SELF-CONFIGURING WIRELESS TRANSMISSION & DECENTRALIZED DATA PROCESSING FOR GENERIC SENSOR NETWORKS

Cornell University

Sponsored by
Defense Advanced Research Projects Agency
DARPA Order No. K257

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the U.S. Government.
This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-IF-RS-TR-2004-204 has been reviewed and is approved for publication

APPROVED: 
/s/
RICHARD C. BUTLER II
Project Engineer

FOR THE DIRECTOR: 
/s/
WARREN H. DEBANY, JR.
Technical Advisor
Information Grid Division
Information Directorate
This report describes the results of the Self-Configuring Wireless Sensor Network effort at Cornell University, funded under the DARPA SensIT program. The primary goal of this effort was the development of core technologies for large sensor networks that are truly self-configuring. Such networks should not depend on fixed emplacements or predefined topologies. Control by a centralized authority should be minimized; all aspects of network management should be handled in a distributed manner, with all network elements sharing a collective responsibility of performance maintenance. Our technology development fell into three basic areas. First, we conducted an in-depth study of phase transitions and complexity in large wireless networks. Our goal in this area was the identification of complexity thresholds, bounding the computational complexity of management protocols for extremely large networks. Second, we considered the use of game theory in the development of truly distributed network control algorithms. Finally, we developed models for the trade-off between energy conservation and robustness in wireless networks.
Contents

1 Introduction .............................................................. 1

2 Phase Transitions, Structure, and Complexity in Wireless Networks .......... 1
   2.1 Background: Wireless Networks .................................. 1
       2.1.1 Cellular Networks .............................................. 1
       2.1.2 Multi-hop Wireless Networks ................................. 2
   2.2 Background: Computational Complexity ............................. 4
   2.3 Results ............................................................... 4

3 Game-Theoretic Approaches to Network Design .................................... 6
   3.1 A Review of Game Theoretic Tools ................................ 6
       3.1.1 A Brief Introduction to Game Theory ....................... 6
       3.1.2 Related Work ................................................... 8
   3.2 A Review of Power Control ........................................ 8
   3.3 A Review of Medium Access Control ................................ 11
   3.4 Carrier Sense Multiple Access (CSMA) ............................... 12
   3.5 Contributions to the SensIT Program ............................... 13

4 The Energy-Robustness Tradeoff for Routing in Wireless Sensor Networks .... 14
   4.1 Illustration ........................................................... 15
   4.2 Experimental Setup ................................................ 16
   4.3 Simulation Results ................................................... 17
   4.4 Conclusions .......................................................... 21

5 Appendix ........................................................................... 22

Reference ............................................................................ 23
List of Tables

1. Energy and Robustness Measures for Alternative Routing Configurations . . . . . . . . . 15

List of Figures

1. Architecture of GSM cellular network . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2
2. Multi-hop wireless network topology . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 3
3. Illustration of Different Routing Schemes for a Sample Configuration of Five Nodes . . . . . 15
4. Number of Node Disjoint Source Sink Routes with respect to the Transmission Radius . . . 18
5. Number of Nodes Transmitting with respect to Transmission Radius . . . . . . . . . . . . . . 18
6. Probability that a route exists with respect to Transmission Radius (5 % failure rate) . . 19
7. Probability that a route exists with respect to Transmission Radius (20 % failure rate) . . 19
8. Probability that a route exists with respect to normalized energy cost (5 % failure rate) . . 20
9. Probability that a route exists with respect to normalized energy cost (20 % failure rate) . . 20
1 Introduction

In our SensIT sponsored research program, we explored several new technologies and their impact on the development of wireless sensor networks. Our efforts fell into three basic areas. First, we considered the complexity of various problems that arise in the design and management of wireless networks. In particular, we focused on the emergence of phase transitions in several NP-hard problems that are inherent in network design and management. Second, we considered game-Theoretic approaches to the development of distributed power control and medium access problems that arise in sensor network engineering. Our goal in this effort was the enablement of distributed control through the modeling of sensors as agents in a resource market. Finally, we considered the energy/robustness tradeoff that arises in the design of routing algorithms for sensor networks.

In this report we provide an overview of each of these three areas, indicating our own published results as appropriate. We have appended a complete bibliography for all three research areas, as well as a list of the participants in this effort and a listing of our related publications.

2 Phase Transitions, Structure, and Complexity in Wireless Networks

While the tetherless nature of wireless communication networks introduces greater flexibility in terms of mobility and set-up time, these networks are often severely resource constrained, particularly when deployed at large scale. The most common resource constraints are those on bandwidth and energy. Many network operations, such as mobility management, routing, and channel allocation, can be formulated as constrained optimization problems or as constraint satisfaction problems. Our research effort in this area consisted of a set of case studies that investigate the computational feasibility of solving such problems in an optimal manner.

2.1 Background: Wireless Networks

We begin with a brief introduction to the two main categories of wireless networks that we considered, namely cellular networks and multi-hop wireless networks.

2.1.1 Cellular Networks

The first-generation wireless systems were developed in the late 1970's and 1980's and were based on analog technology, such as the Advance Mobile Phone Service (AMPS) by AT&T and Nordic Mobile Telephone (NMT) by Ericsson. As demand increased and digital technology matured in the 1980's and 1990's, the second-generation digital wireless systems were designed, such as Global System for Mobile Communications (GSM) in Europe and Digital AMPS in North America. These systems offered higher system capacity and improved quality of service. Third-generation wideband systems based on Code Division Multiple Access (CDMA) schemes are currently under development and deployment. These aim to provide high-quality voice, medium-rate data capability, and increased system capacity. Industrial trends point to much higher rate and higher quality multimedia capability in future generations.

In a cellular network there are base stations that are controlled centrally, and these base stations connect over the radio link to mobile users. The “wireless” element in this kind of network refers mainly to this link. The base stations are connected in a hierarchical manner through high-bandwidth optical or point-to-point microwave links. Communications between any two mobile users must go through the fixed base stations nearest them. Figure 1 shows the architecture of a typical cellular network.

Due to the challenging nature of wireless channels (multi-path, shadow fading etc.), there are a significant number of research problems concerning the physical and link layers of the protocol stack. The
Figure 1: Architecture of GSM, a widely-used cellular standard

primary wireless network management issues in a cellular network are resource allocation and mobility management. There is extensive literature on the problem of allocating channels in an efficient manner allowing for the greatest spatial reuse of channels in non-interfering cells [128]. Associated with channel allocation is the problem of call admission [129, 130]. The problem of locating mobile users in order to route arriving calls has also been widely studied [17, 36].

2.1.2 Multi-hop Wireless Networks

As the name suggests, in this kind of a network, all links are wireless and routes from one node to another can require multiple hops. Distributed multi-hop wireless networks are gaining in importance as a subject of research [57, 62, 126]. Their expected applications range from static environmental sensing to mobile networking for disaster recovery. Many of these applications are likely to involve large-scale operation with hundreds or thousands of wireless communication nodes.

It is worthwhile distinguishing between two broad categories of multi-hop wireless networks: mobile ad-hoc networks and sensor networks.

In wireless mobile ad-hoc networks (MANETs), the mobile nodes can be communication devices or computers. Due to node mobility, there are frequent changes in the network topology. The primary challenge in this domain is to sustain a number of any-to-any flows within the network in the face of such dynamics.

Research on MANETs [85], has evolved from the DARPA packet radio program from the early 1970’s [86]. There has been a renewed interest in this field as wireless technologies are beginning to mature and take hold commercially. Wireless ad hoc networks, which can be deployed rapidly as they do not require much existing infrastructure, are expected to find applications in a number of diverse settings. Examples range from disaster recovery, law enforcement, military communications, distributed computing, and
home/office local area networks, to special events such as conferences and festivals.

Much of the research in the area of ad hoc networks has focused on developing routing protocols. Proactive routing protocols attempt to compute paths in advance and maintain them continuously so that a route is readily available when a packet needs to be forwarded. Examples of proactive routing protocols are destination-sequenced distance vector protocol [87], optimized link state routing protocol [88], and wireless routing protocol [89]. Reactive routing protocols are based on a source initiated query/reply process and discover new routes when a new packet flow is to be initiated. Examples of reactive routing protocols are the temporally ordered routing algorithm [115], dynamic source routing [110], and the ad hoc on demand distance vector routing protocol [90]. Both proactive and reactive strategies are combined in the hybrid zone routing protocol [91]. A number of routing algorithms have also been developed to incorporate more detailed knowledge about the location of nodes: location-aided routing (LAR) scheme [94], the distance routing effect algorithm for mobility (DREAM), the grid location service (GLS) [83], and the greedy perimeter state routing (GPSR) [84].

Wireless sensor networks are envisioned to consist of hundreds to thousands of inexpensive wireless nodes, each with some computational power and sensing capability, operating in an unattended mode [126]. They are intended for a broad range of environmental sensing applications from vehicle tracking to habitat monitoring [98, 117, 123]. The basic hardware technology for these networks – low cost processors, miniature sensing and radio modules – is here today, with further improvements in cost and capabilities expected within the next decade [98, 108, 111, 117, 118]. The applications, networking principles and protocols for these systems are just beginning to be developed [102, 103, 106, 117].

Wireless sensor networks are similar to mobile ad-hoc networks (MANETs) primarily in that both involve multi-hop communications. However, the nature of the applications and routing requirements for sensor networks are drastically different in several respects from MANETs. First, the typical mode of communication in a sensor network is from multiple data sources to a single data recipient/sink,
rather than communication between any pair of nodes. Second, there is likely to be some redundancy in
the data being communicated by the various sources in sensor networks because it is based on common
phenomena. Third, in most envisioned scenarios the sensors are not mobile though the sensed phenomena
may be, hence the nature of the dynamics in the two networks is different.

Further, sensor networks are extremely energy constrained because of their scale and the requirement
of unattended operation. This constraint is worse in sensor networks than in MANETs where the
communicating devices handled by human users can be replaced or recharged.

2.2 Background: Computational Complexity

We review the fundamentals of computational complexity theory in this section, as this is the perspective
from which we will view configuration and optimization problems in wireless networks.

For purposes of analyzing computational complexity, we often consider decision problems. Decision
problems result in simple yes/no answers. Optimization problems can be translated to and from decision
problems with polynomial-time effort. An optimization problem can be converted into a decision problem
by asking the following question: given a number \( y \), does there exist a point \( x \) in the search space \( X \)
such that the cost function \( f(x) \leq y \)?

Two classes of decision problems have been of particular interest to computer scientists: the class P
and the class NP. The class P consists of decision problems that can be solved in time that is at most
polynomial in the input size \( n \). The class NP consists of decision problems for which a truth certificate
can be verified in polynomial time. Loosely, NP-complete problems are the “hardest” problems in the
class NP. An interesting property of NP-complete problems is that they can be mapped in polynomial
time to each other, so that the answer to an instance of one problem is yes if and only if the corresponding
instance of the other problem results in a yes. While it is quite easy to see that P is a subset of NP (since
one can simply ignore any certificate that is provided for a P problem, solve the problem in polynomial
time and determine if it results in a yes solution), it is believed that P \( \neq \) NP although this remains
to be proved. An interesting result is that unless P = NP, no algorithm can solve all instances of an
NP-complete problem in polynomial time. NP-hard problems are problems that are at least as hard to
solve as NP-complete problems.

There are a number of algorithmic approaches for dealing with NP-hard problems: heuristic construc-
tions, local search, and approximation algorithms. In our work, we focus primarily on another mechanism
– identifying special cases of NP-complete problems that are tractable to efficient, polynomial-time algo-

rithms. This allows us to bound the computational complexity of subproblems that may of interest from
an engineering perspective.

Our overview of the rich subject of computational complexity has been necessarily brief. For an
excellent, detailed treatment of NP-completeness and worst-case computational complexity, we refer the
reader to the classic work by Garey and Johnson [9].

In the next section, we begin our investigations into complexity issues in wireless networks with a
case study involving user location in cellular networks.

2.3 Results

The first case study focused on the task of locating mobile users through sequential paging in cellular
networks. In sequential paging probabilistic estimates of user location are used to minimize the average
number of cells that have to be paged upon call arrival. We showed that the problem of minimizing
the paging cost under an average delay constraint, previously believed to be NP-complete, is in fact
polynomial-time solvable. This is because the structure inherent in this problem makes it tractable to
dynamic programming. Further, we derived the conditions under which cluster paging, an even simpler
and faster sequential paging technique, results in provably optimal performance. We also presented a
number of results concerning the performance of these sequential paging mechanisms and their sensitivity to errors in the probabilistic location estimates.

In our next case study, we explored data-centric routing protocols for wireless sensor networks. Data aggregation has been proposed as a mechanism for energy-efficient information routing in such networks. We showed that, although optimal data aggregation is NP-hard in general, there exist polynomial special cases where simple heuristic algorithms can achieve optimum energy savings. These special cases correspond to particular topological arrangements of data sources. We also derived some useful bounds on the energy costs of these protocols.

Finally, we studied emergent structure in multi-hop wireless networks. We showed that in these distributed wireless networks, many tasks such as multi-path routing, conflict-free channel allocation, Hamiltonian cycle formation, coordinated target tracking, and information dissemination through probabilistic flooding, are characterized by zero-one phase transitions. These tasks can be performed with high probability above a critical resource threshold, and with negligible probability below the threshold. This emergent structure can be exploited to simplify problem solving for tasks that can be formulated as NP-complete constraint satisfaction problems. We showed that the average computational complexity for these NP-complete problems decreases to manageable levels beyond the phase transition threshold.

The common theme in all these case studies, which pertain to some of the hardest problems in wireless networks, is the identification of special conditions or emergent structures that help bound the computational complexity. The results of our effort served to demonstrate the usefulness of a computational complexity perspective in analyzing and engineering large-scale wireless networks.

These results were documented in the following publications.

3 Game-Theoretic Approaches to Network Design

The primary contributions of our work in this area were an application of game theory to power control and a study of a game theoretic approach to understanding medium access control. This work encompasses several novel contributions to both game theory and communications theory. We begin by discussing the motivation behind this work.

Engineers have typically approached communications networks as monolithic creations in which the nodes were mere components to be engineered. As a result, the wireless communications literature contains many algorithms designed to be run on communications nodes, and many current communications standards specify the algorithms that standard-compliant nodes must implement. With the increase in the size of networks and the rise of nonproprietary networks, software radios, and the use of unlicensed spectrum, this view of monolithic networks is giving way to a view of networks as an emergent phenomenon that arises through the interaction of independent user agents. The role of the engineer in this new paradigm is to specify mechanisms by which nodes of the network can interact — without presuming to specify the algorithms that those nodes will run.

For example, in a traditional slotted Aloha system, users may be expected to update their retransmit probabilities in accordance with the pseudo-Bayesian algorithm. According to my analysis, however, if everyone else is using the pseudo-Bayesian algorithm then an individual user would prefer the “always transmit” strategy to using the pseudo-Bayesian algorithm. This is obviously a problem, as the user using the always transmit strategy will gain an unfair advantage in using the system. Furthermore, if two users were to adopt the “always transmit” strategy, then no users would have successful transmissions. This is obviously undesirable.

By applying the tools of game theory, we can understand the performance of networks in which the individual users are behaving selfishly to maximize their own performance. Such a game theoretic analysis is valuable for existing systems in that it increases our understanding of which protocols are most vulnerable to exploitation by selfish users. In addition, we shall see that game theoretic analysis can inform the design process itself, allowing for the design of protocols that perform well in the presence of selfish users.

A system optimized for selfish users has several advantages. First, a game theoretic solution is naturally decentralized as it assumes that individual users seek to maximize their own interests. Second, when a game theoretic analysis has been applied from the outset, then there is no advantage for a user to expend energy trying to “game” the system — an assumption of selfish behavior is already included in the analysis.

In the economics literature, pricing mechanisms are often devised in order to extract enough information from players to allocate goods in a manner which is efficient. In other words, an optimization procedure is carried out over the space of possible pricing policies in order to identify those policies that will maximize social welfare. The long term goal of this work is to conduct a similar optimization over the space of protocols. We wish to find networking protocols that, when used by noncooperative users, will maximize social welfare.

3.1 A Review of Game Theoretic Tools

3.1.1 A Brief Introduction to Game Theory

In this section, we provide a brief introduction to game theory and relate existing work in game theory to the new game theory introduced in this report.

The most basic game structure is a normal form game. The normal form game consists of three elements:

- a set of users, \( I \),
- a set of actions available to each user, \( A_i \), \( i \in I \); for convenience, let \( A = \times_{i \in I} A_i \), and
The most common solution concept for such a game is a Nash Equilibrium, named for the Nobel Laureate John Nash. I use the notation \(a = (a_i, a_{-i})\) ∈ \(A\) to denote an action profile where \(a_i\) ∈ \(A_i\) is chosen by player \(i\) and \(a_{-i} \in \times_{j \in I, j \neq i} A_j\) represents the choices of all other players; similarly, \((a'_i, a_{-i})\) represents the action profile where player \(i\) chooses action \(a'_i\) ∈ \(A_i\) and all other players choose the same actions as in the profile \(a\). An action profile, \(a \in A\), is said to be a Nash Equilibrium if for all \(i \in I\) and all \(a'_i \in A_i\), \(u_i(a_i, a_{-i}) \geq u_i(a'_i, a_{-i})\).

The most basic result of game theory concerns the existence of a Nash Equilibrium in finite games (games with finite sets \(I\) and \(A_i\)). Start with any finite game \((I, A, u)\) and define an expanded game in which mixed strategies are allowed. Let \(S_i = \Delta(A_i)\), where \(\Delta(A)\) denotes the set of probability distributions over \(A\). Then \(S = \times_{i \in I} \Delta(A_i)\). Note then, that an element of \(S\) is a probability distribution over \(A\) (although \(S \neq \Delta(A)\) because the probabilities over each \(A_i\) must be independent). Define the utility of a mixed action as the expectation of the utilities of pure actions. For an action \(\sigma \in S\), let \(\sigma(a)\) denote the probability that action profile \(a\) emerges when \(\sigma\) is played. Then \(\hat{u}_i(\sigma) = E(u_i(a)) = \sum_{a \in A} \sigma(a) u_i(a)\).

John Nash proved that given any finite game \((I, A, u)\), there exists at least one Nash Equilibrium in the expanded game which permits mixed strategies, \((I, S, \hat{u})\) [136].

Games which take into account the order of moves and the knowledge of players at each point in the game are called extensive form games. In part of our work, we focused on a specific type of extensive form game called a repeated game. A repeated game is defined by a stage game which is played repeatedly; the stage game is a simple game of the normal form as described above. The players in the repeated game are the same as the players in the stage game.

Consider games of perfect information, where all players’ past actions are revealed at each stage. Over time, a history \(h = (s^0, s^1, s^2, \ldots, s^t)\) of play emerges where \(s^i \in S\) represents the strategy profile played at time \(t\). Let \(H\) be the set of all such finite length histories, including the empty history. A strategy for player \(i \in I\) in the repeated game is then a mapping \(\beta_i : H \rightarrow S_i\); let \(B_i\) be the space of all such mappings, and let \(B = \times_{i \in I} B_i\).

Having defined the set of players and the strategy spaces for the repeated game, all that remains is to define the payoffs. Note that a particular strategy profile \(\beta \in B\) will give rise to a series of stage game strategy profiles forming a particular history \((s^0, s^1, s^2, \ldots)\). Assume that player \(i\)’s payoff from this history is the discounted value of the stream of stage game payoffs which she accrues: \(v_i(\beta) = (1 - \delta) \sum_{t=0}^{\infty} \delta^t u_i(s^t)\). The discount factor \(\delta \in [0, 1]\) represents how patient the players are; a higher discount factor represents players who value the future more highly. The factor \((1 - \delta)\) is included to normalize \(v_i\) to the same scale as \(u_i\).

The definition of a Nash equilibrium for a simple normal form game also applies to the repeated game. A Nash equilibrium of a repeated game is a strategy profile, \(\beta \in B\), such that for all \(i \in I\), \(v_i(\beta_i, \beta_{-i}) \geq v_i(\beta'_i, \beta_{-i})\) for all \(\beta'_i \in B_i\).

A problem with this straightforward application of the Nash Equilibrium concept to extensive form games is that it allows players to make implausible threats. A well-known refinement for Nash Equilibrium to the case of extensive form games is known as subgame perfection. Let \(h \in H\) be any finite history. Then define a new game starting after the history \(h\); if \(G\) is the repeated game, then call this subgame \(G_h\). A subgame perfect equilibrium of \(G\) specifies that the players play a Nash Equilibrium in any subgame \(G_h\) of \(G\).

One criticism of this approach is that the concepts of Nash equilibrium and subgame perfection have very little bite in terms of restricting the payoffs of repeated games. Under weak conditions, one can prove that there exist subgame perfect equilibria with any achievable, individually rational payoffs. This is extremely unfortunate if the game theorist is attempting to use game theory to predict the outcome of a game between rational agents. The designer of a communication system can select an equilibrium
a priori and make an announcement indicating that this is the desired equilibrium. It would then be impossible for an individual to gain by unilaterally deviating from the announced equilibrium. Hence, with most users of a system utilizing the pre-announced equilibrium strategy, it will be impossible for a criminal or hacker to gain any advantage over ordinary users. If the chosen equilibrium is undesirable to groups of users, however, then those groups of users may collude in an effort to switch the system to another equilibrium.

Normal form and repeated games are quite restrictive. All players must choose simultaneously, and it is assumed that all players have complete information about the game being played. Real-life situations modeled by games often have additional details including order-of-moves, opportunities to revise previous choices, incomplete information, and so on. These details are better captured by extensive form games and Bayesian Games. We will not detail these types of structures here; for details, see a complete introduction in a game theory text such as [137].

3.1.2 Related Work

We introduced two types of games in our work. Both are novel primarily in their explicit modeling of user departures and arrivals. In both game types, users arrive via an exogenous random process and departures are endogenous to their game play. There are some economic models which capture the arrival and departure of game players, these are usually referred to as the “overlapping generations” models. These models usually have players who arrive and depart at deterministic times. For instance, one player may arrive in each period and each player may have a lifetime of $T$ periods.

The first game model is that of Games of Population Transition. A game of population transition is a complete information Markov game. (A Markov game is a game where players actions must be functions only of the “state” of the game; for a detailed discussion see [137].) In some sense, such a game could be modeled as a conventional Markov game with a countable collection of players, where each player was assigned an entrance time according to a random process. An equilibrium of such a game might not be symmetric, however — we require that all players of the same type behave identically. In order to impose this requirement, Games of Population Transition have a property denoted “semi-anonymity” by Kalai [138, 139]; this property requires that players’ payoffs depend only on other players’ actions through the proportion or number of other players playing each action. That is, each player views all other players as indistinguishable, except possibly by type. Semi-anonymity itself is a modification of the concept of an anonymous game, utilized in [140, 141, 142] and elsewhere. In an anonymous game, the number of players is infinite; hence individual player choices have no impact on the aggregate distribution of actions chosen.

The second game model is denoted Games of Population Uncertainty and Transition. While maintaining many of the features of games of population transition, games of population uncertainty and transition add an element of imperfect information. In these games, in addition to being uncertain about the state of the game and other players’ knowledge, as is common in games of imperfect information, players are uncertain about the number and type of the other players participating in the game. The only other work of which we are aware in which players are uncertain of the number and type of their opponents is Myerson’s work on games of population uncertainty and Poisson games [143, 144, 145]. Myerson’s games, however, are one-shot games, and the set of players is fixed, though unknown, for the duration of the game.

3.2 A Review of Power Control

In a wireless communications network utilizing a code division multiple access (CDMA) mechanism, power control is an important problem. Each user wishes to maximize her signal-to-noise-and-interference ratio (SINR) while conserving power. When a node increases transmit power, it will increase its SINR,
but it will also increase the interference seen by other nodes, decreasing their SINR. This tradeoff can be modeled as a simple normal-form game.

The game theoretic concept of a Nash equilibrium is important regardless of whether or not the power control problem is addressed in a decentralized manner. For instance, suppose that the operating point selected by a centralized controller in the power control problem is not a Nash Equilibrium. Then, by definition, at least one user can improve her utility by transmitting at a different power level than the power level chosen by the central authority. That user, then, has an incentive to try to “cheat” the system by transmitting at a different power level. (Although many users will not have the expertise and resources to modify their wireless terminal equipment in order to “cheat,” the ability to do so is certainly within the grasp of some individuals and organizations who might then sell “improved” terminals on the black market.) While checks may be put in place to detect and prevent such cheating, the selection of an operating point that is not a Nash Equilibrium has clearly complicated the system. No matter how the operating point is selected, it is desirable that the chosen operating point be a Nash Equilibrium.

Early work on the use of game theoretic models for uplink power control includes [146]. As in many of the papers reviewed in this section, this work assumes that user \(i\)'s utility is a function of her transmit power \(p_i\) and her signal-to-interference-and-noise ratio (SINR), which is defined as

\[
\gamma_i(p_1, p_2, \ldots, p_N) = \frac{h_ip_i}{\sum_{j \neq i} h_j p_j + n}
\]

where \(h_i\) is user \(i\)'s path gain to the base station and \(n\) is the power of the noise. Ji and Huang assume that for each user \(i\), the user's utility function, \(u_i(p_i, \gamma_i)\), satisfies three assumptions:

- For a fixed SINR, \(\gamma_i\), \(u_i(p_i, \gamma_i)\), is monotonically decreasing and concave in \(p_i\).
- For a fixed transmit power, \(p_i\), \(u_i(p_i, \gamma_i)\) is monotonically increasing and concave in \(\gamma_i\).
- For all \(i\),

\[
\frac{\partial^2 u_i}{\partial p_i \partial \gamma_i} \leq 0
\]

Ji and Huang prove that under these conditions \(u_i\) is component-wise concave upward in user \(i\)'s transmit power \(p_i\). Hence, they argue on the basis of a theorem of Rosen [147] that the game has a Nash equilibrium and they propose an algorithm to seek out such an equilibrium. This work does not explore the efficiency of their proposed equilibrium. While Ji and Huang do not use the concept of pricing directly, they discuss the concept of a “shadow price” in the development of their algorithm.

A slightly more recent line of work begins with [148] and is further developed in [149, 150]. This work defines a very explicit utility function. For the moment, however, it will suffice to note that the utility function is of the form

\[
u_i(p_i, \gamma_i) = f(\gamma_i) p_i
\]

where \(f(\gamma_i)\) is a strictly increasing function of the user’s SINR, as defined above. In [148], Shah, Mandayam, and Goodman study the power control game induced by their utility function. They prove that such a game has a unique Nash equilibrium at which all users have the same received power at the base station, but that this equilibrium is not Pareto efficient.

Pareto efficiency is a necessary condition for a socially desirable outcome. A power vector \(p\) is said to be a Pareto improvement over power vector \(q\) if for all \(i\), \(u_i(p_i, \gamma_i(p)) \geq u_i(q_i, \gamma_i(q))\) with strict inequality for at least one \(i\). Power vector \(q\) is said to be Pareto efficient if there does not exist a power vector \(p\) which is a Pareto improvement over \(q\). In other words, an outcome is Pareto efficient if there is no vector of powers which makes some user better off without making some other user worse off.

Shah, Mandayam, and Goodman look briefly at pricing as a means of improving the efficiency of their outcome, but their linear pricing scheme is studied in more detail by Saraydar, Mandayam, and Goodman in [150]. In a linear pricing scheme, the system sets a price \(c\) which users must pay for each unit
of transmitted power. Users then seek to maximize \( u_i(p_i, \gamma_i) - c p_i \). Saraydar, Mandayam and Goodman investigate the best Pareto improvement which can be obtained by such a scheme and prove that even the best Pareto improvement is not Pareto efficient. They also propose an algorithm for finding the price which yields the most Pareto improvement over \( c = 0 \); denote this price \( c_{\text{best}} \).

Curiously, when Saraydar, Mandayam, and Goodman introduce a positive price into their model, they must restrict the player’s transmit powers in order to ensure the existence of equilibrium. Players must transmit at a power such that \( \gamma_i \geq 2 \ln M \), where \( M \) is the number of bits in a frame. Without this restriction, their proposed algorithm may not converge; with the restriction, however, the algorithm sometimes converges to an equilibrium where some users have negative net utility. Such an equilibrium cannot be an equilibrium of the unrestricted game because in the Saraydar, Mandayam, and Goodman framework the net utility of not transmitting is 0; hence, the user would be strictly better off not transmitting. An equilibrium in which users are forced to transmit does not seem to be in the spirit of a game-theoretic approach. The existence of an equilibrium in the unrestricted game with pricing is an open question.

Goodman and Mandayam consider the same basic utility functions in [151], but they consider these utility functions in a cooperative context. That is, the users are no longer assumed to be selfishly seeking to maximize their own utility. This is problematic, as such an operating point must be enforced via external means. In our work we develop an example showing that the same operating point chosen for the cooperative scenario in [151] can be maintained as the equilibrium of a non-cooperative repeated game.

A pricing scheme similar to that of Saraydar, Mandayam, and Goodman, with pricing based on transmit power, is presented in [152]. Heikkinen’s work does not require the assumption of a specific utility function, but it assumes that a user’s utility is a function only of her SINR; the user’s disutility from increasing transmit power enters only in the price that the user is charged for transmitting. Such a model makes it difficult to distinguish the intrinsic cost of transmit power (e.g. battery power) from a cost which might be externally imposed by the system operator. Unlike the other papers considered in this section, Heikkinen does consider the power control problem as a dynamic game which takes place over time. However, Heikkinen primarily considers power control in the framework of a scheduling problem. Specifically, Heikkinen focuses on determining a price at which only one user will transmit per time slot.

Other recent papers on power control and pricing in a game theoretic setting include [153] and [154]. As in Heikkinen, these papers take utility to be a function only of SINR, with power entering only in the price that the user is charged for transmitting. The net utility function in such a model is \( u_i(p_i, \gamma_i) = f(\gamma_i) - c p_i \). (In [154], \( f(\gamma_i) \) is taken to be logarithmic.) With a maximum power restriction and an appropriate monotonic transformation, this utility function can be transformed into an equation of the form \( u_i(p_i, \gamma_i) = f(\gamma_i)/p_i \) (for a different function \( f(\cdot) \)). So, in some sense all of the papers mentioned in this section except for [146], which never explicitly defines a utility function, use variations on the same form of the utility function.

The nature of the utility function of a wireless data user is an open issue. We share with Shah, Saraydar, Mandayam, and Goodman the perception that transmit power is an explicit part of the user’s utility, beyond any penalty which may be imposed in a pricing scheme. However, we are less confident in the specific utility function which they have developed. Hence, our results in power control require relatively weak assumptions on \( u_i(p_i, \gamma_i) \).

Unfortunately, there are problems with several of the pricing schemes discussed in this section. First, schemes which base price on transmit power display the so-called “near-far unfairness.” The primary factor determining the transmit power is the path gain between the user and the base station. Users who are located far from a base station or in a deep fade — who have the lowest utility to begin with — will be penalized by higher prices. This is not surprising as it is the farthest users who are typically causing the most degradation in other users’ performance [155, 150]. Nevertheless, the reason for adding pricing to the model was to encourage users to reach an equilibrium which was a Pareto improvement over the Nash Equilibrium, not to further penalize the users farthest from the base station. [153] suggests that...
this problem can be avoided with a pricing scheme based on received power rather than transmitted power. [154] considers the same solution by proposing that users be charged a price proportional to their path gain.

Second, it would be difficult for a pricing scheme based on transmit power to reliably determine the power at which the user has transmitted. Presumably the system must depend on the user to report the power with which she transmitted. A malicious user has an incentive to underreport her transmit power in order to reduce the “price” that she is charged. Certainly schemes could be developed to mitigate this problem, but it is important to realize that this is a problem which should be addressed if users are to be charged on the basis of their transmitted power. Basing pricing on received power, as suggested in [153] avoids this problem as well.

Saraydar, Mandayam, and Goodman suggest that pricing in their work does not refer to monetary incentives but to “a control signal to motivate users to adopt a social behavior” [150, emphasis in original]. While this is a reasonable view of pricing, the price charged for transmission must result in some actual disincentive from transmitting at high power — otherwise, the user’s incentive problem remains unchanged. Regardless of whether or not the price charged must be paid with real money, an effective pricing mechanism must exact a penalty on the users for high powered transmissions. As a result, implementing a pricing scheme adds considerably to system complexity.

3.3 A Review of Medium Access Control

We focused on two medium access control protocols: Aloha and carrier sense multiple access (CSMA). In this section, we briefly introduce these two protocols, including references to related work.

Aloha is the simplest possible medium access control protocol. Nodes transmit when they have data to send; data is retransmitted after a random delay if there are transmission errors. This protocol was first introduced and analyzed by Norm Abramson at the University of Hawaii (hence the name) in 1970 [156].

Since the game models introduced in this work are discrete time models, we focused exclusively on the discrete-time variant of Aloha, slotted Aloha, which was introduced by Roberts [157, 158]. In slotted Aloha, all nodes are synchronized, and time is divided into packet-sized slots. The basic idea of the algorithm is that when a node has a packet to send, it simply sends it in the next slot. If the packet is not correctly received, a “collision” is said to have occurred, and the node becomes backlogged. A backlogged node must retransmit its packet during a randomly selected slot in the future.

We altered this model slightly in that we sometimes assumed that nodes with newly arrived packets have the same behavior as backlogged nodes. This has little impact on the analysis discussed in this section, however, so set it aside for now.

For analytical simplicity, we adopted the usual assumption that there is an infinite set of nodes. This avoids the question of buffering, as each newly arriving packet arrives at a new node with probability 1. We typically assumed that packets arrive according to a Poisson random process with rate $\lambda$ arrivals/slot, although this assumption is not necessary for most of the development in this work.

There are two additional assumptions in the classical analysis of Aloha. The first assumption relates to the channel model. The classical channel model for Aloha is that of perfect reception or collision. If one packet is transmitted in a given slot, then it is assumed to be received perfectly. If more than one packet is transmitted in a given slot, then all packets are assumed to be completely destroyed. In other words, no packets are received in the presence of an interfering transmitter; that is, there is no capture effect.

The second additional assumption regards feedback. The classical assumption is that of ternary feedback. At the end of each slot, each node obtains feedback from the receiver as to whether 0 packets, 1 packet, or more than one packet were transmitted in that slot. This is also known as 0, 1, $e$ feedback, where $e$ stands for error.
If backlogged nodes use a fixed retransmission probability in each slot, then the slotted Aloha system is said to be unstable for any value of $\lambda > 0$. Unstable in this context means that the Markov chain over the number of backlogged nodes is not ergodic.

Yet slotted Aloha can be stabilized for some values of $\lambda$ by using the feedback from the receiver to adjust the retransmission probability of backlogged nodes. Perhaps the best-known means of stabilizing Aloha is Rivest’s pseudo-Bayesian algorithm [159]. As will be the case for much of our development, Rivest assumes that nodes with newly arrived packets behave identically to backlogged nodes. The pseudo-Bayesian algorithm uses feedback to update an estimate of the number of backlogged users using a reasonable approximation of Bayesian updating. (The pseudo-Bayesian algorithm assumes that the number of backlogged users is a Poisson random variable. If the a priori distribution over the number of backlogged users is Poisson and e feedback is received, then updating via Bayes rule will not produce a Poisson distribution. This distribution can be “reasonably approximated” as Poisson, however. [160])

With the pseudo-Bayesian algorithm, slotted Aloha can be stabilized for $\lambda < 1/e$ where $e$ is the root of the natural logarithm. This is known to be the maximum throughput which can be obtained if all backlogged users use the same retransmit probability, as it is the maximum throughput which can be obtained if the number of backlogged users is known perfectly.

Somewhat better throughput can be obtained via “splitting algorithms” which split the set of backlogged users into groups in order to resolve collisions. The best of these algorithms which is known is stable for arrival rates $\lambda < 0.4878$ [161, 162]. These algorithms, however, rely on users to determine their group membership on the basis of packet arrival time, a randomly generated number, or some other quantity. It is unclear whether self-interested nodes would accurately report their group membership.

For more information about the analysis of slotted Aloha including the derivation of many of the results mentioned in this section so far, see [160].

This study of Aloha uses a more sophisticated channel model. This model was developed by Ghez, Verdu, and Schwartz and allows for both capture and multipacket reception [163, 164]. The only requirement of the model is that the number of successful packets in each time slot depend only upon the number of packets transmitted. Let $\rho_{nk}$ denote the probability that $k$ packets are successfully received in a slot in which $n$ packets are transmitted. Taken together, these values form a reception matrix which defines the channel:

$$R = \begin{bmatrix}
\rho_{10} & \rho_{11} & 0 & 0 & \ldots & 0 & \ldots \\
\rho_{20} & \rho_{21} & \rho_{22} & 0 & \ldots & 0 & \ldots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\rho_{n0} & \rho_{n1} & \rho_{n2} & \ldots & \rho_{nn} & 0 & \ldots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\end{bmatrix}$$

In [163, 164], Ghez, Verdu, and Schwartz define the channel model and demonstrate computations for the maximum throughput which can be obtained for various channels with perfect information about the number of backlogged users. In [165, 166], they provide similar results for imperfect information.

Initially, we study games of perfect information. In terms of Aloha, this means that the number of backlogged users is perfectly known. We later develop a more general game type which admits various types of imperfect information, including the ternary feedback mentioned previously. Obviously, imperfect information is the case of greatest practical interest; unfortunately, the solution of imperfect information games is extremely difficult. We leave the analysis of these games to future work, though we do provide some suggestions for pursuing suboptimal strategies. In some cases, external estimates of the number of nodes transmitting or backlogged may be available, for example via a sensor network. (See, for instance, [167, 168].)

### 3.4 Carrier Sense Multiple Access (CSMA)

In many cases of practical interest, a node can hear the transmissions of some or all other nodes in the system after a small propagation delay. In these cases, it makes sense for the node to exploit this ability
by not transmitting if another node is currently transmitting. This is known as carrier sense multiple access, and was first analyzed by Kleinrock and Tobagi in 1975 [169]. Note that “carrier sensing” does not necessarily imply the use of a carrier per se. All that is required is the ability to detect transmissions and idle periods quickly. As with Aloha, if a collision does occur during the transmission of a packet, that packet must be retransmitted.

We continued to utilize a slotted model, as is typical in the analysis of CSMA, although most CSMA systems are unslotted. The standard development of CSMA shows, though, that most of the lessons learned in the analysis of slotted CSMA can be applied in understanding the unslotted version as well [160]. The length of a packet is normalized to 1 time unit. The most important parameter describing a CSMA system is the length of time required for propagation and detection, denoted $\beta$ and measured in packet length units. Note that for CSMA to be effective, we must have $\beta < 1$. A node in a CSMA system must detect that the channel is idle before transmitting. Hence, the length of a slot containing a packet or a collision will be $1 + \beta$, while the length of an idle slot will be $\beta$.

As for Aloha, it can be shown that if all backlogged nodes attempt retransmission with constant probability $p$ after each idle slot, then with Poisson arrivals of rate $\lambda > 0$, the system will never be stable. As for slotted Aloha, though, the system can be stabilized for some arrival rates. Specifically, the system can be stabilized for maximum arrival rates of approximately $1/(1 + \sqrt{2}/2\beta)$ for small $\beta$ and hence the maximum stable throughput approaches 1 as $\beta \to 0$. Once again, the detailed analysis appears in [160].

Carrier Sense Multiple Access with Collision Detection (CSMA/CD) is a variant of CSMA in which collisions are quickly detected so that collision slots can be shortened — much as idle slots are shortened in CSMA. In CSMA/CD, we assume that a collision can be detected in time $\beta$ after which $\beta$ additional time units will be required to detect that the channel is idle again. Hence, busy slots are of length $1 + \beta$, idle slots are of length $\beta$, and collision slots are of length $2\beta$. For slotted CSMA/CD, the maximum arrival rate for which the system can be stabilized is approximately $1/(1 + 3.31\beta)$ for small $\beta$ [160].

### 3.5 Contributions to the SensIT Program

In our work we applied a repeated game model to the power control problem. Other authors have attempted to model the power control problem game theoretically and have used pricing to attempt to overcome the inefficiency of a straightforward one-shot game model. Our innovation is the use of repeated games to show that patient players are able to enforce efficient operating points themselves.

A great many systems using medium access control protocols such as Aloha and CSMA are deployed. The performance of these systems in the presence of noncooperative nodes is unknown. As such, these systems may be vulnerable to disruption or serious performance degradation by users who either attempt to exploit the system for personal benefit or to maliciously deny service to other users. In this work, we do not focus on malicious users but note that such users can be easily modeled within games of population transition by adding a type of user with a different utility function — a utility function which reflects a preference for disrupting service to other users.

We contributed two game models which differ significantly from those in the existing literature. Both are innovative in featuring players who arrive in the game via an exogenous random process and depart based on strategic interactions. The second model, games of population uncertainty and transition, is also novel in the manner in which it expands Myerson's notion of population uncertainty to include a situation in which players are randomly entering and leaving a game.

Using these game models, we contributed to the understanding of medium access control protocols. We demonstrated a technique for computing throughput bounds of multipacket reception slotted Aloha systems with perfect information in which users are selfish and show how these techniques can be used to design an interaction mechanism with an optimal operating point. In addition, we computed similar bounds for CSMA systems with perfect information. Some other very recent work on Aloha and CSMA appears in [170] and [171].

Our results in this area can be found in the following publications.
The Energy-Robustness Tradeoff for Routing in Wireless Sensor Networks

Wireless sensor networks consisting of large numbers of inexpensive energy-constrained devices are expected to find a wide range of applications from vehicle tracking to habitat monitoring [123], [117], [126], [98]. Although sensor networks are primarily static in nature, the use of inexpensive devices is likely to result in higher rates of failures for individual nodes. It is therefore important for routing algorithms in this space to provide tolerance to such failures in an energy-efficient manner. This is the subject of our work.

One basic solution for robustness that has been proposed with several variations is multipath routing: the use of multiple disjoint or partially disjoint routes to convey information from source to destination.

There is considerable prior literature on multipath routing techniques which date back to work pertaining to the telephone system where they are used to minimize call blocking. The solutions for multipath routing in all kinds of networks have primarily aimed at providing a set of low-cost disjoint paths between the source and destination [29], [7], [8]. In recent years there has been a focus on multipath routing in mobile ad-hoc networks (MANETs), where the primary concern is path failure due to mobility. For example, alternate path routing is investigated in [56] as a mechanism for load balancing and protection against route failure in MANETs. An on-demand multipath routing scheme is proposed in [2] as a means to reduce query floods. An efficient heuristic scheme for selecting multiple reliable paths in MANETs is presented in [4]. Multipath algorithms for wireless networks are also proposed and studied in [1]. In the specific context of wireless sensor networks, the Directed Diffusion algorithm [109] is a routing protocol that allows for multiple alternate paths to be maintained by setting appropriate gradient levels. Partially disjoint multipath routing schemes described as “braided multipath” schemes for wireless sensor networks are studied in [3]. The energy-robustness tradeoff is also studied in [3], but with a focus on distinguishing between complete disjoint multipath routing and the braided multipath scheme.

The basic idea behind multipath routing is fault-tolerance through redundancy. An alternative philosophy for fault-tolerance is to minimize the number of failure modes by reducing the number of intermediate nodes prone to failure. In sensor networks this can be done even with single-path routing algorithms utilizing higher transmission ranges. We will show in what follows that there are in fact situations in which this alternative strategy is superior to multipath routing in terms of energy-efficiency.

In what follows, we consider a simple scenario involving a network of five sensor nodes and perform some analytical calculations on the energy and robustness metrics associated with various routing configurations. The example yields insight into why single path routing with higher transmit powers can potentially be a more energy-efficient mechanism for providing robustness. We then turn more detailed simulations involving 50 sensors. The experimental setup for these simulations is described in section 4.2. The results from these simulations are then presented and discussed in section 4.3. We conclude with a brief discussion in section 4.4.
4.1 Illustration

We begin our exploration of the energy-robustness tradeoffs by considering a simple, small configuration of five sensor nodes. The nodes are placed as seen in figure 3.

Each of the configurations shown in figure 3 represents a possible way to route information from the source to the receiver. If we assume that nodes can only communicate with other nodes within a common radius $R$, then there is a minimum radius required for each routing configuration to be possible. This is shown in the second column of Table 1. We assume that the energy required to transmit on a link is $R^a$, where the path loss exponent $a$ is typically between 2 to 5 (for dense networking situations it is closer to 2).

We define an energy metric for each routing scheme $H$ as follows: if the minimum common transmis-

<table>
<thead>
<tr>
<th>Routing Scheme $H$</th>
<th>Minimum Radius $R_H$</th>
<th>Energy Cost $E_H$</th>
<th>$E_H (a = 2)$</th>
<th>$E_H (a = 4)$</th>
<th>Robustness $\Pi_H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_1$</td>
<td>$d$</td>
<td>$3d^a$</td>
<td>$3d^2$</td>
<td>$3d^4$</td>
<td>$(1 - p)^2$</td>
</tr>
<tr>
<td>$H_2$</td>
<td>$d$</td>
<td>$4d^a$</td>
<td>$4d^2$</td>
<td>$4d^4$</td>
<td>$(1 - p)(1 - p^2)$</td>
</tr>
<tr>
<td>$H_3$</td>
<td>$d\sqrt{2}$</td>
<td>$2(d\sqrt{2})^a$</td>
<td>$4d^2$</td>
<td>$8d^4$</td>
<td>$(1 - p)$</td>
</tr>
<tr>
<td>$H_4$</td>
<td>$d\sqrt{2}$</td>
<td>$3(d\sqrt{2})^a$</td>
<td>$6d^2$</td>
<td>$12d^4$</td>
<td>$(1 - p)$</td>
</tr>
<tr>
<td>$H_5$</td>
<td>$d\sqrt{2}$</td>
<td>$4(d\sqrt{2})^a$</td>
<td>$8d^2$</td>
<td>$16d^4$</td>
<td>$(1 - p)$</td>
</tr>
<tr>
<td>$H_6$</td>
<td>$d\sqrt{2}(1 + \sqrt{2})$</td>
<td>$3(d\sqrt{2}(1 + \sqrt{2}))^a$</td>
<td>$10.2d^2$</td>
<td>$35.0d^4$</td>
<td>$(1 - p^2)$</td>
</tr>
<tr>
<td>$H_7$</td>
<td>$d\sqrt{2}(1 + \sqrt{2})$</td>
<td>$4(d\sqrt{2}(1 + \sqrt{2}))^a$</td>
<td>$13.7d^2$</td>
<td>$46.6d^4$</td>
<td>$(1 - p^3)$</td>
</tr>
<tr>
<td>$H_8$</td>
<td>$d(1 + \sqrt{2})$</td>
<td>$(d(1 + \sqrt{2}))^a$</td>
<td>$5.2d^2$</td>
<td>$34.0d^4$</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1: Energy and Robustness Measures for Alternative Routing Configurations
sion radius required for it is $R_H$, and $m_H$ transmissions are required, then the energy cost for the scheme $H$ is considered to be $m_H R_H^2$. Note that this metric charges each transmitting node the same energy cost irrespective of the number of neighbors that are receiving the information. The corresponding energy metrics for each scheme are shown in the third column of table 1. For clarity, the numerical solutions are provided for $\alpha = 2$ and $\alpha = 4$ in the adjoining columns.

For studying the effect of robustness to node failures we use the following model: we assume that each intermediate node (i.e. a node that is not the source or the destination) is liable to fail independently with probability $p$, and for simplicity we assume that both the source and the destination nodes are guaranteed to be working. The robustness metric $\Pi_H$ corresponding to the routing scheme $H$ is the probability that a message sent from the source can reach the sink given these independent failure probabilities. The calculation of this metric is in general a difficult problem, with no known polynomial algorithm [9]. For the sample configuration, however, this is easy to solve exactly as there are only three intermediate nodes involved. The corresponding robustness metrics for each routing scheme can be calculated to the values shown in the final column of table 1.

Let us now understand the implication of these calculations. First of all it is clear that if intermediate nodes are prone to failure then from a pure robustness perspective it is best to avoid using these nodes entirely. This is why the scheme corresponding to the scheme $H_8$ is clearly the best strategy as far as robustness is concerned. However this does not address the issue of energy-efficiency.

Since we are concerned with two different objectives: minimizing energy while maximizing robustness, it is helpful to make use of the notion of Pareto optimality. A routing scheme $H_i$ is said to dominate a routing scheme $H_j$ if it results in an equal or greater robustness level with strictly less energy cost or if it results in an greater or lesser energy cost with strictly higher robustness level, i.e. if $\Pi_{H_i} \geq \Pi_{H_j}, E_{H_i} < E_{H_j}$, or if $E_{H_i} <= E_{H_j}, \Pi_{H_i} > \Pi_{H_j}$. Routing schemes which are not dominated by others in the set of considered schemes are said to be Pareto optimal and constitute the Pareto set.

From Table 1, we see that for $\alpha = 2$, the Pareto set is $\{H_1, H_2, H_3\}$. What is remarkable is that in this particular case all the Pareto optimal routing strategies are single path routes. The multipath routing scheme $H_2$ is dominated by $H_3$ which provides greater robustness for the same energy level, and the multipath routing schemes $H_4, H_5, H_6, H_7$ are dominated by $H_8$ which provides greater robustness for less energy cost. Since the energy costs depend on the path loss exponent $\alpha$, the Pareto set is also dependent on this parameter. For $\alpha = 4$, as seen in Table 1, the Pareto optimal routing schemes are $\{H_1, H_2, H_3, H_8\}$. Here again, the multipath routing strategies $H_4$ and $H_5$ are dominated by $H_3$ because node $B$ acts as a bottleneck; the additional energy expenditure for multipath does not yield an increase in robustness in these cases. It is also remarkable that even in this higher path loss situation the multipath routing strategies $H_6$ and $H_7$, which result in 2 and 3 node-disjoint paths respectively, result in higher energy consumption than the high transmit power direct transmission in scheme $H_8$.

Although this is a simple analytical example with a small number of nodes and arbitrary placements, it provides an insight into why multipath routing is not always the best solution when the primary concerns are energy efficiency and robustness to intermediate node failures. The Pareto optimal sets we examined in the two cases $\alpha = 2$ and $\alpha = 4$ contain the three possible single path routing schemes.

We now turn to simulation results involving a greater number of nodes with random placement of nodes.

### 4.2 Experimental Setup

For the experiments, 50 nodes are placed in a square area with unit sides. We consider the flow of information from a single source to a single destination. The source is placed at $(0, 0)$, and the destination sink node is placed at $(1, 1)$. The simulation is repeated 100 times with random placements for the remaining 48 intermediate nodes. These nodes are placed at random in the square, independently with a 2D uniform distribution. For each simulation the transmission radius $R$ within which each pair of nodes can communicate is increased in increments of 0.05 from 0.05 to 1.5.
To test out some different single and multipath routing strategies we chose to simulate the family of forward-$k$ routing algorithms that work as follows. The sink first floods a query to all nodes in the network, and the source node responds by routing information in the reverse direction along the paths followed by the initial query. In the forward-$k$ protocol, the source sends its information to the first $k$ neighboring nodes that had sent the sink-initiated query to it. Each intermediate node also forwards this data to its “best” $k$ neighbors. The forward-1 protocol is essentially a single shortest path routing mechanism. The forward-2 and forward-3 protocols that we simulate are examples of braided multipath routing protocols. For comparison we also simulate basic flooding initiated by the source. This is a useful comparison point because of the following result:

**Proposition:** If the transmission radius $R$ for a static wireless network is fixed, the routing scheme $H_{\text{flood}}$ consisting of flooding information outward from the source results in the optimum robustness value. In other words, $\Pi_{H_{\text{flood}}} = \max H \Pi_H$.

This result holds because the directed graph corresponding to the flooding scheme $H_{\text{flood}}$ is maximally dense in that each directed edge and each node of the underlying topology is utilized. This maximizes the robustness to node failures.

### 4.3 Simulation Results

As we mentioned before, much of the prior investigation into multipath routing in wireless networks has focused on providing multiple node-disjoint paths for routing between a source and a destination node. It is intuitive that a forward-$k$ strategy results in greater number of node-disjoint paths as $k$ increases. Figure 4 shows how the number of node disjoint paths varies for the various schemes. It is noteworthy that the flooding is particularly effective as far as this metric is concerned.

We now turn to our robustness metric $\Pi_H$ - the probability that the source is able to send information to the sink in the presence of uniform random node failures. Figures 6 and 7 show how this metric varies with the transmission radius for failure rates of 5% and 20% respectively. We make two observations from these figures. The first is that for a given transmission radius, the single path routing mechanism does indeed provide much lower robustness than the multipath routing schemes. The second is that for low failure rates, the three multipath routing mechanisms all provide nearly the same level of robustness. In essence the additional redundancy provided by having more than 2 node-disjoint paths results in negligible gains in robustness for low levels of node failures. At the failure rate of 20% there is slightly greater differentiation between the different multipath routing schemes but one can again see the law of diminishing returns at play - flooding provides only negligibly greater robustness than the forward-3 routing protocol.

Thus far we have ignored one critical aspect: the energy expenditure. While the multipath routing schemes provide greater robustness for a fixed value of the transmission radius, they do, of course, do so at the cost of a greater number of transmissions. This can be seen in figure 5. Flooding requires an order of magnitude higher number of transmissions than even forward-3, showing it is clearly not an energy-efficient mechanism for providing robustness to node failures. This is still far from a clear picture of the energy-robustness tradeoffs. We have two parameters that we can tune to increase the energy and robustness metrics: one is the value of $k$, which in effect changes the routing structure without affecting the underlying topology. By increasing $k$, we apply energy in the form of greater number of transmissions in order to realize robustness gains through multiple paths. The second tunable parameter is the transmission radius: even if we stick to single path routing, increasing this parameter increases the robustness to node failures because it decreases the number of hops, leaving fewer possible failure nodes.

Hence we plot the robustness metric with respect to the energy metric $E_H = m_H R^\alpha$ which incorporates both the transmission radius $R$ as well as the number of transmission $m_H$. This is shown as scatter plots in figures 8 and 9 for failure rates of 5 and 20% respectively, for $\alpha = 2$. 

17
Figure 4: Number of Node Disjoint Source Sink Routes with respect to the Transmission Radius

Figure 5: Number of Nodes Transmitting with respect to Transmission Radius
Figure 6: Probability that a route exists with respect to Transmission Radius (5% failure rate)

Figure 7: Probability that a route exists with respect to Transmission Radius (20% failure rate)
Figure 8: Probability that a route exists with respect to normalized energy cost (5 % failure rate)

Figure 9: Probability that a route exists with respect to normalized energy cost (20 % failure rate)
Now we have a dramatically different view. The Pareto optimal points are towards the top left hand corner of the scatter plot. Towards the left hand side of the plot in figure 8, where the energy costs are kept low, it is clear that the single path routing mechanism forward-1 provides the best robustness to node failure. While there is a region where the multipath scheme forward-2 dominates, the remaining schemes are all dominated. It is clear from the plot that from both energy and robustness perspectives it is better to transmit directly from the source to sink than to use either forward-3 or flooding, or even forward-2 with a higher transmission range setting. Figure 9 shows the same behavior for higher failure rates as well. If there are severe energy constraints, this figure suggests that it is better to allocate the energy to increasing transmission range, than to transmit along multiple paths. This validates the insight gleaned from the simple example we explored in subsection 4.1.

4.4 Conclusions

Wireless sensor networks with large numbers of inexpensive individual devices are particularly prone to node failures. In several prior studies multipath routing schemes have been proposed in order to provide tolerance to such failures. We studied the issue of robustness to node failures in the particular context of energy-starved sensor networks, and showed that the robustness obtained from multipath routing can sometimes come at too high a cost.

Multipath routing is but one mechanism for trading off energy in order to increase robustness. An alternative to routing through many paths is the use of higher transmit powers with fewer paths, even a single path. We showed through the simple analytical example and a larger set of simulations that, depending on the constraints, this may be a more energy-efficient mechanism for robustness to node failures. In any case, wireless sensor network designers would be well advised to consider using higher transmission powers in order to boost robustness, in addition to multipath routing.

These results were documented in the following publications.

5 Appendix

Publications Resulting from SensIT Sponsored Research


On Selection of Optimal Transmission Power for Ad hoc Networks, accepted by Hawaii International Conference on System Sciences (HICSS-36), Big Island, Hawaii, Jan. 6-9, 2003.


Faculty and Students Supported
1. Lead Pis: Professors Stephen B. Wicker, Terrence Fine, Bart Selman, and Carla Gomes
2. Supported Students:
   • Bhaskar Krishnamachari (Ph.D., Spring 2002 - "Complexity and Phase Transitions in Wireless Networks")
   • Allen MacKenzie (Ph.D. student)
   • Yurong Chen (Ph.D. student, non-US)
   • Yasser Mourtada (Ph.D. student, non-US)

References


