

Energy Aware Routing for Low Energy Ad Hoc Sensor Networks

Rahul C. Shah and Jan M. Rabaey
Berkeley Wireless Research Center
University of California, Berkeley

Abstract - The recent interest in sensor networks has led to a number of routing schemes that use the limited resources available at sensor nodes more efficiently. These schemes typically try to find the minimum energy path to optimize energy usage at a node. In this paper we take the view that always using lowest energy paths may not be optimal from the point of view of network lifetime and long-term connectivity. To optimize these measures, we propose a new scheme called energy aware routing that uses sub-optimal paths occasionally to provide substantial gains. Simulation results are also presented that show increase in network lifetimes of up to 40% over comparable schemes like directed diffusion routing. Nodes also burn energy in a more equitable way across the network ensuring a more graceful degradation of service with time.

I. INTRODUCTION

Recently there has been a lot of interest in building and deploying sensor networks – dense wireless networks of heterogeneous nodes collecting and disseminating environmental data. There is a multiplicity of scenarios in which such networks might find uses, such as environmental control in office buildings, robot control and guidance in automatic manufacturing environments, interactive toys, the smart home providing security, identification, and personalization, and interactive museums.

Crucial to the success of these ubiquitous networks is the availability of small, lightweight, low-cost network elements, which we call PicoNodes [1]. These nodes must be smaller than one cubic centimeter, weigh less than 100 grams, and cost substantially less than one dollar. Even more important, the nodes must use ultra-low power to eliminate frequent battery replacement. We envision a power-dissipation level below 100 microwatts, as this would enable self-powered nodes using energy extracted from the environment, an approach called energy scavenging or harvesting.

Trying to network a large number of such low-power mobile nodes is a challenging problem that has recently been the focus of many researchers. In particular, routing, addressing and support for different classes of service are the primary issues to be tackled at the network layer.

In this paper, we'll concentrate on the routing problem. Current research has focused on protocols that are low power [2][3][11], scalable with the number of nodes [4] and fault tolerant (to nodes that go up or down, or move in and out of

range) [5]. However, we think that a more useful metric for routing protocol performance is *network survivability*. By this we mean that the protocol should ensure that connectivity in a network is maintained for as long as possible, and that the energy health of the entire network should be of the same order. This is in contrast to energy optimizing protocols that find optimal paths and then burn the energy of the nodes along those paths, leaving the network with a wide disparity in the energy levels of the nodes, and eventually disconnected subnets. If nodes in the network burn energy more equitably, then the nodes in the center of the network continue to provide connectivity for longer, and the time to network partition increases. This leads to a more graceful degradation of the network. This is the idea of survivability of networks.

Energy Aware Routing, the protocol that we have developed tries to ensure the survivability of low-energy networks. It is also a reactive protocol such as AODV and directed diffusion; however, the protocol does not find a *single* optimal path and use it for communication. Rather it keeps a set of *good* paths and chooses one based on a probabilistic fashion. As we will show later, this means that instead of a single path, a communication would use different paths at different times, thus any single path does not get energy depleted. It is also quick to respond to nodes moving in and out of the network, and has minimal routing overhead. We present simulation results which show the improved performance obtained by our method, and the increase in network lifetime. While the primary metric of interest is network survivability, we also show the performance results with regard to other metrics mentioned above.

The rest of the paper is organized as follows. Section II gives background information on sensor networks, their requirements and our ongoing research project, PicoRadio. Some important design choices that we made for PicoRadio are detailed in Section III. Section IV compares existing routing schemes for ad hoc networks and their performance issues. We introduce energy aware routing in Section V and provide qualitative arguments as to its performance. The energy metric used for the protocol is detailed in Section VI. Finally, simulations and results are presented in Section VII, concluding in Section VIII.

II. SENSOR NETWORKS AND THE PICORADIO PROJECT

Sensor networks typically consist of hundreds of nodes, deployed for the purpose of environment monitoring and control. Let us consider a Smart Building scenario, one of the

This research was supported by DARPA on grant no. F29601-99-1-0169 entitled, "Communication/Computation Piconodes for Sensor Networks"

key applications for the PicoRadio project at the Berkeley Wireless Research Center. This is aimed at controlling the environment of a typical office environment using a distributed building monitor and control approach.

Thus the three main functions in a sensor network are sensing, controlling and actuating. These functions could be on separate nodes or co-located on the same physical node. In addition, each physical node also has a logical repeater function which helps in multi-hop routing. We thus define three types of nodes – sensors, controllers and actuators. From the system description, it is easily seen that although the formation of the network is ad hoc, most of the sensors and actuator nodes will remain static. Controllers, on the other hand can be mobile, but their speed is low, on the order of 1-5 m/s.

Also on the positive side, the bit rates in sensor networks are fairly low, about a few hundred bits per sec per node. At most, the peak bit rate supported will be about 10 kb/s, which can enable simple voice messaging (not real time). Sensor data is also highly redundant, which means that end-to-end reliability is not a requirement for most data packets.

Finally, most of the communication is fairly periodic in nature – sensor values are sent at regular intervals to the controllers. This means that we can optimize the network for such recurrent communication, while loosely optimizing for less often, one-time cases.

III. DESIGN OF THE PICONODE

The three main layers we concentrate on for designing the PicoNode are the physical, media access control and network layers [6].

A. Physical Layer

Communication between two nodes requires creating a physical link between two radios. The physical layer handles the communication across this physical link, which involves modulating and coding the data so that the intended receiver can optimally decode it in the presence of channel non-idealities and interference.

B. Media Access Control (MAC) Layer

The MAC layer's primary functions are to provide access control, channel assignment, neighbor list management and power control. It also has a location sub-system that computes the x, y and z-coordinates based on the received signal strength of neighboring nodes and the presence of certain anchors in the network that know their exact positions.

The MAC coordinates channel assignment such that each node gets a locally unique channel for transmission, while the channels are globally reused. There is also a global broadcast channel that is used for common control messages and for waking up nodes. Each node has two radio receivers, one of

which runs at 100% duty cycle, but is very low bit rate and consumes very little power. The second radio runs at very low duty cycle (~1%) and is switched on only when the node needs to receive data. This is a higher rate radio (~10kbps) and consumes more power.

To send data, the MAC layer sends a wake-up signal on the broadcast channel. The address of the node to which it needs to send data is modulated with the wake-up. Access to the broadcast channel is CSMA/CA. On receiving this message, the node to which this is addressed powers on its main radio and communication begins. Since each node has a locally unique channel, there is no problem of collisions occurring during data transmission. Thus the MAC layer enables deep sleep of the nodes which leads to substantial power savings.

The MAC layer also keeps a list of its neighbors and metrics such as the neighbor's position and the energy needed to reach it. This list is used heavily by the network layer to take decisions regarding routing of packets. Finally, the MAC layer also performs power control to ensure power savings and maintain an optimal number of neighbors.

C. Network layer

The network layer has two primary functions – routing and addressing of nodes. In this section, we discuss addressing after which the rest of the paper will only be concerned with the routing protocol.

Traditional network addressing assigns fixed addresses to nodes such as in the Internet. The advantage of such schemes is that the addresses can be made unique. However, there is a very high cost associated with assigning and maintaining these kinds of addresses. This problem is exacerbated in mobile networks where the topology information keeps changing. It is very difficult to route packets if the node address does not provide a clue as to the direction in which the packet is to be routed. To solve this problem, there are two approaches. One is to maintain a central server that keeps up-to-date information on the position of every node. Another way is to take the Mobile IP approach, where every node has a Home Agent that handles all the requests for the node and redirects it to the present position of the node.

For sensor networks, however, there is an important property of information flow that can be used to our advantage. Most of the communication in sensor networks is of the form "Give me the temperature of Room 5". Thus nodes can be addressed based on their geographical position. This information is also very useful for the routing protocol, as it can direct communication in the right direction. Thus, for PicoRadio we use *class-based addressing*. These addresses are triplets of the form <Location, Node Type, Node Subtype>. Location specifies a particular point or region in space that is of interest. Node Type defines which type of node is required, such as sensor, controller or actuator. Finally, the Node Subtype further narrows down the scope of the address, such as Temperature Sensor, or Humidity Sensor etc. Thus class based addressing defines the

type of node in the region of space which is needed. In the rest of the paper, we assume class-based addressing within the network layer.

IV. AD HOC ROUTING SCHEMES

Since research on ad hoc routing began with packet radio networks, numerous protocols have been proposed [7]. These protocols have to contend with the wireless medium, i.e. low bandwidth, high error rates and burst losses, as well as the limitations imposed by ad hoc networks such as frequently changing topology and low power devices. Finally, the protocols have to scale well to a large number of nodes in the network. Considering the nature of the protocols, they can be categorized into *proactive* and *reactive* protocols.

A. Proactive Routing Protocols

Proactive routing protocols have the distinguishing characteristic of attempting to maintain consistent up-to-date routing information from each node to every other node in the network. Every node maintains one or more routing tables that store the routing information, and topology changes are propagated throughout the network as updates so that the network view remains consistent. The protocols vary in the number of routing tables maintained and the method by which the routing updates are propagated. Two common proactive protocols are discussed below briefly; there are many others in the literature also.

The Destination-Sequenced Distance-Vector Routing protocol (DSDV) [8] is a typical example of proactive protocols. It is based on the Bellman-Ford algorithm for shortest paths and ensures that there is no loop in the routing tables. Every node in the network maintains the next hop and distance information to every other node in the network. Routing table updates are periodically transmitted throughout the network to maintain table consistency.

Link-state Routing [9] is a proactive protocol in which each node floods the cost of all the links to which it is connected throughout the network. Every node then works out the cost of reaching every other node using shortest path algorithms. In addition, the protocol works correctly even if unidirectional links are present, whereas DSDV assumes bidirectional links.

B. Reactive Routing Protocols

In contrast to proactive routing protocols, reactive protocols create routes only when desired. This means that an explicit route discovery process creates routes and this is initiated only on an as-needed basis. It can be either source-initiated or destination-initiated. Source-initiated routing means that it is the source node that begins the discovery process, while destination-initiated is the opposite. Once a route has been established, the route discovery process ends,

and a maintenance procedure preserves it until the route breaks down or is no longer desired.

Ad-Hoc On-Demand Distance Vector Routing (AODV) [10] is a routing protocol also based on the distance vector algorithm like DSDV, but the difference is that AODV is reactive. It is a source-initiated protocol, with the source node broadcasting a Route Request (RREQ) when it determines that it needs a route to a destination and does not have one available. This request is broadcast till the destination or an intermediate node with a “fresh enough” route to the destination is located. Intermediate nodes record the address of the neighbor from which the first copy of the broadcast packet is received in their route tables, thus establishing a reverse path.

Dynamic Source Routing (DSR) [11] is a reactive protocol that is also source-initiated and is based on the concept of source routing, i.e. the source specifies the entire route to be taken by a packet, rather than just the next hop. If the source node does not have a route, it floods the network with a Route Request (RREQ). Any node that has a path to the destination can reply with a Route Reply (RREP) to the source. This reply contains the entire path recorded in the RREQ packet. The entire path is added to the header of every packet to the destination, thus it is called source routing.

Directed Diffusion [5] is a different communication paradigm specifically for sensor networks. It is a destination-initiated reactive protocol that is data-centric and application-aware. Diffusion works well for sensor networks where queries like “Send me the temperature data in a particular area” and their responses are the dominant form of communication. A destination node (controller) requests data by sending interests for data. This interest is flooded over the network, but each node only knows the neighbor from whom it got the request, and it sets up a gradient to send data to the neighbor. So by this process, the interest reaches the source node (sensor), but each node only knows its neighbor(s) who asked for the data, not the final consumer of the data. Since it is conceivable that each node would receive the same interest from more than one neighbor, data would come down to the controller node along multiple paths. Of these, one high rate path is defined and the rest of the paths remain low rate. This is achieved by sending out positive reinforcements to increase the rate of a particular path. There is also a mechanism for negative reinforcements to change high rate paths to low rate ones; these are used when a better path emerges.

V. ENERGY AWARE ROUTING

The potential problem in current protocols is that they find the lowest energy route and use that for every communication. However, that is not the best thing to do for network lifetime. Using a low energy path frequently leads to energy depletion of the nodes along that path and in the worst case may lead to network partition.

To counteract this problem, we propose a new protocol that we call energy aware routing. The basic idea is that to

increase the survivability of networks, it may be necessary to use sub-optimal paths occasionally. This ensures that the optimal path does not get depleted and the network degrades gracefully as a whole rather than getting partitioned. To achieve this, multiple paths are found between source and destinations, and each path is assigned a probability of being chosen, depending on the energy metric. Every time data is to be sent from the source to destination, one of the paths is randomly chosen depending on the probabilities. This means that none of the paths is used all the time, preventing energy depletion. Also different paths are tried continuously, improving tolerance to nodes moving around the network.

Energy aware routing is also a reactive routing protocol. It is a destination-initiated protocol where the consumer of data initiates the route request and maintains the route subsequently.

Thus, it is similar to diffusion in certain ways. Multiple paths are maintained from source to destination. However, diffusion sends data along all the paths at regular intervals, while energy aware routing uses only one path at all times. But due to the probabilistic choice of routes, it can continuously evaluate different routes and choose the probabilities accordingly. The protocol has three phases:

- Setup phase or interest propagation – Localized flooding occurs to find all the routes from source to destination and their energy costs. This is when routing (interest) tables are built up.
- Data Communication phase or data propagation – Data is sent from source to destination, using the information from the earlier phase. This is when paths are chosen probabilistically according to the energy costs that were calculated earlier.
- Route maintenance – Route maintenance is minimal. Localized flooding is performed infrequently from destination to source to keep all the paths alive.

A. Setup Phase

1. The destination node initiates the connection by flooding the network in the direction of the source node. It also sets the “Cost” field to zero before sending the request.

$$Cost(N_D) = 0$$

2. Every intermediate node forwards the request only to the neighbors that are closer to the source node than oneself and farther away from the destination node. Thus at a node N_j , the request is sent only to a neighbor N_i which satisfies:

$$d(N_i, N_S) \geq d(N_j, N_S)$$

$$d(N_i, N_D) \leq d(N_j, N_D)$$

where $d(N_i, N_j)$ is the distance between N_i and N_j .

3. On receiving the request, the energy metric for the neighbor that sent the request is computed and is added to the total cost of the path. Thus, if the request

is sent from node N_i to node N_j , N_j calculates the cost of the path as:

$$C_{N_j, N_i} = Cost(N_i) + Metric(N_j, N_i)$$

4. Paths that have a very high cost are discarded and not added to the forwarding table. Only the neighbors N_i with paths of low cost are added to the forwarding table FT_j of N_j .

$$FT_j = \{i | C_{N_j, N_i} \leq \alpha \cdot (\min_k C_{N_j, N_k})\}$$

5. Node N_j assigns a probability to each of the neighbors N_i in the forwarding table FT_j , with the probability inversely proportional to the cost.

$$P_{N_j, N_i} = \frac{1/C_{N_j, N_i}}{\sum_{k \in FT_j} 1/C_{N_j, N_k}}$$

6. Thus, each node N_j has a number of neighbors through which it can route packets to the destination. N_j then calculates the average cost of reaching the destination using the neighbors in the forwarding table.

$$Cost(N_j) = \sum_{i \in FT_j} P_{N_j, N_i} C_{N_j, N_i}$$

7. This average cost, $Cost(N_j)$ is set in the “Cost” field of the request packet and forwarded along towards the source node as in Step 2.

B. Data Communication Phase

1. The source node sends the data packet to any of the neighbors in the forwarding table, with the probability of the neighbor being chosen equal to the probability in the forwarding table.
2. Each of the intermediate nodes forwards the data packet to a randomly chosen neighbor in its forwarding table, with the probability of the neighbor being chosen equal to the probability in the forwarding table.
3. This is continued till the data packet reaches the destination node.

VI. ENERGY METRIC

The energy metric that is used to evaluate routes is a very important component of the protocol. Depending on the metric, the characteristics of the protocol can change substantially. As mentioned earlier, the metric can include information about the cost of using the path, energy health of the nodes along the path, topology of the network etc. For energy aware routing, we use a simple metric that has been considered before [12]:

$$C_{ij} = e_{ij}^\alpha R_i^\beta$$

Here C_{ij} is the cost metric between nodes i and j , e_{ij} is the energy used to transmit and receive on the link, while R_i is the residual energy at node i normalized to the initial energy

of the node. The weighting factors α and β can be chosen to find the minimum energy path or the path with nodes having the most energy or a combination of the above. Further study needs to be done as to the best metric as it has a deep impact on the protocol performance.

VII. SIMULATIONS AND RESULTS

Simulations were carried out in Opnet to demonstrate the increased network survivability due to energy aware routing. The simulation consisted of 76 nodes in a typical office setup as in Fig. 1. There were 65 sensors and 7 controllers that were static and 4 nodes that were mobile. Among the sensors, 47 were light sensors and 18 were temperature sensors. The controllers sent out requests for data to the sensors in their region of interest. These requests programmed the light sensors to send data every 10 seconds and temperature data every 30 seconds.

Every node consisted of an application and a network layer. The application layer was programmed to be either a sensor or controller, while the network layer performed the routing operations. Energy aware routing was compared against directed diffusion routing. Both routing protocols used the same energy metrics for path selection. This was the metric function given in Section VI with $\alpha=1$ and $\beta=50$.

The MAC layer was abstracted away by providing for direct transfer of packets from the network layer of one node to the network layer of its neighbor. Thus there was no contention for the medium when sending data. The main purpose of removing the MAC was to orthogonalize the advantages of our network and media access layers and to evaluate the benefits of each separately.

It has been reported previously that the absence of the media access layer may skew the simulations and results might not be totally accurate. However, that is not a problem for PicoRadio. Every node has a locally unique channel assigned to it by the MAC layer, so there is no loss due to collisions during data transmission. And even though the broadcast channel is common to all nodes, it is used only during wake up which is very rare (~1%) due to extremely low data rates in sensor networks and also due to aggregation of data. Thus modeling the radio as being on only during transmission and reception is fairly accurate.

The network layer takes a packet and processes it, sending it either to a neighbor or to the application layer or both. It also maintained the neighbor list. An expanding ring search was used to create the list till it had the minimum number of neighbors (which was 4) or the maximum radio range of the node was reached.

Every node was given an identical amount of initial energy. Transmission used $20 \text{ nJ/bit} + 1 \text{ pJ/bit/m}^3$ (i.e. energy drop-off was r^3 , which is a moderate indoor environment). The energy for reception was 30 nJ/bit . These numbers are typical values for current Bluetooth radios. The packets were 256 bits in size.

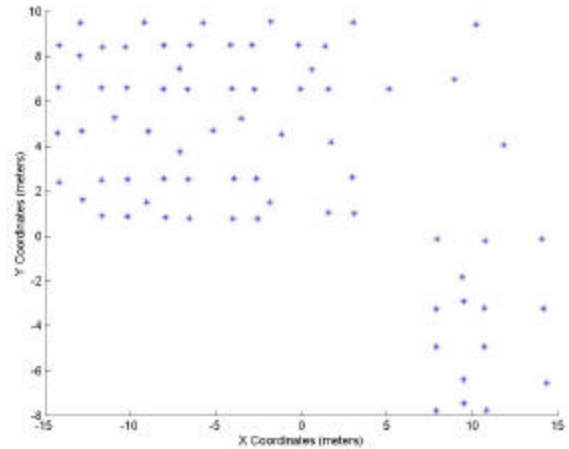


Fig. 1 Layout of static nodes in the network

Fig. 2 shows the results of one of the simulation runs. It shows the energy consumed by the various nodes during a one hour period of the network. This can be compared against the energy consumed by the directed diffusion routing protocol in Fig. 3. As expected, energy aware routing spreads the traffic over the network resulting in a much “cooler” network. As a consequence, the nodes in the center of the network conserve energy for longer and the time till the first node runs out of energy increases.

The simulations show that energy aware routing reduces the average energy consumption per node from 14.99 mJ to 11.76 mJ , an improvement of 21.5%. This is primarily due to the very low overhead of the protocol. At the same time, it reduces the energy differences between different nodes.

In another performance run, the network was simulated till a node ran out of energy. For diffusion routing, this occurred after 150 minutes, while it took 216 minutes for the energy aware routed network to fail. This is an increase in network lifetime¹ of 44%, which agrees with the results of the previous simulation. In that simulation, the maximum energy usage among all nodes was 57.44 mJ for diffusion and 41.11 mJ for energy aware routing. This means that diffusion had a maximum energy consumption of ~1.4 times energy aware routing, hence an increase of ~40% in the network lifetime is expected.

Fig. 4 shows the normalized energy values of the nodes at the end of the one hour simulation run. Normalization was done with respect to the average energy usage across the network, and the purpose of this plot is to show that energy aware routing distributes the load more evenly across the network. Thus the difference in energy usage among nodes is lesser as compared to diffusion in Fig. 5. Also, the bit rate measured by the network is ~250 bits/sec, which demonstrates the extremely low data rate requirements of sensor networks. Thus the results clearly show the improved network health due to energy aware routing.

¹ In this paper, we call the time till the first node runs out of energy as “network lifetime”. There can be many alternate interpretations of this term.

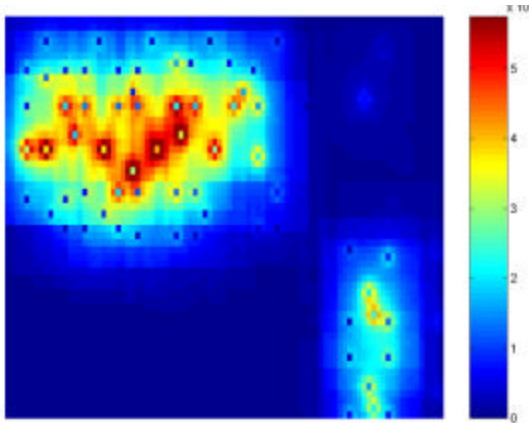


Fig. 2 Energy consumption for energy aware routing (μJ)

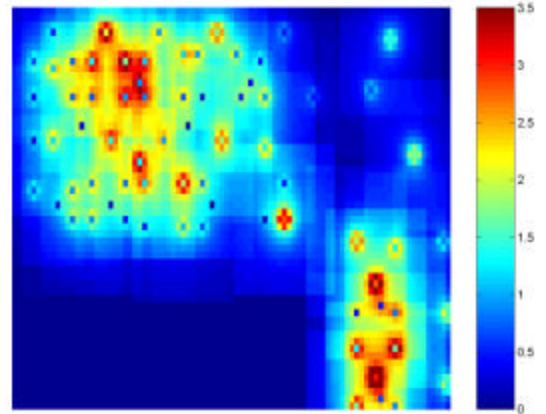


Fig. 4 Normalized energy for energy aware routing

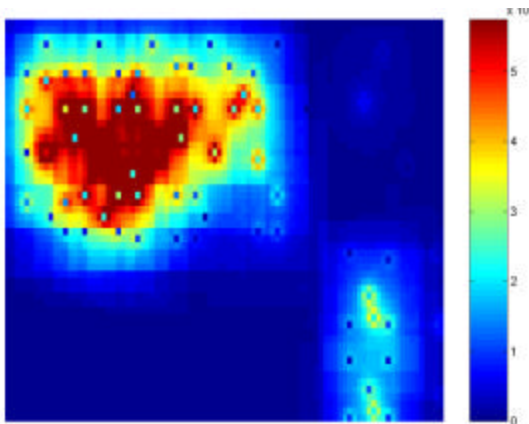


Fig. 3 Energy consumption for diffusion routing (μJ)

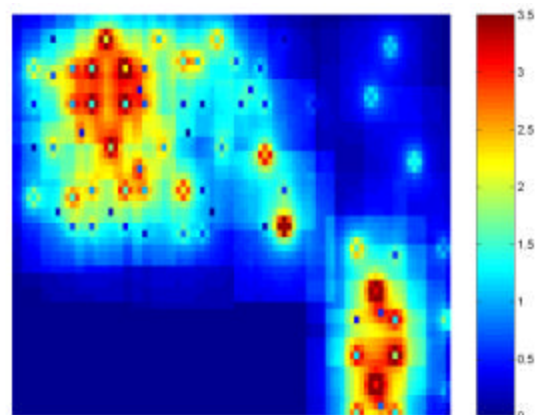


Fig. 5 Normalized energy for diffusion routing

VIII. CONCLUSIONS AND FUTURE WORK

In this paper, we presented a new routing protocol that is suitable for low energy and low bit rate networks. The idea behind the protocol is very simple – using the lowest energy path always is not necessarily best for the long-term health of the network. Thus using a simple mechanism to send traffic through different routes helps in using the node resources more equitably. Using probabilistic forwarding to send traffic on different routes provides an easy way to use multiple paths without adding much complexity or state at a node.

Network survivability is a very important criterion for deciding the efficacy of network protocols. It includes a measure of the network lifetime as well as the kind of service it provides during its life. Both these factors are important in evaluating networks and neither can be considered in isolation. For e.g. a network that does not forward packets at all satisfies the first condition. And one that just burns itself out fast trying to ensure full reliability would satisfy the second one. However, survivability is still a preliminary concept and needs to be worked on further and understood better before using it to judge network protocols.

REFERENCES

- [1] Jan M. Rabaey et. al., "PicoRadio supports ad hoc ultra-low power wireless networking", IEEE Computer, July 2000, pp. 42-48.
- [2] C. K. Toh, "Maximum battery life routing to support ubiquitous mobile computing in wireless ad hoc networks", IEEE Comm. Mag., June 2001, pp. 138-147.
- [3] S. Singh, M. Woo and C. S. Raghavendra, "Power aware routing in mobile ad hoc networks", IEEE/ACM MobiCom, Oct. 1998, pp. 181-190.
- [4] R. Jain, A. Puri and R. Sengupta, "Geographical routing for wireless adhoc networks using partial information", IEEE Personal Comm., Feb. 2001
- [5] C. Intanagonwiwat, R. Govindan and D. Estrin, "Directed Diffusion: A scalable and robust communication paradigm for sensor networks", IEEE/ACM Mobicom, 2000, pp. 56-67.
- [6] J. L. da Silva Jr. et. al., "Design Methodologies for PicoRadio Networks", Design Automation and Test Conference (DATE), March 2001.
- [7] E. Royer, C. K. Toh, "A review of current routing protocols for ad hoc mobile wireless networks", IEEE Personal Comm., April 1999, pp. 46-55.
- [8] C. E. Perkins and P. Bhagwat, "Highly dynamic destination sequenced distance vector routing (DSDV) for mobile computers", Comp. Commun. Rev., Oct. 1994, pp. 234-244.
- [9] P. Jacquet et. al., "Optimized link state routing", Internet draft, <http://www.ietf.org/internet-drafts/draft-ietf-manet-olsr-04.txt>.
- [10] C. Perkins and E. Royer, "Ad hoc On demand distance vector routing", Proc. 2nd IEEE Wksp Mobile Comp. Sys. & Apps., Feb 1999.
- [11] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks", Mobile Computing, Kluwer, 1996, pp. 153-181.
- [12] J. Chang and L. Tassiulas, "Energy conserving routing in wireless ad hoc networks", IEEE Infocom, 2000, pp. 22-31.