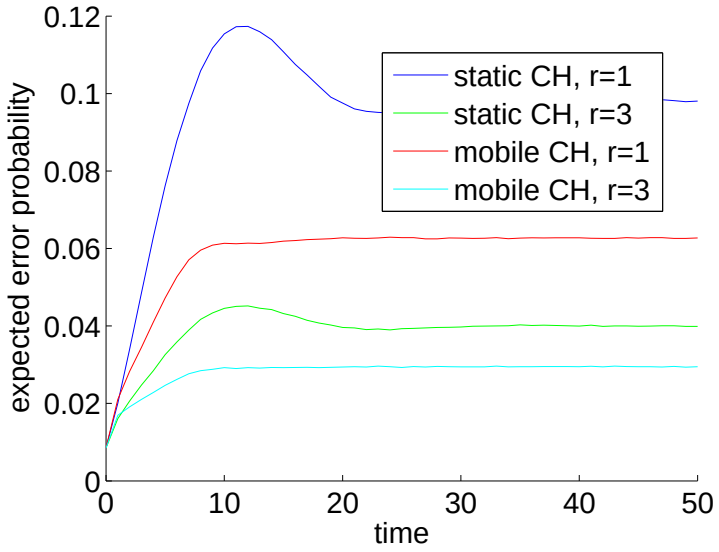


(b) All sensors move according to symmetric CRWs with p_1=0.9, p_2=0.9

Figure 35: Large deviation bound of the error probability when the collection of data from the five mobile sensors begins at time t. The cases considered are when the transmission radius of each sensor is *r=1* or *r=3* with $p_m = 0.10$ and $p_c = 0.05$ and all sensors move according to symmetric or asymmetric CRWs.

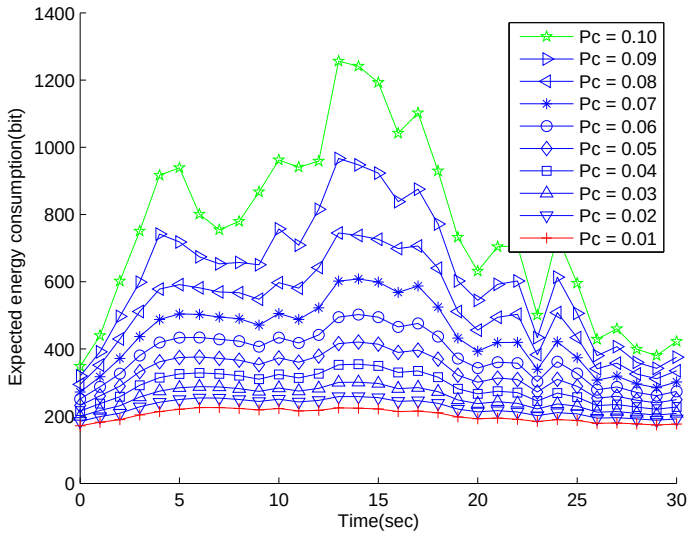


(a) All sensors move according to asymmetric CRWs with p_1=0.8, p_2=0.2

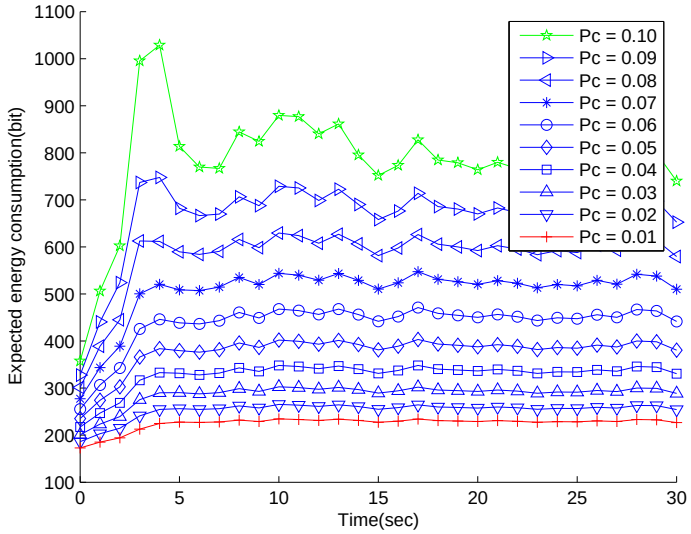


(b) All sensors move according to symmetric CRWs with p_1=0.9, p_2=0.9

Figure 36: Expected error probability when the collection of data from the five mobile sensors begins at time t. The cases considered are when the transmission radius of each sensor is *r=1* or *r=3* with $p_m = 0.10$ and $p_c = 0.05$ and all sensors move according to symmetric or asymmetric CRWs.



(a) All sensors move according to asymmetric CRWs with p_1=0.8, p_2=0.2



(b) All sensors move according to symmetric CRWs with p_1=0.9, p_2=0.9

Figure 37: Minimum energy, as a function of time, required in a mobile sensor network to achieve a specified MSE for different values of the crossover probabilities of the BSC channel. The MSE requirement is 0.015. The The network has 100 nodes that each start at position zero and move independently according to asymmetric CRWs on [0,10] with a mean rate of $\lambda = 1$ unit/sec. The transmission radius of each sensor is $r = 1$; the signal and measurement noise are both uniformly distributed on $[-1, 1]$. This scenario is a reasonable model of a group of people/vehicles who start moving after a traffic light turns and proceed to another traffic light where they again bunch up and wait. Some of them stop for some time, or turn-around occasionally.

4.8 Summary

In this chapter, the solution to the finite state space CRW was provided and its statistical behavior was studied both analytically and numerically. As an illustration of the application-specific performance measures it can help address, we studied the impact of motion on the error probability of the final decision at the CH and the minimum energy required to collect the estimates from all relevant sensors to achieve the prescribed estimation accuracy. Thus the temporary cluster head can decide the most appropriate time to call for data reports to help it make decisions/estimates within the allowed time frame and energy budget. In future research, we will study the impact of motion on the energy required to gather data and the variance/error probability of the final estimate/decision in complex mobile wireless sensor networks.

CHAPTER V

CONCLUSION

5.1 Summary of Contributions

In this dissertation, we studied optimal distributed detection and estimation algorithms in static and mobile ad hoc networks. Our major contributions are summarized as follows:

- Proposed and Analyzed a Local-Decision/Repetition Strategy for Distributed Detection in Clustered Sensor Networks.
- Compared, and Analyzed when Feasible, Hybrid, Hierarchical and Compression Strategies for in Clustered Sensor Networks Distributed Detection in Clustered Sensor Networks.
- Analyzed Dithered Quantization and Proposed and Analyzed Channel Compensation for Distributed BLU Estimation.
- Derived an Algorithm for Optimal Energy Allocation for Distributed Estimation in Clustered Sensor Networks.
- Derived Efficient Numerical Algorithms for Computing the Transient Analysis of Continuous-Time, 1-D, Finite-State-Space, Correlated Random Walks.
- Analyzed the Impact of Motion on System Performance and Energy Consumption for Distributed Detection/Estimation in Mobile Ad Hoc Networks.

Some of the most valuable insights gained from this work include the following: (a) In Chapter II, which was about distributed detection, we have shown when local decision fusion by relay nodes should be used to improve the final global decision at

the CH. Basically, when either p_c is large enough or there are enough hops that the end-to-end communication error becomes large enough, local decision/fusion should be used. (b) In Chapter III, which considered distributed estimation, it is first clear that if p_c becomes 0, then coding is not necessary and the CH need only collect data from enough nodes to overcome measurement error. It also became clear that if the support interval for the measurement noise decreases to zero (V becomes 0) for all nodes, then only one node in ring 1 needs to transmit its measurement to the CH, using sufficient coding to ensure the estimate arrives at the CH with no more than the allowed MSE being sought. (c) The results in Chapter III were combined with new analysis of Correlated Random Walks to show in Chapter IV that large values of the communication bit error probability, p_c, cause the energy requirement to compute the estimate to vary significantly with time. When this is the case, knowledge/models of how/where nodes tend to move are essential for efficiently computing estimates with a desired accuracy.

5.2 Future Research Directions

The objective of this dissertation was development of optimal distributed detection and estimation algorithms in static and mobile wireless sensor networks. Future research may be considered in the following three directions to make the algorithms proposed in the previous chapters more practical and adaptive.

One possible direction is to consider criteria other than energy consumption and detection/estimation accuracy. These may include: average delay or the probability distribution of delay in reaching a decision/estimate; average and peak bandwidth usage; and the contention in the MAC layer. We have analyzed the tradeoff involving accuracy, delay and energy consumption in the scenario of distributed detection/estimation in mobile ad hoc networks in Chapter IV. In most cases, targeting two or more criteria simultaneously will cause the overall optimization problem to

be multi-objective. A simple solution would be to aggregate all the separate objectives via a weighted sum, but more complex scenarios may also be needed because of Stern's effect [77].

Another possible direction is the modification of the system models we used to account for other characteristics of real systems. These could include the routing strategy, the MAC design, and the modulation scheme – i.e., other layers of the protocol stack – to determine the effects they have on performance. Accounting for some or all of these issues is highly likely to make the system model more complicated and thus the overall optimization problem more difficult to solve. For example, [54] is a good effort to incorporate route selection based on the observation that relaying by sensors in inner rings may be infeasible in late stages as their batteries will be exhausted faster than those in outer rings. Hence, simple distributed algorithms with guaranteed good performance will be an interesting topic to look at in the future.

Finally, sequential/adaptive algorithms of distributed detection and estimation are an important method to further reduce energy consumption.

There are thus many possible directions for future work that are made possible by the tools for modeling and analysis that are provided by this dissertation.

APPENDIX A

PROOF OF THEOREM 3 IN CHAPTER 2

Proof: Let $h(M, p)$ be the decision error probability when the median filter is applied to M i.i.d. bernoulli random variables, each with failure probability p. Also define $l(a, b) = a(1 - b) + b(1 - a)$. In the global decision strategy, the error probability of individual detection results received by the CH is then $l(p_d, p_c)$. Also note that in the local-decision/repetition strategy, the multiple copies of the local-decision result sent by the relay mote to the CH are not mutually independent but are conditionally independent given the local decision result. Let E_G and E_L denote the events of the decision error that happens at the CH using the pure-relay and local-decision/repition strategies, respectively. Then:

$$P(E_G) = h(M, l(p_d, p_c)) \tag{120}$$

$$P(E_L) = l(h(M, p_d), h(M, p_c)) \tag{121}$$

We first look at an example. Define $\Delta P = P(E_G) - P(E_L)$. When $M = 3$, $\Delta P = 6p_d p_c(2p_d - 1)(p_d - 1)(2p_c - 1)(p_c - 1)$. It is clear that $\Delta P > 0$, which confirms that making local decisions before forwarding improves performance in this case.

Now consider the case when M is an even number. The definition of the median filter in the even number case is ambiguous. Suppose the motes in a sensor network adopt a binary modulation scheme. In the case that the relay mote receives equal numbers of "0"s and "1"s, instead of sending 1/2 directly it sends "0" or "1" randomly, each with probability 1/2. We partition the M variables into two sets, with set U having $M - 1$ variables and set V having the remaining one. Denote the number of

the error bits in U and V by S_u and S_v, respectively. Therefore,

$$\begin{aligned} h(M,p) &= P(S_u \geq M/2) - \frac{1}{2}P(S_u = M/2)P(S_v = 0) \\ &\quad + \frac{1}{2}P(S_u = M/2-1)P(S_v = 1) \\ &= P(S_u \geq M/2) = h(M-1,p) \end{aligned} \tag{122}$$

It is thus not economical to use an even number of bits because one is wasted.

We now prove the hypothesis in the odd number case. We compare the two strategies when the source code of the repetition code is "1". Denote the number of "1"s received by the relay sensor by S. The decision error probability is,

$$P(E_G) = \sum_{D=0}^{M} P(E_G|S=D)P(S=D) \tag{123}$$

$$P(E_L) = \sum_{D=0}^{M} P(E_L|S=D)P(S=D) \tag{124}$$

where $P(S=D) = \binom{M}{D} p_d^{M-D}(1-p_d)^D$.

In the pure-relay strategy, D "1"s and $M-D$ "0"s are transmitted from the relay mote to the CH. In the local-decision/repetition strategy, the median filter is adopted at the relay mote. Since M is assumed to be odd, when $0 \leq D \leq \lfloor M/2 \rfloor$, the local decision result is "0", and M "0"s are relayed to the CH. When $\lceil M/2 \rceil \leq D \leq M$, the local-decision result is "1", and M "1"s are relayed. Denote the individual detection results received by the relay mote by s_i's, $1 \leq i \leq M$. We partition them into two sets. The received "0"s are assigned to T_0, while the received "1"s are assigned to T_1, i.e., $T_0 = \{i | s_i = 0, 1 \leq i \leq M, i \in N\}$ and $T_1 = \{i | s_i = 1, 1 \leq i \leq M, i \in N\}$. The cardinality of T_1 is D; for T_0 it is $M-D$.

In the pure-relay strategy, after the transmissions from the relay mote to the CH, among the elements in T_1, suppose A ones remain "1" and $D-A$ ones flip to "0". Also, among the elements in T_0, suppose B zeros flip to "1" and $M-D-B$ zeros remain "0". Certainly, all the (A,B) pairs are under the constraints that $0 \leq A \leq D$ and $0 \leq B \leq M-D$. If a pair (A,B) satisfies the condition $A+B \leq \lfloor M/2 \rfloor$, a

decision error will occur at the CH. We call such pairs decision error characteristic pairs and denote the set containing all these (A, B) pairs by F_0. Given M and p_c, the probability of achieving (A, B) from $(D, M-D)$ after transmission is:

$$\begin{aligned} & f_1(M, D, A, B, p_c) \\ &= \binom{D}{A}(1-p_c)^A p_c^{D-A}\binom{M-D}{B}p_c^B(1-p_c)^{M-D-B} \end{aligned} \tag{125}$$

In the local-decision/repetition strategy, partition the detection results received by the relay mote in the same way as pure-relay strategy. Assume $\lceil M/2 \rceil \leq D \leq M$, so all the s_i's that belong to T_0 are converted to "1" to transmit. The probability of achieving (A, B) from $(D, M-D)$ after the transmission is

$$\begin{aligned} & f_2(M, D, A, B, p_c) \\ &= \binom{D}{A}(1-p_c)^A p_c^{D-A}\binom{M-D}{B}(1-p_c)^B p_c^{M-D-B} \end{aligned} \tag{126}$$

Note that the following two equalities always hold,

$$f_2(M, D, D-A, B, p_c) = f_1(M, D, A, B, 1-p_c) \tag{127}$$

$$f_2(M, D, A, M-D-B, p_c) = f_1(M, D, A, B, p_c) \tag{128}$$

Denote the set containing all the (A, B) pairs satisfying both $A+B \leq \lfloor M/2 \rfloor$ and $A-B \leq \lfloor D-M/2 \rfloor$ by F_1. For all such pairs, we can further get $A+(M-D-B) \leq \lfloor M/2 \rfloor$, which implies that the pair $(A, M-D-B)$ is also a decision error characteristic pair. Denote the set containing all such $(A, M-D-B)$ pairs by F_2. Denote the set containing all the (A, B) pairs satisfying both $A+B \leq \lfloor M/2 \rfloor$ and $A-B \geq \lceil D-M/2 \rceil$ by F_3. For all such pairs, we can further get $(D-A)+B \leq \lfloor M/2 \rfloor$, which implies that the pair $(D-A, B)$ is also a decision error characteristic pair. Denote the set containing all such $(D-A, B)$ pairs by F_4. There is a one-to-one mapping between F_1 and F_2, and another one between F_3 and F_4. So $|F_1| = |F_2|$, and $|F_3| = |F_4|$. Apparently, $F_1 \bigcap F_3 = \emptyset$, and $F_1 \bigcup F_3 = F_0$. Thus, F_1 and F_3 are mutually exclusive

and collectively exhaustive subsets in F_0. Assume $F_2 \bigcap F_4 \neq \emptyset$, which means that at least one decision error characteristic pair falls into the intersection of the sets F_2 and F_4. Suppose this pair is denoted as $(A_1, M-D-B_1)$ and $(D-A_2, B_2)$ in F_2 and F_4, respectively. Then A_1, A_2, B_1, B_2 must satisfy both $D - A_2 = A_1$ and $B_2 = M - D - B_1$, which induces $A_1 + A_2 + B_1 + B_2 = M$. Since $A_1 + B_1 \leq \lfloor M/2 \rfloor$ and $A_2 + B_2 \leq \lfloor M/2 \rfloor$, $A_1 + A_2 + B_1 + B_2 \leq M - 1$. But this contradicts our assumption. Thus we have proved that $F_2 \bigcap F_4 = \emptyset$. $F_2, F_4 \in F_0$, and we simply have $|F_2 \bigcup F_4| = |F_2| + |F_4| = |F_1| + |F_3| = |F_1 \bigcup F_3| = |F_0|$. Therefore, we can declare that F_2 and F_4 are also mutually exclusive and collectively exhaustive subsets in F_0.

Define $g(M, D, p_c) = P(E_G|S = D) - P(E_L|S = D)$. For $\lceil M/2 \rceil \leq D \leq M$,

$$
\begin{aligned}
& g(M, D, p_c) \qquad (129)\\
&= \sum_{F_0} f_1(M, D, A, B, p_c) - \sum_{F_0} f_2(M, D, A, B, p_c) \\
&= \sum_{F_1} f_1(M, D, A, B, p_c) + \sum_{F_3} f_1(M, D, A, B, p_c) \\
&\quad - \sum_{F_2} f_2(M, D, A, B, p_c) - \sum_{F_4} f_2(M, D, A, B, p_c) \\
&= \sum_{F_1} f_1(M, D, A, B, p_c) + \sum_{F_3} f_1(M, D, A, B, p_c) \\
&\quad - \sum_{F_1} f_2(M, D, A, M-D-B, p_c) \\
&\quad - \sum_{F_3} f_2(M, D, D-A, B, p_c) \\
&= \sum_{F_3} \left(f_1(M, D, A, B, p_c) - f_1(M, D, A, B, 1-p_c) \right) \\
&= \sum_{F_3} \binom{D}{A} \binom{M-D}{B} p_c^{D-A+B} (1-p_c)^{M-(D-A+B)} \\
&\quad - \sum_{F_3} \binom{D}{A} \binom{M-D}{B} p_c^{M-(D-A+B)} (1-p_c)^{D-A+B}
\end{aligned}
$$

Since $D - A + B \leq \lfloor M/2 \rfloor$ for all (A, B) pairs in F_3, and $0 < p_c < 1/2$, each term

in the above equation is positive, so that the entire sum is positive. Note that for $\lceil M/2 \rceil \leq D \leq M$ and all (A, B) pairs in F_0,

$$f_1(M, M-D, B, A, p_c) = f_1(M, D, A, B, 1-p_c) \tag{130}$$

$$f_2(M, M-D, B, A, p_c) = f_2(M, D, A, B, 1-p_c) \tag{131}$$

Therefore,

$$\begin{aligned} & g(M, M-D, p_c) \\ &= P(E_G|S = M-D) - P(E_L|S = M-D) \\ &= \sum_{F_0} f_1(M, M-D, B, A, p_c) \\ &\quad - \sum_{F_0} f_2(M, M-D, B, A, p_c) \\ &= \sum_{F_0} f_1(M, D, A, B, 1-p_c) \\ &\quad - \sum_{F_0} f_2(M, D, A, B, 1-p_c) \\ &= g(M, D, 1-p_c) = -g(M, D, p_c) \end{aligned} \tag{132}$$

For each $\lceil M/2 \rceil \leq D \leq M$, we have $0 \leq M - D \leq \lfloor M/2 \rfloor$. Now we examine the difference between $P(E_G)$ and $P(E_L)$.

$$\begin{aligned} \Delta P &= P(E_G) - P(E_L) \\ &= \sum_{D=\lceil M/2 \rceil}^{M} P(S = M-D) g(M, M-D, p_c) \\ &\quad + \sum_{D=\lceil M/2 \rceil}^{M} P(S = D) g(M, D, p_c) \\ &= \sum_{D=\lceil M/2 \rceil}^{M} (P(S = D) - P(S = M-D)) g(M, D, p_c) \end{aligned} \tag{133}$$

When $\lceil M/2 \rceil \leq D \leq M$ and $0 < p_c < 1/2$, $P(S = D) - P(S = M-D) > 0$. We also know $g(M, D, p_c) > 0$ from (20). So we have proved that $\Delta P > 0$ for M odd, and thus for M even as well.

■

REFERENCES

[1] A. P. Chandrakasan, A. C. Smith, W. B. H. and Heinzelman, W. B., "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, pp. 660–670, 2001.

[2] Abate, J., Choudhury, G., and Whitt, W., *An Introduction to Numerical Transform Inversion and Its Application to Probability Models*. Kluwer, 1999.

[3] Abate, J. and Valko, P., "Multi-precision laplace transform inversion," *International Journal of Numerical Methods in Engineering*, vol. 60, pp. 979–993, 2004.

[4] Abate, J. and Whitt, W., "The fourier-series method for inverting transforms of probability distributions," *Queuing Systems*, vol. 10, pp. 5–88, 1992.

[5] Akkaya, K. and Younis, M., "A survey on routing protocols for wireless sensor networks," *Elsevier Ad Hoc Networks*, vol. 3, pp. 325–349, 2005.

[6] Akyildiz, I., W. Su, Y. S., and Cayirci, E., "Wireless sensor networks: A survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, 2002.

[7] Amis, A., Prakash, R., Vuong, T., and Huynh, D., "Max-min d-cluster formation in wireless ad hoc networks," in *Proceedings of the International Conference on Computer Communnications (INFOCOM'00)*, (Tel Aviv, Israel), pp. 32–41, IEEE, 2000.

[8] Appadwedula, S., Veeravalli, V., and Jones, D. L., "Decentralized detection with censoring censors," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1362–1373, 2008.

[9] Aravinthan, V., Jayaweera, S., and Tarazi, K., "Distributed estimation in a power constrained sensor network," in *IEEE Vehicular Technology Conference Spring 2006*, (Melbourne, Australia), pp. 1048–1052, 2006.

[10] Ault, A., Krogmeier, J., Dunlop, S., and Coyle, E., "estadium: The mobile wireless football experience," in *Proceedings of the International Conference on Internet and Web Applications and Services (ICIW'08)*, (Athens, Greece), pp. 644–649, IEEE, 2008.

[11] Aysal, T. C. and Barner, K. E., "Blind decentralized estimation for bandwidth constrained wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 1466–1471, 2008.

[12] Aysal, T. and Barner, K., "Constrained decentralized estimation over noisy channels for sensor networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1398–1410, 2008.

[13] Aysal, T., Coates, M., and Rabbat, M., "Distributed average consensus with dithered quantization," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4905–4918, 2008.

[14] Bai, F. and Helmy, A., *A survey of mobility modeling and analysis in wireless ad hoc networks*. Kluwer Academic Publishers, 2004.

[15] Bandyopadhyay, S. and Coyle, E., "Minimizing communication costs in hierarchically clustered networks of wireless sensors," *Computer Networks*, vol. 44, no. 1, pp. 1–16, 2004.

[16] Bandyopadhyay, S. and Coyle, E., "Stochastic properties of mobility models in mobile ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 11, pp. 1218–1229, 2007.

[17] Bandyopadhyay, S., Tian, Q., and Coyle, E. J., "Spatio-temporal sampling rates and energy efficiency in wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 13, no. 6, pp. 1339–1352, 2005.

[18] Blum, R., Kassam, S., and Poor, H., "Distributed detection with multiple sensors: Part ii–advanced topics," *Proceeding of IEEE*, vol. 85, no. 1, pp. 64–79, 1997.

[19] Bohm, W., "The correlated random walk with boundaries: a combinatorial solution," *Journal of Applied Probability*, vol. 37, pp. 470–479, 2000.

[20] Cai, H. and Eun, D.-Y., "Toward stochastic anatomy of inter-meeting time distribution under general mobility models," in *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing (MobiHoc 2008)*, (Hongkong, China), pp. 273–282, 2008.

[21] Camp, T., Boleng, J., and Davies, V., "A survey of mobility models for ad hoc network research," *Wireless Communnication and Mobile Computing, special issue on mobile and ad hoc networking: research, trends, and applications*, vol. 2, no. 5, pp. 483–502, 2002.

[22] Chamberland, J. F. and Veeravalli, V., "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *IEEE Jounral on Selected Areas in Communications*, vol. 22, no. 6, pp. 1007–1015, 2004.

[23] Chen, B., Tong, L., and Varshney, P., "Channel-aware distributed detection in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 23, no. 4, pp. 16–26, 2006.

[24] Chen, B. and Willett, P. K., "On the optimality of likelihood ratio test for local sensor decision rules in the presence of non-ideal channels," *IEEE Transactions on Information Theory*, vol. 51, no. 2, pp. 693–699, 2005.

[25] Chen, J. K., Yao, K., and Hudson, R., "Source localization and beamforming," *IEEE Signal Processing Magazine*, vol. 19, pp. 30–39, 2002.

[26] Chlamtac, I., Conti, M., and Liu, J. J.-N., "Mobile ad hoc networking: Imperatives and challenges," *Elsevier Ad Hoc Networks*, vol. 1, pp. 13–64, 2003.

[27] Dam, T. V. and Langendoen, K., "An adaptive energy-efficient mac protocol for wireless sensor networks," in *International conference on Embedded networked sensor systems(SenSys)*, (Los Angeles, CA), pp. 171–180, 2003.

[28] Demirkol, I., Ersoy, C., and Alagoz, F., "Mac protocols for wireless sensor networks: a survey," *IEEE Communications Magazine*, vol. 44, pp. 115–121, 2006.

[29] El-hoiydi, A., Decotignie, J.-D., and Hernandez, J., "Low power mac protocols for infrastructure wireless sensor networks," in *European Wireless Conference*, (Barcelona, Spain), pp. 563–569, 2004.

[30] Fox, E., Fisher, J., and Willsky, A., "Detection and localization of material releases with sparse sensor configurations," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1886–1898, 2007.

[31] Goldsmith, A. and Wicker, S., "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Communications Magazine*, vol. 9, no. 4, pp. 8–27, 2002.

[32] Goldstein, S., "On diffusion by discontinuous movements, and on telegraph equation," *The Quarterly Journal of Mechanics and Applied Mathmatics*, vol. 4, pp. 129–156, 1951.

[33] Gubner, J., "Distributed estimation and quantization," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1456–1459, 1993.

[34] Guibas, L., "Sensing, tracking, and reasoning with relations," *IEEE Signal Processing Magazine*, vol. 19, pp. 73–85, 2002.

[35] Heinzelman, W. R., Kulik, J., and Balakrishnan, H., "Adaptive protocols for information dissemination in wireless sensor networks," in *Annual ACM/IEEE international conference on Mobile computing and networking(MobiCom)*, (New York, NY), pp. 174–185, 1999.

[36] Huang, Y. and Hua, Y., "Multihop progressive decentralized estimation in wireless sensor networks," *IEEE Signal Processing Letters*, vol. 14, pp. 1004–1007, 2007.

[37] I.D. Schizas, A. R. and Giannakis, G., "Consensus in ad hoc wsns with noisy links–part i: distributed estimation of deterministic signals," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 350–364, 2008.

[38] I.D. Schizas, G. B. Giannakis, S. R. and Ribeiro, A., "Consensus in ad hoc wsns with noisy links–part ii: distributed estimation and smoothing of random signals," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1650–1666, 2008.

[39] Intanagonwiwat, C., Govindan, R., and Estrin, D., "Directed diffusion: A scalable and robust communication paradigm for sensor networks," in *Annual ACM/IEEE international conference on Mobile computing and networking*, (Boston, MA), pp. 56–67, 2000.

[40] Jayaweera, S., "Bayesian fusion performance and system optimization for distributed stochastic gaussian signal detection under communication constraints," *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1238–1250, 2007.

[41] Johnson, D. and Rodríguez-Díaz, H., "Optimizing physical layer data transmission for minimal signal distortion," in *Proceedings of IEEE International Conference Acoustics, Speech and Signal Processing (ICASSP'06)*, (Hong Kong, China), pp. 173–176, IEEE, 2003.

[42] Kailas, A. and Ingram, M. A., "Alternating opportunistic large arrays in broadcasting for network lifetime extension," *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, pp. 2831–2835, 2009.

[43] Kaplan, L., Le, Q., and Molnar, P., "Maximum likelihood methods for bearings-only target localization," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, (Salt Lake City, UT), pp. 554–557, 2001.

[44] Kapnadak, V., Senel, M., and Coyle, E., "Distributed iterative quantization for interference characterization in wireless networks," *Digital Signal Processing*, vol. 22, no. 1, pp. 96–105, 2012.

[45] Khajehnouri, N. and Sayed, A., "Distributed mmse relay strategies for wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3336–3348, 2007.

[46] Krasnopeev, A., Xiao, J.-J., and Luo, Z.-Q., "Minimum energy decentralized estimation in sensor network with correlated sensor noise," *EURASIP Journal on Wireless Communications and Networking*, vol. 5, no. 4, pp. 473–482, 2005.

[47] L. Xiao, S. B. and Kim, S.-J., "Distributed average consensus with least-mean-square deviation," *Journal on Parallel and Distributed Computing*, vol. 67, pp. 33–46, 2007.

[48] Lal, R. and U.N.Bhat, "Some explicit results for correlated random walks," *Journal of Applied Probability*, vol. 27, pp. 757–766, 1989.

[49] Lam, W. and Reibman, A., "Design of quantizers for decentralized systems with communication constraints," *IEEE Transactions on Commun. Theory*, vol. 41, no. 1, pp. 1602–1605, 1993.

[50] Li, D., Wang, K., Hu, Y., and Sayeed, A., "Detection, classification and tracking of targets," *IEEE Signal Processing Magazine*, vol. 19, pp. 17–29, 2002.

[51] Li, J. and AlRegib, G., "Rate-constrained distributed estimation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1634–1643, 2007.

[52] Li, J. and AlRegib, G., "Function-based network lifetime for estimation in wireless sensor networks," *IEEE Signal Processing Letters*, vol. 15, pp. 533–536, 2008.

[53] Li, J. and AlRegib, G., "Distributed estimation in energy-constrained wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 3746–3758, 2009.

[54] Li, J. and AlRegib, G., "Network lifetime maximization for estimation in multihop wireless sensor networks.," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2456–2466, 2009.

[55] Li, W. and Dai, H., "Distributed detection in wireless sensor networks using a multiple access channel," *IEEE Transactions on Signal Processing*, vol. 55, no. 3, pp. 822–833, 2007.

[56] Li, W. and Dai, H., "Energy-efficient distributed detection via multihop transmission in sensor networks," *IEEE Signal Processing Letters*, vol. 15, no. 1, pp. 265–268, 2008.

[57] Lin, Y., Chen, B., and Suter, B., "Robust binary quantizers for distributed detection," *IEEE Transactions on Wireless Communnications*, vol. 6, no. 6, pp. 2172–2181, 2007.

[58] Liu, J., Chu, M., and Reich, J., "Multitarget tracking in distributed sensor networks," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 36–46, 2007.

[59] Liu, K. and Sayed, A., "Optimal distributed detection strategies for wireless sensor networks," in *Proceedings of Annual Allerton Conference on Communication, Control and Computing*, (Monticello, IL), pp. 1–8, IEEE, 2004.

[60] Luo, Z., "An isotropic universal decentralized estimation scheme for a bandwidth constrained ad hoc sensor network," *IEEE Journal on Selected Areas in Commununications (JSAC)*, vol. 23, no. 4, pp. 735–744, 2005.

[61] Luo, Z., "Universal decentralized detection in a bandwidth constrained sensor network," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2617–2624, 2005.

[62] Luo, Z., "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2210–2219, 2005.

[63] Megalooikonomou, V. and Yesha, Y., "Quantizer design for distributed estimation with communications constraints and unknown observation statistics," *IEEE Transactions on Communications*, vol. 48, pp. 181–184, 2000.

[64] Mei, Y., "Asymptotic optimality theory for decentralized sequential hypothesis testing in sensor networks," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2072–2089, 2008.

[65] Neuts, M., *Matrix Geometric Solutions in Stochastic Models*. Johns Hopkins University Press, 1981.

[66] Niu, R. and Varshney, P., "Performance analysis of distributed detection in a random sensor field," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 339–349, 2008.

[67] P. Addesso, S. M. and Matta, V., "Sequential sampling in sensor networks for detection with censoring nodes," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5497–5505, 2007.

[68] Papadopoulos, H., Wornell, G., , and Oppenheim, A., "Sequential signal encoding from noisy measurements using quantizers with dynamic bias control," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 978–1002, 2001.

[69] Polastre, J., Hill, J., and Culler, D., "Versatile low power media access for wireless sensor networks," in *International conference on Embedded networked sensor systems(SenSys)*, (New York, NY), pp. 95–107, 2004.

[70] Pradhan, S., Kusuma, J., and Ramchandran, K., "Distributed compression in a dense microsensor network," *IEEE Signal Processing Magazine*, vol. 19, pp. 51–60, 2002.

[71] Proudfoot, A. and Lampard, D., "A random walk problem with correlation," *Journal of Applied Probability*, vol. 9, pp. 436–440, 1972.

[72] Renshawn, E. and Henderson, R., "The correlated random walk," *Journal of Applied Probability*, vol. 18, pp. 403–414, 1981.

[73] "Wireless sensor network technology trends report summer 2008," 2008. `http://www.researchandmarkets.com/reports/616680/wireless_sensor_network_technology_trends_report`.

[74] Rhee, I., A.Warrier, Aia, M., Min, J., and Sichitiu, M. L., "Z-mac: a hybrid mac for wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 16, no. 3, pp. 511–524, 2008.

[75] Riberio, A. and Giannakis, G., "Bandwidth-constrained distributed estimation for wireless sensor networks–ii: unknown probability density function," *IEEE Transactions on Signal Processing*, vol. 54, no. 7, pp. 2784–2796, 2006.

[76] Riberio, A. and Giannakis, G., "Bandwidth-constrained distributed estimation for wireless sensor networks–part i: Gaussian case," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1131–1143, 2006.

[77] Rosenblatt, M. J. and Sinuany-Stern, Z., "Generating the discrete efficient frontier to the capital budgeting problem," *Operations Research*, vol. 37, no. 3, pp. 384–394, 1989.

[78] Royer, E. M. and Toh, C.-K., "A review of current routing protocols for ad hoc mobile wireless networks," *IEEE Personal Communications Magazine*, vol. 6, no. 2, pp. 46–55, 1999.

[79] Senel, M., Kapnadak, V., and Coyle, E., "Distributed estimation for cognitive radio networks-the binary symmetric channel case," in *Proceedings of the Sensor, Signal and Information Processing Workshop (SenSIP'08)*, (Sedona, AZ), IEEE, 2008.

[80] Sun, X. and Coyle, E. J., "Error analysis for optimal distributed detection in multi-hop sensor networks," in *International Conference on Intelligent Sensors, Sensor Networks, and Information Processing (ISSNIP)*, (Melbourne, Australia), pp. 415–420, 2009.

[81] Sun, X. and Coyle, E. J., "The effects of mobility on applications in mobile ad-hoc sensor networks," in *IEEE Vehicular Technology Conference (VTC) 2010-Spring*, (Taipei, China), pp. 1–5, 2010.

[82] Sun, X. and Coyle, E. J., "Local decisions and optimal distributed detection in mobile wireless sensor networks," in *International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, (Avignon, France), pp. 584–590, 2010.

[83] Sun, X. and Coyle, E. J., "Optimal energy allocation for estimation in wireless sensor networks," in *IEEE International Communications Conference (ICC)*, (Cape Town, South Africa), pp. 1–6, 2010.

[84] Sun, X. and Coyle, E. J., "Optimal energy-aware distributed estimation in wireless sensor networks," in *IEEE International Conference on Computer Communications and Networks (ICCCN)*, (Maui, HI), pp. 1–6, 2011.

[85] Sun, X. and Coyle, E. J., "The effects of motion on distributed detection in mobile wireless sensor networks," *International Journal of Distributed Sensor Networks*, pp. 1–14, 2012.

[86] Sun, X. and Coyle, E. J., "Optimal distributed estimation in wireless sensor networks with spatially correlated noise sources," in *IEEE Wireless Communications and Networking Conference(WCNC)*, (Paris, France), pp. 1–6, 2012.

[87] Sun, X. and Coyle, E. J., "Quantization, channel compensation, and optimal energy allocation for estimation in wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 8, no. 2, pp. 1–25, 2012.

[88] Sun, X. and Coyle, E., "Quantization, channel compensation, and energy allocation for estimation in wireless sensor networks," in *Proceedings of Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt 2009)*, (Seoul, South Korea), pp. 1–10, 2009.

[89] Sun, X. and Coyle, E., "The effects of motion on distributed detection in mobile ad-hoc sensor networks," in *Proceedings of International Conference on Information Fusion*, (Edinburgh, UK), pp. 1–8, IEEE, 2010.

[90] Sun, X. and Coyle, E., "Low-complexity algorithms for event detection in wireless sensor networks," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 28, no. 7, pp. 1138–1148, 2010.

[91] Tian, Q. and Coyle, E. J., "Optimal distributed detection in clustered wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3892–3904, 2007.

[92] Tian, Q.-J. and Coyle, E., "Optimal distributed estimation in clustered sensor networks," in *Proceedings of IEEE International Conference Acoustics, Speech and Signal Processing (ICASSP'06)*, (Toulouse, France), pp. 157–160, IEEE, 2006.

[93] Tsitsiklis, J., "Decentralized detection," *Proceeding of Advanced Statistical Signal Processing*, vol. 2–Signal Detection, pp. 297–344, 1993.

[94] Viswanathan, R. and Varshney, P., "Distributed detection with multiple sensors: Part i–fundamentals," *Proceeding of IEEE*, vol. 85, no. 1, pp. 54–63, 1997.

[95] Willsky, A., Bello, M., Castanon, D., Levy, B., and Verghese, G., "Combining and updating of local estimates and regional maps along sets of one dimensional tracks," *IEEE Transactions on Automatic Control*, vol. 27, no. 1, pp. 799–813, 1982.

[96] Wu, J., Huang, Q., and Lee, T., "Energy-constrained decentralized best linear-unbiased estimation via partial sensor noise variance knowledge," *IEEE Signal Processing Letters*, vol. 15, pp. 33–36, 2008.

[97] Wu, J., Huang, Q., and Lee, T., "Minimal energy decentralized estimation via exploiting the statistical knowledge of sensor noise variance," *IEEE Transactions on Signal Processing*, vol. 56, pp. 2171–2176, 2008.

[98] Xiao, J.-J., Cui, S., Luo, Z.-Q., , and Goldsmith, A., "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 413–422, 2006.

[99] Xiao, J.-J. and Luo, Z.-Q., "Decentralized estimation in an inhomogeneous sensing environment," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3564–3575, 2005.

[100] Xiao, L. and Boyd, S., "Fast linear iterations for distributed averaging," *System and Control Letters*, vol. 53, pp. 65–78, 2004.

[101] Xiong, Z., Liveris, A., and Cheng, S., "Distributed source coding for sensor networks," *IEEE Signal Processing Magazine*, vol. 21, pp. 80–94, 2004.

[102] Yao, Y., "Group-ordered sprt for distributed detection," in *Proceedings of IEEE International Conference Acoustics, Speech and Signal Processing (ICASSP'06)*, (Las Vegas, NV), pp. 2525–2528, IEEE, 2008.

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

[104] Zhang, J. and Coyle, E., "Transient analysis of quasi-birth-death processes," *Communications in Statistics–Stochastic Models*, vol. 5, pp. 459–496, 1989.

[105] Zhang, Y. and Lee, W., "Intrusion detection techniques for mobile wireless networks," in *Annual international conference on Mobile computing and networking(MobiCom)*, (Boston, MA), pp. 275–283, 2000.

[106] Zhang, Y., "Some problems on a one-dimensonal correlated random walk with various types of barriers," *Journal of Applied Probability*, vol. 29, pp. 196–201, 1992.

[107] Zhao, F., Liu, J., Liu, J. J., Guibas, L., and Reich, J. E., "Collaborative signal and information processing: An information-directed approach," *Proceedings of the IEEE*, vol. 91, pp. 1199–1209, 2003.

[108] Zhao, F., Shin, J., and Reich, J., "Information-driven dynamic sensor collaboration," *IEEE Signal Processing Magazine*, vol. 19, pp. 61–72, 2002.

[109] Zhong, X., Chan, H.-H., Rogers, T., Rosenberg, C., and Coyle, E., "The development and estadium testbeds for research and development of wireless services for large-scale sports venues," in *Proceedings of the IEEE/CreateNet International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities (TRIDENTCOM'06)*, (Barcelona, Spain), pp. 340–348, IEEE, 2006.

[110] Zhong, X. and Coyle, E., "estadium: a wireless "living lab" for safety and infotainment applications," in *Proceedings of the International Conference on Communications and Networking in China (CHINACOM'06)*, (Beijing, China), pp. 1–6, IEEE, 2006.

VITA

Xusheng Sun received his B.S. degree in Automation and M.S. degree in System Engineering, both from Department of Automation at Tsinghua University, Beijing, China, in 2003 and 2006, respectively. He received his M.S. degree in Electrical and Computer Engineering in 2007, and is currently pursuing the Ph.D. degree in the School of Electrical and Computer Engineering, Georgia Tech, Atlanta, GA. He has been a Graduate Research Assistant with the Center for Signal and Image processing (CSIP) at Georgia Tech since 2008. He received Dong's Oriental Research Award from Tsinghua University in 2005 and CSIP outstanding Research Award from Georgia Tech in 2011. His research interests include sensor networks, signal processing, wireless communications, and estimation theory.