2011

Cross-layer design for network performance optimization in wireless networks

Xuan Gong

Department: Computer Science

CROSS-LAYER DESIGN FOR NETWORK PERFORMANCE OPTIMIZATION IN WIRELESS NETWORKS

by

XUAN GONG

A DISSERTATION

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

In

COMPUTER SCIENCE

2011

Approved by

Dr. Maggie Cheng
Dr. Frank Liu
Dr. Fikret Ercal
Dr. Wei Jiang
Dr. Abhijit Gosavi

PUBLICATION THESIS OPTION

This thesis consists of the following six articles that have been submitted for publication as follows:

Pages 12-30, “Link Rate Allocation under Bandwidth and Energy Constraints in Sensor Networks” was published in IEEE Globecom 2008

Pages 31-65, “Joint Routing and Link Rate Allocation under Bandwidth and Energy Constraints in Sensor Networks” was published by IEEE Transaction on Wireless Communications.

Pages 66-83, “Interference Modeling, Multipath Routing and Link Rate Control in Multihop Wireless Networks” was accepted by IEEE Globecom 2010 – Wireless Networking Symposium.

Pages 84-97, “Minimum Latency Transmission Scheduling in Multihop Wireless Networks”, was submitted to IEEE ICC’11 AHSM.

Pages 98-122, “Cross-Layer Throughput Optimization with Transmission Power Control in Sensor Networks” was submitted to IEEE INFOCOM, 2011.

Pages 123-143, “Improving Sensor Network Lifetime Through Hierarchical Multihop Clustering” was published in IEEE ICC 2009.

ABSTRACT

In this dissertation, I use mathematical optimization approach to solve the complex network problems. Paper 1 and paper 2 first show that ignoring the bandwidth constraint can lead to infeasible routing solutions. A sufficient condition on link bandwidth is proposed that makes a routing solution feasible, and then a mathematical optimization model based on this sufficient condition is provided. Simulation results show that joint optimization models can provide more feasible routing solutions and provide significant improvement on throughput and lifetime. In paper 3 and paper 4, an interference model is proposed and a transmission scheduling scheme is presented to minimize the end-to-end delay. This scheduling scheme is designed based on integer linear programming and involves interference modeling. Using this schedule, there are no conflicting transmissions at any time. Through simulation, it shows that the proposed link scheduling scheme can significantly reduce end-to-end latency. Since to compute the maximum throughput is an NP-hard problem, efficient heuristics are presented in Paper 5 that use sufficient conditions instead of the computationally-expensive-to-get optimal condition to capture the mutual conflict relation in a collision domain. Both one-way transmission and two-way transmission are considered. Simulation results show that the proposed algorithms improve network throughput and reduce energy consumption, with significant improvement over previous work on both aspects. Paper 6 studies the complicated tradeoff relation among multiple factors that affect the sensor network lifetime and proposes an adaptive multi-hop clustering algorithm. It realizes the best tradeoff among multiple factors and outperforms others that do not. It is adaptive in the sense the clustering topology changes over time in order to have the maximum lifetime.

ACKNOWLEDGMENTS

The author would like to express my deepest gratitude to my advisor, Dr. Maggie Cheng, for her guidance, valuable advice, and encouragement to finish this dissertation. Her knowledge, skills and advice have been of great value. It has been a great experience to work under her.

The author also would like to thanks the member of his advisory committee, Dr. Frank Liu, Dr. Wei Jiang, Dr. Fikret Ercal and Dr. Abhijit Gosavi for their advice and their participation in the defense.

Thanks to all the students in the Network Research Lab at Missouri University Science and Technology for numerous instructive discussions and their kind support of these researches.

Finally, the author would like to thank his family for their help and encouragement. With their help, he has studied hard and lived happily in the peaceful town of Rolla, Missouri.

TABLE OF CONTENTS

Page

LIST OF ILLUSTRATIONS

LIST OF TABLES

1. INTRODUCTION

1.1 WIRELESS SENSOR NETWORK

A wireless sensor network (WSN) is a wireless network, consisting of spatially distributed autonomous sensors. After the initial deployment (typically ad hoc), sensor nodes are responsible for self-organizing an appropriate network infrastructure, often with multi-hop connections between sensor nodes. The onboard sensors then start collecting data, such as acoustic, seismic, infrared or magnetic information about the environment, using either continuous or event driven working modes. The flowing of data ends at special nodes called base stations (sometimes they are also referred to as sinks). When the sensor nodes do the sensing, transmitting, receiving and etc, they will consume their energy, usually the battery. If they run out of battery, these sensor nodes will die and it is very possible that the whole wireless sensor network will be out of service. Since the sensor nodes have limited battery and are hard to recharge or replace, energy efficient routing is important for wireless sensor network to make the network working as longer as possible. The bandwidth, on the other hand, has always been ignored. Actually, in a sensor network where every node transmits towards the sink, the aggregated bandwidth requirement can be surprisingly high. The bandwidth constraint can be used to decide not only the routing topology but also actually data rate on each link.

In this dissertation, the energy constraint and bandwidth constraint are jointly considered for routing and link rate allocation. Sufficient conditions for unidirectional transmission and bidirectional transmission have been discussed separately. Achievable and feasible wireless link rate can be found if the sufficient condition is satisfied.

1.2 LIFETIME

Network lifetime is critical to any wireless sensor network deployment. The goal of both the environmental monitoring and security application scenarios is to have nodes placed out in the field, unattended, for months or years without replacement or battery recharging.

Energy is the primary limiting factor for the lifetime of a sensor network. Sensor nodes have limited battery power. When they do sensing, transmitting and communication, they will consume the battery power. If they are out of power, it is very hard to replace or recharge. In that case, each node must be designed to manage its local supply of energy in order to maximize total network lifetime. In many deployments it is not only the average node lifetime that is important, but rather the minimum node lifetime. In the case of wireless security systems, every node must last for multiple years. A single node failure would create vulnerability in the security systems.

Thus, it is essential to develop protocols that optimize the overall energy utilization of the network, in order to maximize its capability to function for the longest possible time. However, the network lifetime objective in most of these efforts has been centered on maximizing the time until the first node fails. Although the time until the first node fails is an important measure from the complete network coverage point of view, this performance metric alone cannot measure the lifetime performance behavior for all nodes in the network. For wireless sensor networks that are primarily designed for environmental monitoring or surveillance, the loss of a single node will only affect the coverage of one particular area and will not affect the monitoring or surveillance capabilities of the remaining nodes in the network. This is because the remaining nodes in the network can adjust their transmission power (via power control) and reconfigure themselves into a new network routing (relay) topology so that information collected at the remaining nodes can still be delivered successfully to the base-station. Consequently, it is important to investigate how to maximize the lifetime for, not only the first node, but also all the other nodes in the network.

Many previous works addressed network lifetime optimization problem. In [1], it used network coding in multicast traffic and study the trade-off between maximizing the network lifetime and minimizing the network coding operations. Paper [2] divided network into a number of clusters, and improved the network lifetime by periodically choosing higher power node as cluster header to help relay the traffic to sink. The reliability constraint was introduced in [3], and was linked to the average amount of energy consumed by the network. So, it optimized the network lifetime under the reliability constraint (aka energy constraint). [4] also considered the energy efficient routing for maximizing the network lifetime and minimizing the energy multicast problem in ad-hoc network. The tradeoff between throughput and lifetime was discussed in [5], for the case of fixed conflict-free wireless networks. It employed a realistic interference model and provided several insights into interplay between throughputs, network lifetime and transmission power.

How to maximize network lifetime under one or more constraints was also investigated. Paper [6] provided a novel theory to improve the network lifetime of unicast multi-hop wireless sensor networks under the limited bandwidth. A bandwidth allocation scheme was proposed in [7] that used time-frequency slot assignment to reduce the energy consumption to improve the network lifetime. Energy-efficient multi-polling mechanism is discussed in [8] to combine power management strategy with a low overhead MAC protocol is 802.11 MAC. It scheduled the wake-up time slot for wireless stations to reduce the energy consumption, with loss of bandwidth as tradeoff to improve the lifetime. In [9], it constructed a global optimal maximum lifetime multicast tree in wireless static network with distributed manner under limited bandwidth capability. And [10] provided a probabilistic model for route lifetime prediction.

In this thesis, sufficient condition is discussed on link bandwidth that makes a routing solution feasible, then provide mathematical optimization models to tackle both energy and bandwidth constraints. One basic mathematical model is first presented to address using uniform transmission power for routing without data aggregation, and then extend it to handle non-uniform transmission power, and then routing with data aggregation. And two efficient heuristics are proposed to compute the routing topology and link data rate.

1.3 INTERFERENCE

In wireless sensor networks, due to the broadcast nature of wireless transmission, the signal from one sensor could reach many unintended receivers and interfere with the reception of these neighbors. The higher transmission power it uses, the more neighbors it interferes with. As the interference level increases, network throughput decreases. To intuitively understand how transmission power works on network throughput, take a multi-hop wireless sensor network with a fixed number of nodes as an example, if two nodes can hear from each other, a link between them can be built. When one link is active, any other link that interferes with it should not be. When transmission power increases, link density increases, and consequently a wireless link will have many other links interfering with it. All these conflicting links cannot be active at the same time; they must be carefully scheduled to transmit at different time, otherwise their transmissions will interfere with each other. Although the wireless link capacity remains the same, the spatial reuse of the wireless spectrum decreases as the transmission power increases. As a result, network throughput drops.

The question of how to achieve the maximum throughput in sensor networks through cross-layer optimization is addressed by many previous works. [26] used link-directional interference graph to clarify inter-link interference in wireless ad-hoc networks and proposed the coloring algorithm to set the interference domain. In [27], investigated the interaction between MAC protocol and interference in wireless multi-hop network, and jointly introduced the flow rate allocation. The interference-aware flow allocation algorithm was proposed to achieve the fair flow rate. The topology control problem and interference has been discussed in [28]. It formally defined the concept of path interference and designed an algorithm to construct an efficient topology with minimal path interference. In [29], the interference in wireless networks was characterized by using a conflict graph based model. The on-demand routing scheme was proposed to explicitly add the interference model in the route decision process. In the scheme, the nodes can exchange the flow information and compute the available residual bandwidth based on the local information periodically. Many previous works assumed that interference is a binary phenomenon. But in [30], it defined the term named partial interference and presented a framework to characterize the partial interference in a single-channel wireless network under unsaturated traffic condition. And it concluded that by using adapting the partial interference, the gain in capacity can be improved significantly.

In this dissertation, the interaction between interference and network throughput has been discussed. The collision domain and interference model are formally defined. The power control mechanism is used to optimize the interference and a related algorithm is presented to compute the transmission power of each node with objectives of minimizing total interference.

1.4 THROUGHPUT

In general computer networks, throughput is the amount of digital data per time unit that is delivered over a physical or logical link, or that is passing through a certain group of network nodes. In senor network, total amount of data received per second by the sink node is referred while every node except sink node can be a source node and send the data to the sink node.

Specifically, given initial energy for every sensor node in the network, if all nodes are required to satisfy a certain lifetime criterion, what is the maximum amount of data that can be generated by the entire network? Obviously, it appears reasonable to maximize the sum of rates from all the nodes in the network, subject to the condition that each node should meet the network lifetime requirement. Mathematically, this problem can be formulated as a linear programming (LP) problem within which the objective function is defined as the sum of rates over all the nodes in the network and the constraints are: (1) flow conservation is preserved at each node, and (2) the bandwidth constraint at each node can be met for the given network lifetime requirement. However, the solution to this problem shows that although the network capacity (i.e., the sum of bit rates over all nodes) is maximized, there exists a severe bias in the rate allocation among the nodes. In particular, those nodes that that are closer to the base-station will be allocated with much higher bit rates than other nodes in the network. Assume node A and node B are chosen as the source nodes. When the total throughput of the network is considered, it is easy to find if node B send the data as much as it can and node A do not send anything, the network throughput will achieve the maximum. Because node A is far from the sink node, if it want to send data to the sink node, it need many reply node to be the receiver and these nodes will consume the bandwidth, but if node B is the only node which send the data to the sink node (node B is only one hop from sink node), it does not need relay node. Under the bandwidth constraint, node B will send as much as it can and node A will do nothing in the effort to get the maximum throughput.

The fairness issue associated with the network capacity maximization objective calls for a careful consideration in the link allocation among the nodes. In this thesis, this fairness issue has been considered and the center condition has been set to achieve the fairness.

[11] used wireless network coding to improve network throughput and spectrum efficiency. An analytical framework with fairness requirement is proposed to exploit the best coding opportunities to improve the network throughput. And [12] considered throughput and delay problem employing network coding and slotted ALOHA protocol,

and analyzed the performance on relay nodes which used queuing system as buffer. On the other hand, [13] issued the basic limitations for network coding in terms of energy and throughput in multi-hop wireless networks. Two well accepted scenarios: single multicast session and multiple unicast session, are used to illustrate that the gain of network coding is limited in term of throughput and energy saving. In [14], it gave a statistical method to estimation the maximum achievable end-to-end throughput in 802.11 based wireless mesh network. In this method, the 802.11 MAC is adapted to check contention for wireless nodes.

The trade-off between energy and throughput or the trade-off between throughput and lifetime has been further discussed. A network region size threshold is provided in [15]. If network region size is below the threshold, direct transmission routing can be both energy conserving and throughput achieving. Otherwise, energy efficient routing may not achieve the maximum throughput. In [16], it investigated the trade-off between throughput and network lifetime. For a fixed transmission power, relaxing throughput requirement may result in a significant improvement on the network lifetime. It also showed that with fixed throughput requirement, the lifetime is not monotonic with power.

In this dissertation, the questions how to improve the total throughput under the energy constraint and bandwidth constraint and how to achieve fairness have been discussed. The proposed heuristics computes the link-rate allocation and routing path. The simulation results show that they can significantly improve the throughput compared with the previous works.

1.5 END-TO-END DELAY

The end-to-end delay refers to the total time taken for a single packet to be transmitted across a network from source to destination. It is one of the most important and fundamental issue for wireless sensor network. Many applications require an end-to-end latency guarantee for time sensitive data. However, it is hard to bound end-to-end delay for event-driven sensor networks, where nodes produce and deliver data only when an event of interest occurs, thus generate unpredictable traffic load.

How to improve the throughput under delay requirement or how to minimize the end-to-end delay under throughput requirement have been investigated in many previous works. [17] proposed a scheduling algorithm. This algorithm resolved real-time problem of cycle communication task with the character of network topology. In paper [18], delay is investigated in a hybrid wireless network consisting of n randomly distributed normal nodes, and m regularly placed base stations connected via an optical network. With dense networks, the area is fixed and the node density increases linearly as the number of nodes, and [19] assume the whole network is connected. Furthermore, [20] also considered the dense networks, but with area increasing linearly with node. All of three papers give the average packet delay estimation under the per-node throughput capacity constraints. The trade-off between throughput and delay was investigated in [21]. It provided the packet scheduling policy and a method based on queue model for analyzing the packet delay. It also justified that the trade-off remains unchanged with fixed-size packet. Both centralized and distributed algorithms for delay aware routing are proposed in [21] and hybrid architecture which consist wireless sub-network is also introduced. The difference between [22] and [23] is that in [23], the wireless routers are modeled as M/M/1 queue and wireless link states are predicted periodically. The algorithm proposed in [23] also considered load balance and congestion instead of traditional minimum hops. In [24], it minimized the average end-to-end delay by obtaining the optimum link capacity. And a distributed optimization framework is proposed in [25] to improve the end-to-end delay in a multi-hop single-sink wireless sensor network.

How to minimize end-to-end latency in a multi-hop wireless network is addressed in this thesis. The transmission scheduling scheme is presented that minimizes the end-to-end delay along a given route. The link scheduling scheme is based on integer linear programming and involves interference modeling. Using this schedule, there are no conflicting transmissions at any time. Through simulation, the proposed link scheduling scheme can significantly reduce end-to-end latency. By varying different routing policy, the shortest path routing does not necessarily result in minimum delay.

1.6 MAIN CONTRIBUTION

The major contributions of this thesis includes: (1) the energy and bandwidth-constrained routing problem has been formulated as a multi-constraint optimization problem and provided efficient heuristic solutions to it. In addition, a companion time slot assignment algorithm is proposed to support the resulting routing solution at the MAC layer. (2) A linear optimization model has been generated to capture the impact of wireless interference on network delay in multi-hop wireless networks. Compared to previous linear models, this linear model is more accurate; and compared with the exact solution, which is a NP-hard, the solution is more efficient. (3) Another linear model has been proposed to capture the impact of wireless interference on achievable data rates in multi-hop wireless networks. Based on this linear relation, a linear programming model of joint routing and rate control has been presented to achieve both efficiency and fairness in multi-hop wireless networks. This model can be extended to work around loss links in a heterogeneous network to improve throughput performance. The model is not only critical for cross layer optimization, but also useful in a classic separate layer scheme -- it can be used to predict throughput performance, or to control source rate to improve network throughput or fairness when routing information is given. (4) The maximum throughput power control problem has been divided into two sub linear programs and related efficient algorithms have been designed to solve them. The power control algorithms can generate symmetric or asymmetric links as required; (5) for both symmetric links and asymmetric links, we provided mathematical optimization models to compute the maximum achievable throughput on a given topology. Part of it requires to accurately capturing the mutual conflicting relation among wireless links, which is a well-known NP-hard problem. A polynomial-term constraint has been proposed that can sufficiently capture the mutual conflict relation among wireless links and is tighter than all known polynomial-term approximations in previous works; (6) A linear optimization model is presented to capture the impact of wireless interference on network delay in multi-hop wireless networks. Compared to previous linear models, this linear model is more accurate, and compared to the exact solution, which is a NP-hard to compute, it is more efficient.

PAPER

I. LINK RATE ALLOCATION UNDER BANDWIDTH AND ENERGY CONSTRAINTS IN SENSOR NETWORK

Maggie Cheng[1], Xuan Gong[1] and Lin Cai[2]
[1]Department of Computer Science, Missouri University of Science and Technology
Rolla, Missouri, USA
Dept. of Elec. & Comp. Engineering, University of Victoria
Victoria, BC V8W 3P6, Canada
{chengm@mst.edu, xgwt2@mail.mst.edu, cai@ece.uvic.ca}

ABSTRACT. In sensor networks, both energy and bandwidth are scarce resources. In the past, the energy efficient routing problem has been vastly studied in order to maximize network lifetime, but link bandwidth has been optimistically assumed to be abundant. As energy constraint affects not just the routing topology but also the allowed data rate on each link, which in turn affects lifetime. Previous works that focus on energy efficient operations in sensor networks with the sole objective of maximizing network lifetime only consider the energy constraint and ignore the bandwidth constraint. This article shows how infeasible these solutions could be if bandwidth does become a constraint, then provides a new mathematical model to tackle both energy and bandwidth constraints. Two efficient heuristics are proposed based on this model. Simulation results show these heuristics provide more feasible routing solutions than previous works, and provide significant improvement on throughput.

1. INTRODUCTION

Wireless sensor networks are resource scarce, which is manifested in both energy and link bandwidth, as well as computing power etc. While it has been widely accepted that energy constraint limits the total amount of data being transmitted, and plays an important role for sensor network lifetime, bandwidth constraint has long being ignored. In previous work related to energy efficient routing and data aggregation etc., wireless link bandwidth is often optimistically assumed to be large enough. Actually, in a sensor network where every node transmits towards the sink, the aggregated bandwidth requirement can be surprisingly high. Even in a simple chain topology, if the link raw bandwidth is B, the allowed source rate is only 1/3 B as shown in Fig. 1.1, because the transmission of the source node is conflictive with that of its next hop and next next hop. It could be worse in a complicated network topology. If the required bandwidth is higher than link capacity, there won't be a guaranteed end-to-end throughput, nor end-to-end delay, which is devastating to delay-sensitive applications.

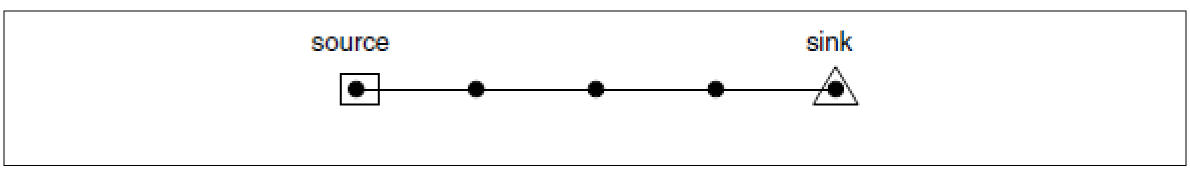

Fig.1.1 In this simple chain topology, link bandwidth B needs to be three times source rate R in order to have a guaranteed data rate R.

In most previous work on energy efficient routing, routing decisions are made to optimize the energy aspect and tend to ignore the bandwidth limitation. In the following example given in Fig. I.2(a), a maximum lifetime routing algorithm would choose any of the routing topologies shown in Fig. I.2(b), (c) and (d) because they all lead to the same lifetime. However, (b) and (c) demand much higher bandwidth than (d). Suppose that there exists an optimal MAC layer solution that requires the minimum bandwidth to support a given routing. If the source is generating 3 units of data per second, (b) requires a bandwidth of 7 units per second by the optimal solution (and 9 units per second by our condition in Section 3); (c) requires 9 units per second by the optimal solution (and 9 units per second by our condition); and (d) only requires 4.5 units per second by the optimal solution (and 4.5 units per second by our condition).

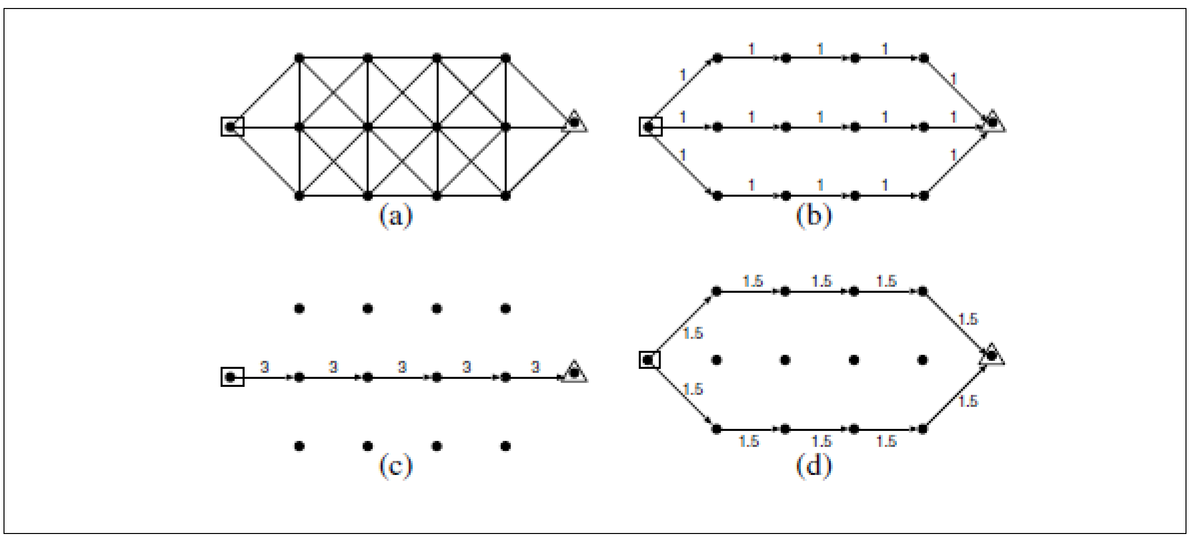

Fig.1.2 For the network shown in (a), the three routing options (b), (c) and (d) lead to the same lifetime, but (b) and (c) demand higher bandwidths than (d).

Yet in a slightly different scenario shown in Fig. I.3, the solution that provides the longest lifetime is actually the worst in terms of bandwidth requirement. A shortest path routing algorithm would choose (b) for the purpose of maximizing lifetime, but the required bandwidth may be too high to accommodate.

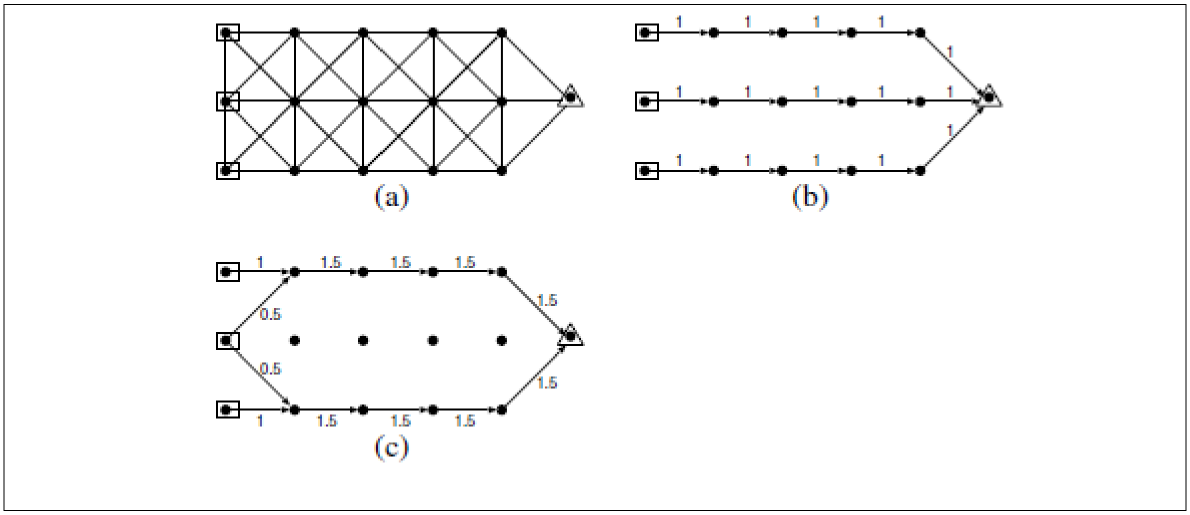

Fig.1.3 For the network shown in (a), both (b) and (c) use shortest paths routing; (b) is optimal in terms of lifetime, but is the worst in terms of bandwidth; (c) is the best in terms of bandwidth, but is suboptimal in terms of lifetime.

From the above two examples, we observed that for a randomly deployed network, usually the one that is likely to be used as a relay node is at the core of the network (if everyone choose what is best for itself selfishly), which unfortunately is also the most highly interfered area due to the broadcast nature of wireless transmissions. Sending a lot of data to the core is likely to congest the network, so it is desirable to detour the traffic before it is congested. However, it is difficult to enforce a generic policy on how traffic should be routed, and sending every packet along the outlier is not the solution either.

This work provides a solution that decides not only the routing topology but also the actual data rate on each link, rather than a generic policy. Link rates are computed by solving an optimization problem that has included both energy and bandwidth constraints.

The above observations lead us to a puzzle: for an arbitrary network topology, what condition(s) should hold in order to ensure all data generated by sources can be put through, with each source generating data at a fixed rate? In this article, we elaborate on the necessary and sufficient conditions on the link bandwidth, and use the bandwidth constraint to decide the actual amount data each node can send, which provides a basis for sensor network lifetime analysis. The major contribution of this work is that we formulated the energy and bandwidth-constrained routing problem as a multi-constraint optimization problem and provided efficient heuristic solutions to it.

The rest of the paper is organized as follows: Section 2 briefly surveys previous work related to transmission scheduling and energy efficient routing; Section 3 formally describes the energy-bandwidth constrained routing problem and provides a mathematical model for the problem; Section 4 presents two heuristics for joint optimization of energy and bandwidth; Section 5 provides numerical simulation results that show the comparison of algorithms in terms of throughput performance and how joint optimization solves lifetime problem differently; Section 6 concludes the article with directions for future research.

2. RELATED WORK

The most related work includes one paper from our previous work on edge coloring for transmission scheduling [1] and one paper by Lall et al. [2]. In [1], we precisely depicted the conflict relation among transmissions with each color corresponding to one time slot at MAC layer. It guarantees conflict-free time slot assignment if each edge carries the same load. However, edge coloring by itself is NP-complete, and it assigns one color to each edge which implies it works best for uniform traffic load. Link rate allocation in this article is an extension from color assignment, but it works well for arbitrary traffic load because the number of time slots each edge gets is proportional to the traffic load on the edge; and furthermore, we consider nodes' energy constraint for link rate allocation. In [2], the authors proposed a distributed algorithm to compute link rates with an objective of maximizing the network lifetime. The major contribution is on the distributed implementation of the optimization algorithm. However, like most previous work on energy efficient routing in sensor networks, bandwidth is not taken into consideration. Similar work along this line includes [3]–[11] and many others.

In [3], the proposed routing algorithms select the routes and the corresponding power levels such that the network lifetime is maximized. In [4], the routing problem is formulated as a linear programming problem, where the objective is to maximize the network lifetime, which is equivalent to the time until the network partition due to battery outage. Packet aggregation techniques were proposed to further reduce the energy consumption rate [5], [6], [8]. In [7], it was proposed to deploy a network clustering scheme and assign a less-energy constrained gateway node to act as a centralized network manager to further improve the energy efficiency and maximize network lifetime. Cui et al. further considered energy-efficient routing, scheduling, and link adaptation strategies together to maximize the network lifetime in [9], but the authors did not explicitly consider the bandwidth constraint in an arbitrary topology as we do. How to arrange the location of base-stations for WSN and select relay paths to maximize the network lifetime was discussed in [10], [11].

3. MATHEMATICAL MODEL

3.1 Problem Definition

Assume that in a sensor network of n nodes, each wireless link has raw capacity B (bits per second), and each node i has initial battery energy E_i (J). Each node i generates sensory data at a rate of R_i bits per second ($R_i > 0$ if node i is a source, $R_i = 0$ if it is a pure relay node, and $R_i < 0$ if it is a sink). Assume that nodes consume energy on transmitting, receiving and sensing (i.e., generating sensory data), and their energy consumption rates are P_t, P_r, and P_s J per bit respectively. Further assume P_t, P_r and P_s are constants in this paper.

The energy-bandwidth constrained maximum lifetime routing problem can be formally stated as follows: Suppose that sources are preselected and each node i's rate R_i is known, but the transmission rate from node i to node j is unknown. Let T be the total network lifetime. The rate allocation problem is to compute the data rate R_{ij} on each link (i, j), given each node i's E_i, R_i and link capacity B, such that the total network lifetime T is maximized and the rate allocation can be accommodated by wireless link capacity and energy reserve.

3.2 Multi-Constraint Optimization Problem

Since every node uses the same transmission power, therefore, links are all symmetric. We define N_i as the neighboring nodes of i excluding i itself. To maximize lifetime T is equivalent to minimize 1/T. For convenience, variables fi is introduced:

$$f_i = 1, if \sum_{j \in N_i} R_{ij} > 0$$

$$f_i = 0, otherwise$$

Thus $f_i = 1$ if node i is a receiver. Thus we can formulate the rate allocation problem as the following.

Table 1. Mathematical Model for Multi-Constraint Optimization Problem

Minimize: 1/T (1)

Subject To:

$$\sum_{j \in N_i} (R_{ij} - R_{ji}) = R_i \qquad \forall i \text{ (2a)}$$

$$P_s R_i + \sum_{j \in N_i} (P_r R_{ji} + P_t R_{ij}) \le E_i / T \qquad \forall i \text{ (2b)}$$

$$\sum_{j \in N_i} R_{ij} + f_i \sum_{j \in N_i} \sum_{k \in N_j} R_{jk} \le B \qquad \forall i \text{ (2c)}$$

$$0 \le R_{ij} \le B \qquad \forall i, \forall j \text{ (2d)}$$

$$f_i = \{0,1\} \qquad \forall i \text{ (2e)}$$

In this formulation, the sensing nodes have source rates $R_i > 0$, the sink nodes have $R_i < 0$, and the pure relay nodes have $R_i = 0$. Equality (2a) indicates that data rates R_{ij} satisfy flow conservation at each node. Inequality (2b) is the energy constraint, and inequality (2c) defines the bandwidth constraint.

In wireless communication, the capacity constraint is different from that in a flow network, where each link (u, v) has a fixed link capacity c(u, v) and flow f(u, v) $\leqslant$ c(u, v) must be satisfied on each individual link. In wireless communications, because of the broadcast nature of transmission, the capacity constraint needs to be considered on a collision domain, rather than on each link separately. In other words, how much can be transmitted over one link depends on not only the link raw capacity B, but also the amount of data transmitted over other links in the same collision domain. Inequality (2c) ensures all transmissions possibly in the same collision domain have a total demand less than B, which is a sufficient but not necessary condition for conflict free transmissions—the sufficient condition guarantees if a TDMA scheme is used at the MAC layer, we can always find a conflict-free transmission schedule.

4. HEURISTICS

The mathematical model defined by objective (1) and inequalities (2a - 2e) considers the bandwidth constraint while optimizing sensor network lifetime, therefore the solution to this model contains the optimal solution to the energy-bandwidth constrained maximum lifetime routing problem. However, it is not linear because fi is also a variable. In the following, we will present two heuristics that both work around the nonlinear problem by using information from the shortest paths (in terms of hops) from sources to the sink. The shortest paths represent the minimum-energy routing topology if data is not aggregated [12]. Heuristic I bears the characteristics of the shortest path routing, and Heuristic II bears the characteristics of the mathematical-programming based optimal solution, but they both include bandwidth constraints for consideration.

4.1 Heuristic I : Scalable Rate Allocation on Shortest Paths

The first heuristic starts from the shortest paths from sources to the sink, but the rate on each link is determined by the available bandwidth.

Table 2. Heuristic I: Scalable Rate Allocation on Shortest Paths

1) Compute the shortest path from each source node to the sink
2) Assume source rate is one unit, check against condition (2c) for each node, and find the most bandwidth-contentious node i. Let LHS=required bandwidth of node i's collision domain. Then compute the scale factor ∂ : ∂ = B/LHS. Set Δf =min{a/2, R_i}
3) Push out Δf amount of flow from each source to the sink then update the remaining input flow $R_i^{'}$=R_i - Δf for each source t

Table 2. Heuristic I: Scalable Rate Allocation on Shortest Paths (Continue)

4) Repeat (5)-(7) until we push through $R_i^{'}$ for each source i or the network is fully saturated
5) Find the shortest paths for nodes with $R_i^{'}>0$ based on the current available nodes and links. Nodes that are saturated on (2c) and their neighbors are not eligible for replaying. In case of a tie, give higher priority to nodes with more remaining energy; if there is still a tie, give higher priority to nodes with smaller degree
6) Decide the scale factor ∂ in a similar manner as in step (2). If pushing min{ ∂, $R_i^{'}$} units does not decrease lifetime, then set Δf =min{ ∂ , $R_i^{'}$}, otherwise, set Δf =min{ $\partial/2$, $R_i^{'}$}
7) Push out Δf amount of flow from each source with $R_i^{'}>0$ then update the remaining input flow $R_i^{'} = R_i^{'} - \Delta f$

In steps 2) and 6), this algorithm uses $\partial/2$ when computing Δf for the purpose of load balancing, which makes the network last longer. A simplified version is to use a instead of $\partial/2$ when we compute Δf . It runs faster but provides shorter lifetime.

4.2 Heuristic II: Optimizing Lifetime With Bandwidth Constraint

Since the mathematical model defined in (1) and (2a- 2e) has an objective of maximizing lifetime, if we can convert it to a linear program in a controlled manner, it is likely to produce a close-to-optimal solution in terms of lifetime. The following describes a heuristic that chooses the likely-to-be relay nodes and sets their f_i = 1 to make the program linear.

It can be observed from the algorithm description and also from the simulation results that if the link bandwidth is abundant, Heuristic II finds the optimal solution for maximum lifetime exactly the same way as MaxLife does in [2]; However, when the bandwidth becomes a limiting factor, Heuristic II can still find feasible routing solutions up to certain point while MaxLife cannot.

Table 3. Heuristic II: Optimizing Lifetime With Bandwidth Constraint

1) Set f_i=1 for the sink, and f_i=0 for all other nodes, solve the linear program, update f_i=1 if $\sum_{j \in N_i} R_{ji} > 0$; if (2c) is satisfied $\forall i$, return link rate R_{ij} for all (i,j), otherwise, go to line 2
2) Compute the shortest path from source nodes to the sink
3) Set f_i=1 for receiving nodes; solve the linear programming; if $\sum_{j \in N_i} R_{ji} > 0$ and f_i=0, update f_i=1
4) Repeat line 3 until there is no update for f_i (converged) or the linear program becomes infeasible
5) If it converges, output link rate R_{ij} for all links (i,j)
6) If it becomes infeasible: if f_i = 1 but $\sum_{j \in N_i} R_{ji} = 0$ set f_i=0 and R_{ji}=0, $\forall j \in N_i$ as input, solve the linear program again; if it is still infeasible, report infeasible.

Heuristic II will either terminate with a valid solution or become infeasible. There won't be endless iterations in line 4. In most of the simulations, it requires solving the linear program two to four times to get a suboptimal solution. If it does become infeasible, it is likely because the given source rates R_i are more than what the network can put through.

5. SIMULATION

In the following simulation study, we use the same energy consumption model as in [2] -- assume that energy consumption is mainly due to transmitting; receiving and sensing consume very small amount of energy and therefore are ignored. But it is worth mentioning that our mathematical model can handle none-zero P_s and P_r as shown in inequality (2b).

In the simulation study, we investigate how the bandwidth constraint can change the routing decision and eventually affect the lifetime of the sensor network. First, we compare the existing algorithms with our two heuristics and observe which algorithm is more likely to cause network congestion and fail to push through the applied load. In a network of 50 nodes with node positions randomly chosen, we randomly select 4 source nodes and apply increasing source rate on them. We ran the optimal solution for maximizing lifetime from [2](labeled as MaxLife), shortest path routing(labeled as SPR), and Heuristic I and Heuristic II proposed in this paper. We found that when each source node's data rate R_i is increased to 12% ~ 13% of the given link bandwidth, MaxLife starts to congest, i.e., some collision domain requires more bandwidth than what is available, and SPR starts to congest when it is increased to 15%. Heuristic I can push through without congestion when the load is increased to 18% and Heuristic II can support as much as 16%. The vertical lines in Fig.1.4.(a) and (b) indicate after this point, increased data rate cannot be put through.

In the second simulation, we compare four algorithms on their contribution toward lifetime. As shown in Fig 1.5, when there is enough bandwidth, MaxLife does not have bandwidth violations and achieves the optimal solution; Heuristic II achieves the same optimal solution; but when bandwidth does pose a constraint, Heuristic II can still push through 33% more data than MaxLife, and Heuristic I can push though 50% more data than MaxLife. Heuristic II achieves the best performance on lifetime and second best on throughput; heuristic I achieves the best performance on throughput, which is consistent with our observation from the first simulation in Fig 1.4.

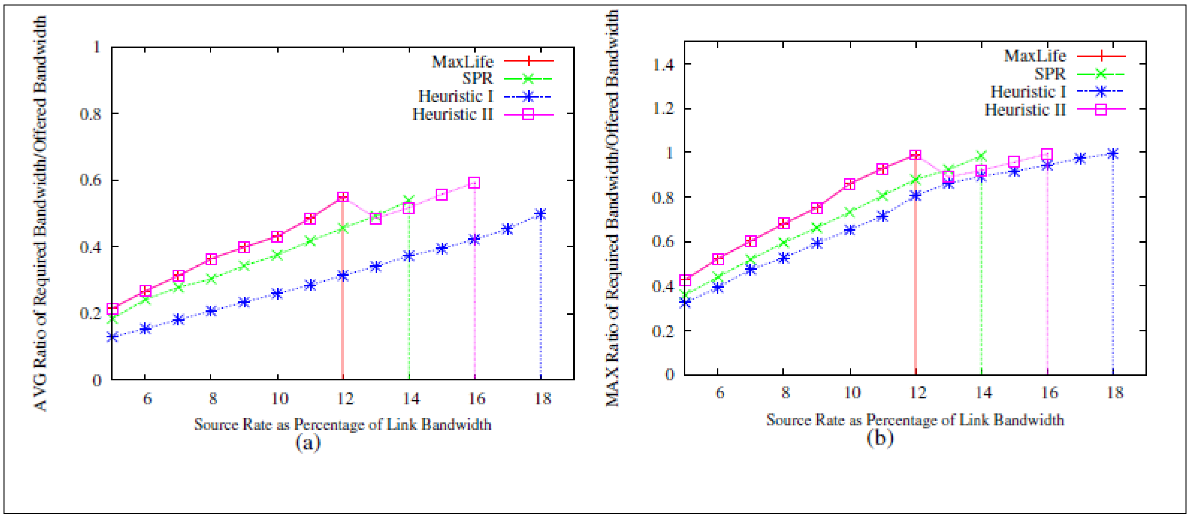

Fig.1.4 (a) The average ratio of required bandwidth/offered bandwidth; (b)the maximum ratio of required bandwidth/offered bandwidth

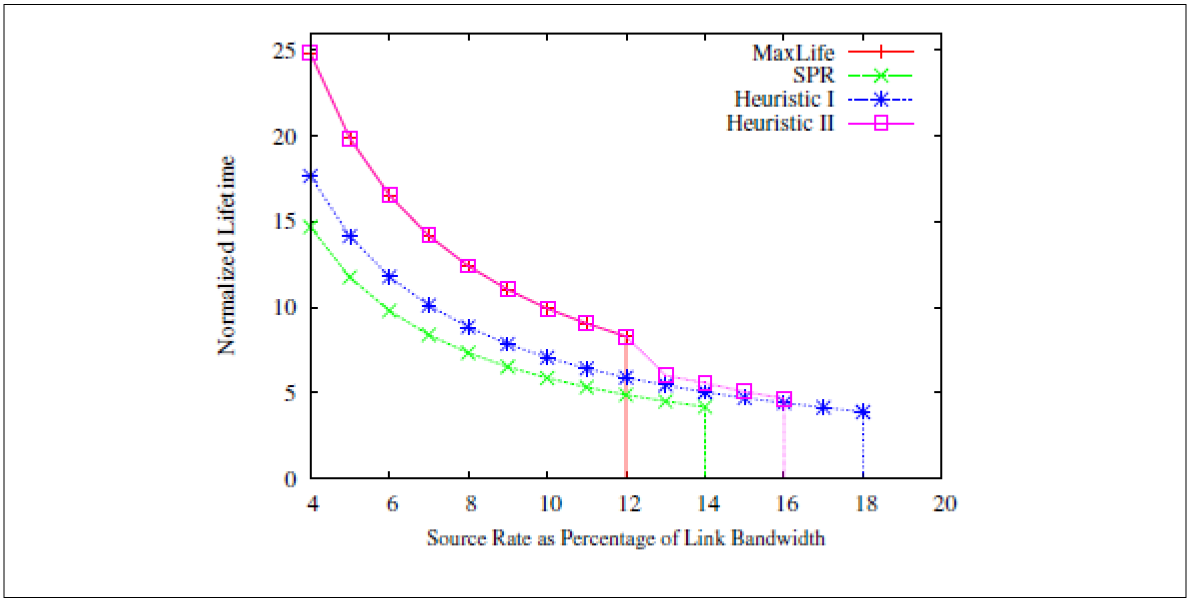

Fig.1.5 Normalized lifetime, assuming sending one unit of data consumes 10% total energy.

6. CONCLUSION

This article provides a generic mathematical model for the optimal routing problem in an energy and bandwidth-constrained sensor network. Using the sole constraint of energy sometimes leads to unrealistic solutions that cannot be accommodated by the link capacity. This work elaborated on the sufficient condition that a given traffic load can be put through a given network and jointly optimized on both energy use and bandwidth allocation. The solution provides not only the routing topology but also the amount of data flow that should be routed to each path. The joint optimization guarantees that there exists a conflict-free time slot assignment to support the given routing solution. To the best of our knowledge, this is the first work that explicitly considers bandwidth constraint in solving an maximum lifetime routing problem in a sensor network with arbitrary topology.

7. REFERENCES

[1] M. Cheng and L. Yin, “Transmission scheduling in sensor networks via directed edge coloring,” in IEEE ICC 2007 proceedings.

[2] R. Madan and S. Lall, “Distributed algorithms for maximum lifetime routing in wireless sensor networks,” in GLOBECOM’04, pp. 748–753.

[3] J. Chang and L. Tassiulas, “Energy conserving routing in wireless ad-hoc networks,” in IEEE INFOCOM 2000, pp. 22–31.

[4] Joongseok Park. Sartaj Sahni, “Maximum lifetime routing in wireless sensor networks,” IEEE/ACM Transactions on Networking, vol. 12, no. 4, pp. 609–619, August 2004.

[5] C. Schurgers and M. B. Srivastava, “Energy efficient routing in wireless sensor networks,” in MILCOM 2001, pp. 357–361.

[6] M. Bhardwaj, T. Garnett, and A. Chandrakasan, “Upper bounds on the lifetime of sensor networks,” in IEEE ICC’01, pp. 785– 790.

[7] M. Younis, M. Youssef, and K. Arisha, “Energy-aware routing in cluster-based sensor networks,” in IEEE MASCOTS 2002, pp. 129– 136.

[8] B. Krishnamachari, D. Estrin, and S. Wicker, “Modelling datacentric routing in wireless sensor networks,” in IEEE INFOCOM’ 02.

[9] S. Cui, R. Madan, A. Goldsmith, and S. Lall, “Joint routing, MAC, and link layer optimization in sensor networks with energy constraints,” in IEEE ICC’05, pp. 725–729.

[10] I. Kang and R. Poovendran, “Maximizing static network lifetime of wireless broadcast ad hoc networks,” in IEEE ICC’03, pp. 2256– 2261.

[11] J. Pan, T. Hou, L. Cai, Y. Shi, and X. Shen, “Topology control for wireless sensor networks,” in ACM MOBICOM’03, pp. 286 299.

[12] M. X. Cheng and L. Yin, “Energy-efficient data gathering algorithm in sensor networks with partial aggregation,” Int. J. Sensor Networks, 2008.

II. JOINT ROUTING AND LINK RATE ALLOCATION UNDER BANDWIDTH AND ENERGY CONSTRAINTS IN SENSOR NETWORK

Maggie Cheng[1], Xuan Gong[1] and Lin Cai[2]
[1]Department of Computer Science, Missouri University of Science and Technology
Rolla, Missouri, USA
[2]Dept. of Elec. & Comp. Engineering, University of Victoria
Victoria, BC V8W 3P6, Canada
{chengm@mst.edu, xgwt2@mail.mst.edu, cai@ece.uvic.ca}

ABSTRACT. In sensor networks, both energy and bandwidth are scarce resources. In the past, many energy efficient routing algorithms have been devised in order to maximize network lifetime, in which wireless link bandwidth has been optimistically assumed to be sufficient. This article shows that ignoring the bandwidth constraint can lead to infeasible routing solutions. As energy constraint affects how data should be routed, link bandwidth also affects not only the routing topology but also the allowed data rate on each link. In this paper, we discuss the sufficient condition on link bandwidth that makes a routing solution feasible, then provide mathematical optimization models to tackle both energy and bandwidth constraints. We first present a basic mathematical model to address using uniform transmission power for routing without data aggregation, and then extend it to handle non-uniform transmission power, and then routing with data aggregation. We propose two efficient heuristics to compute the routing topology and link data rate. Simulation results show that these heuristics provide more feasible routing solutions than previous work, and provide significant improvement on throughput and lifetime.

1. INTRODUCTION

Wireless sensor networks are resource scarce, which is manifested in both energy and link bandwidth, as well as computing power etc. While it has been widely accepted that energy constraint limits the total amount of data being transmitted, and plays an important role for sensor network lifetime, bandwidth constraint has long being ignored. In previous work related to energy efficient routing and data aggregation [1]–[10], wireless link bandwidth is often optimistically assumed to be large enough. Actually, in a sensor network where every node transmits towards the sink, the aggregated bandwidth requirement can be surprisingly high. Even for a single path with three or more hops between a source and a sink, if the link bandwidth is B, the allowed source rate is only 1/3 B, because the transmission of the source node is conflictive with that of the next two hops. It could be worse in a complicated network topology. If the total required data rate is higher than the link bandwidth on any particular link, the source rate cannot be supported, and network congestion is doomed.

In most previous work on energy efficient routing, routing decisions are made to optimize the energy aspect and tend to ignore the bandwidth limitation. For the network shown in Fig.2.1(a), a maximum lifetime routing algorithm would choose any of the routing topologies shown in Fig.2.1(b),(c) and (d) because they all lead to the same lifetime. However, (b) and (c) demand much higher bandwidth than (d). Suppose that there exists an optimal MAC layer solution that requires the minimum bandwidth to support a given routing. If the source is generating 3 units of data per second, (b) requires a bandwidth of 6 units per second by the optimal solution (and 9 units per second by our sufficient condition defined in Section 2); (c) requires 9 units per second by the optimal solution (and 9 units per second by our sufficient condition); and (d) only requires 4.5 units per second by the optimal solution (and 4.5 units per second by our sufficient condition). In a slightly different scenario shown in Fig.2.2, the solution that provides the longest lifetime is actually the worst in terms of bandwidth requirement. A shortest path routing algorithm would choose (b) to maximize lifetime, but the required bandwidth may be too high to accommodate.

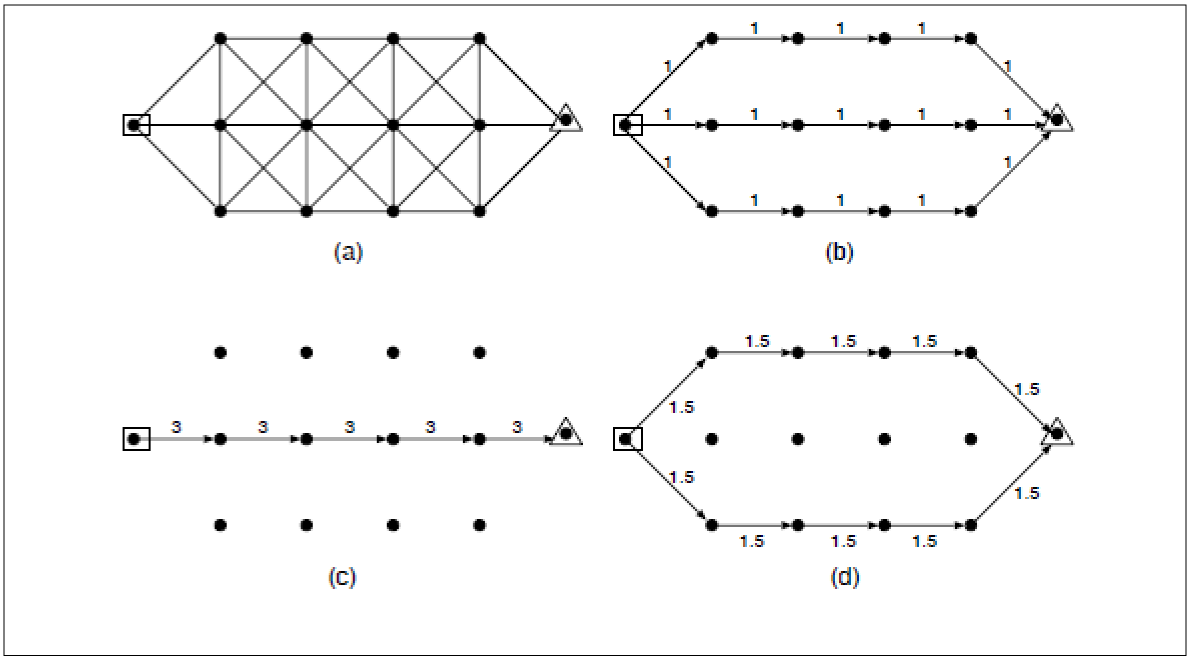

Fig.2.1 For the network shown in (a), nodes within each other's transmission range are connected with a line. The three routing options (b), (c) and (d) lead to the same lifetime, but (b) and (c) demand higher bandwidth than (d)

From the two examples above, we observed that for a randomly deployed network, usually the one that is likely to be used as a relay node is at the core of the network (if everyone chooses what is best for itself selfishly), which unfortunately is also the most interfered area due to the broadcast nature of wireless transmissions. Sending a lot of data to the core is likely to congest the network, but sending every packet along the outlier is not the best solution either.

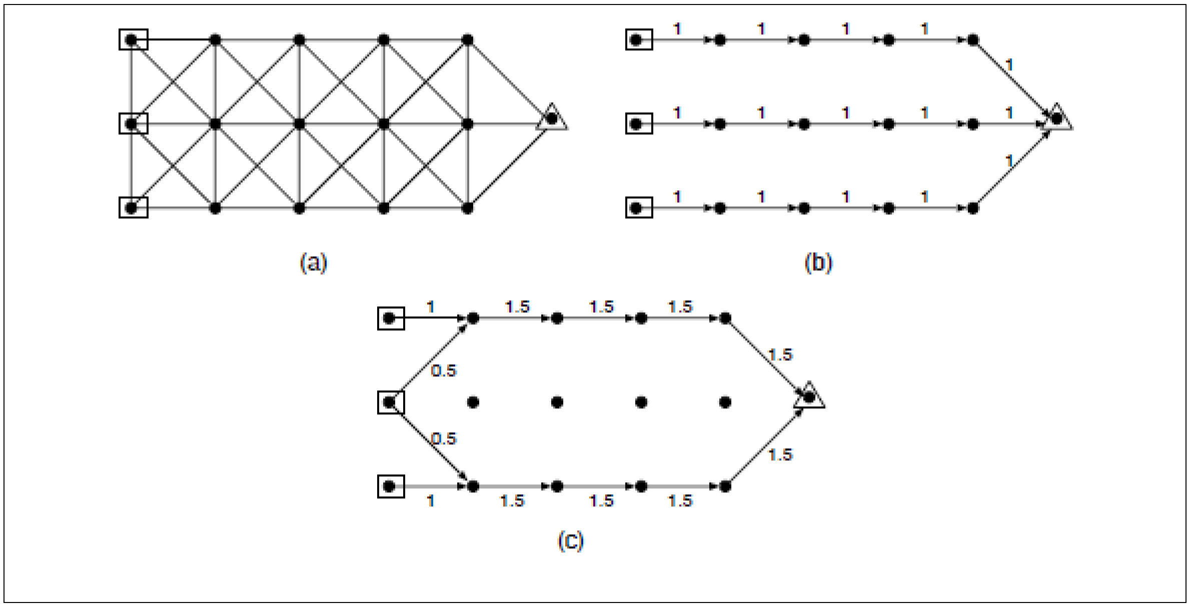

Fig.2.2 For the network shown in (a), both (b) and (c) use shortest paths routing, (b) is optimal in terms of lifetime, but is the worst in terms of bandwidth, (c) is the best in terms of bandwidth, but is suboptimal in terms of lifetime

What should be the maximum lifetime routing solution that is feasible with link bandwidth constraint? Apparently there is no generic policy such as shortest path routing or minimum energy routing that can lead to the maximum lifetime and be accommodated by the link bandwidth. To answer this question, we first consider for an arbitrary network topology, what condition should hold in order to ensure all source data can be put through, with each source generating data at a given rate. In this article, we discuss the sufficient condition on the link bandwidth, and use the bandwidth constraint to decide not only the routing topology but also the actual data rate on each link. The routing topology and link data rate are computed by solving an optimization problem that includes both energy and bandwidth constraints.

The major contributions of this work are that we formulated the energy and bandwidth-constrained routing problem as a multi-constraint optimization problem and provided efficient heuristic solutions to it. In addition, a companion time slot assignment algorithm is proposed to support the resulting routing solution at the MAC layer.

The rest of the paper is organized as follows: Section 2 discusses the sufficient conditions on link bandwidth; Section 3 formally describes the energy-bandwidth constrained routing problem and provides a mathematical model for the problem; Section 4 presents two heuristics for joint optimization of energy and bandwidth; Section 3-E addresses how to use the mathematical model to address in-network data aggregation; Section 5 provides numerical simulation results that show the comparison of algorithms in terms of throughput and lifetime; Section 6 briefly surveys the related work, followed by concluding remarks and further research issues in Section 7.

2. A SUFFICIENT CONDTITION FOR COLLISION-FREE COMMUNICATION

Let R_{ij} denote the data rate from node i to node j. Assume that the MAC layer uses an efficient TDMA scheme in which the number of time slots assigned to link (i, j) is proportional to R_{ij}. For any node i's reception to be successful, the TDMA schedule must satisfy that (1) when node i is receiving, it cannot be sending, and (2) when node i is receiving from j, none of its neighbors except j should be sending. Let N_i denote the neighbors of node i, and B the wireless link bandwidth. These two necessary conditions can be written as:

1). $\sum_{j \in N_i} \left(R_{ij} + R_{ji} \right) \leq B$

2). $\sum_{j \in N_i} R_{ji} + \max_{j \in N_i} \left(\sum_{k \in N_j, k != i} R_{jk} \right) \leq B$

However, these two are only necessary but not sufficient conditions, i.e., satisfying these two conditions does not guarantee that conflicting transmissions can always be assigned to different slots. In this paper, we prove that the sufficient condition to guarantee a global collision-free schedule is: $\sum_{j \in N_i} R_{ij} + f_i \sum_{j \in N_i} \sum_{k \in N_j} R_{jk} \leq B$, $\forall i$ (Sufficient), where f_i = 1 if node i is a receiver, and f_i = 0 otherwise. The proof of the sufficient condition is included in the Appendix.

The sufficient condition may require more bandwidth than necessary, but if this condition is satisfied at each node, it guarantees that a conflict-free time slot assignment can be found, which provides guaranteed data rate for each node. If each node injects data into the network at a rate below the guaranteed source rate, the network will be congestion-free. Moreover, since every node transmits at its scheduled time slot, there will be predictable delay at each hop, and hence bounded delay from the source to the sink. In the following sections, we base our discussion on the sufficient condition only.

3. MATHEMATICAL MODEL

3.1. Problem Definition

We assume that in a sensor network of n nodes, wireless link capacity is B (bits per second), and each node has initial battery energy E_i (J). Each node i generates sensory data at a rate of R_i bits per second ($R_i > 0$ if node i is a source, $R_i = 0$ if it is a pure relay node, and $R_i < 0$ if it is a sink). We assume that nodes consume energy on transmitting, receiving and sensing (i.e., generating sensory data), and their energy consumption rates are P_t, P_r, and P_s J per bit respectively. We further assume that P_r and P_s are constants, but Pt is handled differently in the two models: in the uniform model, each node transmits at the same power level P_t; in the non-uniform model, each node can transmit at different power level from others but the transmission power used by node i is still fixed, denoted by P_{ti}.

The energy-bandwidth constrained maximum lifetime routing problem can be formally stated as follows: Suppose that sources are preselected and each node i's rate R_i is known, but the transmission rate from node i to node j is unknown. Let T be the total network lifetime. The rate allocation problem is to compute the data rate R_{ij} on each link (i, j), given each node i's E_i, R_i and link capacity B, such that the total network lifetime T is maximized and the rate allocation can be accommodated by wireless link capacity and energy reserve.

3.2. With Uniform Transmission Power

In this model every node uses the same transmission power, therefore links are all symmetric. We use N_i to denote the neighbors of i excluding i itself, and f_i as an indicator of the receiver, as defined in Section 2: $f_i = 1$, if $\sum_{j \in N_i} R_{ji} > 0$; $f_i = 0$, otherwise

To maximize lifetime T is equivalent to minimize 1/T. Thus, we can formulate the rate allocation problem as follow:

Table 1. Mathematical Model for Uniform Transmission Power

Minimize: 1/T	(1)
Subject to:	
$\sum_{j \in N_i} \left(R_{ij} - R_{ji} \right) = R_i$	$\forall i$ (2a)
$P_s R_i + \sum_{j \in N_i} \left(P_r R_{ji} + P_t R_{ij} \right) \leq E_i / T$	$\forall i$ (2b)
$\sum_{j \in N_i} R_{ij} + f_i \sum_{j \in N_i} \sum_{k \in N_j} R_{jk} \leq B$	$\forall i$ (2c)
$0 \leq R_{ij} \leq B$	$\forall i, \forall j$ (2d)
$f_i = \{0,1\}$	$\forall i$ (2e)

In this formulation, equality (2a) indicates that data rates R_{ij} satisfy flow conservation at each node; Inequality (2b) is the energy constraint, and inequality (2c) defines the bandwidth constraint. In wireless networks, the capacity constraint is different from that in a flow network, where each link (u, v) has a fixed link capacity c(u, v) and flow f(u, v) $\leq$ c(u, v) is the only capacity constraint on each individual link. In wireless networks, because of the broadcast nature of transmission, the capacity constraint needs to be considered on a collision domain, rather than on each link separately. In other words, how much can be transmitted over one link depends on not only the link capacity B, but also the amount of data transmitted over other links in its collision domain. Inequality (2c) ensures that all links possibly in the same collision domain have a total demand less than B— If node i is a sender but not a receiver, it only needs to satisfy that the sum of the flow going out of i is bounded by B; If node i is a receiver, it needs to satisfy that node i's sending, receiving and other interfering nodes' transmission have a total demand of at most B; If node i is neither a sender nor a receiver, (2c) is automatically satisfied. Inequalities (2d) and (2e) are constraints for the variables.

3.3. With Non-uniform Transmission Power

In this model, we assume that each node still uses fixed transmission power, but node i can use P_{ti} to transmit and node j can use P_{tj} to transmit, and it is possible $P_{ti} \neq P_{tj}$. The inequality (2b) of the above linear program is modified as in (3a) to reflect the individual transmission power.

With this model, network topology is predetermined, but the links can be unsymmetrical. To deal with asymmetrical links, we use N^{+}_{i} to denote the neighbors that can receive from node i; and N^{-}_{i} to denote the neighbors that node i can receive from. Therefore the inequality (2c) is modified as in (3b) to reflect the change on the collision domain.

$$P_s R_i + \sum_{j \in N_i} \left(P_r R_{ji} + P_{ti} R_{ij} \right) \leq E_i / T \qquad \forall i \text{ (3a)}$$

$$\sum_{j \in N_i^+} R_{ij} + f_i \sum_{j \in N_i^-} \sum_{k \in N_j^+} R_{jk} \leq B \qquad \forall i \text{ (3b)}$$

3.4. With Double Disk Model

The models presented in section 3-2 and 3-3 both assume a single disk model, i.e., the effective transmission range is the same as the interference range. In reality, the interference range is usually larger than the effective transmission range. For example, a radio's transmission range is 500 meters, but the nodes located 800 meters away still are interfered by this node's transmission. Between 500 meters and 800 meters, the signal is not strong enough to be decoded, but strong enough to cause interference at others. In this section we modify our model to reflect this phenomenon.

We use the double disk model with the uniform transmission power. In terms of energy constraint, the inequality (2b) remains the same, since the transmission range remains the same; in terms of the bandwidth constraint, the definition of neighbors is changed. We use N_i to denote the nodes that are in the transmission range of node i, N_{iF} to denote the nodes that are in the interference range of node i. Since the interference range is larger than the transmission range, apparently $N_i \subseteq N_{i_F}$. Since all links are symmetrical, if $i \in N_{j_F}$, then $j \in N_{j_F}$. The bandwidth constraint is changed to:

$$\sum_{j \in N_i} R_{ij} + f_i \sum_{j \in N_{i_F}} \sum_{k \in N_j} R_{jk} \le B \qquad \forall i \quad (4a)$$

3.5. Data Aggregation

Section 3-2--3-4 gives a mathematical formulation for a basic data forwarding scheme without data aggregation. In sensor networks, sometimes data aggregation is used to reduce the number of transmissions. In this section we show that this model can be extended to compute the optimal routing and link rate allocation for data aggregation as long as the data aggregation scheme is given.

A well known data aggregation scheme is to aggregate data from different sources when they arrive at a relay node at a close time-frame. The idea is similar to that used in Opportunistic Network Coding [11]. In Fig.2.3, suppose source node i generates data at a rate of 5 packets per second, and input link (j, i) has a rate of 3 packets per second, and (k, i) has a rate of 2 packets per second, then the output flow of node i has a total rate of 5 packets per second, because each packet from the low-rate flows can be combined with a packet of the high rate flow and get a "free ride".

Thus, the flow conservation constraint in equality (2a) is changed to:

$$\sum_{j \in N_i} R_{ij} = \max_{j \in N_j} \{R_i, R_{ji}\} \qquad \forall i \text{ (5a)}$$

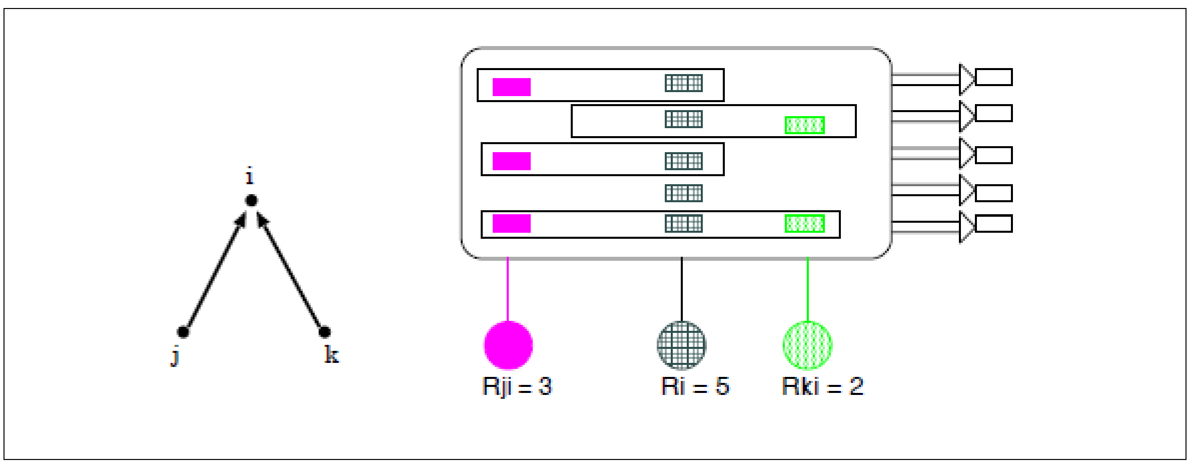

Fig.2.3 Opportunistic Data Aggregation. Low rate streams are aggregated into the high rate stream.

4. HEURISTIC

The mathematical model defined by objective (1) and inequalities (2a) - (2e) considers the bandwidth constraint while optimizing sensor network lifetime, therefore the solution to this model contains the optimal solution to the energy bandwidth constrained maximum lifetime routing problem. However, it is not linear because f_i is also a variable. In the following, we will present two heuristics that both work around the nonlinear problem by using information from the shortest paths (in terms of hops) from sources to the sink. The shortest paths represent the minimum-energy routing topology if data is not aggregated [12]. Heuristic I bears the characteristics of the shortest path routing, and Heuristic II bears the characteristics of the mathematical-programming based optimal solution, and they both consider bandwidth constraints.

4.1. Heuristic I: Scalable Rate Allocation on Shortest Paths

The first heuristic starts from the shortest paths from sources to the sink, but the rate on each link is determined by the available bandwidth.

Table 2. Heuristic I: Scalable Rate Allocation on Shortest Paths

1) Compute the shortest path from each source node to the sink 2) Assume source rate is one unit, check against condition (2c) for each node, and find the most bandwidth-contentious node i. Let LHS=required bandwidth of node i's collision domain. Then compute the scale factor ∂ : ∂ = B/LHS. Set Δf =min{a/2, R_i} 3) Push out Δf amount of flow from each source to the sink then update the remaining input flow $R_i^{'}$=R_i - Δf for each source t

Table 2. Heuristic I: Scalable Rate Allocation on Shortest Paths (Continue)

4) Repeat (5)-(7) until we push through R_i' for each source i or the network is fully saturated

5) Find the shortest paths for nodes with $R_i'>0$ based on the current available nodes and links. Nodes that are saturated on (2c) and their neighbors are not eligible for replaying. In case of a tie, give higher priority to nodes with more remaining energy; if there is still a tie, give higher priority to nodes with smaller degree

6) Decide the scale factor ∂ in a similar manner as in step (2). If pushing min{ ∂ , R_i'} units does not decrease lifetime, then set Δf =min{ ∂ , R_i'}, otherwise, set Δf =min{ ∂/2, R_i'}

7) Push out Δf amount of flow from each source with $R_i'>0$ then update the remaining input flow $R_i' = R_i' - \Delta f$

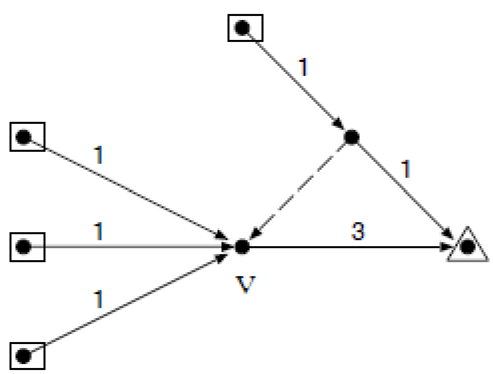

Fig.2.4 The most contentious node v requires 7a units. If link bandwidth is B units, then a=B/7

In steps 2) and 6), this algorithm uses a/2 when computing Δ f for the purpose of load balancing, which makes the network last longer. A simplified version is to use a instead of a/2 when computing Δ f. It finishes faster but leads to shorter lifetime.

4.2. Heuristic II: Optimizing Lifetime With Bandwidth Constraint

Since the mathematical model defined in (1) and (2a) - (2e) has an objective of maximizing lifetime, if we can convert it to a linear program in a controlled manner; it is likely to produce a close-to-optimal solution in terms of lifetime. The following describes an algorithm that chooses the likely-to-be relay nodes and set their $f_i = 1$ to make the program linear.

Heuristic II will either terminate with a valid solution or report “infeasible”. There will not be endless iterations in line 4. If the given source rates R_i are very low, it terminates at line 1. In most of the simulations, it requires solving the linear program two to four times to get a suboptimal solution. If it does become infeasible, it is likely because the given source rates R_i are higher than what the network can support.

Heuristic II is presented as follows:

Table 3. Heuristic II: Optimizing Lifetime With Bandwidth Constraint

1) Set f_i=1 for the sink, and f_i=0 for all other nodes, solve the linear program, update f_i=1 if $\sum_{j \in N_i} R_{ji} > 0$; if (2c) is satisfied $\forall i$, return link rate R_{ij} for all (i,j), otherwise, go to line 2
2) Compute the shortest path from source nodes to the sink
3) Set f_i=1 for receiving nodes; solve the linear programming; if $\sum_{j \in N_i} R_{ji} > 0$ and f_i=0, update f_i=1
4) Repeat line 3 until there is no update for f_i (converged) or the linear program becomes infeasible
5) If it converges, output link rate R_{ij} for all links (i,j)
6) If it becomes infeasible: if f_i = 1 but $\sum_{j \in N_i} R_{ji} = 0$ set f_i=0 and R_{ji}=0, $\forall j \in N_i$ as input, solve the linear program again; if it is still infeasible, report infeasible.

5. SIMULATION

5.1. With Uniform Transmission Power

In the simulation study, we investigate how the bandwidth constraint can change the routing decision and eventually affect the lifetime of the sensor network. First, we compare the existing algorithms with the proposed heuristics and observe which algorithm is more likely to cause network congestion and fail to push through the applied load.

Nodes are randomly deployed in a 100 $\times$ 100 square region, and transmission range is set to 30. In the first simulation (Fig.2.6(a) and (c), we use 50 nodes in total. We randomly select 4 source nodes and apply increasing source rate on them. Source rate is set to be a percentage of link bandwidth. The proposed schemes Heuristic I and II are compared with MaxLife from [1], and shortest path routing (labeled as SPR). The reason we choose MaxLife is because it computes the maximum lifetime without considering bandwidth constraint. When there is enough bandwidth, MaxLife represents the optimal solution. SPR uses the shortest paths from sources to the sink, with link weight representing the transmission power of the node. In the uniform transmission power setup, each link has weight 1.

We found that when each source node's data rate R_i is increased to 12% $\sim$ 13% of the given link bandwidth, MaxLife starts to congest, i.e., some collision domain requires more bandwidth than what is available, and SPR starts to congest when it is increased to 15%. Heuristic I can push through without congestion until the load is increased to 18% and Heuristic II can support as much as 16%. The vertical lines in Fig.2.6(a) and (c) indicate after this point, increased data rate cannot be put through.

Fig.2.6(a) shows the average ratio of the required bandwidth in each collision domain to the offered bandwidth. The lower the average is, the more bandwidth efficient of the scheme will be. Fig.2.6(c) shows the maximum ratio. A scheme stops working when the maximum ratio reaches 1. We can get the maximum throughput of the network at the stop point.

Fig.2.6(a) shows which scheme is more bandwidth efficient from a different angle. If a routing scheme violates the necessary condition, there is absolutely no way to push through the applied traffic load; when it violates the sufficient condition, there is no guarantee we can find a valid transmission schedule at the MAC layer to support the routing.

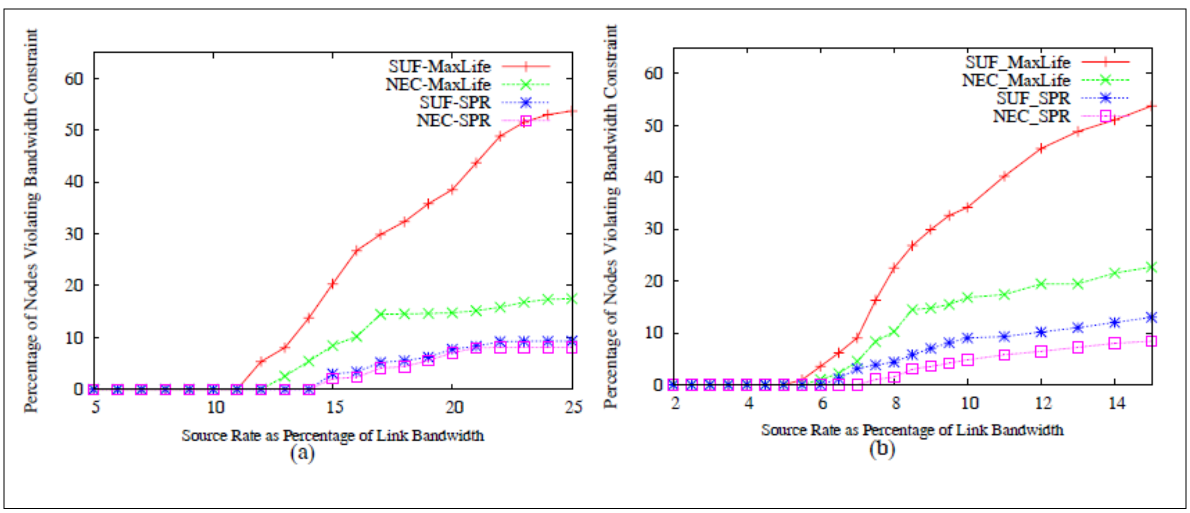

Fig.2.5 Percentage of nodes violating necessary and sufficient conditions.
(a) 50 nodes; (b) 100 nodes

In the second simulation (Fig.2.6 (b) and (d), we show the results with 100 nodes and 10 source nodes. The four algorithms show similar behavior as in the first simulation, except that per node throughput is lower because there are more source nodes. The total throughput of the network is close to that in the first simulation.

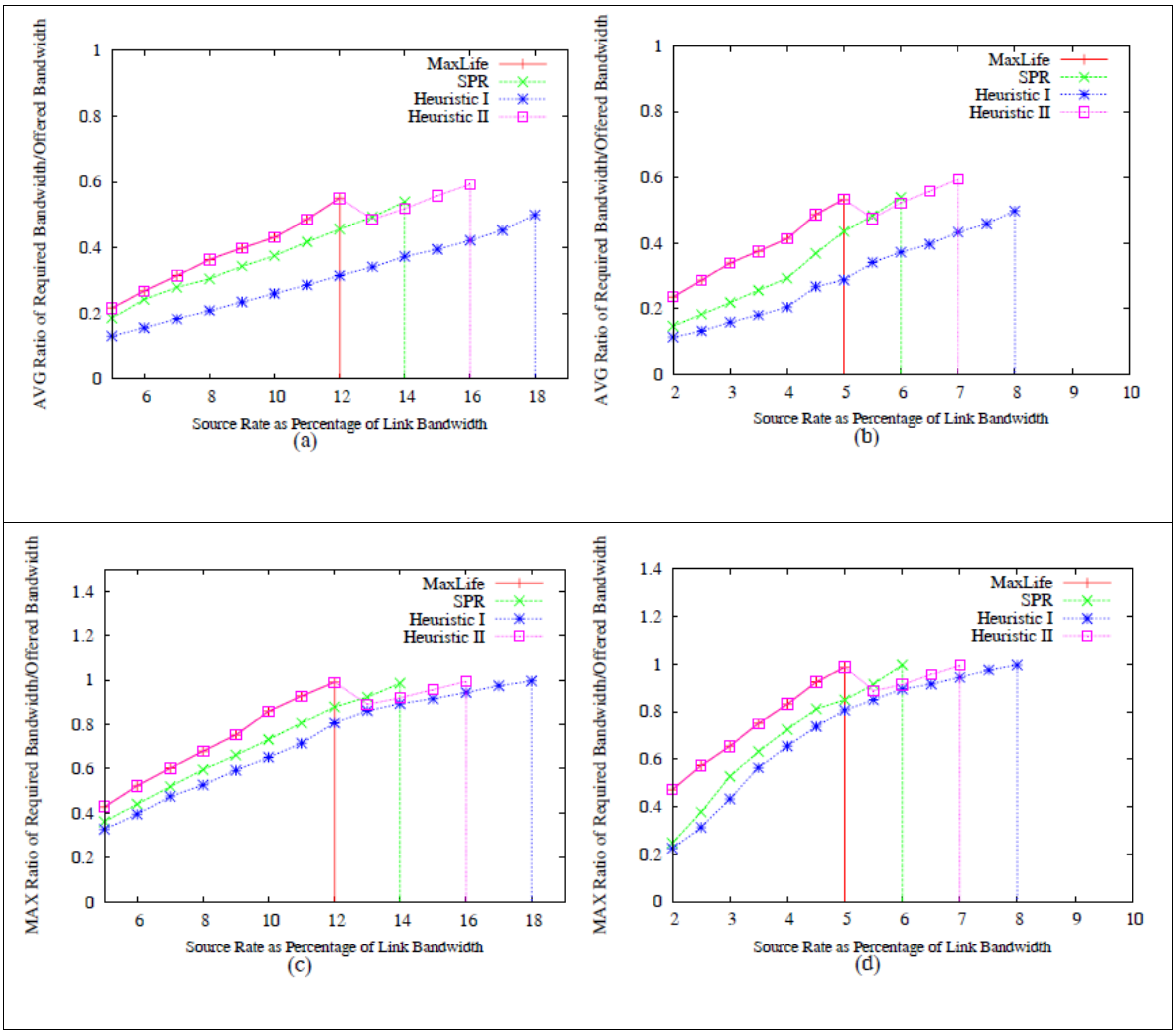

Fig.2.6 (a)-(b) The average ratio of required bandwidth/offered bandwidth; c(c)-(d) The maximum ratio of required bandwidth/offered bandwidth. (a) and (c) for 50 nodes, (b) and (d) are for 100 nodes.

In Fig.2.7 we compare four algorithms on their contribution toward lifetime. The results show when there is enough bandwidth, MaxLife does not have bandwidth violations and achieves the optimal solution, and Heuristic II achieves the same optimal solution; However when bandwidth does pose a constraint, Heuristic II can still push through 33% more data than MaxLife, and Heuristic I can push through 50% more data than MaxLife. Heuristic II achieves the best performance on lifetime and second best on throughput; Heuristic I achieves the best performance on throughput, which is consistent with our observation from Fig.2.6. Networks with 100 nodes can achieve longer lifetime than networks with 50 nodes because the workload is shared among more nodes.

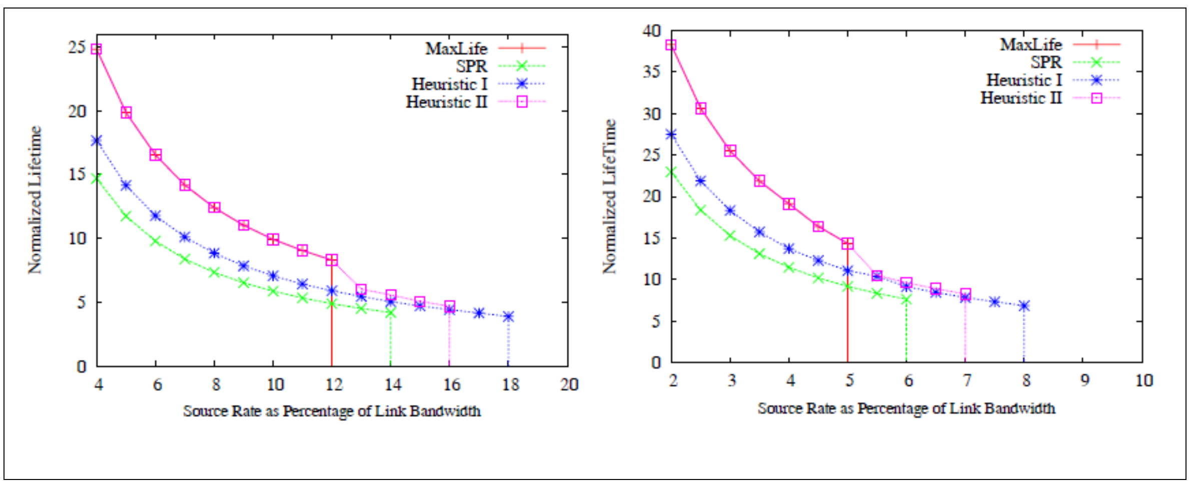

Fig.2.7 Normalized lifetime for data forwarding without aggregation, assuming sending one unit of data consumes 10% total energy. (a) 50 nodes; (b) 100 nodes.

5.2. Non-uniform Transmission Power

In this simulation, transmission range is randomly selected between 25-35. With asymmetrical edges, the performance comparison of the four algorithms in Fig.2.8 is consistent with the uniform power case in Fig.2.6. Network lifetime is reduced because the disparity in energy consumption is severe. Since the non-uniform power distribution is captured in the optimization model given in section 3-C, as a result, Heuristic II shows more performance gain in lifetime over other algorithms.

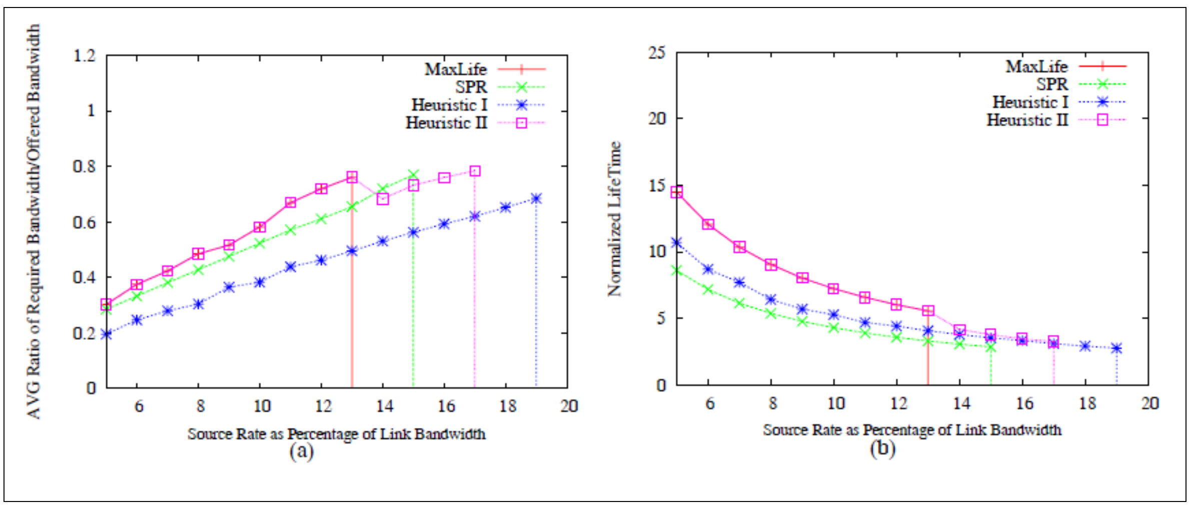

Fig.2.8 With non-uniform transmission power, (a) the average ratio of required bandwidth/offered bandwidth; (b) normalized lifetime.

5.3. With Double Disk Model

In this simulation, we choose transmission range 30, interference range 1.7 × transmission range, with everything else the same as in section 5-A. Fig.II.9(a)-(c) show the throughput performance. With a larger interference range, there is less chance for spatial reuse of channel, therefore the network throughput is less, but the lifetime is increased due to the lower data rate as shown in Fig.II.9(d).

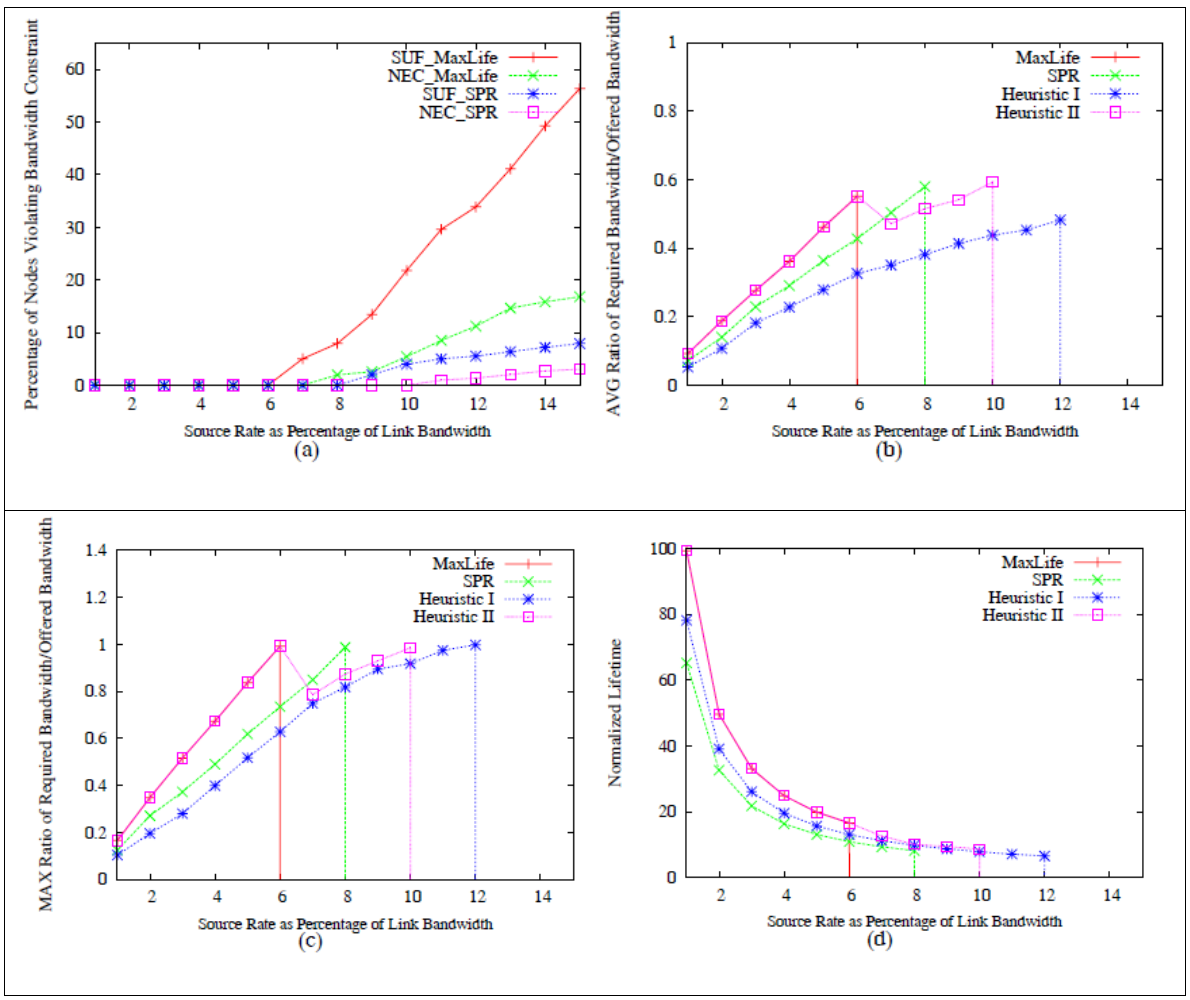

Fig.2.9 With double disk model, (a) percentage of nodes violating necessary and sufficient conditions; (b) the average ratio of required bandwidth/offered bandwidth; (c) the maximum ratio of required bandwidth/offered bandwidth; (d) normalized lifetime.

5.4. Data Aggregation

In this simulation, we test how much improvement we can achieve through mathematical optimization on a chosen data aggregation method. Using the opportunistic aggregation method outlined in section 3-E, we compare our solution with the shortest path tree and the minimum spanning tree, and the results show dramatic improvement on network lifetime as shown in Fig.2.10. LP-SPT results from applying Heuristic II using an initial shortest path tree at step 2, and LP-MST results from applying Heuristic II using an initial minimum spanning tree at step 2. SPT and MST are fixed-route aggregation on the shortest path tree and the minimum spanning tree respectively.

Fig.2.10(a) shows that LP-SPT and LP-MST can push data through until source rate is 20% of link bandwidth, while SPT and MST stop working (due to congestion) when source rates are 15% and 17% of link bandwidth respectively. This indicates a throughput gain of 33% over SPT and 17% over MST.

In Fig.2.10(b), we use networks of different sizes to show the maximum network throughput. Each source sends at a rate $0.01 \leq R_i \leq 10$ with link bandwidth=10, and we try to maximize R_i. Our observation is consistent with that in Fig.II.10 (a) — LP-MST and LP-SPT have the same throughput, and both are better than MST and SPT.

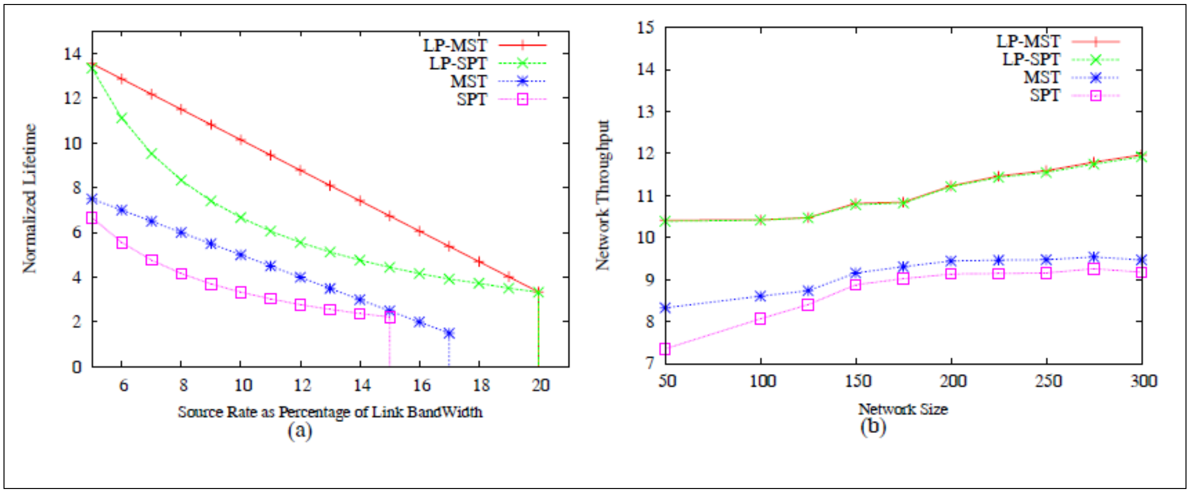

Fig.2.10 Opportunistic aggregation for data forwarding, assuming sending one unit of data consumes 10% total energy. (a) normalized lifetime; (b) throughput.

6. RELATED WORK

The most related work includes our previous work on edge coloring for transmission scheduling [13], maximum lifetime routing [1], and throughput optimization [14]. In [13], we precisely depicted the conflict relation among transmissions with each color corresponding to one time slot at the MAC layer. It guarantees conflict-free time slot assignment if each edge carries the same load. However, edge coloring by itself is NP-complete, and it assigns one color to each edge which implies that it works best for uniform traffic load. Link rate allocation in this article is an extension from color assignment, and it works well for arbitrary traffic load because the number of time slots that each edge gets is proportional to the traffic load on the edge; Furthermore, we consider nodes' energy constraint for link rate allocation. In [14], a linear programming model is used to optimize system throughput subject to the fairness constraint. In this paper, energy is not considered as a constraint, and a network flow model is used that characterizes the capacity constraint: $f(e) \leq c(e)$ on a link e, instead of using the accurate capacity constraint on a collision domain as discussed in this paper. An earlier work [15] also falls in this category and only considers a very simple interference model: when a node sends, it cannot receive. In [1], the authors proposed a distributed algorithm to compute link rates with an objective of maximizing the network lifetime. The major contribution is on the distributed implementation of the optimization algorithm. However, like most previous work on energy efficient routing in sensor networks, bandwidth is not taken into consideration in their model. Similar work along this line includes [2]–[10] and many others.

In [2], the proposed routing algorithms select the routes and the corresponding power levels such that the network lifetime is maximized. In [3], the routing problem is formulated as a linear programming problem, where the objective is to maximize the network lifetime, which is equivalent to the time until the network partition due to battery outage. Packet aggregation techniques were proposed to further reduce the energy consumption rate [4], [5], [7]. In [6], it was proposed to deploy a network clustering scheme and assign a less-energy constrained gateway node to act as a centralized network manager to further improve the energy efficiency and maximize network lifetime. Cui et al. further considered energy-efficient routing, scheduling, and link adaptation strategies together to maximize the network lifetime in [8], but the authors did not explicitly consider the bandwidth constraint in an arbitrary topology as we do. How to arrange the location of base stations for WSN and select relay paths to maximize the network lifetime was discussed in [9], [10].

Along the direction of cross-layer design and optimization, we found [8], [16]–[21] and many others. Optimization problems in multi-hop wireless networks are naturally cross-layer problems ([16], [17]). It involves PHY layer coding, modulation and error control, MAC/link layer resource (both bandwidth and power) management, network layer routing, and transport layer flow and congestion control. Many of the related work in cross-layer design focused on how to minimize energy consumption under various constraints [8], [18]–[20]. Reference [18] proposed to adjust the transmission powers of nodes in a multi-hop wireless network to create a desired topology, aimed to minimize power used while maintaining network connectivity. Cruz and Santhanam studied the problem of joint routing, link scheduling and power control to support high data rates for broadband wireless multi-hop networks in [19]. The main objective is still to minimize the total average transmission power. Since most cross-layer optimization problems are too complex to solve, distributed algorithms with suboptimal (and potentially distributed) scheduling component were studied in [16], [20].

Although this paper aims to provide maximum lifetime routing under energy and bandwidth constraints, the resulting solution naturally satisfies guaranteed data rate for each source and hence guaranteed fairness. Previous works ([21]–[23]) addressed the fairness issue through different mechanisms, such as packet scheduling, distributed layer-2 fairness solution (by modifying the contention and back-off mechanisms of CSMA/CA), joint power allocation and routing etc.

7. CONCLUSION AND DISCUSSION

This article has provided a generic mathematical model for the maximum lifetime routing problem in energy and bandwidth-constrained sensor networks. Using the sole energy constraint sometimes leads to unrealistic solutions that cannot be accommodated by the link capacity. In this paper we have provided a sufficient condition that a given traffic load can be put through a given network and jointly considered energy and bandwidth constraints for routing and link rate allocation. Joint optimization guarantees that there exists a conflict-free time slot assignment to support the given routing solution. To the best of our knowledge, this is the first work that explicitly considers bandwidth constraint in solving a maximum lifetime routing problem in sensor networks. The basic mathematical optimization model can be easily extended to address heterogeneous sensor networks where nodes have different initial energy or different transmission power levels, and to work with various data aggregation schemes.

The proposed heuristics are centralized. To apply mathematical optimization on large scale sensor networks, hierarchical scheme can be used, such as to divide the network into areas or clusters, and then apply the algorithms within the area or cluster. This will compromise the global optimality, but the solution is still better than the pure decentralized algorithms in terms of energy and channel efficiency.

8. REFERENCES

[1] R. Madan and S. Lall, “Distributed algorithms for maximum lifetime routing in wireless sensor networks,” in Proc. IEEE Global Telecommunications Conference (GLOBECOM), 2004, pp. 748–753.

[2] J. Chang and L. Tassiulas, “Energy conserving routing in wireless adhoc networks,” in Proc. IEEE INFOCOM 2000, vol. 1, 2000, pp. 22–31.

[3] Joongseok Park. Sartaj Sahni, “Maximum lifetime routing in wireless sensor networks,” IEEE/ACM Trans. Networking, vol. 12, no. 4, pp. 609– 619, Aug. 2004.

[4] C. Schurgers and M. B. Srivastava, “Energy efficient routing in wireless sensor networks,” in Proc. MILCOM 2001, vol. 1, 2001, pp. 357–361.

[5] M. Bhardwaj, T. Garnett, and A. Chandrakasan, “Upper bounds on the lifetime of sensor networks,” in Proc. IEEE ICC’01, vol. 3, 2001, pp. 785–790.

[6] M. Younis, M. Youssef, and K. Arisha, “Energy-aware routing in clusterbased sensor networks,” in Proc. IEEE MASCOTS 2002, 2002, pp. 129– 136.

[7] B. Krishnamachari, D. Estrin, and S. Wicker, “Modelling data-centric routing in wireless sensor networks.”

[8] S. Cui, R. Madan, A. Goldsmith, and S. Lall, “Joint routing, MAC, and link layer optimization in sensor networks with energy constraints,” in Proc. IEEE ICC’05, vol. 2, May 2005, pp. 725–729.

[9] I. Kang and R. Poovendran, “Maximizing static network lifetime of wireless broadcast ad hoc networks,” in Proc. IEEE ICC’03, vol. 3, 2003, pp. 2256–2261.

[10] J. Pan, T. Hou, L. Cai, Y. Shi, and X. Shen, “Topology control for wireless sensor networks,” in Proc. ACM MOBICOM’03, 2003, pp. 286– 299.

[11] S. Katti, D. Katabi, W. Hu, H. Rahul, and M. Medard, “The importance of being opportunistic: practical network coding for wireless environments,” in Proc. 43rd Allerton Conference on Communication, Control, and Computing, Monticello, IL, Sept. 2005.

[12] M. X. Cheng and L. Yin, “Energy-efficient data gathering algorithm in sensor networks with partial aggregation,” Int. J. Sensor Networks,, 2008.

[13] M. Cheng and L. Yin, “Transmission scheduling in sensor networks via directed edge coloring,” in Proc. IEEE ICC 2007, pp. 3710–3715.

[14] M. Alicherry, R. Bhatia, and L. Li, “Joint channel assignment and routing for throughput optimization in multiradio wireless mesh networks,” in Proc. ACM MobiCom, 2005, pp. 58–72.

[15] M. Kodialam and T. Nandagopal, “Characterizing achievable rates in multi-hop wireless networks: the joint routing and scheduling problem,” in Proc. ACM MobiCom, 2003, pp. 42–54.

[16] X. Lin, N. B. Shroff, and R. Srikant, “A tutorial on cross-layer optimization in wireless networks,” IEEE J. Select. Areas Commun., vol. 24, no. 8, pp. 1452–1463, 2006.

[17] S. Toumpis and A. J. Goldsmith, “Performance, optimization, and crosslayer design of media access protocols for wireless ad hoc networks,” in Proc. IEEE ICC’03, vol. 3, 2003, pp. 2234–2240.

[18] R. Ramanathan and R. Rosales-Hain, "Topology control of multihop wireless networks using transmit poweradjustment," in Proc. IEEE Infocom'00, vol. 2, 2000, pp. 404–413.

[19] R. L. Cruz and A. V. Santhanam, "Optimal routing, link scheduling and power control in multihop wireless networks," in Proc. IEEE Infocom'03, vol. 1, Mar. 2003, pp. 702–711.

[20] X. Lin and N. B. Shroff, "The impact of imperfect scheduling on crosslayer rate control in wireless networks," in Proc. IEEE Infocom'05, vol. 3, Mar. 2005, pp. 1804–1814.

[21] M. Johansson and L. Xiao, "Scheduling, routing and power allocation for fairness in wireless networks," in Proc. IEEE VTC 2004-Spring, vol. 3, May 2004, pp. 1355–1360.

[22] H. Luo, S. Lu, and V. Bharghavan, "A new model for packet scheduling in multihop wireless networks," in Proc. ACM Mobicom 2000, pp. 76– 86.

[23] V. Gambiroza, B. Sadeghi, and E.W. Knightly, "End-to-end performance and fairness in multihop wireless backhaul networks," in Proc. ACM Mobicom'04, pp. 287–301.

APPENDIX: PROOF OF THE SUFFICIENT CONDITION

To prove that the condition in Section II is sufficient for collision free communication, we first introduce a time slot assignment algorithm. The algorithm requires that input link rates satisfy the sufficient condition and outputs a conflict-free schedule.

A. A Slot Assignment Algorithm

Table 4. Slot Assignment Algorithm

SlotAssignment(G(V,E),R)

1) Scale the link rates R_{ij} to integers and scale B proportionally; Let slot size $\Gamma = 1$.

2) Find the most bandwidth-contentious node v according to the sufficient condition, and compute the required bandwidth B_v at node v's collision domain:

$$v = \arg\max_{i \in V} \left(\sum_{j \in N_i} R_{ij} + f_i \sum_{j \in N_i} \sum_{k \in N_j} R_{jk} \right); \text{ and}$$

$$B_v = \sum_{j \in N_v} R_{vj} + f_v \sum_{j \in N_v} \sum_{k \in N_j} R_{jk}$$

3) Let frame size $F=B_v$. Number the slots from 1 to F

4) Create a table of $2 \times F$ associated with each node's sending and receiving schedules, use S row for sending and R row for receiving.

5) Let L= V. Repeat the following until L = $\varnothing$:

   (a) Randomly pick a node i from L;

   (b) For each node j $\in N_i$, if $R_{ij} > 0$, assign R_{ij} slots to link (i,j), starting from the

Table 4. Slot Assignment Algorithm (Continue)

(c) smallest available slot. A slot is available if it is available in both the S row of table[i] and the R row of table[j]; Mark those slots unavailable in the S row of table[j]; For each k $\in N_j$, if k $\neq$ i, mark those slots unavailable in the S row of table[k];

(d) Mark those slots unavailable in the R row of table[i];

(e) For each node j $\in N_i$, mark those slots unavailable in the R row of table[j], if they are not previously assigned;

(f) Remove i from L

6) Update frame size F to be the largest slot number used.

In Fig.2.11, the sufficient condition requires F=14 slots, but actually it only needs 12 slots by allowing the transmissions on (k, w) and (j, u) to occur at the same time. The sloppiness in the sufficient condition guarantees no matter whether there is a link between (j, w) or not, there are always enough slots to use regardless of the order that nodes are picked. This property makes it easy to implement the algorithm in a localized and distributed manner.

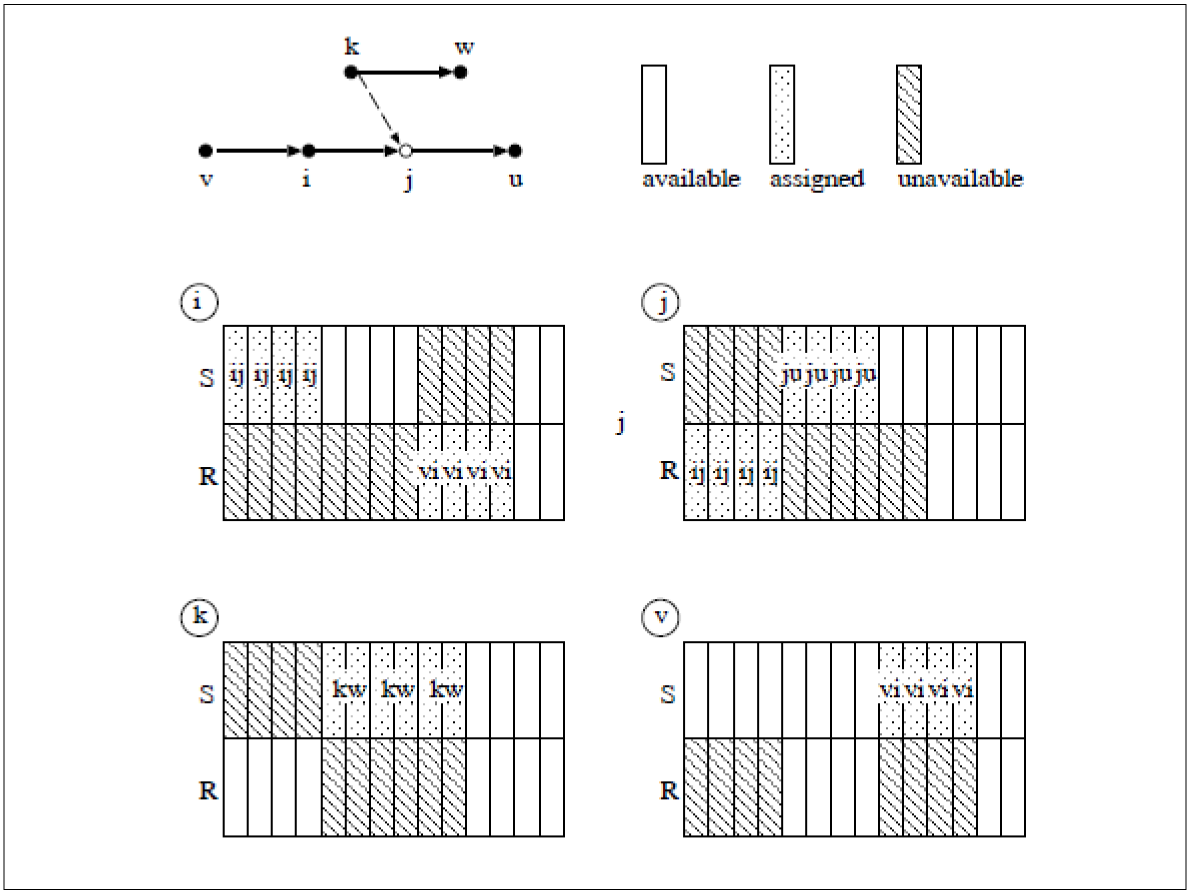

Fig.2.11 A walk-through example for the SlotAssignment algorithm. Suppose $R_{vi} = R_{ij} = R_{ju} = 4$, $R_{kw} = 6$, so node j is the most bandwidth-contentious node; frame size F=14 slots; the order that nodes are randomly picked at step 6 is i, j, k, v.

Lemma 1: The SlotAssignment algorithm generates a collision-free schedule.

Proof: Lemma 1 has two folds:

1) There are always sufficient number of slots to use, i.e., at step 5(b), the number of available slots $\geqslant$ the number of slots needed for any node i being considered, and

2) The resulting schedule is collision-free.

The second statement is obvious because all conflicting transmissions are scheduled at different time— when i is sending to j, j is not sending, and other neighbors of j are not sending, so there is no collision at j according to step 5(b); i is not receiving according to 5(c) so there is no collision at i; other neighbors of i are not receiving according to 5(d) so there is no collision at i's neighbors.

The first statement is proved as follows. Let N_1 be the total number of slots that are needed for sending when a random node i is picked at step 5(a), so $N_1 = \sum_{j \in N_i} R_{ij}$, and let N_2 be the number of slots that are still available for sending at this time.

Case (1), when i is not a receiver ($f_i = 0$): the only reason that i's S row is marked unavailable is when a neighbor l is receiving from another node k (Fig.2.12(a)). Let C = {(k, l)} be the maximum set of such conflicting transmissions, so the total unavailable slots in i's S row is $\sum_{(k,l) \in C} R_{kl}$. Similarly, for each receiver node j of i, the only reason that the R row of j is marked unavailable is because j's neighbor u is transmitting. Transmissions on (k, l) and (u, v), if not conflicting with each other, can be arranged at the same slot. Therefore, as long as the sufficient condition holds at node l with fl = 1 and at node j with $f_j = 1$, the number of available slots N_2 for i's transmission is still $\geq \sum_{j \in N_i} R_{ij}$. Therefore, $N_2 \geq N_1$ is held.

Case (2), when i is a receiver ($f_i = 1$): from case (1) to case (2), there will be $\sum_{l \in N_i} R_{li}$ additional slots marked unavailable in the S row of i, according to step 5(b); others remain unchanged. As long as the sufficient condition holds at node i with $f_i = 1$, the number of available slots N_2 for i's transmission is still $\geq \sum_{j \in N_i} R_{ij}$. Therefore $N_2 \geq N_1$ is held. Because during the iteration in step 5, $F = \max_i \left\{ \sum_{j \in N_i} R_{ij} + f_i \sum_{j \in N_i} \sum_{k \in N_j} R_{jk} \right\}$ so N_2 is sufficient for any node i.

Next we will see that even though the sloppiness of the sufficient condition requires more slots than necessary, the SlotAssignment algorithm itself does not prevent non-conflicting transmissions from happening at the same time.

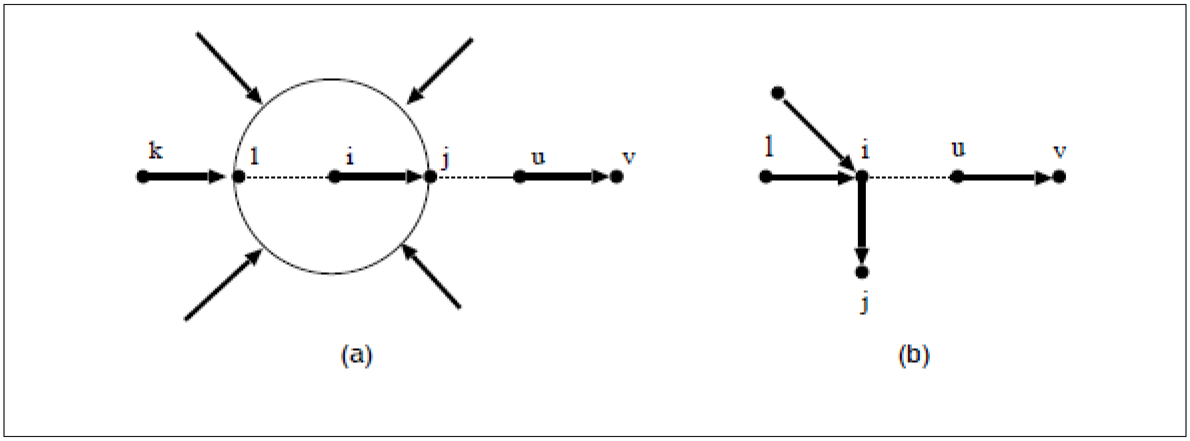

Fig.2.12 (a) with $f_i = 0$; (b) with $f_i = 1$

Lemma 2: The SlotAssignment algorithm can completely avoid the exposed terminal problem.

Proof: In Fig.2.13, if node B is picked first by the algorithm to use the first slot, transmission on (C, D) can still use the first slot because B's transmission in slot 1 only marked the R row of node C unavailable, the S row is still available. If node C is picked first, the result is the same.

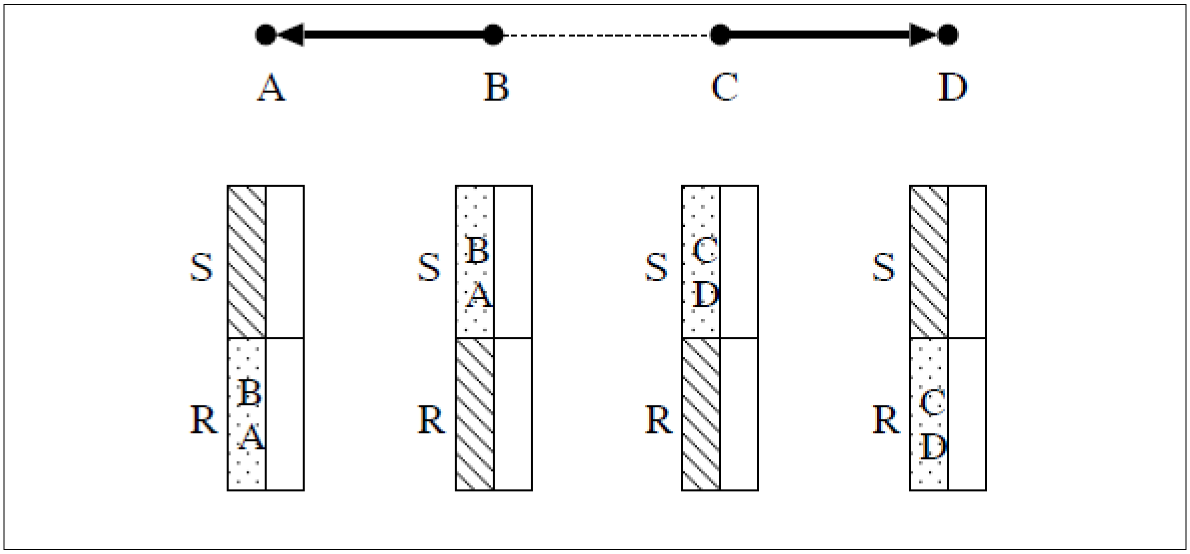

Fig.2.13 The SlotAssignment algorithm would allow C → D and B → A to occur at the same time.

Theorem 1: The following condition is sufficient to have a TDMA schedule that completely avoids collision and the exposed terminal problem in a multi-hop wireless networks with Omni-directional antenna:

$$\sum_{j\in N_i} R_{ij} + f_i \sum_{j\in N_i} \sum_{k\in N_j} R_{jk} \le B$$

Proof: By constructive proof, using the SlotAssignment algorithm described above, we can always find a TDMA schedule that is collision-free (by Lemma 1) and completely avoids the exposed terminal problem (by Lemma 2), as long as the given input Rij satisfies $\sum_{j\in N_i} R_{ij} + f_i \sum_{j\in N_i} \sum_{k\in N_j} R_{jk} \le B$

III. INTERFERENCE MODEL, MULTIPATH ROUTING AND LINK RATE CONTROL IN MULTIHOP WIRELESS NETWORKS

Maggie X.Cheng, Xuan Gong
Department of Computer Science
Missouri University of Science and Technology
Rolla, MO 65401
Email: {chengm@mst.edu,xgwt2@mail.mst.edu}

ABSTRACT. In multi-hop wireless networks, end-to-end throughput is often hard to predict and is even harder to optimize due to the effect of interference. To date there is no precise result other than asymptotic bounds for this question: if there is no routing information given, what is the maximum throughput of a network using uncoordinated transmission such as IEEE 802.11 MAC? This paper attempts to address this question for a given network with specific traffic demand. In this paper we use a cross-layer design scheme to optimize network performance. The paper includes a basic linear programming model, from which the routing paths and link data rates are derived, and then an extended model to consider links with different loss rates. Using ns2 simulation, we show that our joint routing and rate control scheme indeed can predict the maximum throughput and improve network throughput.

1. INTRODUCTION

In a communication network, both channel efficiency and user fairness are important. Fairness means users or applications are receiving a fair share of system resources so that no user with large traffic demand can starve others; Efficiency means the network resource is appropriately allocated so that the network has a high throughput. In multi-hop wireless networks, what mechanisms can we use to provide users with fairness and efficiency? There are different measures for fairness and various metrics for throughput. In this paper, we use one of the fairness measures as an example, and try to answer this question: how should we allocate channel bandwidth so that the network works most efficiently and at the same time guarantees a minimum data rate for all flows? To achieve the maximum throughput in wireless multihop networks with uncoordinated transmission has been a challenging task. Unlike in the wired networks or one-hop wireless networks, the complicated interference from neighboring nodes forbids one flow from achieving the full capacity of wireless links. The achievable data rate on one flow depends on not only its own link capacity but also other flows that are in the same collision domain.

Another important question in wireless network design and planning is: given user traffic demand, how can we estimate the required bandwidth? Bandwidth requirement is hard to estimate compared to wired networks for the same reason. Until wireless interference can be accurately modeled and accounted for, we cannot answer either question.

In this paper, we try to capture the convoluted relationship between wireless transmissions and find out the impact of interference on achievable user data rate. We cast the problem of providing maximum throughput with guaranteed fairness as a mathematical optimization problem. Our approach is cross-layer optimization in the sense that the search space for the optimal solution does not only include data rates of sources but also complete routes from sources to destinations. Using this integrated routing and rate control approach, flows can split or merge at any node, and there is no preselected route or routing policy other than to maximize throughput.

The main contribution of this paper is that we come up with a linear model to capture the impact of wireless interference on achievable data rates in multihop wireless networks. Based on this linear relation, we present a linear programming model of joint routing and rate control to achieve both efficiency and fairness in multihop wireless networks. This model can be extended to work around loss links in a heterogeneous network to improve throughput performance. The model is not only critical for cross layer optimization, but also useful in a classic separate layer scheme— It can be used to predict throughput performance, or to control source rate to improve network throughput or fairness when routing information is given.

The rest of the paper is organized as follows. In Section 2, we briefly survey the most related work in recent years; in Section 3, we present the formal problem definition and then a linear programming based multipath routing and rate control scheme; in Section 4 we validate our model in extensive simulations. Finally, Section 5 concludes the paper.

2. RELATED WORK

The most related work is network performance modeling and optimization with the effect of interference ([1]–[5]). [1] is based on a simplified protocol model, and [3] uses 802.11 interference model. The drawback of this approach is that it used cliques on the conflict graph to capture the interference relation among all links, which is an NP hard problem by itself. [6] focused on estimation of interference and studied the effect of interference on aggregated network throughput based on IEEE 802.11 model. [4] proposed a general interference model to estimate the sender and receiver data rates. The interference model is a physical model based on measured interference, different from the widely used protocol model, which is based on distance between nodes and models interference as a binary variable. [5] proposed a network throughput model to optimize total throughput and fairness among flows. Different from our work, it only applies rate control on flows; traffic demand is limited to one hop traffic, and multihop traffic is first converted to one-hop based on given routing information. In contrast, our work does not presume any routing information; instead, it uses joint routing and link rate control and works for multihop traffic. Our previous work [7] did joint routing and link rate control based on a perfectly controlled TDMA scheduling scheme and a different interference model.

The study of throughput modeling of wireless networks started as early as 1987 ([8]–[10]) for packet radio networks. Since then, many researchers reported their work in throughput modeling and optimization. Some deal with exact solutions ([11], [12]) and some deal with asymptotic results without input on traffic and network topology ([13]–[15]). To deal with the bandwidth constraint, some scholar extended the capacity constraint of flow networks to wireless networks without considering the interference from other links [12], [16]; some attempted to model interference but used global information such as cliques on a conflict graph ([3]). Since to find all cliques in a graph is an NP-hard problem, there is no known solution that is both efficient and accurate. All this motivated a new interference model. Our interference model uses only local information and the algorithm is polynomial time. It can be efficiently applied in practice.

3. MODEL-BASED MULTIPATH ROUTING AND RATE CONTROL

3.1 Problem Definition

We assume that in a multi-hop wireless network of n nodes, each source node generates data at a rate of Di bits per second. Suppose that the source-destination pairs are known. The multipath routing and rate control problem is to compute the routing topology and data rate on each link, such that the total network throughput is maximized, and the achieved throughput by all source-destination pairs satisfies the required fairness requirement. Assume the effective data rate is B after considering protocol overhead (for example, RTS-CTS message exchange in 802.11 MAC).

3.2 The LP Model with Reliable Links

We first address wireless networks with reliable links, i.e., links with zero loss rate. Under this model, the only error condition is collision due to simultaneous transmission from conflicting nodes. We will address the loss links in the next section.

We now consider a wireless network with n nodes. Each node has communication range X. If node j is in node i's communication range, j can successfully receive data from i, we say there is a communication link from node i to node j. Since all nodes have the same communication range, all links are symmetric and all interference relation are mutual. We use N_i to denote the set of nodes in node i's communication range, excluding i itself.

Now we are ready to present the linear programming model. To capture the characteristics of multipath routing, we assume the source data rate D_i can be achieved as the sum of multiple flows originating at node i. We use f(s, d) to denote the flow from source s to destination d, and d(f) denote the destination of flow f, s(f) denote the source of flow f. Let D_i be node i's total data rate in all flows, and $D_{i,f}$ be the rate allocated to flow f. Suppose the applied traffic F = {(s, d)} is given, then the set of source nodes are S = {s(f)|f ∈ F} and sink nodes D = {d(f)|f ∈ F}. Let R_{ij} be the data rate on link (i, j) (from i to j). Apparently R_{ij} = 0 if j is beyond i's communication range.

Now we can formulate the rate allocation problem as the following.

Table 1. The LP Model with Reliable Links

Maximize $\sum_{f \in F} D_{s(f),f}$

Subject to

$$R_{ij} = \sum_{f \in F} R_{ij,f} \qquad \forall link(i,j) \text{ (2a)}$$

$$\sum_{j \in N_i} (R_{ij,f} - R_{ji,f}) = D_{i,f} \qquad \forall i, \forall f \text{ (2b)}$$

$$D_{d(f),f} = -D_{s(f),f} \qquad \forall f \text{ (2c)}$$

$$D_{i,f} = 0 \qquad \forall f, \forall i \neq s(f), d(f) \text{ (2d)}$$

$$r_{ij} + \sum_{l \in N_i, l \neq j} r_{il} + \sum_{k \in N_j, k \neq i} r_{jk} + \sum_{(k,l) \in H3_{ij}} r_{kl} \leq B \qquad \forall link(i,j) \text{ (2e)}$$

$$r_{ij} = R_{ij} + R_{ji} \qquad \forall link(i,j) \text{ (2f)}$$

$$D_{\min} \leq D_{s(f),f} \qquad \forall f \text{ (2g)}$$

$$0 \leq R_{ij}, r_{ij}, R_{ij,f} \leq B \qquad \forall link(i,j) \text{ (2h)}$$

$$0 \leq D_{s(f),f} \leq B \qquad \forall f \text{ (2i)}$$

In inequality (2e), H_{3ij} is the group of links {(k, l)} that satisfy: $1 \leq d(j, k) \leq 2$, $1 \leq d(i, l) \leq 2$, $d(j, k) + d(i, l) \leq 3$, or $1 \leq d(i, k) \leq 2$, $1 \leq d(j, l) \leq 2$, $d(i, k) + d(j, l) \leq 3$, where $k \neq i, j$, and $l \neq i, j$, d(u, v) is the number of hops between node u and node v. By limiting the sum of hops to 3, this condition can capture all conflicting relation but can make the bound tighter than the one that simply includes all links in two hop neighborhood.

In this formulation, equalities (2a-2d) indicate that data rates satisfy flow conservation; inequality (2e) defines the bandwidth constraint. In wireless communication, the capacity constraint is different from that in a flow network, where each link (u, v) has a fixed link capacity c(u, v) and flow(u, v) $\leq$ c(u, v) is the only capacity constraint on each individual link. In wireless communications, because of the broadcast nature of transmission, the capacity constraint needs to be considered on a collision domain, rather than on each link separately. In other words, how much can be transmitted over one link depends not only on the fixed link capacity B, but also the amount of data transmitted over other links in its collision domain. Bandwidth constraint (2e) considers node's own transmitting and receiving, as well as the interference it receives from nearby transmissions. Using IEEE 802.11 MAC, the collision domain includes all links in 2-hop neighborhood. For example, in a chain topology A—B— C—D, all links (AB), (BC), and (CD) are in one collision domain because they all conflict with each other. Inequality (2e) ensures all links possibly in the same collision domain have a total demand less than B. In this paper, we also refer inequality (2e) as the interference model. Inequality (2g) gives the per-flow fairness guarantee to make sure none of the source-destination pair is starving. Finally, inequalities (2h-2i) are the constraints for variables. Note that if node j is beyond the communication range of node i, R_{ij}, $R_{ij,f}$, and r_{ij} are all set to be zero.

Regarding fairness, there are many definitions of fairness. Here we adopt the notion of combined measure of fairness and bandwidth efficiency since the objective is to maximize network throughput. To provide per-flow fairness, we introduce a guaranteed data rate for each flow D_{min}. Once the minimum data rate is satisfied, the remaining bandwidth is allocated to optimize system throughput. It is different from the well known Max-min fairness. However, in order to achieve Max-min fairness, all it takes is to iterate our method for multiple times with null objective function until the network is fully saturated or all sources are satisfied. In general, our linear programming model can combine any fairness measure as long as the fairness relation itself can be presented linearly.

3.3 Remark

To compute the maximum throughput in a multi-hop wireless network is NP-hard ([1], [17]). Previous work has used conflict graph to model the pair-wise conflicting relationship between links. In the conflict graph, vertices represent wireless links, and an edge is created between two vertices if the two corresponding wireless links conflict with each other. Then a clique on the conflict graph is used to represent the group of mutually conflicting links. Based on the conflict graph, the bandwidth constraint can be presented as $\max_Q\{\sum_{j\in Q} r_j\} \le B$ where j is a link in clique Q and rj is the data rate on link j. This approach requires exhaustive search of all cliques. However, to compute all cliques in a graph by itself is an NP-hard problem. Therefore, this approach cannot be used in practice.

To overcome the drawback of the clique approach, we use a local condition rather than a global condition in inequality (2e), which can be computed in polynomial time. This local condition captures all conflicting relation in wireless networks, but does not need to go through the pain of computing all cliques. The inequality (2e) is a sufficient but not necessary condition. Compared to other polynomial-time solutions, which simply include all links within two hops, our solution provides a tighter bound.

3.4 The LP Model with Lossy Links

In a network with lossy links, if a link has 50% loss rate, and the bandwidth is B, then the maximum receiver data rate through this channel is only 0.5B. We use T_{ij} to denote the actual sender data rate, and R_{ij} to denote the receiver data rate at link (i, j). Suppose the loss rate of link (i, j) is l_{ij}, then Rij = $(1-l_{ij})T_{ij}$. We modify the model to reflect the change as follows.

Table 2. The LP Model with Lossy Links

Maximize: $\sum_{f \in F} -D_{d(f),f}$ (3)

Subject to

$$R_{ij} = \sum_{f \in F} R_{ij,f} \qquad \forall link(i,j) \text{ (4a)}$$

$$\sum_{j \in N_i} (T_{ij,f} - R_{ji,f}) = D_{i,f} \qquad \forall i, \forall f \text{ (4b)}$$

$$D_{i,f} = 0 \qquad \forall f, \forall i \neq s(f), d(f) \text{ (4c)}$$

$$r_{ij} + \sum_{l \in N_i, l \neq j} r_{il} + \sum_{k \in N_j, k \neq i} r_{jk} + \sum_{(k,l) \in H3_{ij}} r_{kl} \leq B \qquad \forall link(i,j) \text{ (4d)}$$

$$r_{ij} = T_{ij} + T_{ji} \qquad \forall link(i,j) \text{ (4e)}$$

Table 2. The LP Model with Lossy Links (Continue)

$R_{ij} = (1 - l_{ij})T_{ij}$	$\forall link(i, j)$ (4f)
$D_{\min} \le -D_{d(f),f}$	$\forall f$ (4g)
$0 \le R_{ij}, T_{ij}, r_{ij}, R_{ij,f} \le B$	$\forall link(i, j)$ (4h)
$0 \le D_{s(f),f} \le B$	$\forall f$ (4i)

3.5 Routing Path Reconstruction

By solving the linear programming problem we can get $R_{ij,f}$, the link rate allocated for each flow. From $R_{ij,f}$, we can reconstruct the routing paths. The following algorithm can be used to construct a source-to-destination path.

Table 3. Algorithm for Routing Path Reconstruction

```
PATHRECONSTRUCTION(G(V,E),R)
1.  for each flow f(s, d) ∈ F
2.      do find the minimum value: f_min = min_ij{R_ij,f ≠ 0}
3.              construct a path p (s → d) using links (i, j) with R_ij,f ≥ f_min
4.              update D_s,f = D_s,f - f_min and R_ij,f = R_ij,f – f_min for each link (i, j) on p
5.              P = P ∪ p, datarate(p) = f_min
6.   iterate lines 2–5 until D_s,f = 0
```

4. MODEL VALIDATION THROUGH SIMULATION

The linear programming model defined in section III can be used to provide an end-to-end throughput estimation as well as to control data rates and routing topology so that the system throughput is maximized. We evaluate the model through extensive simulations in this section. All simulations are conducted in ns2 simulator, using IEEE 802.11 MAC. All wireless nodes are equipped with omni-directional antenna with communication range 250m and carrier sense range (a.k.a. interference range) 550m. Wireless channel bandwidth B' is 2 Mbps. The effective data rate B = B'/2.27. Constant 2.27 is due to 802.11 MAC protocol overhead.

In the following, we present our simulation results in two groups. The first group is to find where the optimal operation point occurs in terms of applied traffic load; the second group is to show the effectiveness of joint routing and rate control, and how it improves network throughput.

4.1 Prediction On Optimal Operation Point

The optimal operation point of a network system refers to the applied traffic load under which the network achieves the maximum throughput. Through simulation study, we show that our model can accurately predict the optimal throughput in a range of different network settings.

First, we study how the transmission from a single flow interferes with itself in a multi-hop network (Fig. 3.1). We deploy 5 nodes in a chain topology on a 1500x1500 square. Nodes are 150m apart. Due to the 550m carrier sense range, all four links are conflicting with each other. When source rate increases from 0.02B' to 0.18B', we observed the throughput increases until source rate Di = 0.11B' and then stays flat. Using our interference model (inequality 2e), the optimal operation point is D^*_i = B'/2.27/4 = 0.11B'. This shows in a multi-hop network the optimal throughput is achieved when the network is fully saturated.

Next we extend the one-flow scenario to two flows sharing a path. Two sources S1and S2 each generate data at data rate D_i. When Di increases from 0.01B' to 0.15B', we observed that the highest throughput occurred at D_i = 0.055B'. Using our interference model, the network is fully saturated at D^*_i = 0.055B'. Fig.3.2 shows the network topology and the throughput.

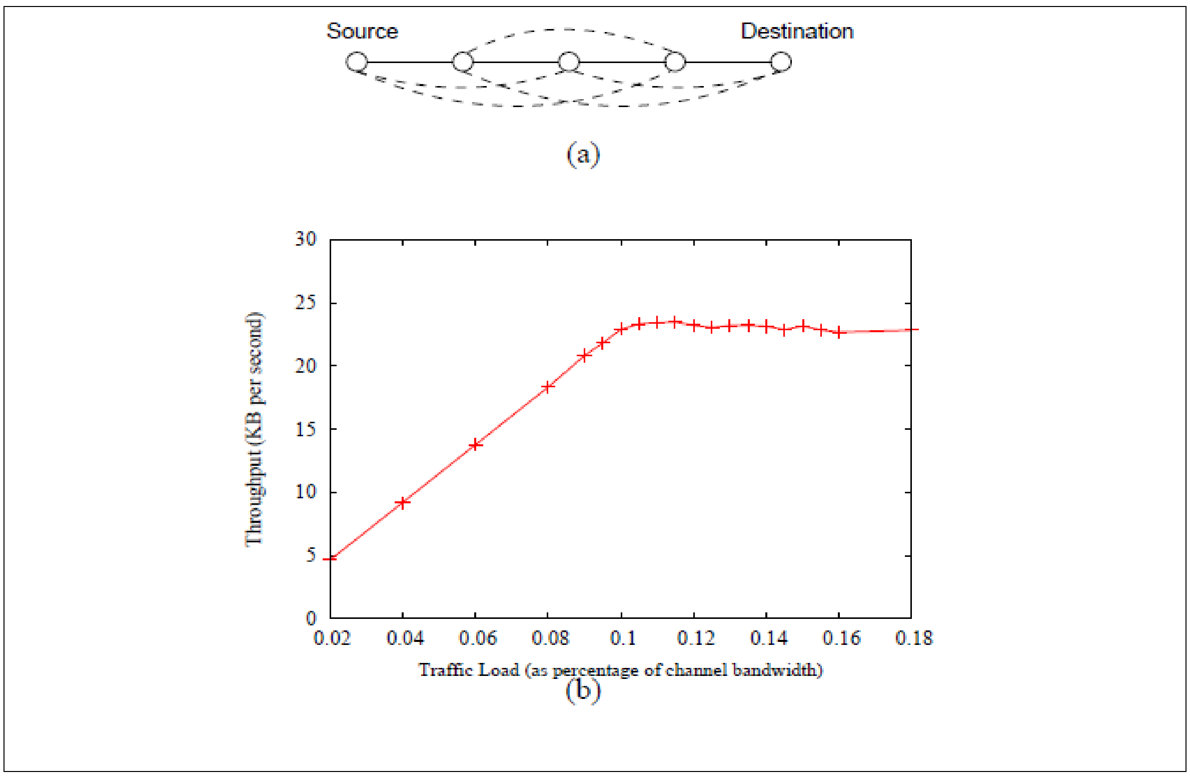

Fig.3.1 Network Throughput for Single Flow

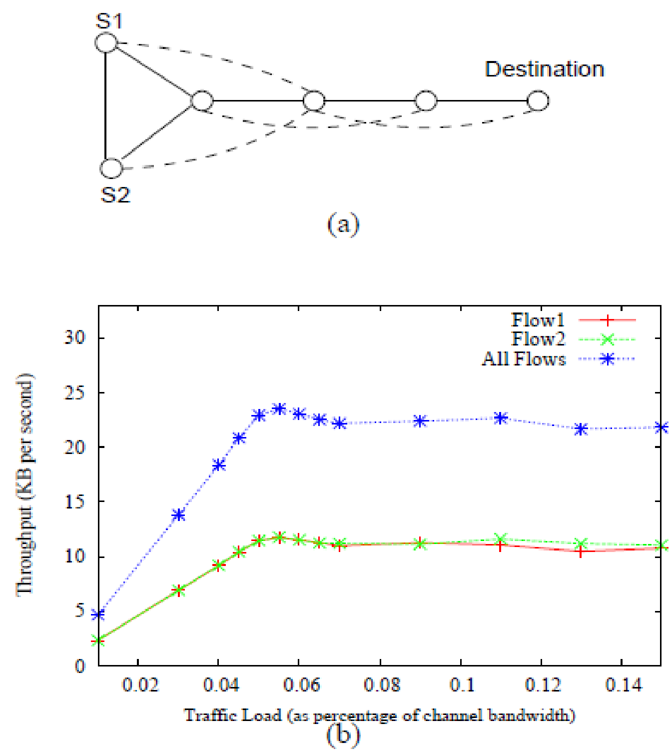

Fig. 3.2 Network Throughput for Two Flows

4.2 Joint Routing And Rate Control

We compare our joint routing and rate control scheme with a routing scheme that does not consider interference in routing. We use shortest path routing here for comparison purpose. Other routing schemes without considering interference will do the same. Wireless testbed results from [18] showed minimum hop-count paths usually have poor throughput performance. Our simulation verifies the observation and our interference model explains why— because shortest path routing tends to select links that are shared by many flows.

We study a 50-node network with random traffic demand. All nodes are randomly and uniformly deployed in a 2500×2500 region. We randomly choose 5 nodes to be sources and 5 nodes to be destinations. All sources have the same data rate D_i. When D_i increases from 0.005B’ to 0.08B’, we observed that shortest path routing reaches a peak throughput value 65.6K bytes per second at D_i = 0.07B’. Then we apply our linear programming model. The linear programming solution indicates network throughput is maximized when the applied traffic load is 0.07B’ for each source. Using our joint routing and rate control scheme, the network achieves the maximum throughput 80.7K bytes per second at D_i = 0.07B’.

5. CONCLUSION AND DISCUSSION

In this paper, we studied an important problem:"How to route data packets and control link transmission rates in order to provide users with communication efficiency and fairness?" We addressed the problem by using a linear programming model, in which wireless interference is effectively accounted for.

Collision is costly in wireless networks. In practice, we should always operate at near but below the optimal operation point. The linear optimization model presented in this paper tells us what the optimal operation point in terms of applied traffic load is, and how to find the routing and link rates to improve network efficiency. The model can be used to address heterogeneous networks with different link quality. Our model is based on static network topology and fixed interference relation. It can be easily extended to address nodes with different communication ranges.

6. REFERENCES

[1] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, “Impact of interference on multi-hop wireless network performance,” in MobiCom ’03: Proceedings of the 9th annual international conference on Mobile computing and networking. New York, NY,USA: ACM, 2003, pp. 66–80.

[2] S. Rangwala, R. Gummadi, R. Govindan, and K. Psounis, “Interference-aware fair rate control in wireless sensor networks,” in SIGCOMM ’06: Proceedings of the 2006 conference on Applications, technologies, architectures, and protocols for computer communications. New York, NY, USA: ACM, 2006, pp. 63–74.

[3] Y. Li, L. Qiu, Y. Zhang, R. Mahajan, Z. Zhong, G. Deshpande, and E. Rozner, “Effects of interference on wireless mesh networks: Pathologies and a preliminary solution,” in HotNets 2007, Nov. 2007.

[4] L. Qiu, Y. Zhang, F. Wang, M. K. Han, and R. Mahajan, “A general model of wireless interference,” in MobiCom ’07: Proceedings of the 13th annual ACM international conference on Mobile computing and networking. New York, NY, USA: ACM, 2007, pp. 171–182.

[5] Y. Li, L. Qiu, Y. Zhang, R. Mahajan, and E. Rozner, “Predictable performance optimization for wireless networks,” in SIGCOMM ’08: Proceedings of the ACM SIGCOMM 2008 conference on Data communication. New York, NY, USA: ACM, 2008, pp. 413–426.

[6] J. Padhye, S. Agarwal, V. N. Padmanabhan, L. Qiu, A. Rao, and B. Zill, “Estimation of link interference in static multi-hop wireless networks,” in IMC ’05: Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement. Berkeley, CA, USA: USENIX Association, 2005, pp. 28–28.

[7] M. Cheng, X. Gong, and L. Cai, “Link rate allocation under bandwidth and energy constraints in sensor networks,” in Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE, 30 2008-Dec. 4 2008, pp. 1–5.

[8] F. Tobagi, “Modeling and performance analysis of multihop packet radio networks,” Proceedings of the IEEE, vol. 75, no. 1, pp. 135–155, Jan. 1987.

[9] A. Kershenbaum, R. Boorstyn, and M.-S. Chen, “An algorithm for evaluation of throughput in multihop packet radio networks with complex topologies,” Selected Areas in Communications, IEEE Journal on, vol. 5, no. 6, pp. 1003–1012, Jul 1987.

[10] R. Boorstyn, A. Kershenbaum, B. Maglaris, and V. Sahin, “Throughput analysis in multihop csma packet radio networks,” Communications, IEEE Transactions on, vol.35, no. 3, pp. 267–274, Mar 1987.

[11] X. L. Huang and B. Bensaou, “On max-min fairness and scheduling in wireless ad -hoc networks: analytical framework and implementation,” in MobiHoc ’01: Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing. New York, NY, USA: ACM, 2001, pp. 221–231.

[12] M. Kodialam and T. Nandagopal, “Characterizing achievable rates in multi-hop wireless networks: the joint routing and scheduling problem,” in MobiCom ’03: Proceedings of the 9th annual international conference on Mobile computing and networking. New York, NY, USA: ACM, 2003, pp. 42–54.

[13] J. Li, C. Blake, D. S. De Couto, H. I. Lee, and R. Morris, “Capacity of ad hoc wireless networks,” in MobiCom ’01: Proceedings of the 7th annual international conference on Mobile computing and networking. New York, NY, USA: ACM, 2001, pp. 61–69.

[14] P. Gupta and P. Kumar, “The capacity of wireless networks,” Information Theory, IEEE Transactions on, vol. 46, no. 2, pp. 388–404, Mar 2000.

[15] M. Gastpar and M. Vetterli, “On the capacity of wireless networks: the relay case,” in INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, vol. 3, 2002, pp. 1577–1586.

[16] M. Alicherry, R. Bhatia, and L. Li, “Joint channel assignment and routing for throughput optimization in multiradio wireless mesh networks,” in ACM MobiCom,2005.

[17] P.-J. Wan, “Multiflows in multihop wireless networks,” in MobiHoc ’09: Proceedings of the tenth ACM international symposium on Mobile ad hoc networking and computing. New York, NY, USA: ACM, 2009, pp. 85–94.

[18] D. S. J. De Couto, D. Aguayo, B. A. Chambers, and R. Morris, “Performance of multihop wireless networks: shortest path is not enough,” SIGCOMM Comput. Commun. Rev., vol. 33, no. 1, pp. 83–88, 2003.

IV. MINIMUM LATENCY TRANSMISSION SCHEDULING IN MULTIHOP WIRELESS NETWORKS

Maggie Cheng[1], Xuan Gong[1] and Yibo Xu[1]
[1]Department of Computer Science, Missouri University of Science and Technology
Rolla, Missouri, USA
{chengm@mst.edu, xgwt2@mail.mst.edu, yxw95@mail.mst.edu}

ABSTRACT. End-to-end delay is an important QoS metric in sensor networks as well as any user application that involves transferring of small files. In this paper, we address how to minimize end-to-end latency in a multi-hop wireless network. End-to-end delay is defined as the total time it takes for a single packet to reach the destination. It is a result of many factors including the length of the routing path and the interference level along the path. In this paper we present a transmission scheduling scheme that minimizes the end-to-end delay along a given route. The link scheduling scheme is based on integer linear programming and involves interference modeling. Using this schedule, there are no conflicting transmissions at any time. Through simulation, we show that the proposed link scheduling scheme can significantly reduce end-to-end latency. By varying different routing policy, we also demonstrate that the shortest path routing does not necessarily result in minimum delay.

1. INTRODUCTION

With the increasing application of wireless mesh networks and sensor networks, multi-hop wireless networking technology is expected to not just provide multi-hop connectivity in locations where wired networks cannot reach, but also to support user traffic with certain service guarantees. Throughout and delay are the two major aspects of quality of service. The user-perceived data transfer speed is a combined effect of both data rate and end-to-end latency. For transferring a small file, the dominating factor is end-to-end latency; for transferring a large file, the dominating factor is data rate. In a typical sensor network, where small packets generated by sensors need to be periodically reported to the base station, delay plays a more important role.

In the past, we have seen many reports regarding how to maximize network throughput in multi-hop wireless networks [1]–[8]. However, the solution that maximizes network throughput often neglects the delay aspect and leads to poor performance in end-to-end latency. For the network in Fig.4.1, a maximum throughput routing algorithm would choose (a) since the two paths do not interfere with each other, and a minimum delay routing algorithm would choose (b) since it is the shortest path and there is no interference from other data flows. Most time the two of them do not choose the same routes.

In the example shown in Fig.4.1, the shortest path happens to have the minimum delay. In this paper, we will demonstrate that it is a misbelieve that the shortest path always leads to the minimum delay. In fact, end-to-end delay is a result of both the number of hops on the path, and the interference level along the path. Shortest path leads to the minimum delay only if the shortest path is the least interfered path.

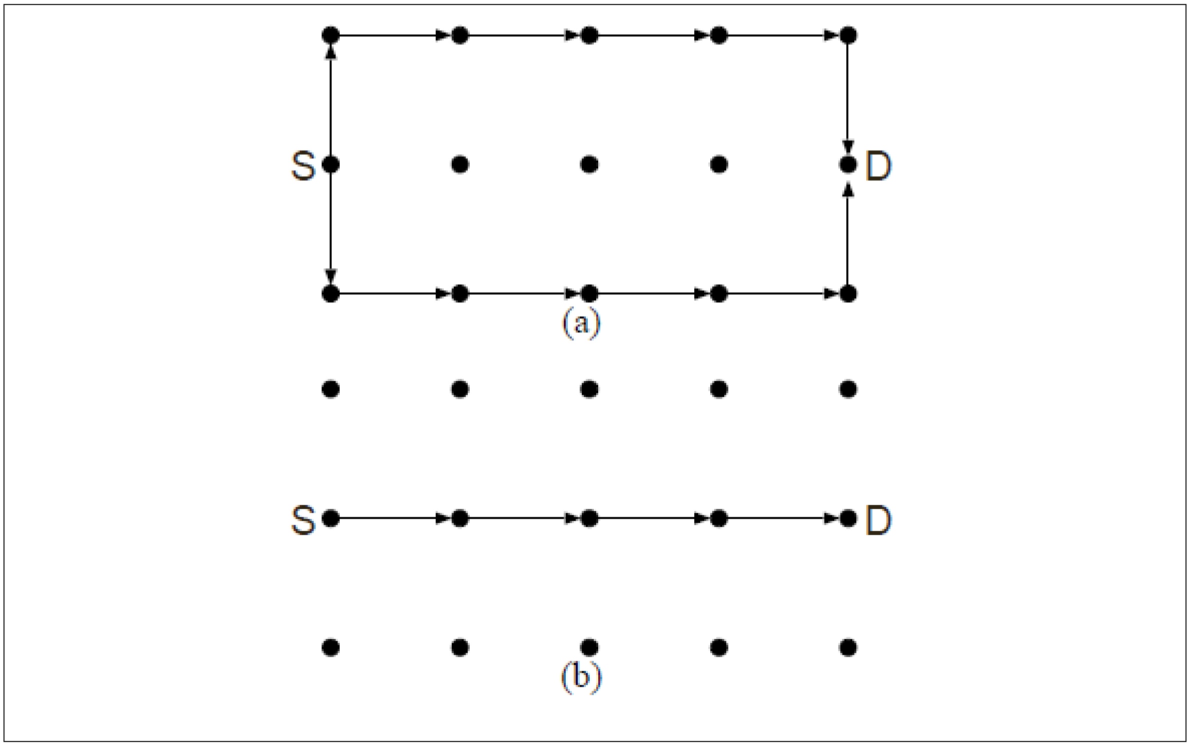

Fig.4.1 (a) With maximum throughput routing, latency is 6 slot time; (b) With minimum delay routing, latency is 4 slot-time.

Interference works adversely for delay the same way it does for throughput. Fig.4.2 shows that if there is only one data flow from source S1 to destination D1, end-to-end latency is 6 slots, assuming each slot is used to transmit one packet. However, if there are other transmissions nearby, the end-to-end latency of the same flow can be increased to 10 slots if we do not use optimization techniques and a packet is scheduled to use the next available slot as soon as it arrives.

When there are multiple data flows in the network, it is not straightforward to find the optimal transmission schedule that leads to the minimum delay. In this paper, we propose a linear programming-based link scheduling scheme that computes time slot assignment such that the end-to-end delay is minimum and at any time there are no conflicting transmissions. This link scheduling scheme can work with any routing scheme. The main contribution of this paper is that we come up with a linear optimization model to capture the impact of wireless interference on network delay in multi-hop wireless networks. Compared to previous linear models, our linear model is more accurate; and compared to the exact solution, which is a NP-hard to compute, our solution is more efficient.

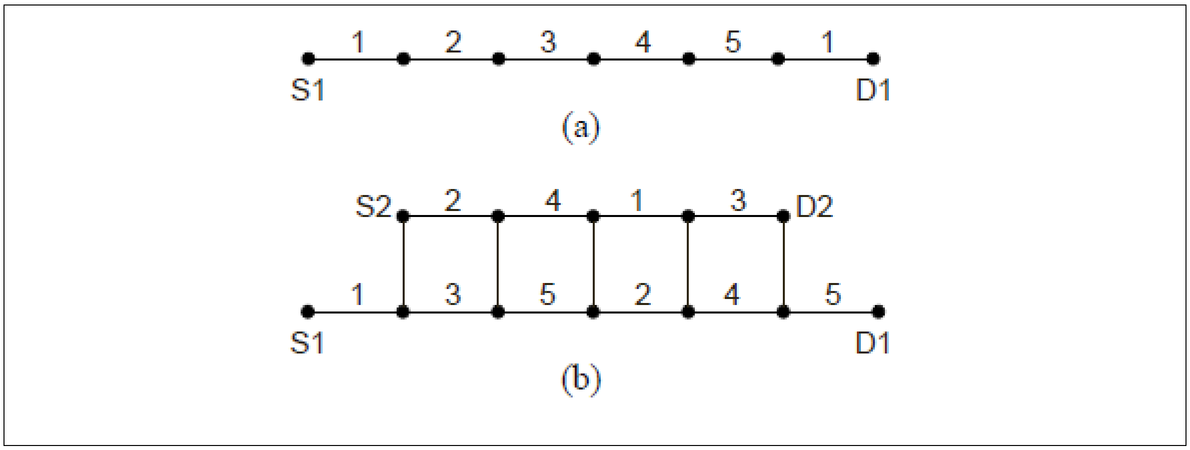

Fig.4.2 (a) With a single data flow, latency is 6 slot time; (b) When other transmitters are active, the latency becomes 10 slot time. Numbers on links are slot numbers. There are 5 distinct slot numbers.

The rest of the paper is organized as follows. In Section 2, we briefly survey the related work on interference modeling and delay optimization in recent years; in Section 3, we present a linear programming-based link scheduling scheme; in Section 4 we validate our model in extensive simulations. Finally, Section 5 concludes the paper.

2. RELATED WORK

We will first survey some papers on interference modeling, then we review some recent work in delay optimization.

For interference modeling, the most related work includes [1]–[5]. [1] first used conflict graphs to model the effect of wireless interference under a simplified protocol model; [3] continued to use conflict graphs to model interference under IEEE 802.11 interference model; [6] focused on estimation of interference and studied the effect of interference on aggregated network throughput based on IEEE 802.11 model; [4] proposed a physical interference model which is based on measured interference rather than distance between nodes. Our previous work [9] did joint routing and link rate control based on a different interference model that is based on directed graphs.

Delay optimization, often very important in sensor networks, has been approached from routing, MAC layer scheduling, or both. [10] presented in sensor networks when the routing tree is given, how to determine the time slot of each node such that the maximum latency to send a packet from a node to the sink is minimized. [11] presented an algorithm to find optimal routing paths between sensor and sink node pairs with the objective of minimizing the total end-to-end delay. [12] presented approximation algorithms for minimum latency aggregation in sensor networks, which computes an aggregation tree as well as time slot assignment for links so that the make span of the schedule is minimum.

3. MODEL-BASED MINIMUM DELAY LINK SCHEDULING

3.1. Scheduling Delay

Given the routing information, we can further reduce end-to-end latency by optimization on link scheduling delay. When a relay node forwards a packet, there is a mandatory store-and-forward delay and a link scheduling delay that is dependent on scheduling policy. Link scheduling delay is introduced when the outgoing link uses a time slot that is not immediately after the slot used by the incoming link. In Fig.IV.3, if the outgoing link uses slot number v, and incoming link uses slot number u, the total delay introduced at relay node r is d_r = u-v if u > v, or d_r = u-v + F if u < v, where F is the total number of distinct slots in a super-frame. If the schedule is conflict-free, it is guaranteed u ≠ v. The end-to-end delay for a path is $\sum_r d_r$. From this formula we can see that end-to-end delay is related to both the total number of hops, and the scheduling delay at each relay node. When routing information is given, the only factor that can be optimized is the scheduling delay.

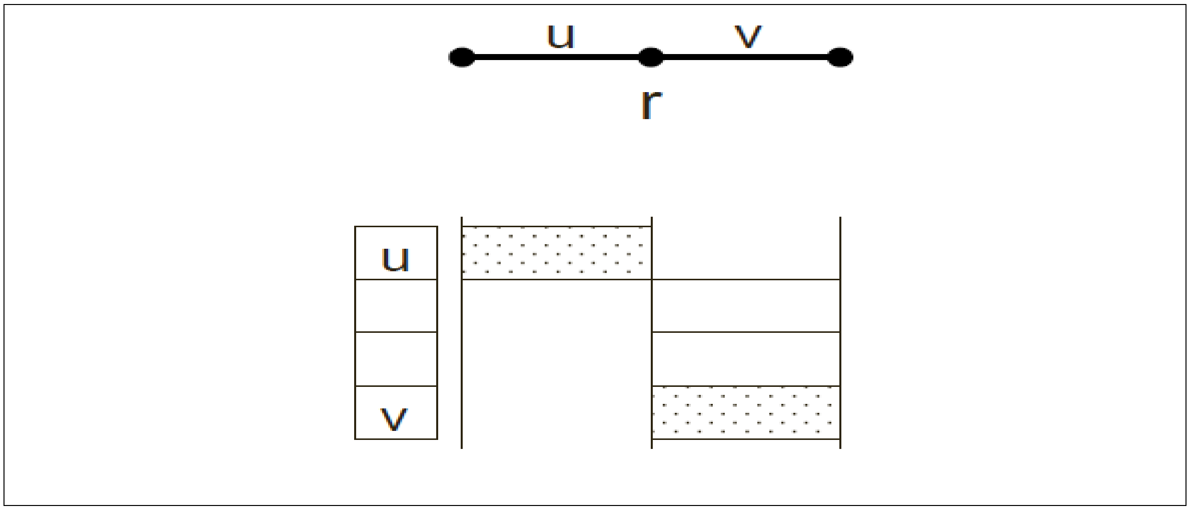

Fig.4.3 Scheduling delay at relay node

3.2. Interference Modeling

To find a conflict-free schedule, it is important that all active links in the same collision domain use different slots. In other words, no two links can use the same slot if they interfere with each other.

The collision domain is defined as a group of links that are mutually conflicting with each other. To list all collision domains in a network requires to build a conflict graph first and then to find all cliques in the conflict graph. The conflict graph is built as follows: we use vertices to represent wireless links, and then add an edge between two vertices if the wireless links they represent interfere with each other. To build the conflict graph can be done in polynomial time, however to find all cliques in the graph is an NP-hard problem. To avoid solving an NP-hard problem, we will find a sufficient set of links that includes all links in a clique and approximates the clique as closely as possible.

Suppose link (k, l) is disjoint from link (i, j) and both endpoints are within 2 hop of i and j respectively. Let H_{3ij} denote the group of links {(k, l)} that satisfy:

1) $1 \leq d(j, k) \leq 2$, $1 \leq d(i, l) \leq 2$, $d(j, k) + d(i, l) \leq 3$, or

2) $1 \leq d(i, k) \leq 2$, $1 \leq d(j, l) \leq 2$, $d(i, k) + d(j, l) \leq 3$.

Where $k \neq i, j$, and $l \neq i, j$; d(u, v) is the number of hops between node u and node v.

The collision domain CD_{ij} of link (i, j) includes: (1) link (i, j), (2) all adjacent links of (i, j), and (3) all two-hop links of (i, j) defined in H_{3ij}.

This set is sufficient in the sense that it captures all conflicting relation; it is also tight compared to previous work that simply includes all links in two-hop neighborhood. Among all polynomial-time solutions, CD_{ij} approximates the maximum clique that includes (i, j) most closely. Using CD_{ij} to describe the collision domain of link (i, j) allows us to address the problem in polynomial time, and at the same time to use the channel resource more efficiently than other polynomial time solutions.

3.3. A ILP Model For Minimum Delay Link Scheduling

To achieve minimum scheduling delay, we first formulate it as an optimization problem. Since the routing information is given, we use $link_{l,s} = 1$ to indicate link l is on the path for flow s. What we need to solve is the slot assignment for links. We introduce a 0-1 variable $sl_{l,f}$ for slot assignment. $sl_{l,f} = 1$ indicates link l uses slot f. If a link l is shared by multiple data flows, only one flow can use the slot f on the same link. $sl_{l,s,f} = 1$ indicates link l uses slot f for sending data from source node S.

Assume for source s, relay node r is on the routing path P_s. Relay node r receives flow from link m and forwards it to link n, the total delay at relay node r is $d_{r,s} = f_n - f_m + xF$, where f_n is the slot number for link n and f_m is the slot number for link m. Each slot time is equivalent to one standard packet transmission time. x is a boolean variable, x = 1 when $f_n < f_m$.

The integer linear programming model is now formulated as follows:

Table 1. ILP model for Minimum Delay Link Scheduling

Minimize: $\max_s \sum_{r \in P_s} d_{r,s}$	(1)
Or to minimize total delay: $\min \sum_s \sum_{r \in P_s} d_{r,s}$	(2)
Subject to:	
$\sum_{l' \in CD_l} sl_{l',f} \leq 1$	$\forall l, f$ (3a)
$sl_{l,f} = \sum_s sl_{l,s,f}$	$\forall l, f$ (3b)
$\sum_{f=1}^{F} sl_{l,s,f} = link_{l,s} R_s$	$\forall l, s$ (3c)
$d_{r,s} = \sum_{f=1}^{F} sl_{n,s,f} . f - \sum_{f=1}^{F} sl_{m,s,f} . f + x_{r,s} F$	$\forall r, s$ (3d)
$0 < d_{r,s} < F$, $x_{r,s} = \{0,1\}$	$\forall r, s$ (3e)

In inequality (3c), Rs is the data rate of source s, given as input. Although our purpose is only to minimize the end-to-end delay of a single packet regardless of the source data rate, the model is general enough to consider sources with different data rates. In simulation, we set $R_s = 1$ for all sources.

3.4. Computing The Slot Assignment

To solve the above integer linear programming problem is NP-hard. We first relax it to a linear programming problem, and then use maximum likelihood rounding to map real numbers to integer slot numbers.

Table 2. Slot Assignment Algorithm

1) Find the optimal solution for the LP problem with slot numbers relaxed to real numbers;
2) Sort $sl_{l,f}$ in non-increasing order, set Th = 0.5;
3) For each non-zero variable $sl_{l,f}$, if $sl_{l,f} \geq$ Th, assign $sl_{l,f} = 1$. Assign $sl_{l',f} = 0$ for other links l' that are conflicting with l. Assign remaining values appropriately to satisfy flow conservation; If Th > the largest $sl_{l,f}$, set Th = the largest $sl_{l,f}$;
4) Repeat step 3) until all variables are rounded to integers.

4. SIMULATION

In this section, we show that the proposed timeslot assignment algorithm can significantly reduce scheduling delay, given the routing path information. Through simulation, we also show that the shortest path does not always lead to the least latency.

In the simulation study, we use 50 nodes deployed on a 150x150 square region, with node transmission range 30. 10 out of the 50 nodes are randomly selected as source nodes, and all source nodes transmit to a common receiver (sink node). We assume routing information is given and we compare the end-to-end latency achieved by using a First-Come-First-Serve(FCFS) schedule with the one achieved by our link scheduling algorithm (call it MinDelay). Each source node generates a packet and we observe the end-to-end latency of the single packet. In FCFS, the packet arrival order is random. A relay node schedules a packet as soon as it arrives; when deciding which slot to use, a relay node chooses the next available slot to transmit the packet if it does not conflict with other transmissions. FCFS is one of the most commonly used scheduling policies in practice. Since the packet arrival order is an important factor to FCFS, we conducted 50 cases on 50 random arrival orders.

In the first simulation, we use the shortest path routing. The simulation results show MinDelay outperforms FCFS by 17% to 25% in total delay. In the second simulation, we use a different routing algorithm presented in [13](call it algorithm T). We compare the end-to-end latency achieved by FCFS and by MinDelay. From this simulation we observed not just MinDelay outperforms FCFS in all scenarios, algorithm T also leads to shorter delay than the shortest path routing. The reason is that shortest path routing does not consider wireless interference. When multiple data flows share the same link, the scheduling delay tends to be increased. On the other hand, algorithm T considers interference in routing and routes data to the less interfered paths. Although sometimes the path length is longer, but the scheduling delay is much shorter. MinDelay outperforms FCFS by 7% to 22%, and algorithm T outperforms shortest path routing by 20%. Fig.4.4.(a) and Fig.4.4.(b) show the total delay of all 10 flows with shortest path routing and interference-aware routing respectively.

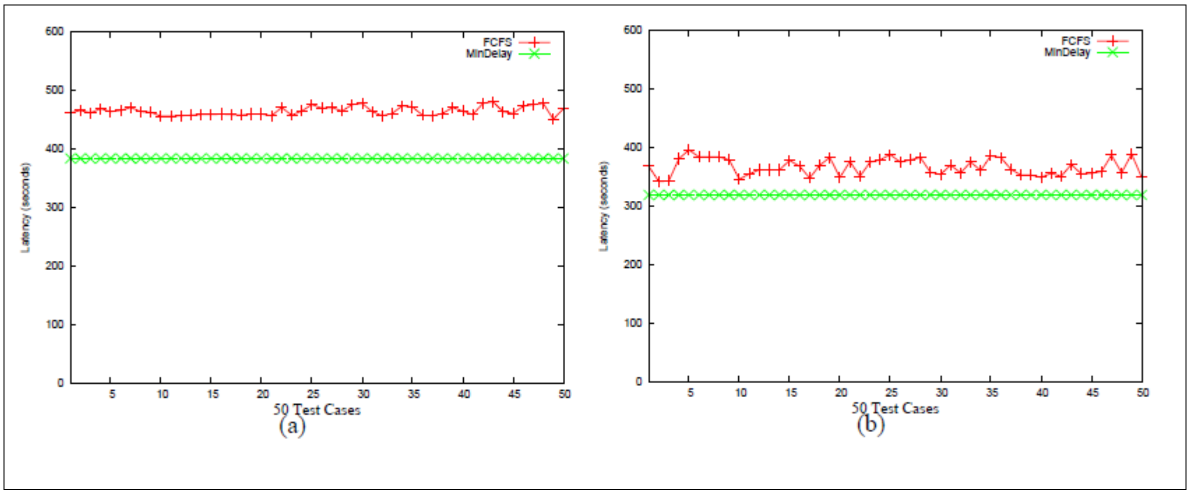

Fig.4.4 (a)Total delay using shortest-path routing; (b) Total delay using interference aware routing [13].

5. CONCLUSION AND DISCUSSION

In this paper, we addressed an important problem in practice: Given a multi-hop wireless network with multiple sources and destinations, how to achieve the minimum end-to-end delay? This paper presented a linear programming-based link scheduling scheme, in which wireless interference is sufficiently addressed.

The optimization model is useful for feasibility analysis given a set of QoS constraints, and it is also useful for predicting the achievable performance of the network and improving delay when routing information is given. The optimization framework can also be used for admission control as part of QoS provisioning in wireless networks. We will address this issue in the future work.

6. REFERENCES

[1] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, “Impact of interference on multi-hop wireless network performance,” in MobiCom ’03: Proceedings of the 9th annual international conference on Mobile computing and networking. New York, NY, USA: ACM, 2003, pp. 66–80.

[2] S. Rangwala, R. Gummadi, R. Govindan, and K. Psounis, “Interference-aware fair rate control in wireless sensor networks,” in SIGCOMM ’06: Proceedings of the 2006 conference on Applications, technologies, architectures, and protocols for computer communications. New York, NY, USA: ACM, 2006, pp. 63–74.

[3] Y. Li, L. Qiu, Y. Zhang, R. Mahajan, Z. Zhong, G. Deshpande, and E. Rozner, “Effects of interference on wireless mesh networks: Pathologies and a preliminary solution,” in HotNets 2007, Nov. 2007.

[4] L. Qiu, Y. Zhang, F. Wang, M. K. Han, and R. Mahajan, “A general model of wireless interference,” in MobiCom ’07: Proceedings of the 13th annual ACM international conference on Mobile computing and networking. New York, NY, USA: ACM, 2007, pp. 171–182.

[5] Y. Li, L. Qiu, Y. Zhang, R. Mahajan, and E. Rozner, “Predictable performance optimization for wireless networks,” in SIGCOMM ’08: Proceedings of the ACM SIGCOMM 2008 conference on Data communication. New York, NY, USA: ACM,2008, pp. 413–426.

[6] J. Padhye, S. Agarwal, V. N. Padmanabhan, L. Qiu, A. Rao, and B. Zill, “Estimation of link interference in static multi-hop wireless networks,” in IMC ’05: Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement. Berkeley, CA, USA: USENIX Association, 2005, pp. 28–28.

[7] M. Kodialam and T. Nandagopal, “Characterizing achievable rates in multi-hop wireless networks: the joint routing and scheduling problem,” in MobiCom ’03: Proceedings of the 9th annual international conference on Mobile computing and networking. New York, NY, USA: ACM, 2003, pp. 42–54.

[8] M. Alicherry, R. Bhatia, and L. Li, “Joint channel assignment and routing for throughput optimization in multiradio wireless mesh networks,” in ACM MobiCom, 2005.

[9] M. Cheng, X. Gong, and L. Cai, “Link rate allocation under bandwidth and energy constraints in sensor networks,” in Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE, 30 2008-Dec. 4 2008, pp. 1–5.

[10] P. Chatterjee and N. Das, “A cross-layer distributed tdma scheduling for data gathering with minimum latency in wireless sensor networks,” in Wireless VITAE 2009, may 2009, pp. 813 – 817.

[11] F. Sivrikaya and B. Yener, “Minimum delay routing for wireless networks with stdma,” Wireless Networks, vol. 15, no. 6, pp. 755–772, October 2007.

[12] P.-J. Wan, S. C.-H. Huang, L. Wang, Z. Wan, and X. Jia, “Minimum-latency aggregation scheduling in multihop wireless networks,” in MobiHoc 09, 2009, pp. 185–194.

[13] M. X. Cheng and X. Gong, “Interference-aware multipath routing and link rate control in multihop wireless networks,” 2010, manuscript.

V. CROSS-LAYER THROUGHPUT OPTIMIZATION WITH TRANSMISSION POWER CONTROL IN SENSOR NETWORKS

Maggie Cheng[1], Xuan Gong[1] and Xiaohua Jia[2]
[1]Department of Computer Science, Missouri University of Science and Technology
Rolla, Missouri, USA
Department of Computer Science, City University of Hong Kong
Hong Kong SAR
{chengm@mst.edu, xgwt2@mail.mst.edu, jia@cs.cityu.edu.hk}

ABSTRACT. In wireless sensor networks, transmission power has significant impact on network throughput as wireless interference increases with transmission power and interference negatively impacts network throughput. In this paper we try to improve network throughput through cross-layer optimization. We first present two algorithms to compute the transmission power of each node with objectives of minimizing total transmission power and minimizing total interference respectively, from which we can obtain a network topology that ensures a connected path from each source to the sink; then in order to evaluate the effectiveness of the power control algorithms, we compute the maximum achievable throughput from the obtained topology. The power control algorithms can generate symmetric links or asymmetric links if so desired. Based on different link models, we use different algorithms to compute the maximum achievable throughput. Since to compute the maximum throughput is an NP-hard problem, we use efficient heuristics that use a sufficient condition instead of the computationally-expensive-to-get optimal condition to capture the mutual conflict relation in a collision domain. The formal proof for the sufficient condition is provided and the proposed algorithms are compared to previous work. Simulation results show that the proposed algorithms improve network throughput and reduce energy consumption, with significant improvement over previous work on both aspects.

1. INTRODUCTION

In wireless sensor networks, due to the broadcast nature of wireless transmission, the signal from one sensor could reach many unintended receivers and interfere with the reception of these neighbors. The higher transmission power it uses, the more neighbors it interferes with. As the interference level increases, network throughput decreases. To intuitively understand how transmission power works on network throughput, we can picture a multi-hop wireless sensor network with a fixed number of nodes. If two nodes can hear from each other, we build a link between them. When one link is active, any other link that interferes with it should not be. When transmission power increases, link density increases, and consequently a wireless link will have many other links interfering with it. All these conflicting links cannot be active at the same time; they must be carefully scheduled to transmit at different time, otherwise their transmissions will interfere with each other. Although the wireless link capacity remains the same, the spatial reuse of the wireless spectrum decreases as the transmission power increases. As a result, network throughput drops.

To increase network throughput, we can address the problem from different layers: at the physical layer, we can adjust transmission power to reduce interference; at the network layer, we can route data packets to the least interfered path; and at the MAC layer, we can schedule transmissions to avoid simultaneous transmissions from interfering links. In order to make sure all transmissions can be scheduled without conflict, we also need to control the transmission data rate to make sure a node's channel occupation time is proportional to its data rate. Overall, it takes a cross-layer design scheme to achieve the maximum throughput.

In this paper, a cross-layer optimization framework is provided. we first try to decide the transmission power of each node towards optimizing throughput, then we use a joint routing and link rate control scheme to achieve the maximum throughput. The second part computes the maximum achievable throughput on a given topology, therefore, can serve as the assessment of the power control schemes.

The main contributions of this paper include: (1) we formulated the maximum throughput power control problem into two linear programs and designed efficient algorithms to solve them. The power control algorithms can generate symmetric or asymmetric links as required; (2) for both symmetric links and asymmetric links, we provided mathematical optimization models to compute the maximum achievable throughput on a given topology. Part of it requires to accurately capturing the mutual conflicting relation among wireless links, which is a well-known NP-hard problem. We proposed a polynomial-term constraint that can sufficiently capture the mutual conflict relation among wireless links and is tighter than all known polynomial-term approximations.

Although the objective of this paper is to achieve maximum throughput, we found that the power control schemes also reduce the total energy consumption of the sensor network. Through cross-layer optimization, we show that it is possible to achieve higher throughput with longer lifetime.

The rest of the paper is organized as follows. In Section 2, we briefly survey the most related work in cross layer optimization; in Section 3, we present the mathematical optimization models and algorithms for power control, and in Section 4 joint routing and link rate allocation; in Appendix we show the theoretical foundation of the optimization model with formal proof; in Section 5 we compare our algorithms with previous work and show the effectiveness of power control on throughput improvement.

2. RELATED WORK

Along the line of maximizing network throughput through transmission power control, the most related work is [1]. In [1] two pruning algorithms were presented to assign transmission power to nodes in order to minimize the maximal interference or total interference respectively, then linear programming models are used for data routing in order to maximize network throughput. We compared our LP-rounding based power control algorithms with the pruning algorithms in [1] and found significant performance improvement. [1] is the most related work since it also crosses three layers that involves power control, routing and transmission rate control.

Most of other cross-layer design schemes only involve two layers, such as joint routing and link rate allocation [2]–[4], and joint power control and scheduling when routing information is given [5]. In [5], links that share a common node are not allowed to transmit in the same slot; for disjoint links, whether a node's reception is interfered by others is decided by a physical model, i.e., if the receiver's signal-to-noise ratio exceeds the threshold, it is considered not interfering. In [5], the interference model is a hybrid of protocol model and physical model. The physical model is applicable only when the routing information is given and traffic demand on each link is given as input. However, in our work, routing information is not given and the traffic demand on each link is unknown, therefore a pure protocol model is used, in which the interfering relation is determined by network link topology rather than the actual signal strength.

Throughput modeling and optimization in wireless networks started as early as 1987 ([6]–[8]), and at that time it was for packet radio networks. In recent years it became a hot topic again when multi-hop wireless networks became popular. Some researchers attempted to give asymptotic results without input on traffic and network topology ([9]– [11]) and most others tried to find the exact solutions ([4], [12]–[18]). To find the exact throughput, part of the effort is to extend the concept of flow networks to multi-hop wireless networks. To come up with the capacity constraint, some scholars used link capacity as the upper bound of the data rate of a single link without considering the interference from other links [13], [19]; some attempted to model interference but used global information such as cliques on a conflict graph ([16]), which is NP hard to get in its first place; and some proposed polynomial-term constraint and simply considered all links within two hops of a common link as conflicting links and required the total data rate of these links be bounded by the wireless link capacity. We have demonstrated in this paper that our polynomial-term constraint is more accurate than this simplified model and can sufficiently capture the interference relation. Our interference model represents the tightest sufficient condition known so far.

3. TRANSMISSION POWER CONTROL

Given a sensor network of N nodes with adjustable transmission power, the objective of power control is to compute the transmission power for all nodes such that network throughput is maximized. Depending on whether DATA packets need to be acknowledged by the next hop, links can be symmetric or asymmetric. The algorithms presented in the following can produce either symmetric or asymmetric links.

Since network throughput is related to interference and interference is related to total transmission power, we use minimum total power and minimum interference as the optimization objectives respectively in the following for transmission power control.

3.1 For Minimum Total Power

a. Linear Programming Model

Variables: Let P_i be the transmission power of node i, let R_{ij} be the data rate on link (i, j), let X_{ij} be the decision variable: $X_{ij} = 1$ if there is a link from i to j and $X_{ij} = 0$ otherwise.

Constants: P_{ij} is the transmission power needed for node i to reach node j, Di is the source rate of node i, and B is the wireless link capacity. At this stage, the objective is to get a connected topology with minimum total power (connected means there is a connected path from each source to the sink), therefore D_i is arbitrarily set. If i is a source node, $D_i > 0$, if i is a sink node, $D_i < 0$, and if i is neither a source, nor a sink, then $D_i = 0$.

Now we can formulate the minimum power topology control problem as the following.

Table 1. Mathematical Model for Minimum Total Power

Minimize: $\sum_i P_i$	(1)
Subject to	
$\sum_j X_{ij} \geq 1$	$\forall i \in source$ (2a)
$\sum_j R_{ij} - R_{ji} = D_i$	$\forall i$ (2b)
$P_i \geq X_{ij} P_{ij}$	$\forall (i, j)$ (2c)
$X_{ij} \geq R_{ij} / B$	$\forall link(i, j)$ (2d)
$X_{ij} = \{0,1\}$	$foralllink(i, j)$ (2e)
$0 \leq R_{ij} \leq B$	$\forall link(i, j)$ (2f)

In the above formulation, equality (2a) requires that each source must have at least one outgoing link; Equality (2b) requires that data rate satisfy flow conservation; Equality (2c) requires that in order to establish a link from i to j, node i must use enough transmission power to reach j; and Equality (2d) requires that if the data rate from i to j is nonzero, there must be a link from i to j.

The solution from the above linear program includes X_{ij} and P_i, from which we can obtain a connected topology with minimum total power. However, since X_{ij} is 0-1 integer variable, the problem remains NP-hard. A LP-Rounding based heuristic is presented in the next section.

The above linear program is for asymmetric links, i.e., X_{ij} and X_{ji} can be different, which implies i can hear j but j cannot hear i or vice versa. If the links are required to be symmetric because DATA packets need to be acknowledged by the next hop, then the following constraint is added: $X_{ij} = X_{ji}$. In practice, P_{ij} could be different from P_{ji}, but to make sure i can reach j and j can reach i, we only need $X_{ij} = X_{ji}$, and apply condition (2c) on both (i, j) and (j, i): $P_i \geq X_{ij}P_{ij}$ and $P_j \geq X_{ji}P_{ji}$.

b. LP-Rounding Based Algorithm

We first relax the integer constraint of X_{ij} and solve the problem as a real-valued linear program. The solution includes fractional values for X_{ij}. We will use rounding based algorithm to construct the network topology.

We introduce two variables C_{ij} and M_i. $C_{ij} = 1$ means link (i, j) is established; $C_{ij} = 0$, otherwise. $M_i = 1$ means node i has a connected path to the sink; $M_i = 0$, otherwise.

Table 2. LP Routing Algorithm for Minimum Total Power

MINPOWER:

1. Sort X_{ij} in non-increasing order into a lost

    Set $C_{ij} = 0$ for all pairs of i,j

    Set $M_i = 0$ for all sources

2. While $\sum_{i \in sources} M_i <| sources |$

3. do remove the largest X_{ij} from the list

    set $C_{ij} = 1$

    set $P_i = \max_j\{C_{ij} \times P_{ij}\}$

    [for symmetric links set $C_{ji} = 1$, set $P_j = \max_i\{C_{ji} \times P_{ji}\}$ remove X_{ji} from the list]

Table 2. LP Routing Algorithm for Minimum Total Power (Continue)

4.	For all j If $P_i \geq P_{ij}$ Set $C_{ij} = 1$ Remove X_{ij} from the list
5.	If there is a connected path from node i to the sink, set $M_i = 1$ Return C_{ij} and P_i

Remark: If there is a tie in choosing the largest X_{ij} in line 3, choose the link (i,j) that leads to the smallest increase in the total power: for symmetric links, $(i,j) = \arg\min_{(i,j)} \left\{ \left(P_{ij} - P_i\right) + \left(P_{ji} - P_j\right) \right\}$; for asymmetric links, $(i,j) = \arg\min_{(i,j)} \left\{ P_{ij} - P_i \right\}$

3.2 For Minimum Total Interference

Since interference has more direct impact on network throughput, we try to use minimum total interference as the objective of power control. Intuitively, we will have a better chance of finding the topology that maximizes network throughput. Simulation results in Section 5 verified the prediction.

The interference model we adopted here is the"protocol model": if node j is in the interference range of node i, then j is interfered by i. We try to minimize the total number of nodes that are subject to interference.

a. Linear Programming Model

Variables: In addition to the variables defined in section 3.1, we define a new variable Y_{ik}: $Y_{ik} = 1$ if node i uses power level k; each node can only choose one transmission power. Ii is the number of nodes interfered by node i's transmission.

Constants: N_{ik} is the number of nodes in node i's interference range when node i uses power level k.

Table 3. Mathematical Model for Minimum Total Interference

Minimize: $\sum_i I_i$ (3)

Subject to:

$$\sum_j X_{ij} \geq 1 \qquad \forall i \in source \quad (4a)$$

$$\sum_j R_{ij} - R_{ji} = D_i \qquad \forall i \quad (4b)$$

$$P_i \geq X_{ij} P_{ij} \qquad \forall (i, j) \quad (4c)$$

$$\sum_k Y_{ik} = 1 \qquad \forall i \quad (4d)$$

$$I_i = \sum_k N_{ik} Y_{ik} \qquad \forall i \quad (4e)$$

$$P_i = \sum_k k Y_{ik} \qquad \forall i \quad (4f)$$

$$X_{ij} = R_{ij} / B \qquad \forall link(i, j) \quad (4g)$$

$$X_{ij} = \{0,1\} \qquad \forall link(i, j) \quad (4h)$$

$$Y_{ik} = \{0,1\} \qquad \forall i, k \quad (4i)$$

$$0 \leq R_{ij} \leq B \qquad \forall link(i, j) \quad (4j)$$

(4d) indicates that each node can only choose one power level. The lowest power level is 0, when node is not transmitting. (4e) defines the number of nodes interfered by node i's transmission. (4f) translates a 0-1 variable Y_{ik} into a discrete-valued power Pi. In this formulation, the constants P_{ij} is also given in discrete power levels.

Similarly, this integer linear program is NP-hard to solve. We will describe a LP-rounding based scheme in the following.

b. Rounding

The rounding algorithm is largely the same as the rounding algorithm for minimum power, except that when there is a tie in choosing the largest X_{ij} in line 3, we will choose the link (i, j) that leads to the smallest increase in the total interference: for symmetric links, if link (i, j) is chosen, $P_i = \max_j \left\{ C_{ij} P_{ij} \right\}$ and $P_j = \max_i \left\{ C_{ji} P_{ji} \right\}$, update Y_{ik} and Y_{jk}, and then calculate the total increase in interference $\delta_i = \left(\sum_k N_{ik} Y_{ik} - I_i \right) + \left(\sum_k N_{jk} Y_{jk} - I_j \right)$, set $(i, j) = \arg\min_{(i,j)} \delta_i$; for asymmetric link, $\Delta_i = \sum_k N_{ik} Y_{ik} - I_i$, set $(i, j) = \arg\min_{(i,j)} \delta_i$

4. MAXIMUM ACHIEVABLE THROUGHPUT

The output from power control algorithms is the transmission power of each node and the resulting topology. It is guaranteed that each source has a connected path to the sink. However, how much throughput can be achieved depends not only on the topology but also on the upper layer protocols such as routing and MAC. Without presumption about what routing and MAC algorithms are used, we calculate the maximum achievable throughput on the resulting topology, which is a measure of the effectiveness of power control algorithms.

4.1 Asymmetric Links for One-Way Communication

If DATA packets do not need to be acknowledged, links do not need to be symmetric. A directed path from source to sink consisting of asymmetric links will suffice.

We define N_i as the group of nodes that i can reach, i.e., $N_i = \{j|C_{ij} = 1\}$, where $C_{ij} = 1$ means there is a directed link from i to j; and we define N^+_i as the group of nodes that can reach node i: $N^+_i = \{j|i \in N_j\}$. N_i and N^+_i are obtained as a result of power control and are given as input to the following optimization model. Let variable R_i be the source rate of node i. If node i is neither a source nor the sink, R_i is set to be zero. We also introduce a decision variable f_i: $f_i = 1$ if i is expected to receive data, i.e., i is a relay node on the routing path or i is a sink. The joint routing and link rate allocation problem can be formulated as follows.

Table 4. Mathematical Model for One-way Communication

Maximize $\sum_{i \in sources} R_i$	(5)
Subject to	
$\sum_{j \in N_i} R_{ij} - \sum_{j \in N_i^+} R_{ji} = R_i$	$\forall i$ (6a)
$\sum_{j \in N_i} R_{ij} + f_i \sum_{j \in N_i^+} \sum_{k \in N_j} R_{jk} \le B$	$\forall i$ (6b)
$0 \le R_{ij} \le B$	$\forall i, \forall j$ (6c)
$f_i = \{0,1\}$	$\forall i$ (6d)

(6a) is for flow conservation, and (6b) is the capacity constraint for wireless transmissions. Inequality (6b) is a sufficient condition to capture the mutual conflict relationship among links. In our previous work [3], a formal proof for its sufficiency is provided.

In order to linearize inequality 6b so that we can solve it as a linear program, we set the initial value of f_i as follows and use iterative approach to find the solution: Initially we set $f_i = 1$ for all nodes that have $\sum_{j \in N_i} C_{ij} \ge 1$, then set $f_i = 1$ if i is the sink, and set $f_i = 0$ if i is a source; it takes 2 to 3 iterations to converge.

4.2 Symmetric Links for Two-Way Communication

If DATA packets must be followed by ACKs, links must be symmetric, i.e., $C_{ij} = C_{ji}$. In the following, we assume links are symmetric and communication on a link is two-way, therefore, all links within two hops of each other interfere.

Table 5. Mathematical Model for Two Way Communication

Maximize: $\sum_{i \in sources} R_i$	(7)
Subject to:	
$\sum_{j \in N_i} \left(R_{ij} - R_{ji} \right) = R_i$	$\forall i$ (8a)
$r_{ij} + \sum_{l \in N_i, l \neq j} r_{il} + \sum_{k \in N_j, k \neq i} r_{jk} + \sum_{(k,l) \in N_{2ij}} r_{kl} \leq B$	$\forall link(i, j)$ (8b)
$r_{ij} = R_{ij} + R_{ji}$	$\forall link(i, j)$ (8c)
$0 \leq R_{ij}, r_{ij} \leq B$	$\forall link(i, j)$ (8d)

In this linear program, equalities (8a) is for flow conservation, and inequality (8b) defines the capacity constraint. Capacity constraint is the reason for not being able to further increase throughput. Inequality (8b) ensures all links possibly in the same collision domain have a total demand less than B.

In inequality (8b), N_{2ij} is defined as: N_{2ij} = {(k, l)| link (k, l) is a two-hop neighbor of link (i, j), and the sum of distance from k to (i, j) and from l to (i, j) via a different path is ≤ 4}. If there is no other path, the distance is counted as ∞.

For example, in Fig.5.1.(a), link (k_1, l_1) and (k_2, l_2) belong to N_{2ij} , but (k_2, l_1) does not, because (k_2, l_1) is not a 2-hop neighbor of link (i, j); in fig.5.1.(b), link (k, l) does not belong to N_{2ij} , since there is only one path to reach link (i, j) from k and l; the distance from k to (i, j) is 1 and the distance from l to (i, j) is 1. In this case, the mutual conflicting relation among (i, j), (j, k), and (k, l) is captured when we apply the constraint (8b) on link (j, k): we make sure the data rate satisfy $r_{jk} + r_{ij} + r_{kl} \leq B$.

Inequality (8b) is a sufficient but not necessary condition for capturing all conflict relation in wireless communication. The accurate condition, which is both sufficient and necessary condition, includes no more than necessary links in the left hand side of the inequality. However, it is an NP-hard problem to identify these links. To identify these links, we need to first construct a conflict graph [4], in which a link is represented as a node, and a pair of wireless links that are mutually conflicting with each other is connected by an edge. Then we need to compute all cliques on the graph and make sure all nodes in a clique have total data rate no more than B. However, it takes exponential time to list all cliques. To the best of our knowledge, inequality (8b) is so far the most accurate polynomial time solution. For links within 2 hops of link (i, j), we only include the links that belong to N_{2ij} in the inequality. Compared to previous work in which all links within 2 hops of (i, j) are included in the left hand side of the inequality ([1]), our solution provides a tighter bound therefore the enables higher throughput.

Consider the topology in Fig 5.2.(a), the conflict graph is in Fig 5.2.(b), in which each wireless link is represented as a node, and links that are conflicting with each other are connected by an edge. The optimal solution requires the total data rate on any clique be bounded by B, therefore, the following conditions must be satisfied: $r_{ij} + r_{il} + r_{jk} \le B$ (9a), $r_{il} + r_{ij} + r_{lx_1} + ... + r_{lx_n} \le B$ (9b) and $r_{jk} + r_{ij} + r_{ky_1} + ... + r_{ky_n} \le B$ (9c)

Our solution, derived from inequality (8b), requires the following conditions be satisfied. It is the same as the optimal solution.

for (i, j) : $r_{ij} + r_{il} + r_{jk} \leq B$ (10a)

for (i, l) : $r_{il} + r_{ij} + r_{lx1} + ... + r_{lxn} \leq B$ (10b)

for (j, k) : $r_{jk} + r_{ij} + r_{ky1} + ... + r_{kyn} \leq B$ (10c)

However, the previous work that simply includes all links that interfere with (i, j) requires the following condition be satisfied ([1]), even though links $l_{x1},\ldots l_{xn}$ have no conflict with links $k_{y1}, \ldots k_{yn}$.

$r_{ij} + r_{il} + r_{jk} + r_{lx1} + ... + r_{lxn} + r_{ky1} + ... r_{kyn} \leq B$

Apparently the above condition introduces larger performance gap than condition (8b). (8b) can sufficiently capture all conflicting relation and is the most accurate polynomial-term condition known so far. The formal proof for sufficient condition is included in Appendix VII.

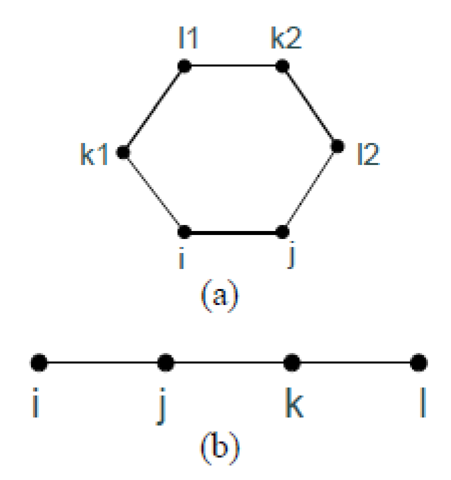

Fig. 5.1 Capacity Constraint

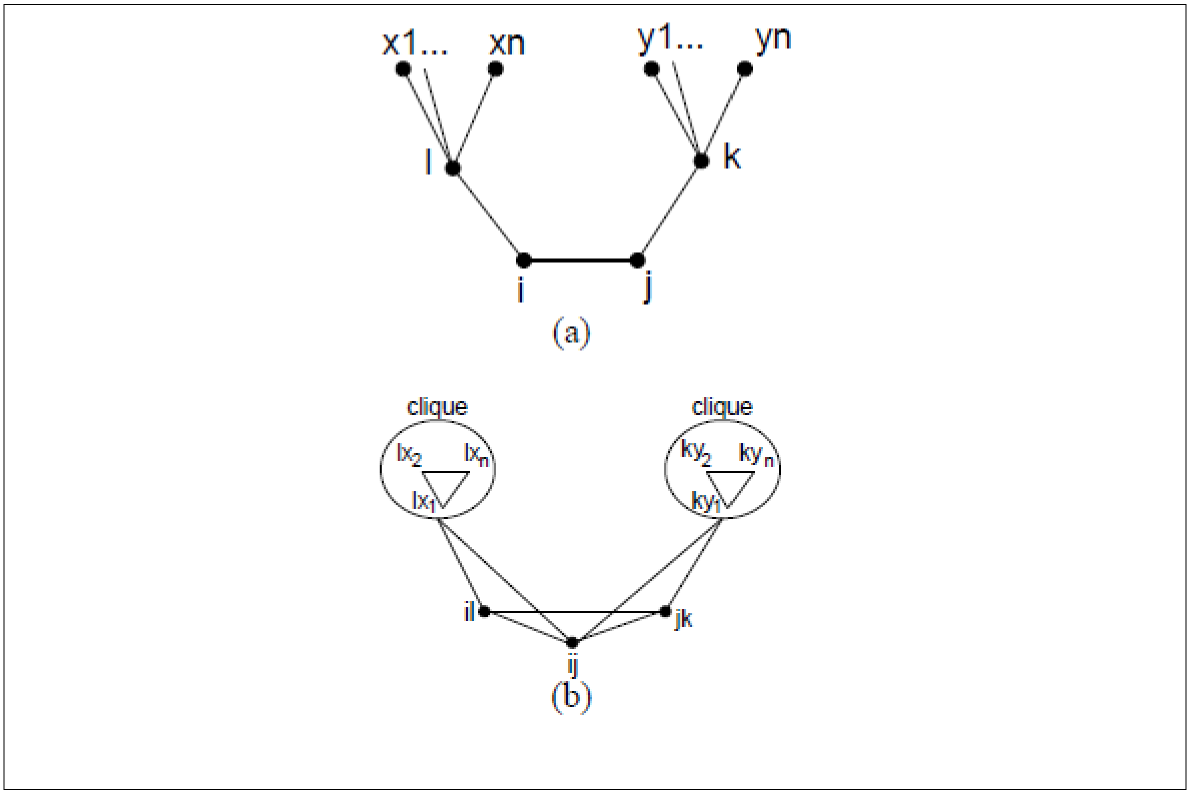

Fig. 5.2 (a) A simple network (b) the conflict graph of (a)

5. SIMULATION

We first evaluate the effect of power control algorithms on total energy savings and throughput improvement, and compare our algorithms with the ones that do not use power control(referred to as uniform model), then we compare our algorithms with previous work in [1] on total energy consumption and throughput.

The network consists of 50 sensor nodes and one sink node. All nodes are randomly deployed in a 250×250 region. One node is randomly chosen as sink and other 50 nodes are source nodes. Each node has 10 different power levels (K=10) and the difference in transmission range of adjacent power levels is 5, while the minimal transmission range (power level 1) is also 5. In addition, the interference range is assumed to be 2 times of the corresponding transmission range. The link capacity is assumed to be 30 (normalized B=30).

In the uniform model without power control all nodes transmit at the same power level, therefore links are symmetric. For comparison purpose, we ensure links are symmetric in our power control algorithm. Once the topology is determined, we run the maximum throughput algorithm on the symmetric model. Fig.5.3(a) shows the total power consumed by all nodes, and Fig.5.3(b) shows the throughput achieved. We compare two of our power control algorithms, LP-MinPower and LP-MinInterference with the uniform models with transmission range 35(at power level 7) and 45(at power level 9). The results show that our algorithms use less energy and achieve better throughput. LP-MinPower has the lowest total power consumption and LPMinInterference has the highest throughput.

The second simulation is to compare the performance of our algorithms with previous work in [1]. The network setup is the same. Since the algorithms in [1] produces topology with asymmetrical links, for comparison purpose, we also use asymmetric model in our algorithms. It is observed that LP-MinPower uses the least energy, and LP-MinInterference achieves the highest throughput. Both LP-MinPower and LP-MinInterference achieved higher throughput than the previous work.

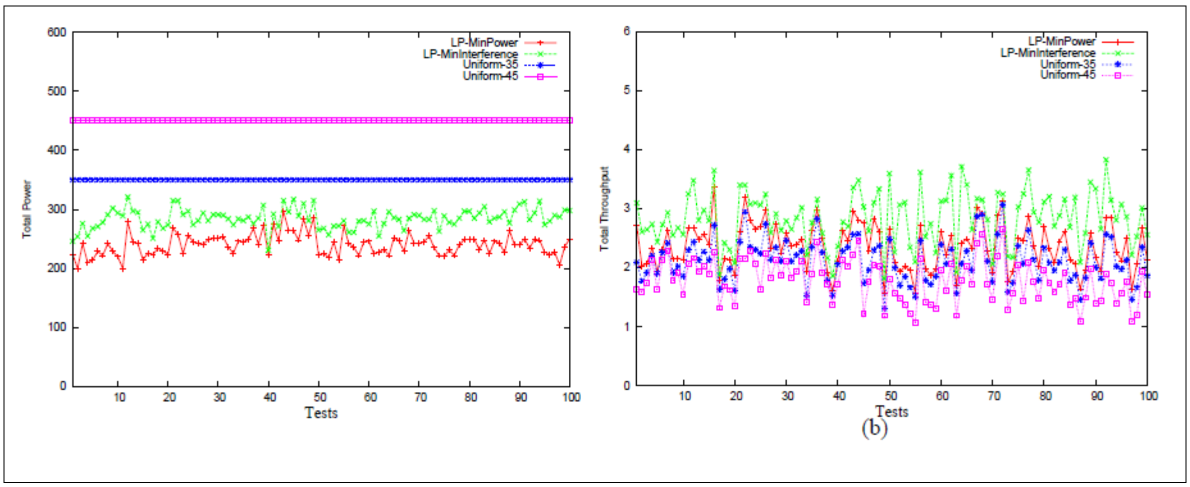

Fig.5.3 LP-MinPower and LP-MinInterference compared to the uniform model with symmetric links. (a) total power (b) total throughput.

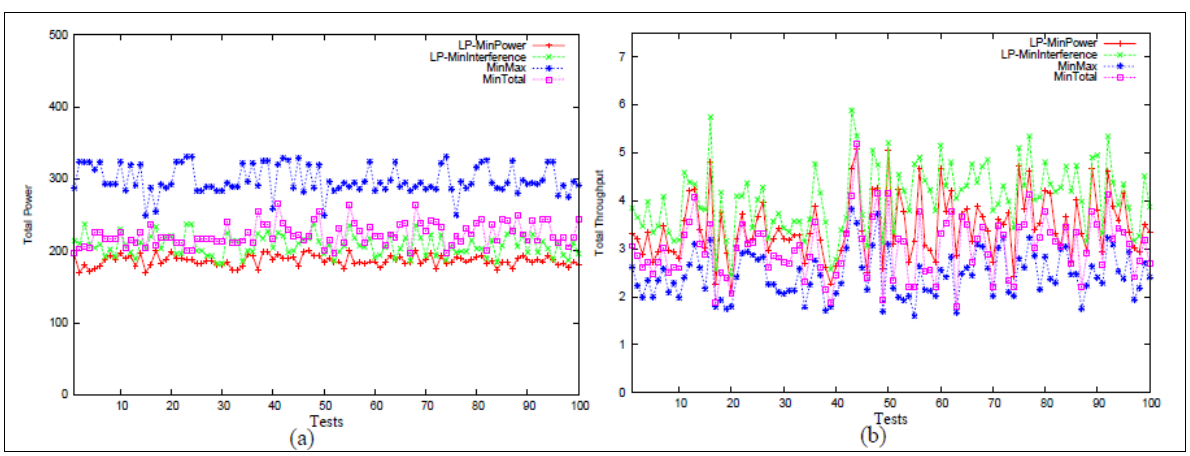

Fig.5.4 LP-MinPower and LP-MinInterference compared to MinMax and MinTotal in previous work with asymmetric Links. (a) total power (b) total throughput.

6. CONCLUSION

In this paper, we addressed the question of how to achieve the maximum throughput in sensor networks through cross-layer optimization. We first use transmission power control to decide the link topology and then use joint routing and link rate control to decide the maximum achievable throughput on the topology. We provided optimization models and efficient algorithms for power control as well as for joint routing and rate control. To effectively estimate the impact of wireless interference on throughput, we proposed to use a sufficient condition in the linear program, and also provided vigorous mathematical proof that the condition is sufficient to capture the interfering relation among wireless links. Although the proposed algorithms aim to optimize throughput only, they also reduce energy consumption of sensor networks. For future work, we will consider the joint optimization of throughput and energy with specific requirement on energy or lifetime.

7. REFERENCES

[1] D. Li, B. Wang, and X. Jia, “Topology control for throughput optimization in Wireless mesh networks,” in Mobile Ad-hoc and Sensor Networks, 2008. MSN 2008. The 4th International Conference on, 10-12 2008, pp. 161 –168.

[2] M. X. Cheng, X. Gong, and L. Cai, “Link rate allocation under bandwidth and energy constraints in sensor networks,” in IEEE Globecom, 2008, pp. 1–5.

[3] M. X. Cheng, X. Gong, and L. Cai, “Joint routing and link rate allocation under bandwidth and energy constraints in sensor networks,” IEEE Transactions on Wireless Communications, vol. 8, no. 7, pp. 3770–3779, 2009.

[4] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, “Impact of interference on multi-hop wireless network performance,” in MobiCom ’03: Proceedings of the 9th annual international conference on Mobile computing and networking. New York, NY, USA: ACM, 2003, pp. 66–80.

[5] J. Tang, G. Xue, C. Chandler, and W. Zhang, “Link scheduling with power control for throughput enhancement in multihop wireless networks,” Vehicular Technology, IEEE Transactions on, vol. 55, no. 3, pp. 733 –742, may 2006.

[6] F. Tobagi, “Modeling and performance analysis of multihop packet radio networks,” Proceedings of the IEEE, vol. 75, no. 1, pp. 135–155, Jan. 1987.

[7] A. Kershenbaum, R. Boorstyn, and M.-S. Chen, “An algorithm for evaluation of throughput in multihop packet radio networks with complex topologies,” Selected Areas in Communications, IEEE Journal on, vol. 5, no. 6, pp. 1003–1012, Jul 1987.

[8] R. Boorstyn, A. Kershenbaum, B. Maglaris, and V. Sahin, “Throughput analysis in multihop csma packet radio networks,” Communications, IEEE Transactions on, vol. 35, no. 3, pp. 267–274, Mar 1987.

[9] J. Li, C. Blake, D. S. De Couto, H. I. Lee, and R. Morris, “Capacity of ad hoc wireless networks,” in MobiCom ’01: Proceedings of the 7th annual international conference on Mobile computing and networking. New York, NY, USA: ACM, 2001, pp. 61–69.

[10] P. Gupta and P. Kumar, “The capacity of wireless networks,” Information Theory, IEEE Transactions on, vol. 46, no. 2, pp. 388–404, Mar 2000.

[11] M. Gastpar and M. Vetterli, “On the capacity of wireless networks: the relay case,” in INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, vol. 3, 2002, pp. 1577–1586.

[12] X. L. Huang and B. Bensaou, “On max-min fairness and scheduling in wireless ad-hoc networks: analytical framework and implementation,” in MobiHoc ’01: Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing. New York, NY, USA: ACM, 2001, pp. 221–231.

[13] M. Kodialam and T. Nandagopal, “Characterizing achievable rates in multi-hop wireless networks: the joint routing and scheduling problem,” in MobiCom ’03: Proceedings of the 9th annual international conference on Mobile computing and networking. New York, NY, USA: ACM, 2003, pp. 42–54.

[14] J. Padhye, S. Agarwal, V. N. Padmanabhan, L. Qiu, A. Rao, and B. Zill, "Estimation of link interference in static multi-hop wireless networks," in IMC '05: Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement. Berkeley, CA, USA: USENIX Association, 2005, pp. 28–28.

[15] S. Rangwala, R. Gummadi, R. Govindan, and K. Psounis, "Interference-aware fair rate control in wireless sensor networks," in SIGCOMM '06: Proceedings of the 2006 conference on Applications, technologies, architectures, and protocols for computer communications. New York, NY, USA: ACM, 2006, pp. 63–74.

[16] Y. Li, L. Qiu, Y. Zhang, R. Mahajan, Z. Zhong, G. Deshpande, and E. Rozner, "Effects of interference on wireless mesh networks: Pathologies and a preliminary solution," in HotNets 2007, Nov. 2007.

[17] L. Qiu, Y. Zhang, F. Wang, M. K. Han, and R. Mahajan, "A general model of wireless interference," in MobiCom '07: Proceedings of the 13th annual ACM international conference on Mobile computing and networking. New York, NY, USA: ACM, 2007, pp. 171–182.

[18] Y. Li, L. Qiu, Y. Zhang, R. Mahajan, and E. Rozner, "Predictable performance optimization for wireless networks," in SIGCOMM '08: Proceedings of the ACM SIGCOMM 2008 conference on Data communication. New York, NY, USA: ACM, 2008, pp. 413–426.

[19] M. Alicherry, R. Bhatia, and L. Li, "Joint channel assignment and routing for throughput optimization in multiradio wireless mesh networks," in ACM MobiCom,2005.

APPENDIX

The optimal solution to the maximum throughput problem defined in Section 3 requires the total data rate of all links represented by any clique be bounded by B: $\sum_{(i,j)\in Q} r_{ij} \leq B$, $\forall$ clique Q on the conflict graph. This is a sufficient and necessary condition. Since to list all cliques in a graph is an NP-hard problem, hereby we use a sufficient condition in its place. Inequality 8b is a sufficient condition and it takes polynomial time to compute.

Theorem.1: If inequality (8b) is satisfied on every wireless link, then the following constraint is satisfied: $\sum_{(i,j)\in Q} r_{ij} \leq B$, $\forall$ clique Q on the conflict graph.

Proof: We show that for any clique found on the conflict graph, the left hand side of inequality (8b) includes the data rate of all links represented in the clique.

We take an arbitrary clique of size n. When n=2, there are only two links concerned. Call them link i and link j. Inequality (8b) requires $r_i + r_j \leq B$ when i and j are 1-hop neighbors, or $r_k+r_i+r_j \leq B$ when i and j are 2-hop neighbors (see Fig.5.5). So the sufficient and necessary condition $r_i+r_j \leq B$ (from the clique approach) is trivially satisfied.

When $n \geq 3$, we distinguish two cases: case (1), the n links are on a network that does not have closed cycles (see Fig.5.6(a)); and case (2), the n links are on a network that has closed cycles with zero or more open tails (see Fig.5.6(b) and (c)). We assume wireless links i, j and k are on a clique. In case (1), since all links on a clique are within two hops of each other, and there is no cycle, choosing the link with the check mark to apply condition (8b) can ensure $r_i+r_j+r_k \leq B$. In case (2), apparently if i, j, and k are on a single cycle of 7 or more links (”single” means it does not contain any other cycles in it), then they must be connected head-totail in order to have mutual conflicts and form a clique (Fig.5.6(b)). This trivial case can be easily solved by applying condition (8b) on the middle one. Otherwise, if the cycle has at most 6 links, from Fig. 5.6(c), it can be shown that by applying condition (8b) on the link with the check mark, we can ensure $ri+rj+rk \leq B$. Therefore, the inequality (8b) is a sufficient condition.

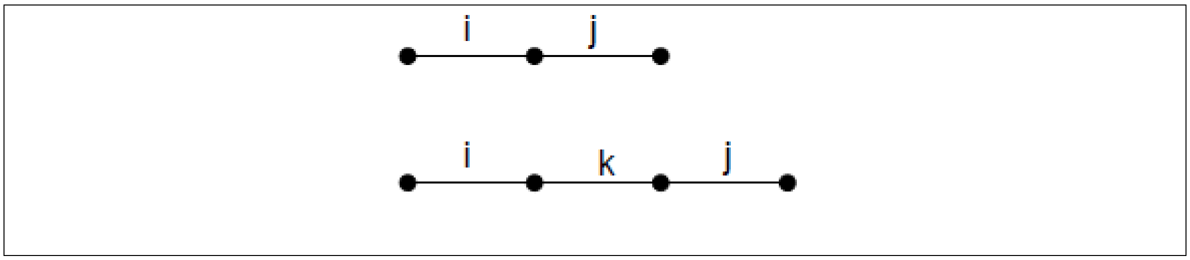

Fig.5.5 For a clique of size n = 2

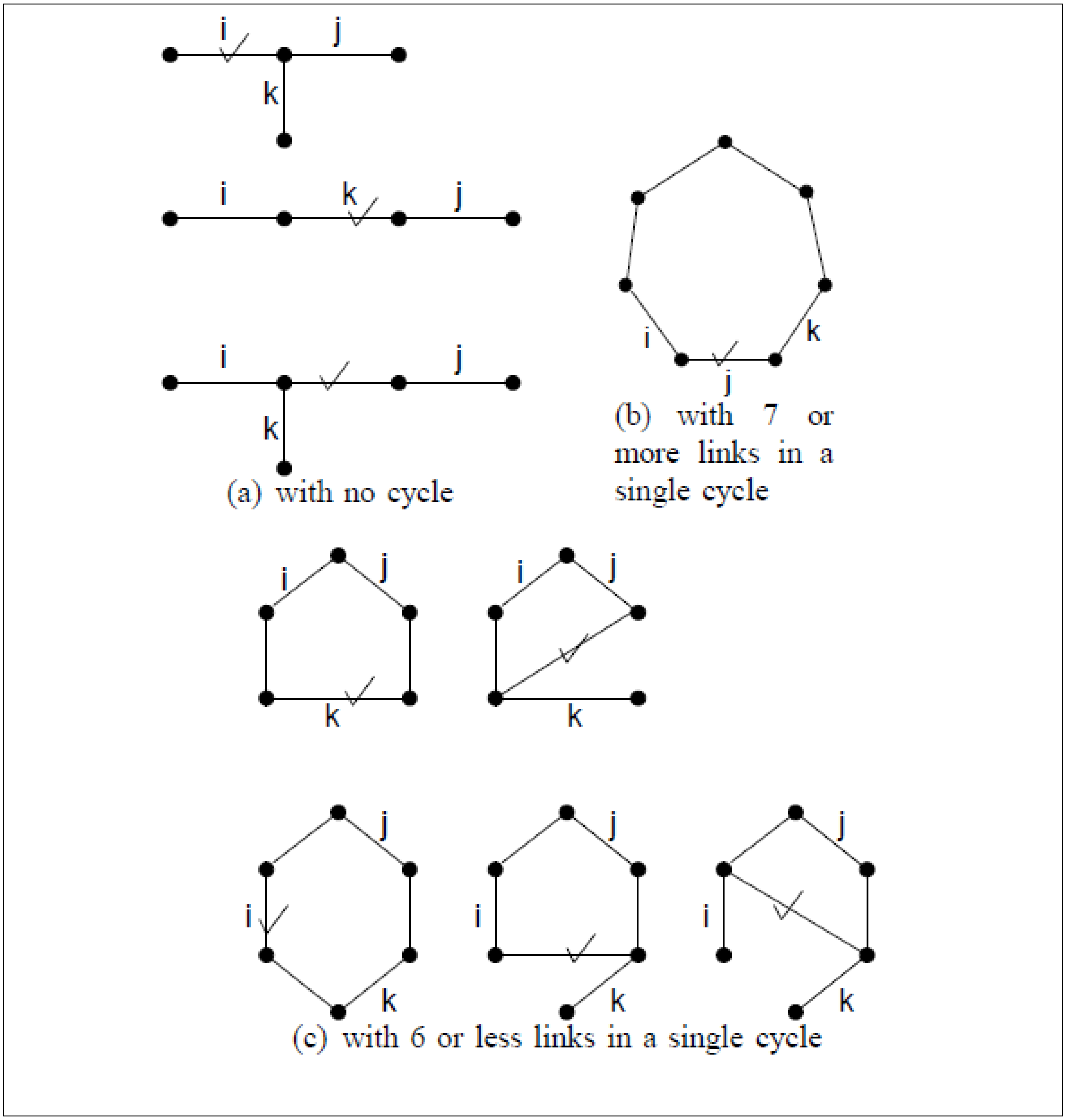

Fig.5.6 For a clique of size n > 3

VI. IMPROVING SENSOR NETWORK LIFETIME THROUGH VII.HIERARCHICAL MULTIHOP CLUSTERING

Maggie Cheng[1], Xuan Gong[1] and Scott C.-H. Huang[2]

[1]Department of Computer Science, Missouri University of Science and Technology
Rolla, Missouri, USA

[2]Department of Computer Science, City University of Hong Kong

Kowloon, Hong Kong

{chengm@mst.edu, xgwt2@mail.mst.edu, shuang@cityu.edu.hk}

ABSTRACT. In this project, we developed an adaptive multi-hop clustering algorithm MaxLife for sensor networks. MaxLife significantly improves sensor network lifetime by balancing energy dissipation and minimizing energy consumption at the same time. The algorithm is compared to Random and MinEnergy algorithms and shows great performance gain. Random is extended from its original design of single hop clustering in [1] to multi-hop clustering, which elects cluster heads with absolute fairness. However, the idea of rotating the role of cluster heads does not work well in a multi-hop environment, because relay nodes can also drain out energy quickly. MinEnergy chooses cluster heads to minimize total energy consumption, which leads to large energy disparity and hurts long-term performance. MaxLife on the other hand, uses global optimization techniques and directly maximizes network lifetime. Simulation results verified that MaxLife achieves the best tradeoff between fairness and energy efficiency, and the clustering topology computed from it has significantly longer lifetime than those from the other two algorithms.

1. INTRODUCTION

A typical sensor network features limited energy supply, limited wireless transmission range, and large amount of data to process. Sensor networks have large potential in habitat monitoring, health care, as well as military applications, which renders them a hot research topic in the past few years. One of the major issues dominating the literature is energy efficiency. Much work has been done to improve energy efficiency with the ultimate goal of having a long lifetime without replacing sensor nodes. Hierarchical routing via cluster head is one of the approaches to improve energy efficiency.

In general, hierarchy improves scalability. For the same reason hierarchical routing is implemented in OSPF, clustering is used in sensor networks. Many operations in sensor networks such as routing and query dissemination can be more efficient if they are confined within the boundary of a cluster. Moreover, clustering avoids direct communication between every single sensor node and the base station (BS), and therefore effectively prolongs the sensor network lifetime.

To achieve the maximum lifetime through clustering, three factors need to be considered: energy cost within a single cluster, called interior cost, energy cost from heads to base station, called exterior cost, and the balance of energy consumption over time.

The three factors have complicated tradeoff relation in terms of their contribution to network lifetime. First, there is a tradeoff between single round minimum energy and the balance of energy consumption over time. Focusing on either one alone will not get the maximum lifetime. Second, In order to minimize the total energy cost of a single round, there is a tradeoff between the interior energy cost and the exterior energy cost — the more cluster heads, the more nodes using long radio range, but the fewer hops from member nodes to cluster heads; the fewer cluster heads, the less energy spent on head-to-BS transmission, but the more energy spent on member-to-head communication.

In previous work, we have seen schemes that focus on mainly one or at most two of the three factors and the results from these schemes are far from being optimal. The maximum lifetime clustering problem is yet to be solved. In this paper, we study the complicated tradeoff relation among multiple factors that affect the sensor network lifetime and propose an adaptive multi-hop clustering algorithm to simultaneously evaluate the role of each factor. The algorithm successfully realizes the best tradeoff among the three factors and outperforms others that do not. It is adaptive in the sense the clustering topology changes over time in order to have the maximum lifetime.

The rest of the paper is organized as follows. In Section 2, we briefly survey the closely related work to ours and point out how our approach distinguish itself from others; In section 3, we formally state the maximum lifetime clustering problem; In section 4, we describe our MaxLife algorithm, its Integer Linear Program formulation and heuristic solution; Following this section are its counterparts that are compared to it in the simulation; In section 5, we show the lifetime results from our approach and other approaches. In section 6, we conclude the paper with directions for future work.

2. RELATED WORK

The most related work to ours is cluster-based hierarchical routing. Li et al. [2] proposed HPAR, a hierarchical power-aware routing protocol that divides the network into clusters. Each cluster/zone is allowed to decide how to route a message hierarchically across the other clusters such that the battery lives of the nodes in the system are maximized. Estrin et al. [3] discussed a hierarchical clustering method with emphasis on localized behavior and the need for asymmetric communication and energy conservation in sensor networks. Jiang et al. [4] proposed CBRP, a cluster based routing protocol for mobile ad-hoc networks. It divides the network nodes into a number of overlapping or disjoint two-hop diameter clusters in a distributed manner. Manjeshwar and Agrawal proposed two hierarchical, energy-efficient routing protocols: TEEN [5] and APTEEN [6] for timecritical applications. Heinzelman et al. [1] proposed LEACH protocol, which was originally designed for single-hop clustering. In LEACH protocol, the duty of being a cluster head is evenly distributed among all sensors in a network. LEACH randomly selects sensor nodes as cluster-heads and rotates this role to evenly distribute the energy load among the sensors. LEACH works for a small network where every node can reach every other node in the network, because there is only one hop between member nodes and their cluster head. In this paper, LEACH has been extended to multiple hop clustering and compared to our scheme. Simulation results show that it does not work well for multi-hop clustering. Lindsey et al. [7] improved LEACH and designed the PEGASIS algorithm. In order to extend network lifetime, nodes only communicate with their closest neighbors and take turns to communicate with the base-station. PEGASIS increases the lifetime of each node by using collaborative techniques, and it allows only local coordination between nodes that are close together so that the bandwidth consumed in communication is reduced.

3. PROBLEM DEFINITION

The purpose of clustering is to find the best way to organize sensor nodes into disjoint groups and to designate a head for each group, which communicates with the base station directly, so that the lifetime of the network is maximized. However, clustering in sensor networks can be done differently under different assumptions about whether data is aggregated, whether transmission power is adjustable, etc. In this work, we adopted the following widely accepted assumptions:

First, we assume a non-head node uses a constant transmission power P1 to route data to its cluster head by using the shortest path routing, and the sole metric is the number of hops; and a cluster head uses a larger constant power P2 to reach the base station. In this way the performance of the algorithm will not be influenced by the position of base station.

Second, how data is aggregated makes significant difference. We assume a cluster head will aggregate data from all its members and itself, and then send only one aggregated packet to the base station; non-head nodes can serve as relay nodes but do not aggregate data. Otherwise the problem degenerates to a trivial case-- (1) If a cluster head only forwards data for its members without data aggregation, there is no need to use the head, because routing to cluster heads only increase the total energy consumption, and the total data transmitted using long radio range is still n units for n nodes. (2) On the other hand, if non-head nodes also aggregate data, then each node sends out one unit of data in each round by using either P_1 or P_2, and in every round, there are $n - 1$ nodes using P_1, and exactly one node using P_2, thus the optimal solution can be easily found by just rotating the role of cluster heads.

Third, we define lifetime as the functional lifetime of the network. In the literature, first-node-die lifetime, or p-percentage-of-nodes-die lifetime have been used. We think the functional lifetime can better depict how long the sensor network can function. We assume the operation of a clustered sensor network is broken into rounds, and in each round, members send to their heads and heads send to BS. During the functional period, each member node ships one unit of data in each round, to a node that can be reached by using transmission power P_1. Since a cluster head consumes more energy than its members, a node will not be eligible for being a cluster head when its remaining energy is below a specified threshold. The functional lifetime is the time period from when the network is deployed until the occurrence of the first case in which data from some node cannot eventually be routed to a base station.

Now we formally introduce the problem:

Definition 3.1: Maximum Lifetime Clustering Given a sensor network of n nodes, each non-head node uses transmission power P_1 to transmit, and each cluster head uses transmission power P_2 to transmit with $P_2 >= P_1$, and data from each source node is routed along the shortest path to a closest cluster head and aggregated at the cluster head, how to form clusters in the sensor network such that the total functional lifetime is the longest?

To address the problem, minimizing total energy cost and minimizing energy disparity both play important roles. Previous works have been focusing on either the fairness or the minimum energy aspect alone. But in fact, neither of the two approaches addresses the maximum lifetime problem directly; as a result neither of them leads to the maximum lifetime. In section 4 we propose a new approach that directly addresses the maximum lifetime problem. For comparison purpose, we describe the two indirect approaches in section 5.

4. A NEW APPROACH: MAXIMIZING LIFETIME DIRECTLY

Ideally for load balancing purpose, a sensor network will have several clustering topologies and each will operate for certain amount of time and together they achieve the maximum lifetime. The functional lifetime is broken into sessions of multiple rounds. During each session, one clustering topology is used, and the topology is adjusted at the beginning of each session. We assume there are K sessions and hence K cluster topologies through its lifetime. For each clustering topology, a group of nodes serve as heads and non-head nodes use the shortest path routing to reach the closest head. We call this algorithm MaxLife algorithm.

4.1 ILP Formulation

To cast the maximum lifetime clustering problem into an Integer Linear Programming (ILP) problem, we use the following notations:

Table 1. Notations of ILP Formulation

N	input, the total number of nodes
K	input, the total number of topologies
f_{vji}	input, $f_{vji}=1$ if node i is on the shortest path from v to j and $i \neq v$, $i \neq j$, otherwise $f_{vji}=0$
E_i	input, initial energy reserve at node i
n_k	variable, number of rounds for the k^{th} topology
X_{ik}	variable, $X_{ik} = 1$ if node i is a head in the k^{th} topology, otherwise $X_{ik} = 0$
e_{ijk}	variable, $e_{ijk} = 1$ if node i is node j’s head in the k^{th} topology and $i \neq j$, otherwise $e_{ijk} = 0$

Recall that a non-head node consumes $P_1 \times 1$ amount of energy to send one unit of data for interior cluster communication, and a head node consumes $P_2 \times 1$ amount of energy to send one unit of data to the base station. The problem is formulated as the following.

Table 2. Mathematical Model of ILP

Maximize: $\sum_{k=1}^{K} n_k$ (1)

Subject to

$$f_{vji} \times e_{jvk} \leq 1 - X_{ik} \qquad \forall v, j, i, k \ (2)$$

$$\sum_{k=1}^{K} n_k P_1 \sum_{j=1}^{N} (e_{jik} + \sum_{v=1}^{N} e_{jvk} f_{vji}) + \sum_{k=1}^{K} n_k P_2 X_{ik} \leq E_i \qquad \forall i \ (3)$$

$$\sum_{i=1}^{N} e_{ijk} = 1 - X_{jk} \qquad \forall j, k \ (4)$$

$$e_{iik} = 0 \qquad \forall i, k \ (5)$$

$$\sum_{i=1}^{N} X_{ik} \geq 1 \qquad \forall k \ (6)$$

$$X_{ik} \in \{0,1\} \qquad \forall i, k \ (7)$$

$$e_{ijk} \in \{0,1\} \qquad \forall i, j, k \ (8)$$

Unfortunately, with the variable n_k, the program is not linear. To remove n_k from the inequality (3), we change the definition of K— we use K as the upper bound of the total number of rounds. Since in each round a node needs to consume at least P_1, so K <= E_i/P_1; and a node can be a head for at most E_i/P_2 rounds, and together that is at most $\sum_{i=1}^{N} E_i / P_2$ rounds. Thus

$$K = \min\left\{\frac{E_i}{P_1}, \sum_{i=1}^{N}\frac{E_i}{P_2}\right\} \qquad \forall i \; (8)$$

And $n_k \in \{0, 1\}$, which means each round is either in operation or totally off. To get rid of the 0-1 variable n_k, we introduce a constant n: $0 < n < 1$, then the inequality (3) is changed to the following:

$$\sum_{k=1}^{K} n\left(P_1 \sum_{j=1}^{N}\left(e_{jik} + \sum_{v=1}^{N} e_{jik} \times f_{vji}\right) + P_2 X_{ik}\right) \le E_i \qquad \forall i \; (9)$$

Inequality (9) suggests if the energy consumption of each round is scaled down by a factor of n, the network can last for K rounds. Fig.6.1 shows the conversion. The real lifetime L in terms of the total number of rounds has the following relation with K: $1 \times L = n \times K$.

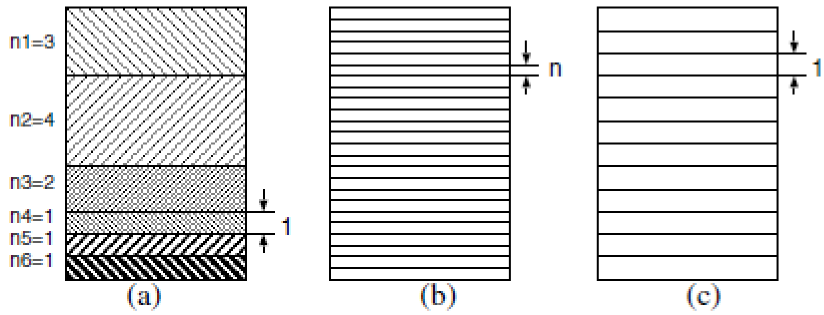

Fig.6.1 (a) Using K as the total number of topologies, n_k is the number of rounds used with the k^{th} topology. Example shows lifetime= $\sum_{k=1}^{K} n_k = 12$; (b) Put on a finer grid of size n — the time for each round is scaled down to n : 0 < n < 1, and there are K such rounds. K is an estimated upper bound of rounds; (c) n × K = 1 × L, stretching the grid size to 1, there will be L rounds. L is the actual lifetime.

So L = nK. To address the rounding error due to the conversion from a continuous problem to a discrete problem, we round it down to the largest integer smaller then nK. Now the objective is to maximize n. We define q = 1/n, the objective function (1) is replaces with:

Minimize: q (10)

The inequality (3) is finally replaced with

$$\sum_{k=1}^{K}\left(P_1\sum_{j=1}^{N}(e_{jik}+\sum_{v=1}^{N}e_{jvk}\times f_{vji})+P_2X_{ik}\right)\le qE_i \qquad \forall i \text{ (11)}$$

The solution to the above ILP only provides a relatively accurate lifetime L a sensor network can last. In the following, we discuss how to get the clustering topology in each round. We introduce new variables ER_i, the remaining energy of node i after L rounds. Note that the second constraint is updated and the objective function becomes to maximize the minimum remaining energy.

Table 3. Mathematical model of Maximize Remaining Energy

Maximize: E_{min}	(12)
Subject to	
$f_{vji} e_{jvk} \le 1 - X_{ik}$	$\forall v, j, i, k$ (13)
$\sum_{k=1}^{L} P_1 \sum_{j=1}^{N} \left(e_{jik} + \sum_{v=1}^{N} e_{jvk} f_{vji} \right) + \sum_{k=1}^{L} P_2 X_{ik} = E_i - ER_i$	$\forall i$ (14)
$\sum_{i=1}^{N} e_{ijk} = 1 - X_{jk}$	$\forall j, k$ (15)
$e_{iik} = 0$	$\forall i, k$ (16)
$\sum_{i=1}^{N} X_{ik} \ge 1$	$\forall k$ (17)
$X_{ik} \in \{0,1\}$	$\forall i, k$ (18)
$e_{ijk} \in \{0,1\}$	$\forall j, k$ (19)
$ER_i \ge E_{\min}$	$\forall i$ (20)

The solution to the ILP problem defined by (12)–(20) provides the cluster topology in each round: if $X_{ik} = 1$, then i is a head in round k; if $e_{ijk} = 1$, then i is j's head in round k. We can reconstruct clusters from this solution.

The ILP problem is NP-complete. In our implementation, we specify a timeout interval to bind the running time. The ILP solver (lp_solve v5.5) sometimes finds the optimal solution, which can be directly used to construct clusters; sometimes it finds a suboptimal integer solution or a real solution and times out. In case the optimal solution is not available, we round up X_{ik} to 1 if $X_{ik} \geq 0.5$, and round down X_{ik} to 0 if $X_{ik} < 0.5$. Then we enlist the non-head nodes with the closest head.

The final lifetime we computed from the above algorithm is the same as L most of the time, occasionally $L-1$ due to the sub-optimality of the solution returned from the ILP solver and rounding errors.

Remark: A similar approach is to maximize the minimum remaining energy after each round, and iterate until the network is no longer functional. This approach involves solving a smaller sized ILP problem multiple times, and the lifetime result is not as good as MaxLife because it does not directly optimize lifetime. Moreover, if we use coarse-grained timeout to obtain the solution, this approach is actually slower because it involves solving more ILP problems.

To use the MaxLife algorithm in practical sensor networks, a Link-State type protocol is needed at the initial stage. The number of messages and time needed is the same as a typical Link-State protocol to get the network topology. After each node has learned the network topology, there is no additional message overhead in running this algorithm, only computational overhead. But this overhead is well paid off by the long lifetime it achieves later. Alternatively, the computation for clustering can be done at a more powerful node such as the base station and broadcast to sensor nodes.

In the following, we describe two indirect approaches that focus only on one aspect of the tradeoff relation. Similar ideas have appeared in the literature for different optimization objectives or in different routing topology. We now apply them to the multi-hop clustered hierarchy and compare their performance with our approach.

5. INDIRECT APPROACHES

5.1 Minimizing Energy Consumption

In this algorithm, both the interior and exterior energy costs are considered, but it focuses on minimizing the energy consumption of a single round.

Let d_{ij} denote the number of hops from node i to node j along the shortest path. We formulate the problem into an ILP problem as follows:

Table 4. Mathematical Model of minimizing energy consumption

Minimize: $$\sum_{j=1}^{N}\sum_{i=1}^{N} X_{ij} d_{ij} P_1 + \sum_{i=1}^{N} X_{ii} P_2 \qquad (21)$$

Subject to

$$X_{ii} \geq X_{ij} \qquad \forall i, j \ (22)$$

$$\sum_{i=1}^{N} X_{ij} = 1 \qquad \forall j (23)$$

$$\sum_{i=1}^{N} X_{ii} \geq 1 \qquad \forall i (24)$$

$$X_{ij} \in \{0,1\} \qquad \forall i, j (25)$$

The only variables are X_{ij} : $X_{ij} = 1$ if node i is node j’s head, otherwise $X_{ij} = 0$. If node i is a head then $X_{ii} = 1$. Other notations are used the same way as in section 4.

To solve this ILP problem is NP-complete. Sometimes the ILP solver returns with the optimal solution, in which case the solution itself suggests the clustering topology;

sometimes the ILP solver fails to find the optimal integer solution. We accept suboptimal solution and use the same technique as in section 4 to round a real-numbered solution to an integer solution. From the rounded integer solution we can construct the cluster topology as follows: if $X_{ii} = 1$, then node i is a head; then non-head nodes use the shortest path routing to enlist with the closest head.

5.2 Rotating the Role of Cluster Heads

In this algorithm, each node should have the same opportunity to be a cluster head. The idea is to choose a probability P of being a head upfront, and this probability is consistently used by all nodes. In every round, each node chooses a random number and feeds it in a predefined threshold function to decide whether it will be a head. The nodes that are not selected as heads will use the shortest path routing to associate with the closest head. The duty of being a cluster head is perfectly rotated among all nodes, so in the long run, every node will act as a head for the same number of times. The threshold function is the key to this algorithm. We adopt the threshold function from [1] and apply it in multiple hop clustering.

The pseudo code of the algorithm is skipped here due to space limit. Because it depends on random numbers at each round, we call it Random algorithm.

This algorithm provides each node equal chance to be a head in the long term; however, the randomness of the algorithm does not provide optimality of energy consumption in each single round. As we will see in the simulation, sometimes the minimum remaining energy $E_{min} > P_2$, but the Random algorithm fails to find the optimal topology that utilizes the remaining energy of all nodes to make one more round. Therefore the total lifetime is not maximized.

The idea of rotating the role of cluster heads works well with a single hop clustering hierarchy as in [1], where each node either uses P1 or P2 energy in each round, so the heads always drain out faster than nonhead nodes; but in a multi-hop clustering hierarchy, a non-head relay node could drain out energy very fast, even faster than the cluster head if P1 is not negligible.

6. SIMULATION

In this section, we compare the proposed MaxLife algorithm with the Random algorithm and the MinEnergy algorithm in section 5.

In all experiments, initial energy E is set to 1 unit for all nodes; P_1 and P_2 are also normalized to E. 20 nodes are randomly deployed on a 100 $\times$ 100 square region. The transmission range of a node changes with the transmission power P_1. After the initial deployment, network connectivity is checked and only connected networks are selected for study. Network lifetime is counted as the number of rounds. We use functional lifetime in the simulation, which can be interpreted as follows: suppose after L rounds, the minimum remaining energy is still ≥ 0, but after L+1 rounds, the minimum remaining energy becomes < 0, then the lifetime is L rounds.

In the first setting, we use a fixed value $P_2 = 0.4$, and vary the value of P_1. When P_2/P_1 goes from 1 to 64, we observed that the lifetime L achieved by MaxLife is significantly longer than those by the other two schemes as shown in Fig.6.2.(a). In terms of the adaptability, MinEnergy is the worst in this experiment, because the relative large value of P_2, head node(s) can only last for two rounds, and the topology does not change from round to round. Random shows better adaptability than MinEnergy, but the balance of energy consumption is achieved by the random selection of heads, not through an optimized design, so MaxLife still beats Random. From Fig.6.2(b) we can see MaxLife makes the best use of available energy as it ends with the least remaining energy. Because we used functional lifetime, so after L rounds, each node still has non-zero remaining energy. Random tends to terminate with more remaining energy. The reason is even though there is still energy to make another round, but due to the random nature of the algorithm, the algorithm fails to find a topology that can make more rounds.

In the second setting, we still use the normalized E = 1, but with a fixed P_1 = 0.01, and we vary P_2 to have P_2/P_1 going from 1 to 64. We compare lifetime and the minimum remaining energy after lifetime. MaxLife shows the best adaptability again, achieving the longest lifetime and ending with the least non-negative remaining energy, as shown in Fig.6.3.

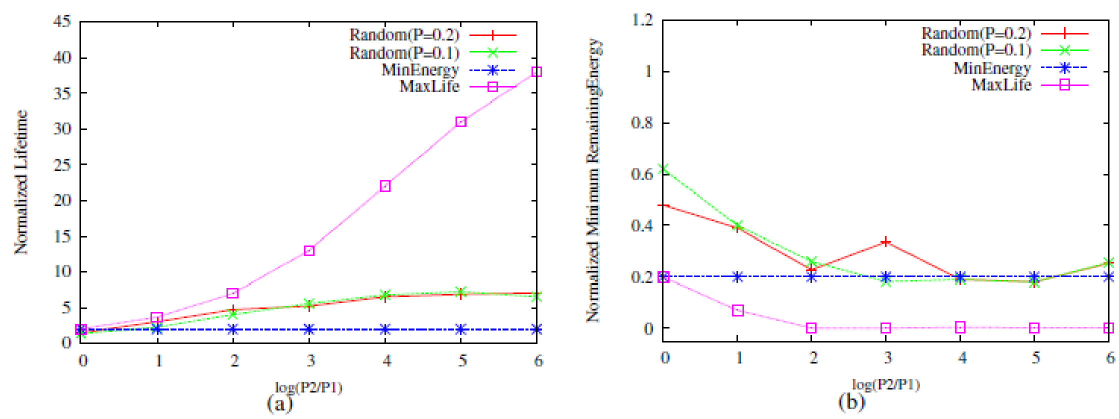

Fig.6.2 With initial energy reserve E = 1, BS-to-head P_2 = 0.4, node-to-head P_1 varies. (a) Functional lifetime (b) Remaining Energy after lifetime

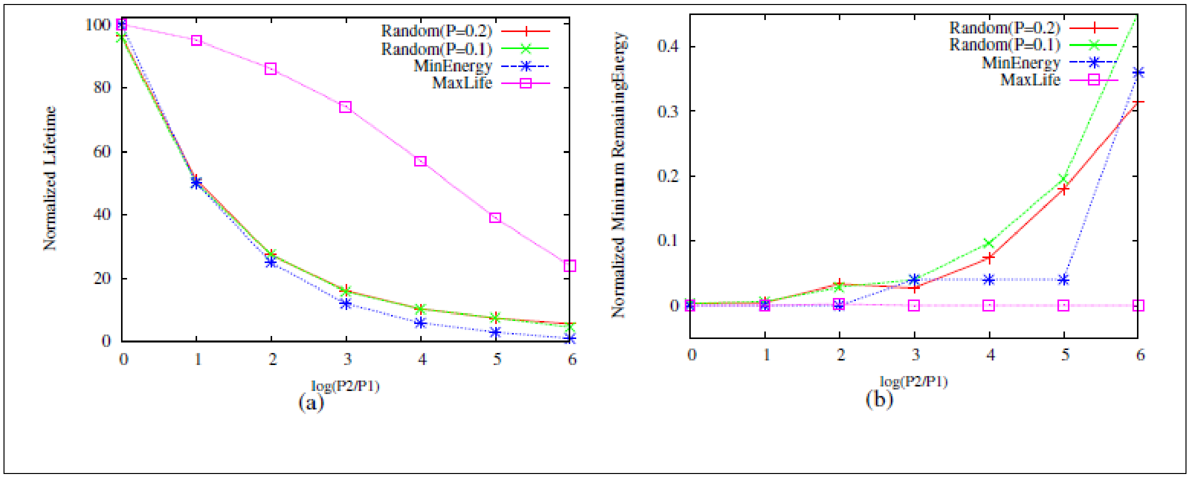

Fig.6.3 With initial energy reserve E = 1, node-to-head P_1 = 0.01, head-to-BS P_2 varies. (a) Functional lifetime (b) Remaining Energy after lifetime

7. CONCLUSION

In this paper, we addressed the problem of the maximum lifetime clustering in a multi-hop environment, analyzed the tradeoff relation among the major factors that contribute to the lifetime of sensor networks, and proposed a new algorithm MaxLife, in which we formulated the maximum lifetime clustering problem as an Integer Linear Program and provided a heuristic to select cluster heads. The idea is to break network lifetime into sessions of multiple rounds, and the clustering topology is adjusted at the beginning of each session to ensure energy efficiency and the balance of energy dissipation. The simulation results show that this algorithm performs significantly better than those that only focus on optimizing one aspect of the tradeoff relation.

8. REFERENCES

[1] Wendi Rabiner Heinzelman, Anantha Chandrakasan, and Hari alakrishnan, “Energy-efficient communication protocol for wireless microsensor networks”, in HICSS, 2000.

[2] Q. Li, J. Aslam, and D. Rus, “Hierarchical power-aware routing in sensor networks”, in DIMACS Workshop on Pervasive Networking, 2001.

[3] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, “Next century challenges: Scalable coordination inwireless networks”, in MobiCom’99: 5th IEEE/ACM Int’l Conf on Mobile Computing and Networking, 1999, pp. 263–270.

[4] M. Jiang, J. Li, and Y. C. Tay, “Cluster based routing protocol”, 1999, Internet Draft.

[5] A. Manjeshwar and D. P. Agarwal, “Teen: a routing protocol for enhanced efficiency in wireless sensor networks”, in 1st Int’l Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing, 2001.

[6] A. Manjeshwar and D. P. Agarwal, “Apteen: A hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks”, in IPDPS’02: Parallel and Distributed Processing Symposium, 2002, pp. 195–202.

[7] S. Lindsey and C. Raghavendra, “Pegasis: Power-efficient gathering in sensor information systems”, in IEEE Aerospace Conference, 2002, vol. 3, pp. 1125–1130.

CONCLUSION

This dissertation has provided generic mathematical models for several optimization problems in energy and bandwidth-constrained sensor networks. It sufficiently considered the impact of wireless interference on network performance including throughput and delay. The basic mathematical optimization models can be easily extended to address heterogeneous sensor networks where nodes have different initial energy or different transmission power levels, and to work with various data aggregation schemes. Four important problems have been further investigated: (1) “How to route data packets and control link transmission rates in order to provide users with communication efficiency and fairness?” This problem is addressed by using a linear programming model, in which wireless interference is effectively accounted for. The linear optimization model presented in this thesis tells what the optimal operation point is in terms of applied traffic load, and how to find the routing and link rates to improve network efficiency. (2) “How to achieve the maximum throughput in sensor networks through cross-layer optimization?” Transmission power control is used to decide the link topology and then use joint routing and link rate control to decide the maximum achievable throughput on the topology. The optimization models and efficient algorithms are proposed for power control as well as for joint routing and rate control. (3) “How to achieve the minimum end-to-end delay given a multi-hop wireless network with multiple sources and destinations?” A linear programming-based link scheduling scheme is proposed. The optimization model is useful for feasibility analysis given a set of QoS constraints, and it is also useful for predicting the achievable performance of the network and improving delay when routing information is given. (4) “How to achieve maximum lifetime under energy and bandwidth constraints?” A sufficient condition is presented that makes a routing solution feasible and the mathematical optimization model based on this sufficient condition is proposed to maximize the network lifetime using uniform transmission power. Then this optimization model is extended to handle non-uniform transmission power and routing with data aggregation. This dissertation also shows that this joint optimization can guarantee that there exists a conflict-free time slot assignment to support the given routing solution.

REFERENCES

[1] V. Shah-Mansouri and V.Wong, "Lifetime-resource tradeoff for multicast traffic in wireless sensor networks," IEEE Transactions on Wireless Commnunications, (9)6: 1924-1934, June,2010

[2] M.Dhanaraj and C. Siva Ram Murthy, "On Achieving Maximum Network LifeTime Through Optimal Placement of Cluster heads in Wireless Sensor Network," IEEE International Conference on Commnuications, 3142-3147, June,2007

[3] M.Esseghir and H.Perros, "Optimizing the wireless sensor network lifetime under reliability constraint," 1st IFIP Wirelss Days, 1-5, Nov.2008

[4] Weifa Liang, "Minimizing energy and maximizing network lifetime multicasting in wireless ad hoc networks", 2005 IEEE International Conference on Commnuications, 3375-3379 Vol 5, May,2005

[5] J.Luo, A. Iyer and C.Rosenberg, "Throughput-Lifetime Tradeoffs in Multihop Wireless Networks under an SINR-Based Interference Model," IEEE Transactions on Mobile Computing, (99):1-1, 2010

[6] Joseph C.Dagher, Michael W. Marcellin, and Mark A. Neifeld, "A Theory for Maximizing the Lifetime of Sensor Networks," IEEE Transactions on Commnuications, 55(2) February, 2007

[7] SeongHwan Cho and Kee-Eung Kim, "Variable bandwidth allocation scheme for energy efficient wireless sensor network," ICC 2005, 3314-3318, May, 2005

[8] Jing-Rong Hsieh, Tsern-Huei Lee and Yaw-Wen Kuo, "Energy-efficient muti-polling scheme for wireless LANS", IEEE Transactions on Wireless Commnuications, 8(3): 1532-1541, March,2009

[9] Song Guo and V.M Leung, "A Distributed Min-Max Algorithm for Maximum-Lifetime Multicast in Resource limited Wireless Ad Hoc Networks," 2006 3rd Annual IEEE Commnuications Society on Sensor and Ad Hoc Commnuications and Networks, 673-677, September 28th, 2006

[10] Yu-Chee Tseng, Yueh-Feng Li and Yu-Chia Chang, "On route lifetime in multihop mobile ad hoc networks," IEEE Transactions on Mobile Computing, 2(4): 366-376, December, 2003

[11] Hang Su and Xi Zhang, "Throughput-Gain Analysis of Network Coding in Multi-Channel Multi-Radio Wireless Networks," IEEE International Conference on Communication, 1-5, June, 2009

[12] D Umehara, T Hirano, S Denno and M Morikura, "Throughput Analysis of Wireless Relay Slotted ALOHA Systems with Network Coding," 2008 IEEE GLOBECOM, 1-5 November, 2008

[13] A Keshavarz-Haddadt and R. Riedi, "Bounds on the Benefit of Network Coding: Throughput and Energy Saving in Wireless Networks" , The 27th Conference on Computer Communications, 376-384, April, 2008

[14] G Venkatesh and Kuang-Ching Wang, "Estimation of maximum achievable end-to- end throughput in IEEE 802.11 based wireless mesh networks", IEEE 34th Conference on Local Computer Networks. 1110-1117, October, 2009

[15] L.L Dai and V.W.S Chan, "Energy and throughput tradeoff in wireless networks with processing energy considerations," IEEE GLOBECOM'05, 2522-2527, December, 2005

[16] J Luo, A Iyer and C Rosenberg, "Throghput-Lifetime Tradeoffs in Multihop Wireless Networks under an SINR-Based Interfernece Model," IEEE Transactions on Mobile Computing, (99)1-1, September, 2010

[17] Jun Wang and Wei Gao, "Scheduling Algorithm of Great Tolerance Wireless Communication Delay for Wireless Network Control System," Wireless Communication, Networking and Mobile Computing, 2009. WiCom'09.5th International Conference on Beijing, Sept. 2009

[18] Pan Li, Chi Zhang and Yuguang Fang,"Capacity and delay of hybrid wireless broadband access networks," Selected Areas in Communications, 27(2):117-225,February 2009

[19] P. Gupta and P. Kumar. The capacity of wireless networks. IEEE Trans. Inform. Theory, 46(2):388–404, March 2000.

[20] O. Dousse, M. Franceschetti, and P. Thiran. On the throughput scaling of wireless relay networks. IEEE Trans. Inform. Theory (Joint special issue with IEEE/ACM Trans. Networking), 52(6):2756–2761, 2006.

[21] A. EI Gamal, J. Mammen, B. Prabhakar and D. Shah, "Throughput-delay scaling in wireless networks wirh constant-size packets," Information Theory, 2005, 1329-1333, Sept.2005

[22] Zeyu Zheng, Jianping Wang and Jin Wang, "Interference Aware and Delay Bounded Routing in Hybrid Wireless-Optical Access Network" Commnuications and Mobile Computing, 2009, 461-465, January 2009

[23] S.Sarkar, Hong-Hsu Yen, S. Dixit and B.Mukherjee, "DARA: Delay-Aware Routing Algorithm in a Hybrid Wireless-Optical Broadband Access Network," ICC'07, 2480- 2484,June 2007

[24] C.S. Taggart, Y.Viniotis and M.L.Sichitiu, "Minimizing average network delay for ultrawideband wireless networks," wireless communication and Networking Conference, 2276-2280,March 2005

[25] M.F. Munir, A.A Kherani, F.Filali, "Distributed algorithm for minimizing delay in multi-hop wireless sensor networks," Modeling and Optimization in Mobile, Ad Hoc, and wireless Networks,1-9, June 23-27, 2009

[26] Ping Chuang Ng, D.J. Edwards and Soung Chang Liew, "Colouring Link-Directional Interference Graphs in Wireless Ad Hoc Networks," IEEE GLOBECOM'07, 859-863, November, 2007

[27] Kun-Da Wu and Wanjiun Liao, "Impact of Interference and Medium Access Control on Flow Allocation in Multi-Hop Wireless Networks," ICC'07, 5270-5275, June,2007

[28] Hongli Xu, Liusheng Huang, Wang Liu, Ben Xu and Mingjun Xiao, "Topology control for minimal path interference in wireless sensor networks," ISCC'08, 988-993, July, 2008

[29] Xiaolin Cheng, P.Mohapatra, Sung-Ju Lee and Banerjee S, "MARIA: Interference- Aware Admission Control and Qos Routing in wireless Mesh Networks", ICC'08, 2865-2870, May, 2008

[30] Ka-Hung Hui, Wing-Cheong Lau and On-Ching Yue, "Characterizing and Exploiting Partial Interferenece in Wireless Mesh Networks," ICC'07, 102-108, June,2007

VITA

Xuan Gong was born in Nan Chang, P.R.China on June 19th, 1984. In July,2005, he received his B.S in Computer Science from Nan Chang University, P.R.China. Xuan joined the former University of Missouri-Rolla, currently named Missouri University of Science and Technology in order to get his Masters degree in Computer Science, 2006. On May, 2008, Xuan received his M.S. in Computer Science and enrolled at Missouri University of Science and Technology to get his Phd in Computer Science June, 2008. His areas of interests include Wireless network, Mobile computing and combinatorial optimization algorithms. He will receive his Phd degree on May, 2011.