The aim of this research is to develop new theory and techniques that significantly improve performance of ad hoc networks by advanced signal processing and medium access control (MAC). In particular, we are interested in signal processing algorithms that are capable of separating signals from multiple users therefore provide multipacket reception (MPR) at the receiver, and MAC protocols that take advantage of the MPR capability of the physical layer.

During the reporting period, we investigated signal processing techniques with efficient implementation to strengthen the multiuser physical layer of ad hoc networks. We have extended our previous results to wireless sensor networks for which we considered the problem estimating the number of operating sensors and energy distribution. We also considered the problem of reconstructing random signal field using data samples collected from the sensor field. Blind and semi-blind channel estimation and symbol detection algorithms are developed for long code wide-band CDMA systems, including systems with multirate and multicode transmissions. A new identifiability condition is established, which suggests that channels unidentifiable in short code CDMA systems are almost surely identifiable when aperiodic spreading codes are used. Applications in large scale sensor networks are also considered.
Wireless Ad-Hoc Networks with Receiver Multipacket Reception
Performance Analysis and Signal Processing

Final Report Submitted to
The US Army Research Office
Contract No. ARO-DAAD19-00-1-0507

Sept 2000–Sept 2004

Lang Tong, PI
School of Electrical and Computer Engineering
Cornell University, Ithaca NY 14853
Tel: (607) 255-3900. Fax. (607) 255-9072.
Email: ltong@ece.cornell.edu
Contents

1 Forward 1
  1.1 Objectives and Overview ........................................ 1
  1.2 Honors and Awards ............................................. 1
  1.3 Transition and Synergistic Activities .......................... 2

2 Research Statements and Significant Findings 3
  2.1 Capacity and Throughput Analysis of Networks with MPR Nodes ...... 3
  2.2 Medium Access with MPR: Protocols and Signal Processing ............ 6
  2.3 Applications in Sensor Networks ................................ 14

3 Publications 17
  3.1 Journal Publications ........................................... 17
  3.2 Conference Proceedings ......................................... 18
  3.3 Manuscripts Under Review ...................................... 20

4 Participating Personnel 21

5 Report of Inventions 21

6 Bibliography 22
1 Forward

1.1 Objectives and Overview

This research investigates receiver-MultiPacket Reception (MPR) in wireless ad hoc networks. The objective of the research is twofold. First, a general theoretical framework is developed for the analysis of code division multiple access (CDMA) ad hoc networks whose nodes are capable of simultaneously detecting multiple packets. Potential improvements due to MPR in network throughput, delay, and stability are measured against the increased complexity of receiver implementations. Tradeoffs among spread spectrum bandwidth, error control, and packet error probability are examined. Comparisons are made between ad hoc and cellular networks. Also investigated are network capacity, effects of bandwidth expansion, receiver selectivity, and transmission protocols. Second, multipacket reception techniques are developed and analyzed for slow frequency hopping code division multiple access (SFH-CDMA) ad hoc networks. Packet formats, spatial and propagation diversities are exploited for signal processing based receiver MPR techniques. Resolvability of multipacket reception techniques is analyzed as a function of signal-to-noise ratio, multiaccess interference, and transmission protocols. Beyond what was proposed in the original proposal, we have also investigated the application of MPR techniques in sensor networks.

1.2 Honors and Awards

- 2001 IEEE Signal Processing Society Best Young Author Award.

- 2004 IEEE Signal Processing Society Best Paper Award for paper supported in part by this grant:

- IEEE Fellow. The PI is elected to the grade of Fellow by IEEE for his contribution to Statistical Signal Processing and its Applications in Communications and Wireless Networks.

- 2001 Cor Wit Visiting Professor, Delft University of Technology.
1.3 Transition and Synergistic Activities

In 2002, the PI spent 3-month sabbatical leave at ARL working with Dr. A. Swami and B. Sadler on applications of signal processing in ad hoc networks. Dr. Swami, Dr. Sadler and Mr. Misra visited the PI in July 2002 for a workshop focusing on adaptive signal processing and communications for wireless networks. With Dr. Swami, the PI co-chaired a workshop, cosponsored by NSF, ARO, and ONR, on future challenges of signal processing and communications in wireless network.

Professor Tong has extensive collaboration with ARL as an outside researcher of the ARL Collaborative Technology Alliances (ARL-CTA). His research focuses on novel cross layer techniques for mobile ad hoc networks (MANET) and signal processing for large scale wireless sensor networks. As part of the CTA rotation program, Cornell hosted three ARL scientists Dr. Ananthram Swami, Dr. Brian M. Sadler, and Mr. Saswat Misra during the summer of 2004.
2 Research Statements and Significant Findings

We outline in this section major contributions in this research. A position paper that articulate the scientific thrust is presented in [1].

2.1 Capacity and Throughput Analysis of Networks with MPR Nodes

Capacity and Stability Region of Ad Hoc Networks [2, 3, 4, 5, 6, 5]  The network capacity problem deals with finding the fundamental limits on achievable communication rates in wireless networks. Computing the network capacity requires an optimization with respect to medium access and routing policies. For general networks, this task is typically prohibitive due to the excessive dimensionality of the problem. As a result, one has to contend with certain asymptotics and order computations (e.g. [7, 8]).

In this research, we aim to provide an exception to this rule. Namely, we analyze certain regular networks for which we can compute the capacity explicitly. We obtain analytical expressions for the capacity, and find the leading coefficient besides the scaling law. The knowledge of the coefficient enables us to make comparisons between various design choices that affect the coefficient but not the scaling law. For example, we quantify the loss incurred by suboptimal protocols and the effects of increasing connectivity. We also find the optimal (i.e., capacity achieving) medium access and routing policies.

Before analyzing the capacity of regular networks, we provide a characterization of the capacity region of arbitrary networks. This characterization shows that every rate in the capacity region can be achieved by a class of policies that do randomized medium access and routing. Our formulation also suggests a natural outer bound on the capacity region in terms of the transport capacity of the network. To be able to consider regular networks, we extend the original definition of transport capacity [7] to networks with time-varying topology and arbitrary distance metric. The generalized transport capacity is used extensively in proving upper bounds on the capacity of regular networks.

The capacity formulation assumes that there are always packets waiting to be delivered at the source nodes. However, in reality data packets arrive randomly in time, and the node buffers should be kept stable for proper network operation. Using a general network model with time-varying topology, we find a mild condition on the reception channel under which the proposed randomized policies stabilize the node buffers for all arrival rates in the capacity region.

We obtained a complete characterization of the stability and capacity of networks with probabilistic multipacket reception models. We analyzed regular networks for which we
obtain closed-form expressions for the capacity of Manhattan networks (two-dimensional grid) and ring networks (circular array of nodes). We analyzed the performance loss due to suboptimal medium access and routing. We also investigated the impact of link fading, link state information, and variable connectivity on achievable rates in Manhattan networks. See [2, 3, 4, 5, 6, 5].

Stability and Delay of Finite User Slotted ALOHA with Multipacket Reception [9, 10, 11] It has been more than three decades since Abramson’s landmark work on ALOHA [12]. Much of what we know about slotted ALOHA is based on the so-called collision model: a transmission is successful if and only if a single user transmits. While a deterministic collision model is accurate for wire-line communications, it is inadequate to model probabilistic receptions in wireless multiple access. Furthermore, advances in multiuser detection and space-time processing make it necessary to have a multipacket reception model that captures the ability of the receiver to decode simultaneous transmissions and the probabilistic nature of reception.

Insights into the effect of MPR on ALOHA can be gained by examining two extreme cases: the collision channel and the orthogonal channel. For the collision model, the stability region is not convex; an increase in the maximum rate of one user implies a disproportionate decrease of the other. As a random access protocol, ALOHA is inferior to centralized TDMA since its stability region is contained inside that of TDMA. To stabilize any point in the rate region, transmission control is necessary by choosing transmission probabilities carefully. The onus of handling multiuser interference rests entirely with the random access protocol. The orthogonal channel, in contrast, models a physical layer that nullifies multiuser interference. As a result, the stability region takes the simple form of a unit square. There is no need for transmission control, and the rate for one user is independent of that of the other; ALOHA is optimal.

The orthogonal channel, of course, is not interesting for random access. What would be interesting are those cases when the multiuser interference affects the reception but not as severely as in the collision model. Can a distributed random access protocol such as ALOHA still be optimal? Is transmission control necessary? Is the stability region convex? A positive answer to the last question implies that given two stable rate pairs, all rate pairs on the line joining them are stable as well. What can we say about the performance of ALOHA for the general \(N\)-user system?

We consider in this work a general multipacket reception model. For each scheduled transmission, this model specifies a probability measure on the event space. We first give a
complete characterization of MAC capacity region. By MAC capacity we mean the maximum throughput achievable by any MAC protocol without considering queue stability. We show that this region is a convex hull of a set of marginal probabilities. In particular, the MAC capacity region is specified only by the marginal probabilities of success of individual users.

We consider next the ALOHA stability region. Obviously, the ALOHA stability region is contained in the MAC capacity region. We give a complete characterization for the two-user ALOHA system. We show that the stability region undergoes a distinct phase-transition, from a nonconvex region to a convex polyhedral region, from a strict subset of the capacity region to the exact capacity region (thus ALOHA is optimal). Furthermore, there is no need for transmission control once ALOHA is optimal. The same results hold for the symmetrical $N$-user system which has indistinguishable users with equal arrival rates. An inner bound for the general asymmetric $N$-user system is provided.

For a given rate vector, there are usually many transmission probabilities that stabilize the system. It is thus interesting to find the transmission probability that minimizes the average delay. We provide a complete delay characterization for the capture model in a symmetrical two-user system. Any nonzero probability of capture leads to a set of rates for which no transmission control minimizes the delay. As the probability of capture increases, the region of rates for which no transmission control minimizes the delay increases. As soon as the stability region becomes convex, no transmission control is delay-optimal for all stable arrival rates.

**Ad Hoc Networks vs. Centrally Controlled CDMA Wireless LAN: A Performance Comparison [13]** With the proliferation of the Internet, there is an increasing demand for providing broadband wireless access to offices and homes. Wireless Local Area Networks (LANs) require minimum infrastructure requirements and are becoming an attractive choice-of-technology for the emerging home and office networking market. A widely used architecture in wireless LAN which is a network centrally controlled by the base station (BS) where every user communicates with others through the BS. An alternative is the ad hoc architecture where each user communicates with others directly. In this paper, we present a performance comparison between these two architectures.

Wireless LANs in applications such as home/office networking often cover a small area with a relatively small number of nodes. This makes it reasonable to assume a one-hop topology that requires no dynamic reconfiguration. We will also restrict ourselves to packet-switched CDMA systems employing slotted Aloha random access protocols\(^1\). Again, the

\(^1\)If demand-assignment medium access control (MAC) protocols are used, the comparison here reflects
restriction to a small coverage area allows us to ignore the near-far effect. We also assume that the spreading code of every node is known to every other node after initialization.

A defining characteristic of CDMA is the possibility of receiving multiple packets at the same time. As demonstrated in this paper, the multipacket reception (MPR) capability of the network nodes can have a significant impact on the network performance, which affects the selection of network architecture. In contrast, for narrowband transmissions without using spatial and temporal diversity techniques, simultaneously transmitted packets are destroyed. Under this classical collision model, Aloha behaves the same in either a centrally controlled network or a fully connected ad hoc network.

While the behavior of the slotted ALOHA in a centrally controlled CDMA network is understood, little has been reported for ad hoc CDMA systems where nodes transmit directly to each other and any node can be a potential transmitter or receiver. One reason, as suggested in [14], is that the analysis becomes intractable due to the uncoordinated behavior of nodes in such networks. Therefore, the effect of the ad hoc architecture on the network performance has not been investigated. Furthermore, the effects of spreading gain and error control coding on the network performance are not fully understood.

In this research, we present analysis that overcomes this difficulty. Our contribution includes: (1) the performance of ad hoc systems employing the transmitter-based CDMA scheme is analyzed; (2) the effect of the network architecture on the network performance is studied by comparing the performance of an ad hoc system with that of a centrally controlled system; (3) impacts of the spreading gain and error control coding on the network performance of both systems, which also lead to the understanding of efficiency of bandwidth utilization in both systems, are investigated. Specifically, based on a finite population model, the network throughput, average packet delay and the First Exit Time (FET) of the two systems are derived and performance comparisons are evaluated. Furthermore, effects of spreading gain and error control coding on the network performance of both systems are quantitatively analyzed.

2.2 Medium Access with MPR: Protocols and Signal Processing

MAC for Wireless Networks with Multipacket Reception [15, 16, 17] In multiaccess wireless networks where a common channel is shared by a population of users, a key issue, referred to as medium access control (MAC), is to coordinate the transmissions of all users so that the common channel is efficiently utilized and the Quality-of-Service (QoS) requirement
of each user is satisfied. The schemes for coordinating transmissions among all users are called MAC protocols.

The conventional assumption on the channel is that any concurrent transmission of two or more packets results in the destruction of all the transmitted information. Based on this assumption, numerous MAC protocols, such as ALOHA [12, 18], the tree algorithm [19], the first-come first-serve (FCFS) algorithm [20], the window random access algorithm [21], and a class of adaptive schemes [22, 23, 24, 25], have been proposed. However, with the development of spread spectrum, space-time coding, and new signal processing techniques, this collision channel model does not hold in many important practical communication systems where one or more packets can be successfully received in the presence of other simultaneous transmissions. For instance, the capture phenomenon is common in local area radio networks. Other examples include networks using CDMA and/or antenna array, multiuser detection techniques, and signal processing based collision resolution algorithms [26].

In this research, we propose a MAC protocol designed explicitly for MPR channels. A slotted network with a finite population of users is considered. Users may have different QoS requirements which are characterized by their average packet delay at the heaviest traffic load. Since, in general, packet delay increases with the traffic load, this delay constraint specifies the worst case performance of the network. The proposed protocol maximizes the per-slot throughput (the expected number of successfully received packets in each slot) while ensuring each user’s QoS requirement. The key to maximizing per-slot throughput is an optimal estimate of the state of users. By fully exploiting the information provided by previous channel outcomes, the state of each user is updated at the beginning of each slot. Based on the inferred user state, an appropriate access set which consists of users who gain access to the channel is chosen to maximize the expected number of successfully received packets in each slot under the heterogeneous delay constraints. The proposed protocol achieves the maximum possible throughput among all protocols at heavy traffic load and has small delay when the traffic load is light.

A dynamic medium access control (MAC) protocol is proposed for a finite-user slotted channel with multipacket reception (MPR) in [16]. This protocol divides the time axis into transmission periods (TPs) where each TP is dedicated to the transmission of the packets generated in the previous TP. At the beginning of each TP, the state (active or idle) of each user is estimated based on the length of the previous TP and the incoming traffic load. By exploiting the information on the state of users and the channel MPR capability, the number of users who can simultaneously access the channel in the current TP is chosen so that the expected length of this TP is minimized. As a result, the MPR capability is more efficiently
utilized by the proposed protocol as compared to, for example, the slotted ALOHA with optimal retransmission probability. Furthermore, the proposed protocol requires little online computation. Its simplicity is comparable to that of slotted ALOHA. It can be applied to random access networks with spread spectrum and/or antenna array.

**Pilot Assisted Transmissions [27, 28]** We developed a number of signal processing algorithms aimed at providing multipacket reception.

Channel estimation plays a critical role in packet-switched wireless systems where it is often necessary to acquire the channel state for each packet. To facilitate channel estimation and synchronization, pilot symbols are usually embedded in a data stream. Consequently, it is important to fully utilize these symbols to obtain optimal estimation performance, and the placement of these pilot symbols can affect significantly the overall performance of a wireless system.

The optimization of pilot symbols and their placement has not been investigated until recently, although the design of optimal pilot sequence for training based channel estimators is an old problem and has been investigated by many. See [29] for a comprehensive survey.

The paper [27] received the 2004 IEEE Signal Processing Society Best Paper Award. In [27, 28], we consider the optimal design and placement of pilot symbols for channel estimation. Since mobile users may choose different channel estimators, in searching for the optimal placement, it is desirable to use a criterion that is independent of any specific estimation technique used by individual receivers. A natural choice is the CRB on the MSE of channel estimators, and the objective of designing the pilot sequence and its optimal placement is to minimize the CRB.

The main contributions of this work are as follows. For both single-input single-output (SISO) and MIMO finite impulse response (FIR) random channels, under the assumption of independent and identical distributed (i.i.d.) channel taps, we first obtain an expression of the CRB as a function of pilot symbols and their placement. It is then shown that the CRB is minimized by placing pilot symbols with smaller magnitudes closer to two ends of a packet and those with larger magnitudes closer to the center while satisfying certain orthogonality conditions. We show that, although the CRBs are functions of channel distributions, the optimal pilot placements are independent of probability distribution of the channel. This is especially important in broadcasting applications where the pilot design should be optimal for channels of all users. We further consider estimation of channels with correlated taps, and show that the previous placement strategy is also optimal among orthogonal pilot sequences.

For constant modulus pilot symbols with sufficient power, we show that the optimal
strategy is to place pilot symbols, possibly in multiple clusters, in the middle of a packet. While this result confirms the advantage of using the midamble placement as in the Global System for Mobile communication (GSM), it also suggests that some other placements are also optimal. One of such optimal placements is the QPP-α (Quasi Periodic Placement) scheme that, under mild conditions, was shown to be optimal for DFE and also optimal in terms of maximizing channel capacity.

**Channel Estimation under Asynchronous Packet Interference [30]** This work considers the channel estimation problem in the presence of asynchronous packet interference. Asynchronous interference arises in ad-hoc networks, wireless LANs, and even in cellular networks where packet collisions can not be avoided and packet transmission is asynchronous or packet synchronization is not perfect. Moreover, the interference packets can have a wide range of power levels due to the near-far effect.

Channel estimation is crucial in coherent symbol detection, optimal scheduling and power allocation, and in the design of the medium access control protocol in random access networks. Typically, channel estimation is performed by including a certain number of training symbols in the data packet. When the channel is memoryless, the placement of these training symbols does not affect the performance and is designed to simplify the receiver implementation. When the channel has memory, however, the placement of training symbols can affect the performance significantly. Optimal training placement for intersymbol interference channels has been considered.

Asynchronous packet interference introduces a different kind of channel memory. The event that a symbol of a data packet is hit by an interfering packet affects the chance that its adjacent symbols are also hit. The effect of packet interference on the training symbols, however, is somewhat subtle. If we assume that an interference packet arrives randomly, and its position relative to the packet of interest is uniformly distributed, then the average number of training symbols hit by the interference is the same regardless how training symbols are placed in the packet. However, the distribution of the number of training symbols that are hit by the interference is a function of the placement. It is the distribution of the number of training symbols survived the interference—not the average number—that determines the performance of the channel estimator.

We assume that the receiver uses the training-based minimum mean square error (MMSE) channel estimator, i.e., only those observations corresponding to the training symbols are used in the estimation. If there are \(N\) training symbols in a packet of size \(B\), the brute-force approach to finding optimal placement requires comparing \(\binom{B}{N}\) possible placements. The
lack of a simple expression for the MMSE coupled with the enormous number of possibilities makes the brute-force approach unappealing. Also, it is unlikely that such an approach will lead to useful insights.

In searching for the optimal placement, we first obtain an expression of MMSE as a function of the Fisher information matrix (FIM) of the received signal. This crucial step allows us to exploit the convexity of the Fisher information functional and therefore reduce the number of searches from $\binom{B}{N}$ to $\lceil \frac{N}{2} \rceil + 1$, which depends only on the number of training symbols and not on the size of the packet. Furthermore, independent of the system parameters such as signal-to-noise ratio, the optimal placement belongs to a fixed set of placements with either one or two clusters.

The main difficulty involved in obtaining the optimal placement in closed-form comes from the nonlinearity of the MMSE estimator. One way to overcome this problem is to consider a lower bound given by the MSE of a linear estimator that knows the position of the interference with respect to the data packet. This estimator and its MSE are called the genie estimator and the genie bound respectively. The genie estimator can only be approximated by a detect-then-estimate scheme where the receiver first detects the presence of the interference. What we gain in considering the genie estimator is that the relation between its MSE and the training placement can be obtained explicitly. We show that the placement that minimizes the genie bound has two clusters of equal or quasi-equal length at the two edges of the data packet, which is in contrast to widely accepted single cluster placement (such as that in GSM) and the uniformly distributed periodic placement. We further show that the genie bound is tight when the interference power is high, which implies that if the interference level is high, the two equal sized clusters placed at the two ends of the packets is optimal. In general, we can only conjecture that this placement is optimal for all values of the parameters involved; this conjecture is supported by simulations.

Channel Estimation for Space-Time Orthogonal Block Codes [30] A major challenge in wireless space-time communications is coping with channel uncertainties. While Shannon theory does not mandate channel estimation, the idea of acquiring the channel state before decoding, either blindly or through the use of pilot symbols, is entrenched in practice and has also been proposed for space-time systems. The use of pilot symbols, however, may impose an unacceptable overhead that limits the effective data throughput. Here system designers must consider two contradictory goals. On the one hand, it is desirable to minimize the number of pilot symbols in a data packet so that more information carrying symbols can be transmitted. On the other hand, more pilot symbols result in better channel estimation
hence reducing the symbol error rate and the need for packet retransmissions.

Conventionally, each transmitted symbol is either a pilot or a data symbol. Furthermore, pilot symbols are clustered so that training-based techniques which use received samples corresponding only to the pilot symbols can be applied. For such schemes, observations affected by the unknown data are discarded. Although training-based techniques simplify receiver design, they may carry a substantial penalty in performance for two reasons. First, the received samples corresponding to the unknown data contain valuable information about the channel. It was first established by de Carvalho and Slock that the channel estimation errors can be reduced significantly by using semiblind techniques which utilize all observations for channel estimation. The second reason comes from the placement of pilot symbols in clusters suitable only for training-based techniques. It has been established recently that placing and designing pilot symbols optimally provides gain in channel capacity.

In this work, we consider the channel estimation problem for multiple-input multiple-output (MIMO) systems that use the orthogonal block codes proposed by Tarokh, Jafarkhani and Calderbank. In addition to the placement of pilot symbols in time, we must now take the spatial domain into consideration. Within the framework of semiblind channel estimation that utilizes all observations for channel estimation, and using the Cramér-Rao Lower Bound (CRLB) as the performance measure, we examine general training strategies that allow the superposition of pilot and data symbols. In particular, we consider the effect of number of training symbols, specific training signal used, and power allocation of training symbols on CRLB. To this end, we characterize general training schemes by the power allocation matrices that specify, for each transmitted symbol in the space-time coordinate, the amount of power used for training and data respectively.

The challenge of finding the optimal (even a good) training strategy is twofold. First, one needs an expression of CRLB as a function of the power allocation matrices. Although conceptually simple, such an expression is in general complicated and not easy to optimize. Fortunately, by exploiting special properties of the orthogonal codes, we are able to simplify the CRLB expression to the point that equivalence among certain power allocation schemes can be established. The second challenge is to minimize, under a certain power constraint, the CRLB with respect to the power allocation matrices. This is again intractable in general. For the orthogonal codes presented in [31], however, we are able to show a convexity property of the CRLB. This leads to an optimal power allocation strategy under the per-symbol power constraint among those training schemes that have one pilot symbol transmitted in each block. It turns out that superimposing training with data is not optimal for channel estimation, although, with other considerations such as channel tracking and capacity enhancement,
such a technique may be an appropriate compromise between accuracy in channel estimation and high rate in data transmission. While the optimal power allocation for the most general case is still unknown, our investigation reveals power allocation patterns that favor channel estimation in the acquisition stage and the optimal allocation once the channel has been acquired.

Finally, one must question whether the CRLB is the appropriate measure. The use of CRLB as the performance measure is motivated by the consideration that training placement is a transmitter technique, and its design should not be affected by the specific technique used at the receiver. Furthermore, the asymptotic efficiency of maximum likelihood (ML) technique lends support for the use of CRLB. In this paper, we have also implemented the ML estimator and found that, for the case of using finite data samples, the performance of the ML estimator is still close to the CRLB.

**Blind Decorrelating Rake Receiver for Long Code WCDMA [32, 33, 34, 34]** We consider the problem of joint channel estimation and symbol detection in a long-code wide-band CDMA system that has features of third generation wireless: the scrambling sequences are aperiodic; data and control information may be modulated separately onto the in-phase and quadrature parts of the signal using different channelization codes with different spreading gains; pilots are often part of the control symbols; users may have different spreading gains, or multiple channelization codes may be assigned to the same user. For uplink applications, users are asynchronous, and their multipath channels may have delays longer than the symbol period. Multiple antennas may be used.

RAKE receivers are widely used in both up-link and down-link CDMA systems. If the spreading codes have good cross- and auto-correlation properties, the matched filter frontend suppresses multiaccess interference, and the RAKE receiver captures multipath diversity through its diversity branches (or the RAKE fingers). For high rate CDMA under frequency selective fading, however, code orthogonality can not be guaranteed, and the conventional RAKE receiver that uses a bank of matched filters may perform poorly. The loss of code orthogonality has adverse effects on both channel estimation and symbol detection, and the performance degradation is especially pronounced when the network is heavily loaded and power control imperfect.

In this work, we propose a joint channel and symbol estimation scheme for RAKE receivers. A decorrelating matched filter projects the received chip-rate sequence $y[n]$ into the signal space of each user whose channel and data sequence can be estimated jointly and independent of other users by least squares via a rank one decomposition. The decorrelating
matched filter does not depend on channel coefficients and may be precomputed for certain applications. The proposed scheme imposes no conditions on channel parameters and is capable of dealing with rapid multipath fading. We also establish a new identifiability condition that depends only on the spreading codes used in the system but not on channel parameters. Implied by this identifiability condition is that aperiodic spreading codes enhance channel identifiability; channels not identifiable in short-code CDMA are almost surely identifiable in a long code system.

A key contribution of this work is an efficient implementation of the decorrelating matched filter. The idea of using the decorrelating matched filter for short code CDMA is known [35], but applying it to long code CDMA presents a daunting task in terms of both computational complexity and storage requirements. A direct implementation for a ten user system—each has three multipath fingers with a 100-symbol slot and a spreading gain of 64—amounts to inverting a code matrix of size around $6,400 \times 3,000$. The code matrix, fortunately, is highly structured and sparse; only 1% of its entries are nonzero. The inverse of the code matrix, however, will in general lose the structure and the sparsity. In this paper, we use the extensive theory and algorithms developed by Dewilde and Van der Veen [36] who considered the inversion of infinite size structured matrices. The idea is to replace the code matrix by a time-varying state-space realization and implement the inversion locally in state space. For cases where pre-computation of the code matrix is possible, our approach has the same level of on-line computations and storage requirements as that of the conventional matched filter. If the inversion of code matrix must be performed on-line, we are able to reduce the computational complexity to the same level as that required in the short code case.

In [33], we propose a new blind tracking technique for fast-fading long code CDMA systems with slotted transmissions where the slot size is larger than the channel coherence time. Based on a linear interpolation model, the proposed method exploits the multipath diversity of mobile channels and estimates the channel coefficients at the selected estimating points which are located arbitrarily within a block. The use of interpolation model converts the time-varying channel parameters to fixed parameters associated with the block and often makes blind estimation approaches tractable in time-varying channels [37]. The proposed scheme consists of front end processing and blind identification of a parameter matrix which is associated with the estimating points. The elimination of unknown symbols is based on the cross referencing [38]. A new identifiability is established for noiseless case and new detection schemes are proposed based on the estimated parameter matrix. For multiuser scenario, decorrelating or regularized decorrelating front end can be used. The fast inversion of code matrix can be implemented with a state-space technique with a comparable amount
of complexity with short code CDMA cases [?].

2.3 Applications in Sensor Networks

MAC Protocols for Optimal Information Retrieval Pattern in Sensor Networks with Mobile Access [39, 40] In many applications of sensor network, the sensor network operates in three phases: sensing, information retrieval, and information processing. As a typical example, in physical environmental monitoring, sensors first take measurements of the signal field at a particular time. Then, data are collected from individual sensors. Finally, data from sensors are processed centrally to reconstruct the signal field.

An appropriate network architecture for such applications is SEnor Networks with Mobile Access (SENMA) [41]. SENMA has two types of nodes: low-power low-complexity sensors randomly deployed in large number and a few powerful mobile access points that communicate with sensors. The use of mobile access points enable data collections from specific areas of the network.

We focus on the latter two phases of operation in the SENMA architecture: information retrieval and processing, which are strongly coupled. In SENMA, medium access control is needed to regulate data retrieval from sensors to the access points. A multiple access control (MAC) protocol in SENMA governs how information is retrieved from the sensor field. With a specific MAC, packets collected by mobile access points form a sampled signal field with a specific pattern. This sample pattern directly affects the performance of the signal reconstruction at the last data processing stage. Therefore, it is important to design a MAC protocol that results in the desired retrieval pattern and hence, the optimal signal reconstruction performance. If the mobile access point can schedule transmissions from sensors, it is natural to use a centralized scheduler to retrieve data from equally spaced sensors locations.

While it may appear obvious that collecting data from optimally chosen locations using centralized scheduler gives better performance, there are several nontrivial practical complications. For sensor networks with finite density, there may not exists a sensor at the desired location. The optimal scheduler must find sensors locations closest to the desired sampling pattern. Such scheduler needs to have the additional information of each sensor locations, which usually is not available at the access point. Furthermore, it comes with nontrivial complications of centralized control and large communications overhead. Decentralized MAC requires much less intervention from the mobile access point and is simple to implement. In order to take advantage of decentralized access but not lose much reconstruc-
tion performance, our problem is how to design decentralized MAC protocols to achieve the desired retrieval pattern.

We design MAC protocols for the desired data retrieval pattern. We consider a one-dimensional problem for simplicity which can be extended easily to two-dimensional problem. Taking both performance and implementation complexity into consideration, besides the optimal centralized scheduler, we propose three MAC protocols. We first propose a decentralized scheduler via carrier sensing which, under the no processing delay assumption, provides minimum performance loss comparing to the optimal scheduler. Then, to simplify the implementation, we describe a MAC scheme which uses Aloha-like random access within a resolution interval centered at the desired retrieval location. Finally, to improve the performance of the previous MAC, we propose an adaptive Aloha scheduling scheme which adaptively chooses the desired retrieval location based on the history of retrieved locations. Design parameters are optimized for the proposed schemes. The performance comparison under various sensor density conditions and the size of desired sampling locations is also provided.

**Non-Parametric Approach to Change Detection and Estimation in Large Scale Sensor Networks [42]** We consider the problem of detection and estimation of changes in a large scale sensor network in which each sensor is in either an “excited state” or a “baseline state”. One example is that each sensor detects locally the presence of targets or certain chemical/biological agents. A sensor is RED if has detection and GREEN otherwise. We assume the architecture of SENMA—Sensor Networks with Mobile Access [41]—where a mobile Access Point (AP) collects data directly from sensors via random access (such as ALOHA). Each sensor transmits its local detection in the form of a packet, and the process of data collection by the mobile AP can then be modeled as a random sampling of the sensor field. We assume that at the mobile AP, the location of each received local decision is known. This can be implemented either by requiring sensors transmit their location information or letting the mobile AP poll sensors at specific locations.

The problem of change detection and estimation in SENMA, illustrated in Figure ??, is to decide, from two consecutive records, whether there is a change of the underlying probability distribution (detection) and the distribution of the change (estimation) if changes occur. Since random access is used by the mobile AP, there is no guarantee that a measurement of one sensor in the first data collection will appear in the second, nor can we be assured that the same number of samples are collected. For many applications, there is no prior knowledge about the the distribution of the sensor states. The change detection and estimation is
therefore non-parametric in which we make no assumption about the specific form of the probability distribution of the binary random field.

Two factors must be considered for change detection and estimation in a large scale sensor networks: the time required or the number of packets required for data collection and the complexity of the detection and estimation algorithm. In practice, minimizing the number of samples required for detection reduces the collection time of a mobile AP and, more importantly in terms of energy consumption, the number of transmissions from sensors. In this paper, we aim to provide a mathematical characterization of the required sample size for which, with high probability, the distance of any two probability distributions can be estimated accurately from empirical distance between samples generated by these distributions. The basic tool used here is the Vapnik-Chervonenkis Theory. We are also concerned with the detection and estimation algorithm. We present three algorithms with varying complexity, scalability, and levels of confidence.

Non-parametric change detection is a classical problem [43]. Classical techniques include permutation tests, Kolmogorov-Smirnov test, and Cramér-von Mises tests [44]. In the context of large scale sensor networks, however, it is not obvious that these classical techniques are applicable. Specifically, the samples are two dimensional, and we are interested in not only whether changes have occurred but also possible locations of changes. Furthermore, we are also interested in the asymptotic behavior as the number of samples increases and whether the detection-estimation algorithm scales.
3 Publications

3.1 Journal Publications


### 3.2 Conference Proceedings


### 3.3 Manuscripts Under Review


4 Participating Personnel

Senior Personnel
Lang Tong, Professor, Cornell University, Ithaca, NY 14850.
Shai Ben-David, Visiting Professor (currently with University of Waterloo)

Graduate Students
Gokhan Mergen, Received the PhD degree in 2004. Now with Qualcomm Inc..
Petru Budianu, Received the PhD degree in 2004. Now with Qualcomm Inc..
Srihari Adireddy, Received the PhD degree in 2003. Now with Silicon Labs Inc..
Atul Maharshi, Received MS degree in 2002. Now with Flarion Inc.
Youngchul Sung, expected to graduate in May, 2005.
Zhiyu Yang, expected to graduate in September, 2005.
Parvathinathan Venkitasubramaniam, expect to graduate in 2006.

5 Report of Inventions

There is no invention disclosure filed in this project.
6 Bibliography


