

# Adaptive QoS Routing by Cross-Layer Cooperation in Ad Hoc Networks

**Hongxia Sun**

*Department of Computer Science, University of Calgary, Calgary, AB, Canada T2N 1N4  
Email: sunh@cpsc.ucalgary.ca*

**Herman D. Hughes**

*Computer Science and Engineering Department, Michigan State University, East Lansing, MI 48824-1027, USA  
Email: hughes@cse.msu.edu*

*Received 30 June 2004; Revised 8 April 2005*

QoS provisioning is a complex and challenging issue in mobile ad hoc networks, especially when there are multiple QoS constraints. In this paper, we propose an adaptive QoS routing scheme supported by cross-layer cooperation in ad hoc networks. The cross-layer mechanism provides up-to-date local QoS information for the adaptive routing algorithm, by considering the impacts of node mobility and lower-layer link performance. The multiple QoS requirements are satisfied by adaptively using forward error correction and multipath routing mechanisms, based on the current network status. The complete routing mechanism includes three parts: (1) a modified dynamic source routing algorithm that handles route discovery and the collection of QoS-related parameters; (2) a local statistical computation and link monitoring function located in each node; and (3) an integrated decision-making system to calculate the number of routing paths, coding parity length, and traffic distribution rates. Simulation results are presented to illustrate the overall performance of our scheme. Our results indicate that our adaptive routing scheme provides suitable QoS performance that is less sensitive to network conditions (i.e., node mobility, transmission power, channel characteristics, and the traffic pattern) than a nonadaptive routing strategy.

**Keywords and phrases:** QoS routing, ad hoc network, multiple path, end-to-end delay, packet loss.

## 1. INTRODUCTION

A wireless ad hoc network consists of a collection of mobile nodes interconnected by multihop wireless paths with wireless transmitters and receivers. Such networks can be spontaneously created and operated in a self-organized manner, because they do not rely upon any preexisting network infrastructure.

There are numerous applications (e.g., military, rescue) for this type of network. The emergence of multimedia applications in communications has generated the need to provide quality-of-service (QoS) support in mobile ad hoc networks, and such applications require a stable path to guarantee QoS requirements. However, the topology of ad hoc networks is highly dynamic due to the unpredictable node mobility. In addition, wireless channel bandwidth is limited. So, QoS provisioning in such networks is complex and challenging.

There are many routing algorithms proposed for wireless ad hoc networks; a good survey is provided in [1]. Prior work on ad hoc network routing can be categorized based on how the state information is maintained and how the search for feasible paths is carried out. General approaches include source routing [2], distributed routing [3], and hierarchical routing [4]. There are some hybrid methods [5, 6] reported in the literature, and these schemes have been shown to enhance network performance.

QoS routing usually involves two tasks: collecting and maintaining up-to-date state information about the network and finding feasible paths for a connection based on its QoS requirements. There are currently several main approaches for QoS routing in ad hoc networks. These approaches can be classified as network level only and combined network-data-link level [7]. An example of a network-level-only approach is given in [3]. However, it may suffer from several potential problems. For example, while the path is being discovered, only the link bandwidth between neighboring nodes is considered. Because transmissions between neighboring nodes also affect other nodes in ad hoc networks, neglecting

---

This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

the physical properties of the transmission channel can lead to problems such as possible failure at high data loads. The network-data-link level approach is more promising because it combines information from both the network and data-link layer [7, 8]. However, due to dynamic changes in network topology and the difficulties in predicting link states, indirect or estimation approaches are often used (e.g., using signal strength and link lifetime as routing parameters [9]). The obvious problem with such approaches is that the impacts on QoS performance are hard to quantify, since the cross-layer behaviors of mobile networks are not considered. Therefore, most of the proposed routing schemes for mobile networks are only QoS aware, but do not guarantee QoS. To address this problem, appropriate cross-layer cooperation is required. We propose an adaptive scheme to provide QoS information by factoring the impacts of node mobility and lower-layer link parameters into QoS performance.

There are many proposals for QoS routing in the literature [7, 10, 11]. Most approaches tend to focus on only one QoS parameter (e.g., packet loss, end-to-end delay, and bandwidth). For example, while many of the QoS-related schemes are successful in reducing packet loss by adding redundancy in the packet [12, 13, 14], they do this at the expense of end-to-end delay. Because packet loss and end-to-end delay are inversely related, it may not be possible to find a path that simultaneously satisfies the delay, packet loss, and bandwidth constraints. Some proposed QoS routing algorithms [15, 16, 17] do consider multiple metrics, but without considering cross-layer cooperation. Multipath routing is another type of QoS routing that has received much attention, since it can provide load balancing, fault tolerance, and higher aggregate bandwidth [12, 18, 19]. Although this approach decreases packet loss and end-to-end delay, it is only efficient and reliable if a relationship can be found between the number of paths and QoS constraints.

In this paper, we propose a cross-layer cooperation mechanism to support adaptive multipath routing with multiple QoS constraints in an ad hoc network. The cross-layer mechanism provides information on link performance for the QoS routing. It treats traffic distribution, wireless link characteristics, and node mobility in an integrated fashion. That is, it reflects the impacts of lower-layer parameters on QoS performance in higher layers, with emphasis on translating these parameters into QoS parameters for the higher-layer connections. A multiobjective optimization algorithm is used to calculate routing parameters using the cross-layer mechanism. These parameters are adapted to the current network status, determining the number of routing paths and code parity lengths for FEC. In addition, a traffic engineering strategy is used to evenly distribute traffic over multiple paths.

The remainder of the paper is organized as follows. Section 2 provides an overview of our cross-layer routing architecture. Three functions (the routing, the local statistic computation, and the integrated decision-making functions) in the routing mechanism are introduced in Section 3. The layered network models used to support the implementation of these functions are also presented in this part. Simulation

and numerical results are discussed in Section 4. Finally, the summary is presented in Section 5.

## 2. SCHEME OVERVIEW

We propose an adaptive routing algorithm for supporting QoS in hybrid mobile ad hoc networks. The computation of the parameters in the routing is adaptive with respect to the current network status. This distributed routing utilizes the most up-to-date local information at each node, where local states are maintained by a cross-layer mechanism. QoS requirements are satisfied by adaptively using forward error correction (FEC) at packet level and multipath routing mechanisms based on the current network status.

Due to the effects of changes in network topology and wireless link, QoS performance on a node becomes complex. Therefore, an adaptive QoS routing mechanism needs several cross-layer functions cooperating harmoniously to deal with changes in different layers. Firstly, a local QoS performance prediction mechanism is needed. It should include local information collection and local QoS performance computations. Once this prediction mechanism is built, the second step is to construct a distributed routing strategy based on the predicted QoS performance along selected paths. This includes routing discovery and routing maintenance. A hybrid asynchronous local information update mechanism is also introduced.

In order to implement our adaptive multipath routing scheme, three functions distributed in different parts of the network are needed. First, a modified dynamic source routing function is needed. It handles route discovery and collecting the local QoS-related information along the selected routes. Second, there is a local statistical computation and link monitoring function located in each node. This function is used to support the above routing function. It will manage and build the local routing information in each node, which includes a QoS-related table. The third function will be in charge of the final decision-making process. The adaptive routing parameters are derived from the decision-making algorithm based on the QoS constraints. They are the number  $N$  of selected paths, parity length  $k$  of the FEC, code and the set  $\{R\}$  of the traffic distribution rates on each path. With these functions, adaptive multipath QoS routing is implemented.

Obviously, this adaptive routing is a hybrid approach because it includes both a local QoS status precomputation and an on-demand multipath routing algorithm. Routing parameters, such as the number of paths, the forward error correction (FEC) parity length, and the packet distribution rate on each path, are finally determined by the integrated decision-making system.

The link local status depends mostly on the lower-layer parameters such as the wireless channel characteristics and the nodes' mobility, which are provided by our local cross-layer mechanism. Since on-demand routing finds feasible paths, given a specified request, it can operate by using either the regular method [20, 21] or an improved method [9, 12]. For example, signal strength and link lifetime constraints

will decrease the available paths, but have the advantage of facilitating the location of a more reliable link for routing [9]. We simply select an on-demand multipath routing protocol to use in our system, but modify the request packet or reply packet, which depends on where the decision making is located (source or destination). If decision making takes place in the source, the reply packet structure is modified to piggyback the local QoS information along the path. If this occurs at the destination, the request packet structure will be modified to carry the QoS requirements from the user. This packet also piggybacks local QoS information along the path. The reply packet sends the final decision back to the source. We design an iterative algorithm to calculate routing parameters for QoS guarantees. These QoS requirements can be based on either a delay or a delay and bandwidth requirement, or a packet loss requirement. FEC parity length is derived from the difference between the QoS delay requirement and the average delay on selected paths under the packet-loss constraint. Average packet loss under this FEC scheme is achieved by using multiple routing paths. At the same time, the packet distribution rate on each path is determined under fair packet-loss and load-balance principles. Routing maintenance under the same QoS guarantees is achieved without increasing its computational complexity.

### 3. CROSS-LAYER COOPERATIVE FUNCTIONS

#### 3.1. Routing function

Generally, ad hoc routing protocols can be classified into proactive and reactive protocols [1]. We propose the use of a distributed dynamic source routing. Split dmultipath routing (SMR) [19] is modified to fulfill the multipath routing function in our adaptive routing scheme. It is an on-demand routing protocol that builds multiple disjoint routes using request/reply cycles. For QoS considerations in our scheme, we extended the structure of a request or a reply packet to include three new fields in the packet. These fields will keep three parameters defined as follows.

*Definition 1.* Assume that  $D(i, j)$  is a metric for link  $(i, j)$ , for a path  $p = (s, i, j, \dots, k, d)$ . Let

$$D(p) = D(s, i) + D(i, j) + \dots + D(k, d). \quad (1)$$

*Definition 2.* Assume that  $L(i, j)$  is a metric for link  $(i, j)$ , for a path  $p = (s, i, j, \dots, k, d)$ . Let

$$L(p) = L(s, i) \cdot L(i, j) \cdot \dots \cdot L(k, d). \quad (2)$$

*Definition 3.* Assume that  $B(i, j)$  is a metric for link  $(i, j)$ , for a path  $p = (s, i, j, \dots, k, d)$ . Let

$$B(p) = \min \{B(s, i), B(i, j), \dots, B(k, d)\}. \quad (3)$$

Upon doing this, the receiver on the path will know the accumulated value of  $D(p)$ ,  $L(p)$ , and  $B(p)$ . If we have  $D(p)$  represent the accumulated value of the delay, then  $1 - L(p)$  represents the accumulated value of the packet loss and  $B(p)$

represents the minimum bandwidth on the path. Whenever the request or reply packet proceeds for another link  $(i, j)$ , let

$$\begin{aligned} D(p) &= D(p) + D(i, j), & L(p) &= L(p) \cdot L(i, j), \\ B(p) &= \min \{B(p), B(i, j)\}. \end{aligned} \quad (4)$$

These QoS parameters are brought into the function of the integrated decision making located in the destination or source node. They will also be used in the calculations of the adaptive routing parameters.

The end-to-end delay of a path is the sum of the node delay at each node plus the link delay at each link on that path. Node delay includes the protocol processing time and the queuing delay at node  $i$  for link  $(i, j)$ . Link delay is the propagation delay on link  $(i, j)$ . The delay metric is defined as

$$\text{delay}(i, j) = \text{nodedelay}(i, j) + \text{linkdelay}(i, j). \quad (5)$$

The end-to-end packet-loss rate of a path is an accumulation of the packet loss caused by buffer overflow, link failure, and packet discard caused by channel error. The packet loss rate metric is defined as

$$\text{pkloss}(i, j) = \text{pklbo}(i, j) + \text{pkllf}(i, j) + \text{pkldc}(i, j). \quad (6)$$

The residual bandwidth metric in link  $(i, j)$  is

$$\text{residual } B(i, j) = \text{capacity}(i, j) - \text{aggregatetraffic}(i, j). \quad (7)$$

Since reply/request packets travel at speeds based on the delay, a reply/request packet traveling along the path with the smallest delay will arrive first. So, the source/destination node always discovers available routes according to the arriving order of the reply/request packet.

#### 3.2. Local statistical computation and monitoring function

The second function of our adaptive scheme is the local information statistics and monitoring. It is used to support the routing function stated above. Local QoS parameters used in routing are obtained from this function.

In order to show how to collect and use information in different layers, we divide our discussion into two parts. One part is the architecture of the mechanism for monitoring and gathering statistics; another part presents the models used to support the architecture.

##### 3.2.1. Architecture

Each node in the network has a monitoring mechanism to collect and exchange its local information periodically. In our scheme, a node is assumed to keep up-to-date local information, including all outgoing links and neighbors. There are two tables in each node: one has the link state information, the other has the QoS-related and the link weight information. This QoS-related table is combined with the original

routing table in the routing function. Since a distributed dynamic source routing is used, an arbitrary link weight can be defined based on our objective. The state information of one outgoing link  $(i, j)$  may include (1) average signal-to-noise ratio  $S(i, j)/N_0$ , (2) link capacity  $C(i, j)$ , (3) average aggregate traffic  $\lambda(i, j)$ , (4) link stability  $\psi(i, j)$ , and (5) neighbor nodes' mobility characteristics sorted by several types  $\xi_j$ .

The local information table is built from the monitoring and statistical mechanism in the node. It consists of two parts: (1) an exchange information part (i.e.,  $\xi_j$ ), which comes from its neighbors, and (2) a statistical information component (i.e.,  $\lambda(i, j)$ ), which comes from its monitoring mechanism. Such information may be located in both the lower and higher layers. For example, average signal-to-noise ratio  $S(i, j)/N_0$  is a physical layer parameter. Because the wireless device drivers normally provide signal-to-noise ratio information, the receiver can read such information from the driver periodically. At the same time, link stability  $\psi(i, j)$  can be obtained from the statistics of the beacons exchanged between wireless devices. Although each node's mobility level has different characteristics from others, it is assumed that once end nodes of the type  $\xi_j$  are found, the local wireless channel model and the nodes' mobility will be known. Since the aforementioned data is already available at different layers in the network, there is no need for an additional mechanism to gather it. As such, the network overhead associated with message exchanges will not increase.

The local QoS-related table will be produced from local state information. This phase could be implemented by two different approaches. We call them pure statistical methods and prediction methods.

A statistical method provides a simple approach to collect local QoS-related parameters. It is based on an assumption that all factors affecting those QoS-related parameters of a node do not change as frequently as other previously discussed statistical values. From this statistical mechanism, a QoS-related table is built. Residual bandwidth can be obtained from the information of link capacity  $C(i, j)$  and the average aggregate traffic,  $\lambda(i, j)$ . Average delay and average packet loss can be obtained directly from the historical information in the node. Other information, such as  $S(i, j)/N_0$ ,  $\psi(i, j)$ , and  $\xi_j$ , can be used to calculate the link weight. Different methods about how to build link weight are proposed in literature [9, 22]. We proposed the method in [9], since its selected paths have higher QoS properties.

A prediction method is another approach to get the information in the local QoS-related table. Packet delay and packet-loss rate are estimated by predicting link states. These predictions are based on current local information. The link model discussed below will be a bridge to bring lower-layer parameters (e.g.,  $S(i, j)/N_0$ ,  $\xi_j$ , etc.) into higher layers. As such, they can be combined with the information in higher-layers (e.g.,  $C(i, j)$ ,  $\lambda(i, j)$ , etc.) to predict the local QoS performance.

The difference between the statistical method and prediction method is that the former uses historical network information. This previous information is inherently imprecise in an ad hoc network because the network state and topology

may change at any time. The latter uses the forecasting information, as it collects all possible information affecting future network states to calculate the QoS-related information. So, the prediction method is more precise for predicting network states in the next transmission interval, though the former is less complex.

### 3.2.2. Models

In order to support the collection of the desired network information (especially, for the collection of the lower-layer information, i.e., node mobility and wireless channel characteristics), we consider the network with three models. First, the network model defines the properties of the network components, including node classification and channel classification. Second, the node model provides node mobility characteristics (e.g., relative speed, direction, etc.) and the capacity of the node with respect to its communication properties (e.g., transmission power, receiving sensitivity, etc.). Third, the link model characterizes the changes of the wireless channel with features of mobile ad hoc networks, which allows the effects of the lower-layer parameters to be factored into the computation of network performance at higher layers. All these models quantify values associated with the messages (e.g., node and wireless channel types, etc.) in different layers. These quantified values are further used in the statistics or prediction regarding local QoS performance. A brief description of the three models follows.

#### (a) Network model

We propose a flexible network model with either mobile or stable nodes. It is a set  $V$  of nodes that are interconnected by a set  $L$  of wireless communication links.  $V$  and  $L$  change over time since nodes join and leave the network. Nodes in the network can be classified as belonging to several groups according to node mobility level, which is represented by a set  $G$ . Each group has its unique mobile characteristics that will be later defined in the node's model, and each node,  $n$ , has its unique identifier. Wireless channels are also classified according to their local communication environments, which are represented by a set  $H$ . The channel characteristics (e.g., average signal-to-noise ratio  $S/N_0$ , channel gain  $\eta$ , etc.) are grouped according to several typical radio channels. So, a hybrid network can be represented by the following two sets:

$$V = \{G(n)\}_{n \in G}, \quad L = \left\{ H \left( \frac{S}{N_0}, \eta \right) \right\}_{S/N_0, \eta \in H}. \quad (8)$$

#### (b) Node model

We assume that each node communicates with its neighbors, and a link is available when two nodes are in the transmission range of each other. We consider the communication range of the nodes individually in this paper, because we do not want to miss the information of individual nodes coming from the physical layer (e.g., transmission power, transmission rate, signal-to-noise ratio, mobility, etc.). Since we are discussing the link state in the mobile network, two nodes' features are defined: one is the mobility  $\xi_i$  of the node  $i$ , which

is characterized by the relative speed  $v(i, j)$  with its neighbor  $j$  on that link, and the direction  $\theta$  and the movement function  $f(\cdot)$ . These characteristics can be modeled by using random functions (i.e., a mobility model proposed by [23]). This node's mobility affects the link failure function directly. Another node's feature is the node's communication coverage range, which is presented by a two-dimension circle areas. The maximum communication distance  $r_{i,j}(t)$  from node  $i$  to node  $j$  is used to represent these areas, and it is determined by the parameters in the transmitter and the receiver (i.e., transmission power  $S_i$ , receiving sensitivity  $\beta_j$ ), physical channel fading  $\alpha_c$ , and the background noise power  $N_0$ . This coverage range finally determines the neighbors of the node. So, a node in the mobile network can be modeled as follow:

$$\xi_i = \{v(i, j), \theta, f(\cdot)\}, \quad r_{i,j} \leq \alpha_c \sqrt{\frac{S_i}{(\beta \cdot N_0)}}. \quad (9)$$

Generally, a link state is dominated by two factors: one is the lower-layer status, such as node mobility and radio channel characteristics; the other is the higher-layer traffic. By exchanging neighbor's local information in the network model, a link's lower-layer characteristics can be moved up to a higher layer. Combining these lower-layer parameters with the higher-layer parameters in a link model, we are able to predict the local performance in mobile networks.

### (c) Link model

A mobile wireless channel can be modeled by a multiple-state Markov chain. For simplicity, we use a two state Markov model. State one represents a good channel state and state two is the bad channel state. States are defined by the range of signal-to-noise ratio on this mobile wireless link.

We use  $p$  to denote the transition probability from state one to state two and use  $q$  to denote the transition probability from state two to state one. Due to the nodes' mobility, both  $p$  and  $q$  are functions of the nodes' mobility profile and the channel parameters. So, the steady-state probabilities follow:

$$P_1(t, \rho_{i,j}) = \frac{q}{p+q}, \quad P_2(t, \rho_{i,j}) = \frac{p}{p+q}, \quad (10)$$

$\rho_{i,j}$  represents the combined mobility parameter of node  $i$  and node  $j$ ,  $t$  is a time variable.

### 3.3. Integrated decision-making function

This function is in charge of the final decision-making process. The adaptive routing parameters are derived from the decision-making algorithm. They are the number  $N$  of selected paths, the parity length  $k$  of FEC code, and the set  $\{R\} = \{r_1, r_2, \dots, r_N\}$  of the traffic distribution rate on each path. With this function, adaptive multipath QoS routing will be implemented.

Usually, a QoS request can be a delay constraint, a packet-loss constraint, or a minimum-bandwidth constraint. Many routing algorithms support QoS by guaranteeing only

one of those constraints [8]. We discuss a scheme supporting multiple constraints. It is assumed that a QoS request may include delay constraint, packet-loss constraint, and minimum-bandwidth constraint, all at same time or just one or two of them.

A packet-level FEC coding with lower coding rate may decrease packet loss [13, 24], but would increase packet delay in the network. In our scheme, the coding rate is determined under both end-to-end delay constraint and packet-loss constraint, which are guaranteed by using multiple paths. The increase of the delay caused by coding is also compensated by the gain from the parallel transmission mechanism in the multipath routing.

Our strategy is to build a multiobjective optimization function. We set two objective functions under the constraints in our adaptive multipath routing scheme to satisfy the QoS requirements. One objective function is to minimize the differences of the actual end-to-end delay in our scheme with its QoS requirements. Another is to minimize the difference of the actual end-to-end packet loss in our scheme with its QoS requirements. Both are constrained by the multipath transmission and traffic balance mechanisms in the scheme. Bandwidth requirement will be satisfied as a constraint of this optimization function if there is an option for it.

Following are the functions that will be used to derive the adaptive routing parameters  $N$ ,  $k$ ,  $\{R\}$  in our multiobjective optimization algorithm. The multiobjective functions are

$$\begin{aligned} \min \{D - \bar{D}(N, k, \{R\})\}, \\ \min \{P - \bar{P}(N, k, \{R\})\} \end{aligned} \quad (11)$$

subject to

$$\bar{D}(N, k, \{R\}) \leq D, \quad (12)$$

$$\bar{P}(N, k, \{R\}) \leq P, \quad (13)$$

$$B \leq \sum_{i=1}^N \min B_i, \quad (14)$$

$$\sum_{i=1}^N r_i = 1, \quad (15)$$

$$r_i \cdot \frac{P_i}{\min(B_i)} = r_{i+1} \cdot \frac{P_{i+1}}{\min(B_{i+1})}, \quad i = 1, 2, \dots, N-1, \quad (16)$$

where

$$\bar{D}(N, k, \{R\}) = \frac{\sum_{i=1}^N (D_i \cdot r_i)}{N} \cdot \frac{M}{k},$$

$$\bar{P}(N, k, \{R\}) = \bar{P} \cdot \left\{ 1 - \sum_{j=0}^k \binom{M-1}{j} \cdot \bar{p}^j \cdot (1-\bar{p})^{M-j-1} \right\},$$

$$\bar{P} = \sum_{i=1}^N (r_i \cdot P_i). \quad (17)$$

$D$ ,  $B$ , and  $P$  are the delay, bandwidth, and packet loss, respectively, requested by the QoS requirements.  $M$  is the FEC code length.  $D_i$  is the average delay on path  $i$ .  $P_i$  is the average packet-loss rate on path  $i$ .  $B_i$  is the minimum bandwidth on path  $i$ .  $\bar{P}$  is the average packet loss rate on the  $N$  paths before FEC decoding and  $\bar{P}(N, k, \{R\})$  represents the reconstructed packet loss probability after the FEC decoding.  $\bar{D}(N, k, \{R\})$  is the actual average end-to-end delay of the data packet after  $N$  parallel transmissions with FEC coding.

The above equations include several functions we considered in the algorithm. Expressions (12) and (13) present delay and packet-loss constraints under the end-to-end QoS guarantee, respectively.

For a general BCH code, the generator polynomial is a function of the code length,  $M$ . So, the parity length  $k$  will be derived from the highest power of the generator polynomial. One example of the BCH code is the Reed-Solomon (RS) code, which is the popular one used in packet level application [14]. The direct relations between  $M$  and  $k$  in RS code are

$$k = 2\gamma, \quad c(\gamma) = \frac{M - k}{M}, \quad (18)$$

where  $\gamma$  is correctable length in an RS code, and  $c(\gamma)$  is the code rate.

The recursive equation (16) is constructed to accommodate load balancing and fair packet loss in each path. We implement such a policy by two steps: (1) ordering our pre-computed paths from the largest to the smallest ratio of the packet loss and the minimum available bandwidth; (2) distributing the data stream to those ordered paths. The distribution rate is inversely proportional to the ratio of packet loss and minimum available bandwidth. Through this method, the data stream is spread evenly along the paths. More packets are switched to the lightly loaded and the least packet loss paths than the heavily loaded and large packet loss paths. Once a path fails during a rerouting interval, the packet loss rate and the distribution rate on this path are switched to new routes. The number of the new routes is derived from the same decision algorithm using previous parameters. The existing routes do not need any changes at this time. So, routing maintenance is realized without increasing computational complexity of algorithm.

To simplify, we use an iterative algorithm (see Algorithm 1) to look for appropriate adaptive parameters  $N$ ,  $k$ , and  $\{R\}$ . If one or two of the QoS requirements for delay, bandwidth, or packet loss are not requested, the initial values of the QoS requirement is set to infinity or zero, resulting in this requirement being skipped because it is always satisfied.

If integrated decision making is done in the source, the request and reply packets are built, respectively, as follows:

$$\begin{aligned} \text{request} \{ & \text{source, destination, sequenceID, QoSdelay, } D(p) \}, \\ \text{reply} \{ & \text{pathID, } D(p), L(p), B(p) \}. \end{aligned} \quad (19)$$

Step 1: initialize the number of the paths selected,  $N$ , and set the initial values of the QoS requirements.  
 Step 2: calculate the traffic distribution rate  $\{R\}$  of each path, based on its constraints (15), (16). These constraints are based on the current link status in the network.  
 Step 3: choose the parity length  $k$  of the FEC code from the calculations based on the delay and packet loss along each path according to (12) and (13).  
 Step 4: if the QoS requirements are not satisfied,  $N = N + 1$ , return to Step 2, a new path is identified to join the calculations.  
 Step 5: if link failed, return to Step 4.  
 Step 6: output  $N$ ,  $k$ , and  $\{R\}$ .

#### ALGORITHM 1

If integrated decision making is made in the destination, the request and reply packets are built, respectively, as follows:

$$\begin{aligned} \text{request} \{ & \text{source, destination, sequenceID, } D(p), L(p), B(p) \}, \\ \text{reply} \{ & \text{pathID, } N, k, \{R\} \}. \end{aligned} \quad (20)$$

If the final decision is made in the destination, the reply message cannot be sent out until  $N$  path request messages are all received, since  $N$  paths are needed to guarantee the packet loss and the packet delay. If the final decision is made in the source, the integrated decision-making algorithm can be started, once the first reply message is received. But the final decision still needs to be made after  $N$  path replies are received. It should be noted that the two decision-making schemes are equivalent at a decision-making node except for a slight difference in the overhead. Both cases will induce a processing delay (although this may be very small), due to different path delays.

## 4. SIMULATION AND DISCUSSION

This section presents some numerical and simulation results, showing the performance of the adaptive multipath routing. We present results in two categories. One part is for investigating the performance benefits obtained from the adaptive multipath routing. This discussion includes two aspects: improvement of the QoS performance regarding (1) the end-to-end delay and (2) packet loss, with comparing each feasible adaptive multipath routing scheme (AMPR) to dynamic source routing (DSR) and bandwidth-aware (BAR) QoS routing. Another part is for evaluating integrated performance of the scheme through defining three performance metrics (network control overhead, QoS redundancy, and QoS balance effect). The routing parameters (i.e., path number  $N$ , traffic distribution rate  $\{R\}$ , and FEC parity length  $k$ ) are also given in the numerical and simulation results.

During the numerical computation regarding the coding, we do not consider the relationship between forward error correction (FEC) parity length and the code length. We just assume that packets with fixed lengths are sent from sources, and then we calculate the parity length under the proposed scheme. This parity length will guarantee that the packet loss

TABLE 1: Parameters value.

Parameters	Value
Integrated traffic rate	0.5–3.0 Mbps (1 Mbps = 2604 packets/s)
Bandwidth	2000–6000 packets/s
Node coverage range	100–1000 m
Radio frequency	2.4e-9
Channel average SNR	1–20 dB
Radio-Tx power	10–20 db
FEC code length	13 + $k$ (packets)

requested by QoS requirements is satisfied. If certain coding techniques (e.g., Reed-Solomon code) were used at the packet level, then only selected parity length, calculated by the algorithm (i.e., those that form a valid code), can be used as the real parity length of the code. Different coding techniques have different relationships between the code lengths and their parity lengths. These relationships are determined by different generator polynomials. In practice, this coding constraint must be added to the algorithm.

We use two tables to describe our computation and simulation environments. Table 1 lists the parameter value used in computation and simulation. Table 2 illustrates the scenario used in GloMoSim2.0 simulation environment.

Figures 1, 2, 3, 4, 5, and 6 show two groups of results regarding the improvement of the QoS performance. Both compare the end-to-end QoS performance of the adaptive multipath routing scheme (AMPR) on the multiple paths to the performance of DSR and the scheme BAR in which only bandwidth constraint is considered without other protections (i.e., packet loss, delay, and traffic balance). The first group (Figures 1, 2, and 3) shows the performance changes with increasing node mobility when the different schemes are used. The second group (Figures 4, 5, and 6) shows performance changes with increasing average signal power (represented by Tx power) when the different schemes are used. To simplify the illustration, we fix the paths number, (i.e.,  $N = 2$ ) and only keep the  $k$  and  $\{R\}$  adjustable. Then, we observe one pair of the source and destination nodes in the network scenario. The  $k$  in these figures represents the FEC parity length selected by each feasible AMPR scheme when the network is in a different state, either node speeds are changed (i.e., Figures 2 and 3) or transmission powers are changed (i.e., Figures 5 and 6).

Figures 1 and 4 highlight the variability of the traffic distribution on the selected two paths. For example, in Figure 1, when the nodes' speed increases from 15 m/s to 20 m/s, the distribution rates on the two paths change because of the changes of QoS-related parameters in the selected paths. Obviously, the parity length and the distribution rate are adaptive to the network changes. As the communication parameters change, the adaptive nature of our scheme guarantees the end-to-end QoS performance.

Figure 2 presents the relationships between average end-to-end packet loss probability with the mobility of the nodes, where the line marked as DSR is the performance on one of the selected paths by using dynamic source routing, and that

TABLE 2: Simulation scenarios.

Terrain dimensions	(1200 m, 1200 m)
Simulation time	600 s
Mobile model	Random waypoint (speed 0–20 m/s)
Propagation path loss	Two-ray
Propagation fading model	Rayleigh, Ricean
Radio-Rx-type	SNR bounded
MAC protocol	802.11
Routing protocol	DSR
Node numbers	100
Traffic	CBR, 512 bytes

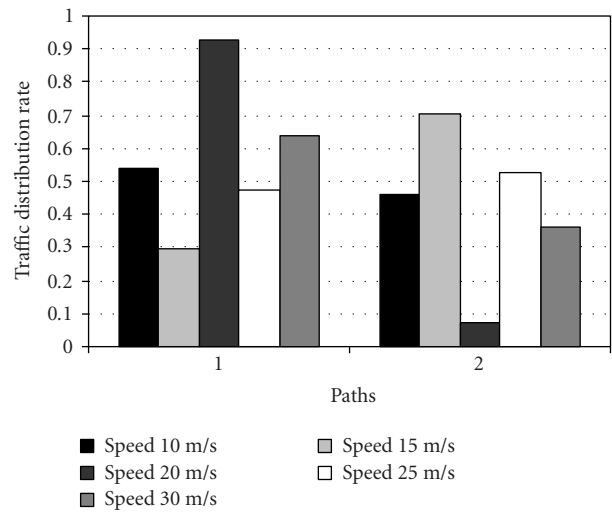


FIGURE 1: Traffic distribution rate on each path.

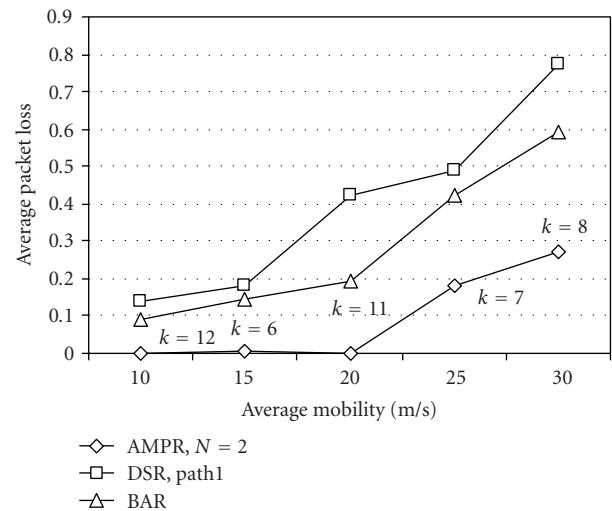


FIGURE 2: Average packet loss on paths with mobility.

marked as BER is a case where only bandwidth constraint is involved in the routing. The end-to-end packet loss is defined as the complementary value of the packet delivery ratio.

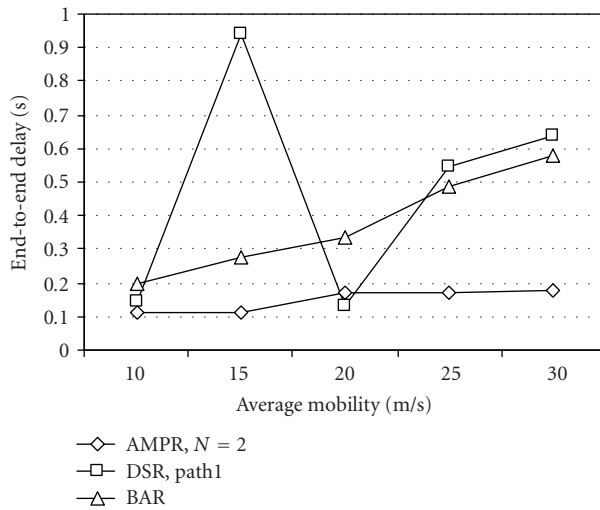


FIGURE 3: End-to-end delay on paths with mobility.

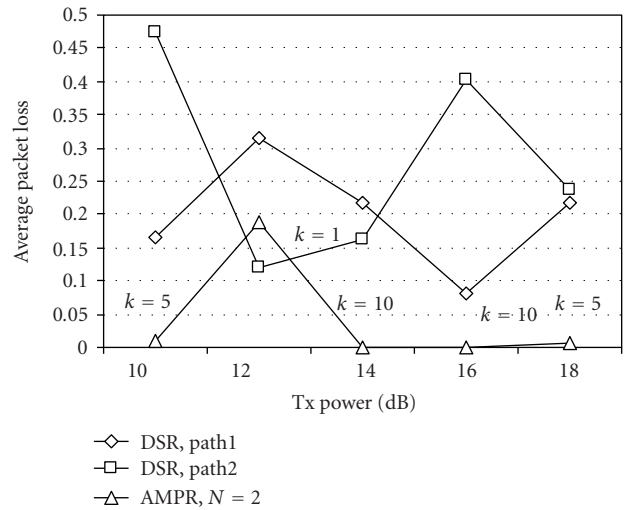


FIGURE 5: Average packet loss on paths with Tx power.

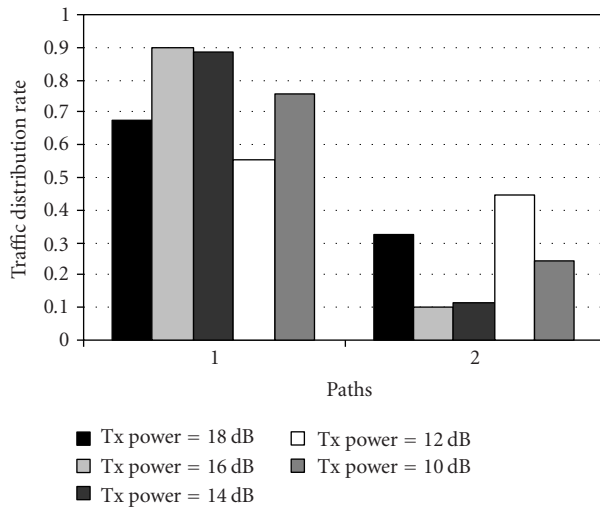


FIGURE 4: Traffic distribution rate on each path.

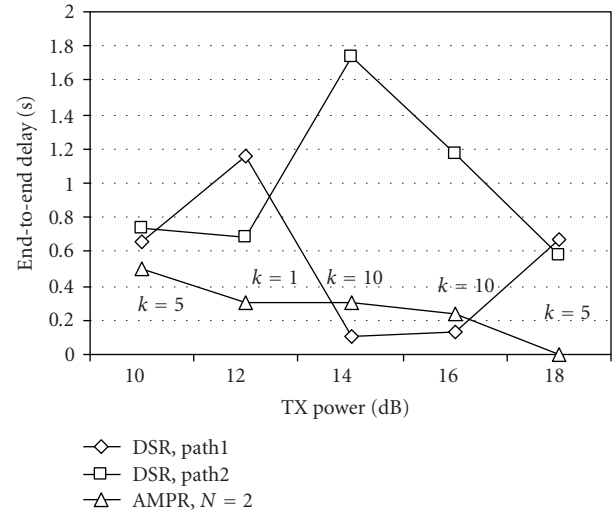


FIGURE 6: End-to-end delay on paths with Tx power.

This packet delivery ratio is obtained by dividing the number of data packets correctly received by the destination by the number of data packets originated by the source. Obviously, mobility increases the instability of the links so that mean packet loss increased with increasing of the nodes' velocity. However, due to the FEC protection of our AMPR scheme, packet loss rate is much lower than those in DSR and BAR schemes. The induced delay was complemented by the optimization algorithm in our QoS routing algorithm.

Figure 5 shows the variations of end-to-end delay of these schemes with the changes of nodes mobility. Same as in Figure 2, we can see performance goes worse when velocity goes higher. The DSR case shows large variability of delay on the path, and BAR presents an adaptive feature with the mobility but is not as good as the AMPR performed because AMPR balances QoS requirements and link performance by adjusting its routing parameters. As it can be seen in the

figures, parity length and traffic distribution rate dynamically change with the link state (including mobility, paths and power etc.) and QoS requirements.

Corresponding to the above figures, Figures 3 and 6 illustrate the relationships between average end-to-end delay or packet loss with the varying of signal power when using AMPR and DSR. It is clear that AMPR performs better than the others. Also it can see the redundancy on QoS performance (discussed in the following part) shown by the high protected packet loss. This is benefit of the knowledge about QoS requirement and the link status balanced by the distributed traffic and packet-loss protection.

The integrated performance of the AMPR scheme is studied as another aspect to show the performance benefit. We define three performance metrics to evaluate it. They are network control overhead, QoS redundancy, and QoS balance effect.



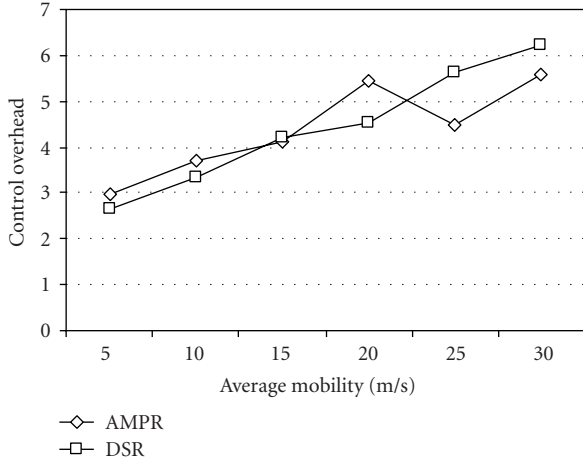


FIGURE 7: Network control overhead with node mobility.

Network control overhead is used to show the efficiency of the AMPR scheme. It is the ratio of the number of control messages propagated by every node and the number of the data received by the destinations. The definition is

$$\text{Overhead} = \frac{\text{Sum(CTRL/EachNode)}}{\text{Sum(ReceivedData/DesNode)}}. \quad (21)$$

Figure 7 shows the comparison of the adaptive multipath routing and the dynamic source routing. From this figure, we can observe that the overhead caused by DSR is less than that caused by multipath routing. This result is expected because searching for diverse multiple paths is more costly than searching for a single path using DSR. However, as mobility increases, multipath routing shows better than single path routing. The reason is that more route reconstructions are required for DSR than AMPR, due to more link failures resulting from higher node mobility.

In order to show the overhead caused by our extra QoS considerations, we redefine (21) in a way so that it is the actual ratio of the control bytes to the received bytes. This new ratio is required because the sizes of the request/reply packets are increased to include the local QoS information. Figure 8 presents the difference between the QoS multipath routing and the multipath routing without the QoS local information. Obviously, the QoS support in our scheme slightly increases the network's overhead.

QoS redundancy and QoS balance effect are defined to describe the integrated balance of the performance improvement. They are presented in Figures 9 and 10, respectively. Let

$$\text{diff}\{x\} = \text{QoSrequirement}\{x\} - \text{QoSperformance}\{x\}. \quad (22)$$

The following is the definition of the QoS redundancy:

$$\text{QoSredundancy} = \frac{\text{diff}\{\text{Delay, packetloss, bandwidth}\}}{\text{QoSrequirement}}. \quad (23)$$

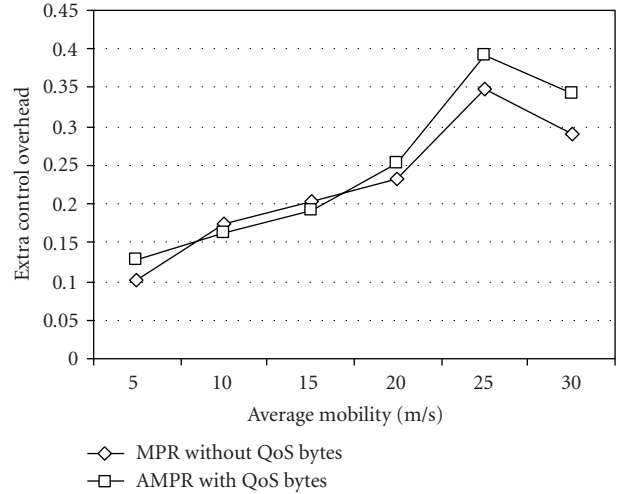


FIGURE 8: Extra control overhead from QoS information.

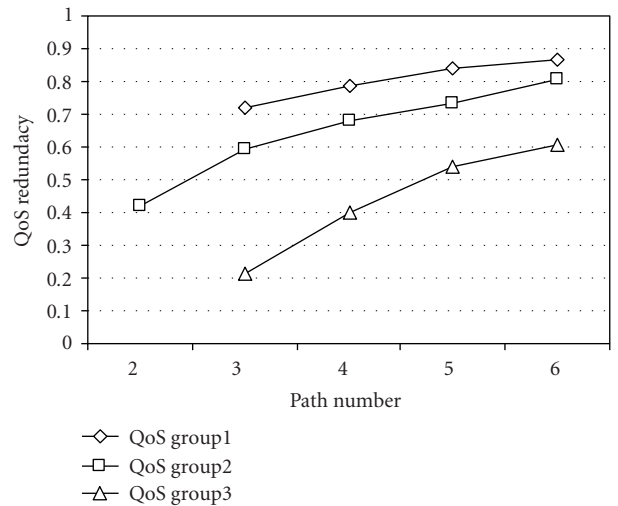


FIGURE 9: QoS redundancy versus path number.

This equation may be described by either delay or packet loss. In Figure 9, the delay parameter is used. QoS requirements vary from user to user. Obviously, QoS redundancy increases when path number increases under the same QoS requirement, because increasing the number of routing paths means decreasing the average packet loss and end-to-end delay due to the parallel transmissions when traffic load is light.

QoS balance effect is expressed as

$$\text{QoSbalanceeffect} = |\text{diff}\{\text{Delay}\} - \text{diff}\{\text{Packetloss}\}|. \quad (24)$$

This equation shows how adaptive multiple path routing ensures the balance between QoS performances without wasting network resources. From Figure 10, we can see that the difference of the effects between different QoS requirements is small. The QoS balance effect only increases slightly with increasing the number of selected paths. This means

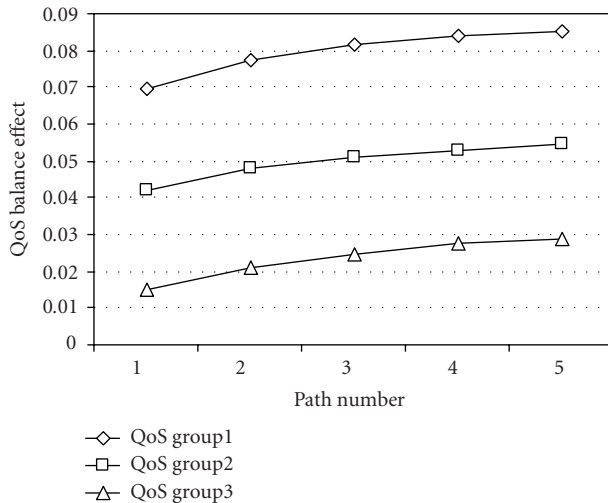


FIGURE 10: QoS balance effect versus path number.

that the adaptive multipath routing scheme will keep the balance between  $\text{diff}\{\text{Delay}\}$  and  $\text{diff}\{\text{Packetloss}\}$ .

## 5. SUMMARY

For QoS support in mobile ad hoc networks, we proposed an adaptive multipath routing scheme supported by a cross-layer cooperation mechanism. Using this scheme, performance satisfying QoS requirements is realized. Additionally, the forward error correction coding technique, along with a multiple-path routing algorithm, is implemented to satisfy the multiple QoS requirements. The adaptive routing is completed in a distributed manner based on local QoS performance provided by cross-layer mechanism. Three functions (routing function, local statistic computation and monitoring function, and integrated decision-making function) are implemented in the different parts of the mobile network. Due to the distributed structure, the computation and implementation complexity of the routing scheme are reduced. Also, since routes are discovered based on the up-to-date local information and selected by the optimization computation, routing parameters (e.g., number of paths, FEC parity length, and traffic distribution rate) are dynamic and optimized. In addition to supporting multiple QoS requirements, traffic balancing and bandwidth resources are factored into our decision-making process. The distributed structure of the local QoS statistics used in the routing enables this QoS support mechanism to be scalable in mobile networks. Our simulation results indicate that the performance (i.e., packet loss and end-to-end delay) are much better and less susceptible to the state changes (i.e., node mobility, transmission power, channel characteristics, and the traffic pattern) of the network, compared to a nonadaptive routing strategy.

## REFERENCES

- [1] E. M. Royer and C.-K. Toh, "A review of current routing protocols for ad hoc mobile wireless networks," *IEEE Pers. Commun.*, vol. 6, no. 2, pp. 46–55, 1999.
- [2] D. H. Lorenz and A. Orda, "QoS routing in networks with uncertain parameters," in *Proc. 17th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '98)*, vol. 1, pp. 3–10, San Francisco, Calif, USA, March–April 1998.
- [3] S. Chen and K. Nahrstedt, "Distributed quality-of-service routing in ad hoc networks," *IEEE J. Select. Areas Commun.*, vol. 17, no. 8, pp. 1488–1505, 1999.
- [4] B. Das and V. Bharghavan, "Routing in ad-hoc networks using minimum connected dominating sets," in *Proc. IEEE International Conference on Communications (ICC '97)*, vol. 1, pp. 376–380, Montreal, Quebec, Canada, June 1997.
- [5] S. H. Shah and K. Nahrstedt, "Predictive location-based QoS routing in mobile ad hoc networks," in *Proc. IEEE International Conference on Communications (ICC '02)*, vol. 2, pp. 1022–1027, New York, NY, USA, April–May 2002.
- [6] L. Xiao, J. Wang, and K. Nahrstedt, "The enhanced ticket-based routing algorithm," in *Proc. IEEE International Conference on Communications (ICC '02)*, vol. 4, pp. 2222–2226, New York, NY, USA, April–May 2002.
- [7] I. Gerasimov and R. Simon, "Performance analysis for ad hoc QoS routing protocols," in *Proc. IEEE International Mobility and Wireless Access Workshop (MobiWac '02)*, pp. 87–94, Fort Worth, Tex, USA, October 2002.
- [8] C. R. Lin and L. Chung-Ching, "An on-demand QoS routing protocol for mobile ad hoc networks," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM '00)*, vol. 3, pp. 1783–1787, San Francisco, Calif, USA, November–December 2000.
- [9] R. Dube, C. D. Rais, K.-Y. Wang, and S. K. Tripathi, "Signal stability-based adaptive routing (SSA) for ad hoc mobile networks," *IEEE Pers. Commun.*, vol. 4, no. 1, pp. 36–45, 1997.
- [10] Q. Xue and A. Ganz, "Ad hoc QoS on-demand routing (AQOR) in mobile ad hoc networks," *Journal of Parallel and Distributed Computing*, vol. 63, no. 2, pp. 154–165, 2003.
- [11] C. Zhu and M. S. Corson, "QoS routing for mobile ad hoc networks," in *Proc. 21st Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '02)*, vol. 2, pp. 958–967, New York, NY, USA, June 2002.
- [12] A. Tzirigios and Z. J. Haas, "Multipath routing in mobile ad hoc networks or how to route in the presence of frequent topology changes," in *Proc. IEEE Military Communications Conference (MILCOM '01)*, vol. 2, pp. 878–883, McLean, Va, USA, October 2001.
- [13] C. Huitema, "The case for packet level FEC," in *Proc. 5th IFIP International Workshop on Protocols for High-Speed Networks (PfHSN '96)*, pp. 109–120, Sophia Antipolis, France, October 1996.
- [14] A. E. Mohr, E. A. Riskin, and R. E. Ladner, "Unequal loss protection: graceful degradation of image quality over packet erasure channels through forward error correction," *IEEE J. Select. Areas Commun.*, vol. 18, no. 6, pp. 819–828, 2000.
- [15] H. Badis and K. A. Agha, "A distributed algorithm for multiple-metric link state QoS routing problem," in *Proc. International Workshop On Mobile and Wireless Communications Networks (IFIP MWCN '03)*, Singapore, Singapore, October 2003.
- [16] M. M. Al-Fawaz and M. E. Woodward, "Fast quality of service routing with multiple constraints," in *Proc. 8th IFIP Workshop on Performance Modelling & Evaluation of ATM & IP Networks*, Ilkley, UK, July 2000.

- [17] L. H. M. K. Costa, S. Fdida, and O. C. M. B. Duarte, "Distance-vector QoS-based routing with three metrics," in *Proc. Broadband Communications, High Performance Networking, and Performance of Communication Networks (Networking '00)*, pp. 847–858, Paris, France, May 2000.
- [18] Z. Ye, S. V. Krishnamurthy, and S. K. Tripathi, "A framework for reliable routing in mobile ad hoc networks," in *Proc. 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '03)*, vol. 1, pp. 270–280, San Francisco, Calif, USA, March–April 2003.
- [19] S.-J. Lee and M. Gerla, "Split multipath routing with maximally disjoint paths in ad hoc networks," in *Proc. IEEE International Conference on Communications (ICC '01)*, vol. 10, pp. 3201–3205, Helsinki, Finland, June 2001.
- [20] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*, pp. 153–181, Kluwer Academic, New York, NY, USA, 1996.
- [21] A. Nasipuri and S. R. Das, "On-demand multipath routing for mobile ad hoc networks," in *Proc. 8th IEEE International Conference on Computer Communications and Networks (ICCCN '99)*, pp. 64–70, Boston, Mass, USA, October 1999.
- [22] S.-T. Sheu and J. Chen, "A novel delay-oriented shortest path routing protocol for mobile ad hoc networks," in *Proc. IEEE International Conference on Communications (ICC '01)*, vol. 6, pp. 1930–1934, Helsinki, Finland, June 2001.
- [23] A. B. McDonald and T. F. Znati, "A mobility-based framework for adaptive clustering in wireless ad hoc networks," *IEEE J. Select. Areas Commun.*, vol. 17, no. 8, pp. 1466–1487, 1999.
- [24] M. Podolsky, C. Romer, and S. McCanne, "Simulation of FEC-based error control for packet audio on the Internet," in *Proc. 17th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '98)*, vol. 2, pp. 505–515, San Francisco, Calif, USA, March–April 1998.

recent research emphasis is on quality-of-service (QoS) issues in mobile ad hoc networks (MANETs). He is an Adjunct Professor in the Department of ECE at Georgia Institute of Technology and involved in research and consulting. He is affiliated to Georgia Centers for Advanced Telecommunications Technology, housed in the School of Electrical and Computer Engineering.

**Hongxia Sun** received the Ph.D. degree in computer and communication engineering from Southwest Jiaotong University, Chengdu, China, in 1999. From March 2000 to February 2002, she worked in Telecommunications Research Center at Arizona State University, and from February 2002 to August 2003 worked in Wireless and High Speed Networks Lab in Michigan State University. She is currently working in Computer Science Department at University of Calgary, Canada, as a Research Associate. Her research interests are in the areas of CDMA system, admission and access control, QoS provisioning and performance modeling, and evaluation for wireless networks.



**Herman D. Hughes** received his B.S. degree, M.S. degree, and Ph.D. degree from Stillman College, Tuskegee University and University of Louisiana at Lafayette, respectively. He began his career as a high school mathematics teacher in 1959, followed by faculty positions in mathematics at Tuskegee University and Grambling State University. During the summer months, he worked in both industry and government (IBM, American Oil Company, Dow Chemical, Advanced Avionics Research Lab). He joined the MSU faculty as an Assistant Professor in the Department of Computer Science and Engineering, 1973. He works in the area of network traffic management and modeling with respect to wireless and high-speed networks. His most

