A Comprehensive Energy Conservation Solution for Mobile Ad Hoc Networks

John A. Stine and Gustavo de Veciana
Department of Electrical & Computer Engineering, The University of Texas at Austin, Austin, TX 78712

Abstract- Multiple energy conserving approaches have been proposed for wireless networks that are exploited by the link layer and network layer protocols. Unfortunately, integrating these approaches in ad hoc networks is difficult. Due to the temporally random nature of access protocols, methods based on entering low energy states cause severe degradation of network capacity and also degrade the performance of routing protocols. Meanwhile, methods used by routing protocols that give preference to shorter links or attempt to balance load to prolong the longevity of the plurality of nodes require commitment to one or the other of these metrics without regard to link layer approaches. In this paper, we show that through the integrated use of our access and routing protocols, Synchronous Collision Resolution (SCR)[1] and Node State Routing (NSR)[1], that these types of energy conservation mechanisms can be managed simultaneously. We conclude with a simple simulation of the integrated use of these protocols. The simulations demonstrate that these protocols reduce the rate of energy consumption by the network but that in determining their effectiveness, the end-to-end throughput of the network must be considered.

I. Introduction

Ad hoc networks have been proposed as a solution to wireless networking where nodes are mobile, the range of their mobility exceeds the transmission range of any single transceiver, and there is no existing network infrastructure. Mobile nodes in these networks frequently rely on batteries for energy and therefore have a finite lifetime. Conserving energy is important to extending the lifetime of both individual nodes and the network. This is especially difficult in ad hoc networks since energy conserving actions must be made in a distributed manner. In fact, the continuous participation of the mobile nodes to create a fabric of connectivity is critical to the overall performance of the network. Typically, this results in a choice of either operating at peak performance at the expense of a shortly lived network or choosing sub optimum performance for network longevity. Additionally, most energy conserving protocols focus on the implementation of a single energy conserving approach. In this paper we demonstrate that through the novel features of our access and routing protocols, Synchronous Collision Resolution (SCR)[1] and Node State Routing (NSR)[1], that we can manage the use of most known energy conserving approaches and without the problem of sacrificing performance for longevity. The energy conservation mechanisms of the MAC layer are fully integrated into the algorithms of the routing protocol. Meanwhile, the routing protocol independently implements the conservation mechanisms that are managed exclusively at its level.

We start this paper in Section II with a review of the mechanisms that have been proposed for protocols to use to conserve energy. Then in Sections III and IV we describe how these same mechanisms are implemented in SCR and NSR. In Section V we present the results of a simulation that integrated the use of access level and routing level conservation mechanisms. Finally, we conclude the paper with Section VI.

[1] Patent Pending

RADIO	TRANSMIT	RECEIVE	STAND-BY/ DOZE
WaveLAN Turbo 11 Mb Card [1]	285mA	185mA	9mA
RoamAbout 915 MHz DS/ISA [2]	600mA	300mA	36mA
RoamAbout 2.4 GHz DS/ISA [2]	365mA	315mA	30mA
Nokia C020/C021 LAN Card [3]	1.7W	1.3W	0.2/0.1W
Aironet PC4800B [4]	350mA	250mA	<10mA

Table 1 Digital Radio Power States

II. Energy Conservation Mechanisms

Protocols may use four sets of mechanisms to reduce energy consumption: 1. Help nodes enter low energy states. 2. Choose routes that consume the least energy. 3. Selectively use nodes based on their energy status. 4. Reduce overhead.

The potential for conserving energy using low energy states is made most apparent by the relative energy consumption of transceivers at different states. Table 1 presents the rates of consumption for some commercial transceivers. As seen, the rate of consumption in the receive state is more than 50% of that consumed in transmitting. We note that the default state of nodes is receiving since signal processing is required to detect and to synchronize to an incoming signal. Entering a low energy state requires the node to cease sensing the channel and to stop participating in the network's activities. The objective of type 1 energy conserving protocols is to assist nodes that are not participating in data exchanges to enter a low energy state without degrading overall performance of the network. Proposed methods for managing nodes entering the doze state may be one of two kinds. In the first, nodes doze and then wakeup on a periodic basis according to network wide parameters. The 802.11 standard [5] provides this kind of mechanism. The second requires the node desiring to doze to specifically coordinate a dozing cycle with another supporting node that agrees to act as a surrogate destination for the dozing node's traffic while it is dozing. The ETSI HIPERLAN standard [6] uses this approach.

In both the 802.11 and HIPERLAN protocols, the decision to doze is initiated by the individual nodes desiring to conserve energy. In the ad hoc version of an 802.11 network, the node that first forms the network decides whether it permits energy conservation by establishing an "ATIM Period." A node that desires to conserve energy may doze so long as it wakes each ATIM Period to listen for ad hoc traffic indication messages (ATIM). ATIMs are transmitted during a short window at the beginning of each ATIM period, called an ATIM Window. If the node wakes and hears an ATIM directed to itself, it acknowledges the ATIM and remains awake for the rest of the ATIM period prepared to receive traffic. If it receives no ATIM directed to itself, the node returns to the doze state at the conclusion of the ATIM window. Note that there is no method for a node's intent to doze to be disseminated. Other nodes assume this state after failing to transfer data through regular contention.

The energy conserving mechanism in HIPERLAN requires a node desiring to doze, a "p-saver," to coordinate with another to serve as its surrogate, a "p-supporter".[2] As part of this coordination the two nodes agree to a period at which the p-saver will awaken to receive unicast messages and a period at which the p-supporter will transmit multicast messages. The p-supporter node collects transmissions intended for the p-saver and then attempts to forward them to the p-saver during the coordinated transmission periods.

The 802.11 mechanism was studied in [7] and an ATIM "window to period" ratio of 1:4 was recommended. The authors provided the intuition that as ATIM periods become longer more nodes need to transmit ATIMs and, in turn, these nodes remain awake during the ATIM period. Alternatively, as the ATIM window becomes longer, more ATIMs are transmitted also resulting in more nodes remaining awake during the ATIM period and, in turn, reduced throughput on account of a greater number of nodes contending with each other.

We are aware of no study of the HIPERLAN energy conserving mechanisms. Such a study would be difficult since it would be scenario dependent. Intuitively, HIPERLAN's approach is disconcerting since it does not make the dozing states known throughout the network. Node in ad hoc networks depend on each other to route and distribute packets to each other. The arrangement of having a surrogate node collect data for another may defeat many routing protocols. The p-supporter node may not be in a location to collect data from a relaying node in the opposite direction to the p-saver. Additionally, the p-saver may be a critical next hop in a route.

The critical deficiency of both the 802.11 and HIPERLAN techniques is that they do not account for the repercussions of a single node's decision to enter the doze state. These repercussions are more congestion as nodes attempt to send traffic to nodes that are dozing and complications for other protocols higher in the stack such as routing. To minimize these adverse effects, access protocols must be able to make dozing more predictable and to integrate the occurrence of dozing with the activities of the routing protocol.

Routing protocols conserve energy by identifying routes based on energy consumption. From the protocol perspective, energy is consumed in transmission and in reception. The energy consumed in transmission can vary based on the range between a source and its destination.[3] The energy consumed in reception is constant. Due to the power law relation of energy consumed to the distance transmitted a route with more shorter hops may consume less energy than a route with fewer longer hops. The log-distance path loss model illustrates the energy consumption dependence on distance.

$$P_t(d) = Kd^n \quad (1)$$

$P_t(d)$ is the power required to successfully transmit a packet to a destination separated from the transmitter by the distance d. K is a constant and the variable n is referred to as the path loss exponent. Typical path loss exponents provided by [8] range from 1.6 for indoors line of sight to as high as 6 when obstructed in a building. A path loss exponent of 4 is used in most literature concerning ad hoc networks. With this exponent, a route that used two equidistant hops to a destination could require as little as 1/8th the transmission energy of the direct one hop route.

A low energy route uses a series of hops that consume the least energy. A simple method to select the next hop is described in [9]. This paper demonstrates that all traffic from a source should be forwarded through a subset of the neighbors that surround it. This subset includes all nodes for which a single hop exchange is the most energy efficient method of delivering a packet. It demonstrates that about these nodes a relay boundary can be drawn that defines the relay region to which each of these nodes could be used as an energy conserving intermediate hop. It then shows that the combination of these relay boundaries from these single hop neighbors forms an enclosure of the source. All next hop neighbors for low energy routing are included in this enclosure.

A node j is an energy conserving next hop to node k from node i if the following inequality is true.

$$d_{ik}^n > d_{ij}^n + d_{jk}^n + c\,. \quad (2)$$

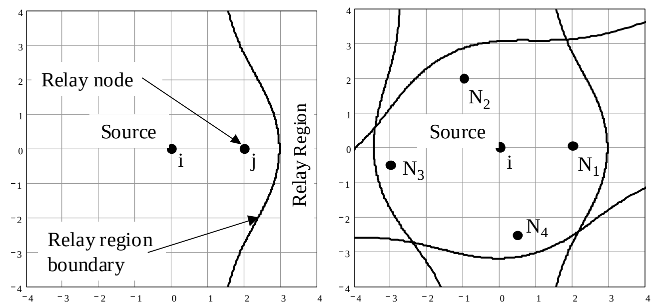
The variables d_{ik}, d_{ij} and d_{jk} are the distances between nodes i and k, i and j, and j and k respectively, n is the power law exponent, and c accounts for the energy consumed by a node receiving a packet. In Figure 1a we illustrate a possible orientation of the nodes i and j and graph the boundary across which node k must be located for the inequality in (2) to be true. Then in Figure 1b, we illustrate an enclosure formed by 4 nodes that surround a source node i.

The application of this approach using standard link state and distance vector protocols requires the development of an energy consumption metric for links. Since path loss exponents can vary they must be measured. Also, since propagation conditions can change quickly, it is very risky to commit to a minimum energy transmission on account of possible failure. We are aware of no application of this metric to a routing protocol.

Routing protocols may prolong the lifetime of a network by preferring the use of nodes that are not energy constrained and by balancing the use of nodes that are energy constrained. One approach to solve this problem is Power-Aware Routing [10]. It uses an energy cost metric for links that is obtained by weighting the energy consumption on the path by the energy reserve on each node of the path. This has a load balancing characteristic that steers traffic away from low energy nodes. The conclusions in [10] state that the effectiveness of this approach is dependent on the load. This metric is most effective in large moderately loaded networks. It had a negligible effect in networks with low or high loads. The conclusions of [11] corroborate this observation and notes that routing protocols that use this metric tend to prefer shorter routes that load intermediate nodes with relay

[2] We assume that the nodes that serve as p-supporters are not energy constrained and do not need to conserve energy themselves.

[3] We assume that all nodes know each other's location and that a source can adjust its transmission power to the minimum required for a successful exchange with a destination.



a. Relay boundary of node *j* for source *i* b. Enclosure of node *i* formed by nodes N_1, N_2, N_3, and N_4

Figure 1. Relay boundaries for energy conserving routing

traffic. Although, power aware routing may increase the time until the first failure, the average lifetime of the nodes decreases.

The energy conservation benefits of reducing overhead are obvious but it is rather difficult to quantify. Goodput and overhead are correlated and energy consumed per goodput is the more revealing energy consumption statistic. In most cases, however, differences in goodput performance are more significant in ranking the energy consumption of different protocols as we reveal in our results later.

The challenge of implementing energy conservation mechanisms is their interlayer dependence. The success of a mechanism based on a MAC mechanism can greatly affect the routing protocol (e.g. dozing can remove potential routes) and vice versa (e.g. using shorter hops can increase congestion and preclude dozing). Energy conservation mechanisms must be integrated across layers. Such integration can be achieved only if the dozing methods are made known to the routing protocol and if the routing protocol does not cause congestion. Our MAC and routing protocols, SCR and NSR, achieve both of these. SCR makes dozing very predictable and since it is a spatial protocol as opposed to temporally random protocol it benefits from routing choices that choose shorter hops. NSR is based on the dissemination of node states. The dissemination of dozing states and periods is easily included in the state information. These states can be considered in creating a metric for route calculations. These protocols and their energy conservation features are described below.

III. SYNCHRONOUS COLLISION RESOLUTION

Figure 2 illustrates the organization and operation of SCR. SCR requires all nodes with packets to send to contend simultaneously and synchronously. Then SCR uses a signaling protocol similar to that used by HIPERLAN followed by an RTS-CTS handshake similar to that used in 802.11 protocol to resolve the contentions. Resolving collision in a synchronous manner using signaling provides several benefits. The signaling itself allows the contending nodes to fairly resolve a set of dispersed nodes that can transmit simultaneously. In essence, after the signaling, the remaining nodes constitute a random cellular-like network. The RTS-CTS handshake that follows insures that there are no hidden node collisions during the data transmission. A comprehensive description of the

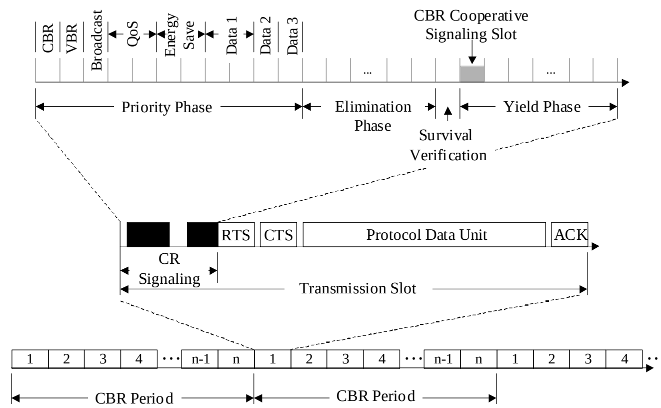


Figure 2. The Synchronous Collision Resolution Protocol

protocol and its many other benefits can be found in [12]. The predictability of when contentions occur and the ability to use signaling to identify the types of traffic that are present make SCR perfectly suited to support the coordination of dozing states. The signaling scheme consists of three signaling phases and two access signals. The first signal starts at some point in the first phase and ends in the second. The second access signal starts at some time within the third phase and ends at the phase's end when a node starts to transmit a packet. A node wins the contention by being among the first to start transmitting in the first phase, among the last to stop transmitting in the second phase and the first to start transmitting in the third phase. Nodes that recognize that they have lost the contention in any one of the phases will defer from attempting to gain access. All nodes at the conclusion of the signaling will know which types of packets are being transmitted since they will know which priority signaling slot was used to gain access. Figure 2 illustrates the types of packets that are differentiated in the priority phase.

In our previous work on energy conserving protocols, [13], we explored energy conservation in wireless networks that use a central controller. This work argues and provides evidence that the most significant characteristic of an energy conserving protocol is its ability to promptly assist nodes not participating in data exchanges to enter the doze state. The key feature of wireless protocols that enables nodes to promptly enter the doze state is their ability to schedule the dissemination of network state information when energy conserving nodes first wake up from dozing periods. We are aware of no distributed access protocol that achieves this goal. SCR meets these requirements since contentions are synchronous and take a finite amount of time. Nodes can wake-up prior to the contention signaling and then immediately return to the doze state after the contention if they will not participate in a data exchange. Signaling not only identifies which nodes win a contention but also whether dozing nodes need to remain awake.

The default energy conservation mode of SCR is for nodes to doze on a slot-by-slot basis. Nodes wake prior to each slot, listen to the signaling and the RTS-CTS exchanges and can enter a low energy state as soon as they determine they are not participating in the following data exchanges. The effectiveness of this technique is dependent on the transition times required to enter low energy states. 802.11 transceivers can transition into a doze state on the order of 5 µsec but then take upwards of 200 µsec to return to a receiving state. The latter transition time limits the usefulness of the doze state in this mode. This transition time corresponds to the time to send 275 bytes on a 11 Mbps channel. Other faster transitioning low energy states may be provided to take advantage of this mechanism. An empirical study of the operation of a WaveLAN card, [14], demonstrated that such a fast transitioning low energy state occurred during the process of dropping packets. Dropping packets consumed less than 80% of the energy that was consumed when the transceiver was in the receive state waiting for a contention. The availability of this mechanism may motivate the design of fast transitioning low energy states in future transceivers, especially since using it has no effect on access performance.

SCR provides two additional dozing modes. The first, which we call extended dozing, is used in low load networks and is similar to the 802.11 scheme. Nodes doze and wake on a periodic basis according to a network wide schedule. Nodes enter this mode when they identify a no load network, i.e. no nodes contend in a transmission slot. They remain in the doze state until the specified slot when all nodes are required to wake up. They wake-up and remain awake from that slot on returning to the doze state only after a slot where no nodes contend. In lightly loaded networks this method allows all nodes to doze the entire dozing period except for the brief signaling portion of the first transmission slot. Such low load conditions are not expected to be the norm so a third mechanism is made available. This mechanism is modeled after the HIPERLAN scheme and we call it coordinated dozing. Here we require nodes to coordinate a dozing schedule with a neighboring node. As in HIPERLAN these p-supporter nodes collect packets for the p-saver nodes. The p-supporter nodes and other neighboring nodes[4] attempt to transmit data to the p-saver node when it wakes up. To enhance the exchange of data to these p-saver nodes, the p-supporter nodes use the energy conservation slots of the priority phase to gain access. The use of these slots or higher priority slots for gaining access is an indication to the dozing nodes that they should remain awake. These energy conserving nodes then use the default energy conserving state until the energy conservation and higher priority slots are no longer used. At that time they return to using their original dozing schedule.

IV. Node State Routing

The NSR protocol uses nodal as opposed to link status to build routes. There are two routing constructs for which state are disseminated, a node and a wormhole. The node construct is modeled as a point in space and is assumed to have connectivity with other nodes using wireless links. We do not expect wireless networks to be connected entirely by wireless links. In many cases nodes may be connected using a dedicated link such as a cable. To use these links within the node state routing protocol

ADD	LOC	PE	IN	PD	RF	ES	SS
CLS	V	D	DS	PS	DO	DP	TS

a. Node state information

ADD	LOC1	PE1	IN	SS1	CLS1	V	D
LOC2	PE2	IN	SS2	CLS2	V	D	TS
COST							

b. Wormhole state information

ADD	Address
CLS	Class
D	Direction
DO	Dozing Offset
DP	Dozing Period
DS	Dozing State
ES	Energy State
IN	Infrastructure
LOC	Location
PD	Preferred Direction
PE	Path Loss Exponent
PS	P-Supporter Address
RF	Receive Fraction
SS	Spread Spectrum Code
TS	Timestamp
V	Velocity

Figure 3. State information provided for nodes and wormholes

we define a second routing construct called a wormhole. The wormhole gets its name from popular science fiction where it is conceptualized as an accelerator tube between two points in space that catapults whatever goes into it to the distant end using minimum energy in minimum time. Similarly, we define our wormhole construct as a directed path between two points in the network across which packets traverse with minimum energy. The basic algorithm used to select which routing constructs to use in a route considers the cost of sending a packet to a construct, the cost of using the construct, and the cost of sending the packet from the construct. These costs are derived from the states of the nodes and the wormholes. Figure 3 lists the proposed states that are disseminated for each construct.

The protocol defines two processes, the process of disseminating node state information and the process by which routes are calculated. NSR uses a diffusion process to disseminate states. We have shown that regulating the rate at which node states diffuse through the network provides a win-win situation of lower overhead and higher goodput. [15]. The process of calculating routes consists of three steps. First, all possible links are inferred from the node states. Links between nodes up to two hops away are tracked in the conventional manner since nodes identify their one hop neighbors in their node state updates. Links further away are considered to exist if a threshold signal to noise ratio can be achieved using the equation

$$SNR = \frac{P_t / d^n}{N} \tag{3}$$

where P_t is the effective radiated power from a transmitter, N is the background noise power[5], d is the distance that separates the source from the destination, and n is the largest path loss exponent of the two nodes. Using the largest path loss exponent results in choosing symmetric links. In the second step a weight is assigned to each link. The link metric between node i and j is given by

$$w_C(i,j) = \frac{c_{1j} + p_r \cdot d_{ij}^n}{e_{ri} / e_m} \tag{4}$$

[4] Dozing periods are disseminated through the NSR protocol so all neighbors are aware of the p-saver nodes dozing schedule and its p-supporter's identity.

[5] The noise N is just the background noise level and assumes no interference by any adjacent nodes. We assume this is measured at a node and then assumed the same throughout the network

where c_{1j} is the energy consumed by a node receiving a packet[6], p_r is the required signal power at a destination to receive a packet, d is the distance that separates the two nodes on the link, n is the largest path loss exponent of the two nodes on the link, e_{ri} is a measure of the energy reserve and e_m is a constant larger than e_{ri} that weights the influence of the energy reserve on the link metric. Thus, in practice, (4) changes based on whether either the source or the destination is energy unconstrained. When the source is energy unconstrained the denominator in (4) is 1 (i.e. we are not concerned about energy reserves) and if the destination is energy unconstrained then c_{1j} is 0 (i.e. we are not concerned with how much energy the destination consumes receiving a packet). We can penalize dozing nodes in the metric by increasing c_{1j} is using (4). Appropriately weighting dozing nodes can preclude their use in routes unless they form a critical link. Knowledge of the dozing methods allows any neighbor to identify the optimum time to relay data to a dozing node. Finally, the third step is to use Dijkstra's algorithm [15] to calculate the routes. We note that this approach achieves both the objectives suggested in [9] and [10] of choosing shorter hops, giving preference to energy unconstrained nodes, and balancing load across energy constrained nodes.

We note that NSR and SCR complement each other in the energy conservation process. Clearly, NSR enables the dissemination of the dozing parameters to insure they have the least effect on the routing calculations but the SCR protocol also makes the energy conservation mechanisms of NSR possible. First, since SCR uses a spatial mechanism that exploits capture to enhance access success, shorter hops can increase capacity. This counters the adverse effect of increased relay load. Second, the RTS-CTS exchange provides a conservative closed loop mechanism to assist nodes in adjusting their transmission powers. Nodes signal and transmit the RTS and CTS packets using a maximum allowed transmission power, so, power adjustments are based on the worst-case interference conditions. The subsequent power adjustments and the reduction in nodes transmitting improve interference conditions during the payload portion of each transmission slot.

VI. Simulations

We conducted two distinct sets of simulations. In the first we attempted to determine the effect of choosing shorter hops on the capacity of SCR. A description of the simulation and more detailed results are reported in [12]. Of interest is that a next hop selection policy based on (2) nearly doubles the capacity of SCR resulting in a level of goodput that is slightly better than that which is achieved using a minimum hop strategy. Shorter hops can be used without the effect of further congesting the network that is the concern of [11].

In the second set of simulations we combined the use of SCR with. The primary goal of this study was to test the effect of diffusion rates and a load balancing technique on the performance of NSR but we also measured the energy consumption of the network using the default dozing method. Nodes were able to spend more than 50% of their time in that low energy state. Interestingly, the quantity of energy consumed was only slightly affected by load and the node state dissemination parameters. Using a metric based on energy consumed per goodput demonstrated that the effectiveness of the routing protocol to achieve goodput is a significant energy conservation metric as it varied more than 2 to 1 for different protocol parameters.

V. Conclusion

In this paper we have reviewed several energy conservation mechanisms that have been proposed for access and routing protocols. We described the problems in their implementation especially in an integrated application. Our contribution is the introduction of access and routing protocols that work together to manage multiple energy conservation mechanisms. Specifically, dozing methods are accounted for in the routing protocol and the access protocol counters the deleterious effect of power aware routing, i.e. increased relay load, with increased capacity. We have cited the results of two sets of simulations. The first corroborates the correlation of increased capacity with power aware routing and the second demonstrates the effective use of dozing together with a routing protocol. Futher research is necessary to determine how to optimize the routing metrics for the best energy conservation approach for different scenarios. We show that such an optimization cannot be blind to the overall effectiveness of the routing protocol as goodput per energy consumed is very sensitive to goodput capacity.

References

[1] Lucent Technologies WaveLAN High-Speed Wireless Network Interface Card IEEE Turbo 11 Mb Specification, 1999.

[2] http://www.networks.digital.com/dr/npg/deina-ch.html, February 1999.

[3] Nokia C020/C021 Wireless Lan Card Technical Specifications, 1998

[4] http://www.aironet.com/products/in_building/PC4800B.asp.

[5] P802.11, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Press, May 1997.

[6] ETSI, EN300 652 V 1.2.1, "Broadband Radio Access Networks (BRAN); High Performance Radio Local Area Network (HIPERLAN) Type 1; Functional Specification, July 1998.

[7] H. Woesner, J. Ebert, M. Schlager, and A. Wolisz, "Power-Saving Mechanisms in Emerging Standards for Wireless LANs: The MAC Level Perspective," *IEEE Personal Communications*, June 1998, pp. 40-48.

[8] T. Rappaport, *Wireless Communications, Principles and Practice*, Prentice-Hall, Inc, Upper Saddle River, NJ, 1996.

[9] V. Rodoplu and T. Meng, "Minimum Energy Mobile Wireless Networks," *IEEE JSAC*, Vol. 17, No. 8, August 1999, pp. 1333-1344.

[10] S. Singh, M. Woo, and C. Raghavendra, "Power-Aware Routing in Mobile Ad Hoc Networks," *The Fourth Annual ACM/IEEE International Conference in Mobile Computing and Networking*, 1998, pp. 181 – 190.

[11] C.K. Toh, "Maximum Battery Life Routing to Support Ubiquitous Mobile Computing in Wireless Ad Hoc Networks," *IEEE Communications*, June 2001, pp. 138-147.

[12] J.A. Stine, "Energy Conserving Protocols for Wireless Data Networks," Ph.D. Dissertation, University of Texas at Austin, 2001.

[13] J.A. Stine and G. de Veciana, "Energy Efficiency of Centrally Controlled Transmission of Fixed Sized Packets," *WCNC 2000*, 2000.

[14] L.M. Feeney and M. Nilsson, "Investigating the Energy Consumption of a Wireless Network Interface in an Ad Hoc Networking Environment," *INFOCOM 2001*, April 2001.

[15] D. Bertsekas and R. Gallagher, *Data Networks*, Prentice Hall, Inc., Upper Saddle River, NJ, 1992.

[6] The energy consumed in receiving a packet also includes the energy consumed in calculating routes to forward the packet. In the case of a wormhole construct c_2 is the cost to traverse the wormhole.