Real-Time Traffic Support in Large-Scale Mobile Ad hoc Networks

Yuan Sun Elizabeth M. Belding-Royer
Department of Computer Science
University of California, Santa Barbara
{suny, ebelding}@cs.ucsb.edu

Xia Gao James Kempf
DoCoMo Communications Laboratories USA
San Jose, CA
{gao, kempf}@docomolabs-usa.com

Abstract

Ad hoc networks have been proposed for a variety of applications where support for real-time multimedia services will be necessary. This requires that the network is able to offer quality of service (QoS) appropriate for the latency and jitter bounds of the real-time application constraints. In this paper, we analyze the primary challenges of realizing QoS in large scale mobile ad hoc networks and propose a QoS framework for real-time traffic support. Specifically, our proposed QoS framework first utilizes a call setup protocol at the IP layer to discover paths for real-time flows, as well as to perform admission control by accurate service quality prediction. We then use a prioritized MAC protocol to provide priority access for flows with real-time constraints to reduce interference from unregulated non-real-time traffic. We foresee the utility of our proposed solution in large-scale ad hoc networks, such as campus or community-wide wireless networks. In these environments, fixed wireless routers may further be leveraged to achieve better service quality when node movement is significant. Through experimental results, we demonstrate the utility and efficiency of our approach.

1 Introduction

Wireless networking and multimedia content are two rapidly emerging technological trends. Among types of wireless networks, multi-hop ad hoc networks provide a flexible means of communication when there is little or no infrastructure, or the existing infrastructure is inconvenient or expensive to use. With the development of ad hoc networks, we can anticipate that multimedia applications will become popular in personal networks or other collaborative scenarios. For instance, large-scale wireless networks with thousands of mobile users have received an increase in deployment, where popular applications include VoIP, streaming multimedia, and peer-to-peer file sharing [11].

We target the support of real-time traffic with latency and jitter constraints in large-scale mobile networks, such as campus or community-wide networks, where students or residents move freely within the network and perform peer communications through PDAs and laptops. In these environments, fixed wireless routers may be placed, for instance, at classrooms, kiosks, etc., in a multi-hop mesh to serve as a network backbone. Such multi-hop wireless networks with stationary nodes have witnessed a recent increase in deployment [3, 13, 15, 19, 22] and are poised to become even more prevalent.

One major challenge of providing multimedia services is that certain quality of service (QoS) metrics must be satisfied. There has been significant research on providing QoS in wired networks. For instance, Intserv [25] and Diffserv [9, 20] are two well-known approaches. In wireless networks, however, several unique characteristics make QoS provisioning more challenging. These characteristics include the shared wireless medium, mobility, and the distributed multi-hop communication.

Most of the QoS solutions for wired networks rely on the availability of precise link utilization information. In ad hoc networks, however, all traffic within a mobile node's transmission range contends for medium access. Hence the shared nature of wireless communication channels makes resource estimation more difficult. Multi-hop interference introduces further challenges to the problem, making it complex to accurately determine the available resources. However, without sufficiently accurate resource prediction, it is difficult to provide multimedia services with satisfactory quality.

Node mobility also brings new obstacles to QoS ensurance in ad hoc networks. In general, node mobility has two impacts on network performance. First, the movement of a node on an active path often leads to a link break, and subsequently loss of packets. This effect is even more severe in large-scale networks with long communication paths. The packet loss after the link break, accompanied by increased packet transmission delay during a consequent route repair, significantly impacts the QoS of multimedia services. The second effect is the load increase due to node movement. Movement brings new traffic to the communication area, due in part to the moving node's ongoing traffic and also due to the temporary surge of control packets in the network during the route repair. Therefore, mitigating the impact of node mobility is important to maintain the service quality.

Finally, communication within an ad hoc network primarily occurs in a distributed fashion. There is no centralized node that can provide resource coordination for the network; every node is responsible for its own traffic and is unaware of other traffic flows in the network. Non-real-time traffic, i.e., those without service requirements, may be injected into the network and then interfere with ongoing real-time traffic. Minimization of the interference from non-real-time traffic is needed to ensure the service quality.

Bearing in mind these objectives, we propose a cross-layer QoS framework between the MAC and IP layers. In particular, the framework first utilizes a call setup protocol at the IP layer to discover paths for real-time flows, as well as to perform admission control by accurate service quality prediction. We then use a prioritized MAC protocol to provide priority access for flows with real-time constraints to reduce interference from unregulated non-real-time traffic. We foresee the utility of our proposed solution in large-scale ad hoc networks, such as campus or community-wide wireless networks. In these environments, anticipated applications such as Instant Messaging, IP telephony, and interactive distance learning lectures all require quality of service provisioning. If the ad hoc network includes Internet access points, real-time services from or to the Internet can also be provided with the needed quality.

The remainder of this paper is organized as follows. Section 2 describes related work. Section 3 presents our proposed framework. Specifically, Section 3.1 first describes our targeted environment and the needed routing modification for this environment. Section 3.2 presents the call setup process for real-time flows through accurate service quality prediction. Section 3.3 explains the prioritized access protocol used to reduce interference and alleviate the impact of node mobility. The performance of our proposed approach is evaluated in Section 4 and finally Section 5 concludes the paper.

2 Related Work

Many routing schemes and frameworks have been proposed to provide QoS support for ad hoc networks [2, 7, 8, 18, 26]. Among them, INSIGNIA [18] uses an in-band signaling protocol for distribution of QoS information. The information is included in the IP headers of the data packets, and the available resources are calculated at each station the packet traverses based on “soft-state” traffic reservation information. SWAN [2] improves INSIGNIA by introducing an Additive Increase Multiplicative Decrease (AIMD)-based rate control algorithm. Both [7] and [8] utilize a distance-vector protocol to collect end-to-end QoS information via either flooding or hop-by-hop propagation. CEDAR [26] proposes a core-extraction distributed routing algorithm that maintains a self-organizing routing infrastructure, called the “core”. The core nodes establish a route that satisfies the QoS constraints on behalf of other nodes.

None of these approaches significantly diverge from QoS approaches for wired networks, and they do not completely address the differences between wired and wireless networks. Specifically, they often do not consider the contentious nature of the MAC layer, nor the neighbor interference on multi-hop paths. This leads to inaccurate path quality prediction for real-time flows. Additionally, most of the work does not consider the fact that a newly admitted flow may disrupt the quality of service received by ongoing real-time traffic flows. Furthermore, service differentiation is often desired in ad hoc networks. Most of the solutions do not provide an accurate quality estimation when flows of multiple priorities exist.

Recently, other work has proposed the performance improvement of MAC protocols and the support of service differentiation. Many of these approaches specifically target IEEE 802.11 [27]. For example, studies in [1, 6, 12, 16] propose to tune the contention windows sizes or the inter-frame spacing values to improve network throughput, while studies in [1, 4, 14, 23, 31] propose priority-based scheduling to provide service differentiation. Most of this work utilizes different backoff mechanisms, different DIFS lengths, or different maximum frame lengths, based on the priority of the traffic. Based on this previous work, we propose a priority access mechanism in our framework that considers the current network status, i.e., the current channel collision probability, in determining the backoff behavior of different priority traffic. This enables us to achieve adaptive service differentiation.

Our model of utilizing a fixed wireless backbone is similar to recent commercial deployments of multi-hop wireless networks, such as “rooftop” and “community wireless networks” [3, 15, 22]. Other companies are field-testing multi-hop wireless networks that use stationary or minimally mobile nodes to provide broadband Internet access [13, 19]. In these solutions, all traffic flows only through these designated wireless routers. In our work, however, traffic may flow through the fixed wireless routers, or direct connections between the mobile users can be leveraged. The decision of which path is selected is application-dependent. Furthermore, our work focuses on the QoS aspect of providing real-time services in ad hoc networks, which is not specifically addressed in these solutions.

3 QoS Framework

3.1 Targeted network environment and routing modification

Our targeted large-scale ad hoc network consists of mobile users such as students and residents carrying PDAs and laptops. Fixed or minimally mobile nodes may be placed at strategic locations to provide network backbone access.

The majority of proposed ad hoc routing protocols place no preference on the selection of paths with regard to node mobility, i.e., highly mobile nodes have the same likelihood of inclusion on a communication path as stationary nodes. Routes consisting of highly mobile nodes, for instance, students bicycling on campus, will change frequently. Consequently, a path that satisfies a flow’s QoS requirements may not last for the entire data session due to link breaks. When links break, packet loss occurs, and packet delivery latency

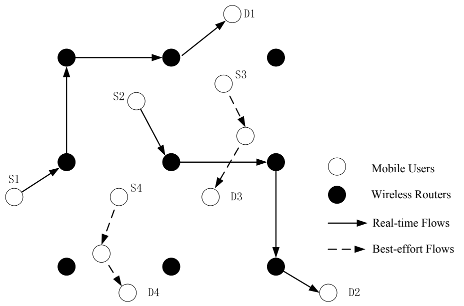


Figure 1. An example of the routes for different traffic.

increases during route discovery. This also results in significant packet jitters due to both delay during the route discovery and delay variations between the new and old routes. This effect is more severe in large ad hoc networks where the longer communication paths are more prone to break when mobility is present. While best-effort traffic may be more tolerant to these events, the quality of real-time traffic will be significantly degraded and is likely to become unacceptable.

The utilization of fixed wireless routers in these networks will greatly improve the quality of real-time traffic by the elimination of intermediate link breaks. These routers can provide a valuable backbone that is crucial to QoS support in large mobile networks. To leverage these stationary wireless routers, an ad hoc routing protocol must be modified to prefer routes that include fixed nodes. Otherwise, there is no guarantee that the fixed nodes will be selected, and so the benefit will not be obtained.

Using the AODV routing protocol [21] as an example, we propose modifications to the routing protocol to reflect the selection of stationary routes for real-time traffic. Specifically, when a source node initiates route discovery for real-time traffic with strict quality requirements, only the fixed routers respond to the control packets by either forwarding the RREQ, or unicasting a RREP. The mobile nodes do not respond to these packets, unless they are the destination. In this case they reply with a RREP. This modification to the routing protocol can be achieved by marking the data packet with a real-time traffic label. For instance, the ToS field in the IP header can be set. Routing control packets, such as RREQs and RREPs, need to include this information accordingly. This way, combined with the path quality prediction method described in Section 3.2, real-time traffic will flow through a stable route that consists of fixed nodes only. This increases the likelihood that the QoS requirements can be satisfied. If a route through fixed nodes is not discovered on the first attempt, the second route discovery can alleviate the fixed node requirement. For non real-time traffic, i.e., best-effort traffic, both the fixed routers and the mobile nodes can participate in route discoveries as in unmodified AODV. Figure 1 illustrates an example network where real-time and best-effort traffic utilize different routes.

It can be argued that only the fixed wireless routers should be used for all traffic within the ad hoc network, since they provide the most stable paths. This, however, will not result in high spatial reuse throughout the network and will consequently result in less utilization of the network capacity. Note that the addition of fixed relay nodes may not be necessary for small ad hoc networks. Applications without real-time constraints also may not need the benefits of the stable routes provided by fixed wireless routers. However, for large ad hoc networks, the introduction of fixed relay points will greatly improve the quality of service by the avoidance of path breaks, and the subsequent reduction of packet loss. This will result in low packet jitter, which is essential for real-time traffic.

Given the targeted large-scale ad hoc networks, we now explain how real-time traffic can be supported using the proposed QoS Framework. Specifically, we first describe how to predict the flow quality for real-time traffic and subsequently accomplish call admission. We then present our solution to reduce the interference from non-real-time traffic to provide the needed quality to real-time traffic.

3.2 Call setup for real-time traffic

When a real-time flow is requested, a call setup process is needed to acquire a valid transmission path with satisfied QoS requirement. Call setup also enables effective admission control when the network utilization is saturated. This requires accurate estimation of channel utilization and prediction of flow quality, i.e., throughput or transmission delay.

The proposed QoS approach is based on our previous work of a model-based resource estimation mechanism, called *MBRP* [29]. By modeling the node backoff behavior of the MAC protocol and analyzing the channel utilization, MBRP provides both per-flow and aggregated system-wide throughput and delay. The basic premise of MBRP is to provide quality prediction for both ongoing traffic and new flows so that a correct flow admission decision can be made according to the quality of service policy of the network. The input of the analytical model in MBRP is a set of flows $A = \{a_1, a_2, ..., a_s\}$ in the network, where s denotes the total number of priority classes supported by the system, and $\forall a_i \in A$, a_i is the number of flows of priority class i. The output of the model is the average throughput or delay calculation for each flow. To mitigate the effects such as hidden terminals, unexpected collisions, etc., the MBRP model further utilizes measurement results as run-time feedback to improve the estimation accuracy. A detailed description of the service quality prediction using analytical modeling can be found in [29].

Call setup

The above-mentioned MAC layer estimation function provides channel statistics for a node's local contending area. Such a resource estimation based on the MAC layer modeling can better capture the packet scheduling behavior of the wireless access medium and provide more realistic predic-

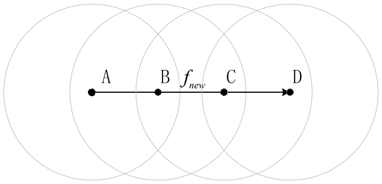


Figure 2. An example topology.

tion than can reservation information at the IP layer. However, because paths typically consist of multiple hops, a local decision is not sufficient for the setup of an entire transmission path. Furthermore, due to interference from neighboring nodes, the resources available to a new flow consist of the minimum of the available resources in the neighborhood of nodes on the path.

For example, in Figure 2, the circles indicate the transmission range, and hence neighborhood, of each node. Suppose node A requests a new flow using the path $A \rightarrow B \rightarrow C \rightarrow D$, and the resource consumption of the flow is x. In this case, the resource consumption is actually $2 \times x$ at nodes A and C, and $3 \times x$ at node B. This is because all nodes within transmission range of each other contend for the shared medium access. Therefore, the actual resource consumption is not just the requirement of the flow, but the resources consumed in the neighborhood of all the nodes along the transmission path.

Hence, the functionality of the call setup process must be to first analyze the interference relationship among the nodes in the potential transmission path, as well as to disseminate the requirements of the flow along the path. Then, based on the potential flow set information, the estimated throughput or delay can be calculated using the analytical model described in [29]. Finally, once the information is propagated to the source, the source can choose the path that best meets the flow's QoS requirement.

In our solution, we base the call setup process on the modified AODV routing protocol described in Section 3.1, which can be divided into a Request and a Reply phase, as shown in Figures 3 and 4. In the request phase, the source node sends Route Request messages (RREQ) for the new flow. RREQ messages, indicated by solid arrows, include QoS information such as the traffic class of flow, the required quality, and the minimum throughput or accumulated delay through previous hops. Each intermediate node, upon reception of the RREQ packet, adds a pending record for this flow in its routing table and rebroadcasts the RREQ if the flow is locally admissible. This indicates that the predicted quality of the new flow meets the needed service requirement, i.e., the minimum available bandwidth through the previous hops is larger than the flow's throughput requirement, or the accumulated delay over previous hops is smaller than the latency requirement. If the requirement cannot be satisfied, the RREQ packet is dropped. Intermediate nodes notify neighbors about the potential load through the broadcast of Neighbor Reply messages (NREP), indicated by the dotted lines in Figure 3. The flow information, disseminated by NREP packets, is needed to determine the input to the model. The RREQ packet finally reaches the destination if a path with the needed quality exists.

During the reply phase, the destination node sends a Route Reply message (RREP) along the reverse path to the source node, as shown in Figure 4. In this phase, intermediate nodes have updated neighbor load information through the NREP packets transmitted in the Request phase. They can now more accurately recompute the predicted quality of the flows and forward the RREP if the new flow is locally admissible. The source node then selects an optimal path based on the path quality. The nodes along the selected path also send NREP packets to confirm the admitted flow status with their neighbors. In this way, all nodes that are affected by the new flow receive updated channel utilization information.

Handling call setup failures

A source node may not be able to find a valid path using the fixed routers if the ongoing traffic over those routers results in a nearly congested wireless medium. In this case, there are several possibilities to still accommodate a new QoS session.

1. The new flow backs off for a given interval and tries to set up the call after the waiting time. This is a simple solution whereby nodes passively wait for the network resource until some ongoing session ends.

2. The requested flow lowers its service requirement level. For instance, it can request a lower bit rate. In some cases, this will allow the needed quality of the flow to be met by the network.

3. If more intelligence is provided by the network, the mobile user with the new flow can leverage its mobility capability to improve its quality of the service by productively moving toward some less congested area [24]. Since the wireless routers are placed in fixed locations, the network can indicate the location of another less congested wireless router if the resource utilization information about other peers is known.

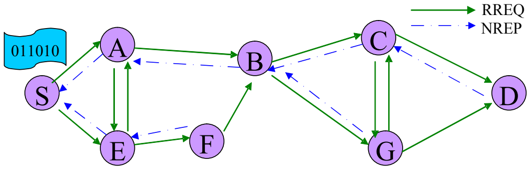


Figure 3. Request phase.

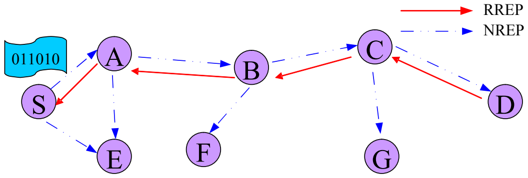


Figure 4. Reply phase.

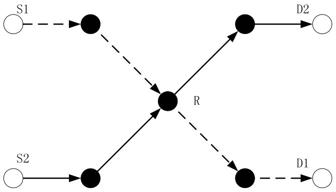


Figure 5. Intra-node interference at fixed wireless router R.

The mobile user can then choose to move toward that router to complete its service via the new route.

In our simulation, we explore the performance of our proposed approach using the first solution. We plan to investigate the other two solutions in future work.

3.3 Prioritized medium access

The use of fixed wireless routers, together with accurate resource estimation, allows service calls for real-time traffic to be established if valid paths exist. However, communication in ad hoc networks occurs in a distributed fashion. There is no centralized point that can provide resource coordination for the network; every node is responsible for its own traffic and is unaware of other traffic. Consequently, best-effort traffic, which traverses both mobile nodes and fixed routers, will interfere with the real-time traffic.

The impact of interference becomes more severe as mobility increases. As described in Section 1, the first effect of node mobility on intermediate link breaks can be mitigated through the utilization of fixed wireless routers. However, the second effect of the load increase due to node movement is more difficult to alleviate. Movement brings new traffic to the communication area, e.g., the moving node's ongoing best-effort traffic, thereby resulting in increased contention and degraded quality for real-time traffic.

Interference among different types of traffic

There are two general scenarios where best-effort traffic significantly interferes with real-time traffic, resulting in reduced quality for real-time flows.

The first scenario occurs when the best-effort traffic and the real-time traffic share the same fixed wireless routers, thereby causing *intra-node contention*. Best-effort flows may select routes via the fixed wireless routers if there is available capacity. However, this can result in a new real-time flow being unable to be admitted because there is not enough remaining capacity in the paths through the routers. Figure 5 illustrates such an example, where a best-effort flow (indicated in the dotted line) traverses the fixed wireless router R. Later, a real-time flow (indicated in the solid line) is unable to be admitted because the router R does not have enough capacity to support it. To prevent this from occurring, it would be beneficial to reserve routers for the real-time traffic. The best-effort traffic does not require use of the fixed routers since it is more tolerant to delays. Hence it can choose alternate paths through mobile nodes. By distributing the traffic load, network capacity is increased.

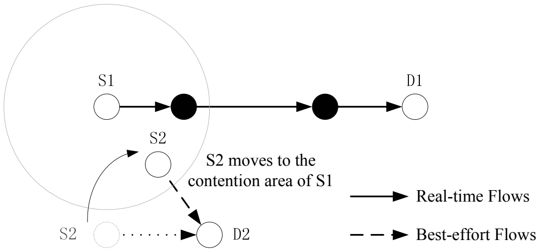


Figure 6. Mobility caused inter-node interference.

The second scenario occurs due to node mobility. Mobile users with unregulated best-effort traffic may move around freely. A pre-established high-quality path will experience degraded performance when nodes with ongoing flows move into the contention area. We call this *mobility-induced inter-node contention*. An example is shown in Figure 6. Since it is not realistic to prevent users from moving, a more flexible medium access mechanism is needed.

To reduce the contention caused by unregulated best-effort traffic and alleviate its impact on the quality of real-time traffic, we propose an adaptive service differentiation mechanism at the MAC layer to achieve prioritized medium access.

Adaptive priority scheduling

Our proposed priority scheduling algorithm is based on IEEE 802.11 [27]. Currently, there are several approaches that propose to provide service differentiation based on 802.11, by either assigning different minimum contention window sizes (CW_{min}), Arbitrary Inter Frame Spacings (AIFS), or backoff ratios, to different types of traffic. These approaches can all provide differentiation; however, the parameters are typically statically assigned and cannot adapt to the dynamic traffic environment. This reduces the usage efficiency of the network. For instance, if low priority traffic is configured to use a pre-defined large backoff window, it will experience longer service latency even when there is no competing high priority traffic presented in the network. On the other hand, a small static value for high priority traffic may result in more collisions and backoffs when multiple high priority flows compete for channel access, thereby reducing the channel efficiency. Hence, it is difficult to find suitable static values to achieve a good trade-off between the needs of service differentiation and global efficiency, given an unknown and dynamically changing traffic composition in the network.

To this end, we propose an adaptive scheme to address this trade-off. The basic idea is that, because the state of ad hoc networks can vary greatly due to mobility and channel interference, it is advantageous to adjust the backoff behavior according to the current channel condition. Specifically, mechanisms for avoiding collisions can be developed. Given a high traffic load in the network, the number of collisions and subsequent packet retransmissions significantly affects the throughput and packet delivery latency [16]. Hence, it is beneficial to consider the collision rate in the backoff scheme.

To achieve service differentiation, as well as to adapt to the current network usage, we combine the collision rate with the exponential backoff mechanism in IEEE 802.11. We have:

$$f_{pri} = Rand[0, (2^r + R_{col} \times pri) \times CW_{min}] \times T_{slot} \quad 0 \leq r \leq m \tag{1}$$

where R_{col} denotes the collision rate between a station's two successful frame transmissions, and pri is a variable associated with the priority level of the traffic. By applying Eq. (1), traffic with different priority levels will have different backoff behavior when collisions occur. Specifically, after a collision occurs, low priority traffic will backoff for longer, and subsequently high priority traffic will have a better chance of accessing the channel.

For intra-node interference, a new real-time flow can be serviced when the prioritized access mechanism is applied, even when there are ongoing best-effort flows sharing the same path. For mobility-induced inter-node interference, when a best-effort flow contends with a real-time flow, the channel collision rate will increase. By applying the prioritized mechanism, the best-effort traffic will experience a longer backoff time than the real-time traffic and will become less likely to access the channel. Therefore, the impact of the best-effort traffic on the real-time flows is mitigated.

Simulation results in Section 4 indicate that adaptive priority scheduling can provide better quality of service and reduce interference. Detailed comparison results to current approaches such as EDCF can be found in [28].

3.4 Summary of the framework

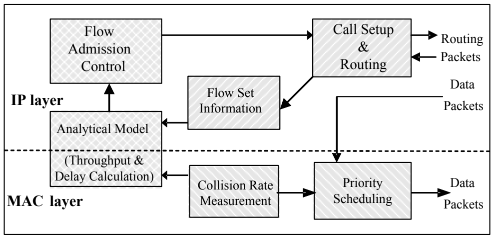


Figure 7. Functionalities of the framework at IP and MAC layers.

We now briefly summarize the functionality of our proposed framework, as well as our implementation architecture. Figure 7 depicts the corresponding modifications at the IP and MAC layers inside a node, as well as the high-level interaction between them. Upon reception of a flow request, the call setup process builds an end-to-end path using the routing process described in Section 3.1. If the request is for real-time traffic, a flow admission control component checks whether the quality requirement of the new flow can be met without bringing quality degradation to ongoing traffic through the fixed wireless routers in the network. Specifically, this is achieved based on the resource estimation information of the path, which is the output of our analytical model. The input of the model is the flow information disseminated along the route setup, as well as the run-time collision rate measurement results. Note that the model calculation for admission control is only performed at fixed wireless routers. After a route with satisfied quality for a new flow is set up, the MAC layer priority scheduling component described in Section 3.3 schedules data packet transmissions according to the packet's priority level.

4 Experimental Study

We have implemented the proposed approach in the NS-2 [10] simulator and conducted experiments to verify our resource estimation model and evaluate the effectiveness of our QoS routing protocol. The experimental studies focus on examination of the efficiency of our framework in real-time traffic support in large-scale ad hoc networks.

4.1 Experiment setup

Table 1. Parameters for Simulated Networks

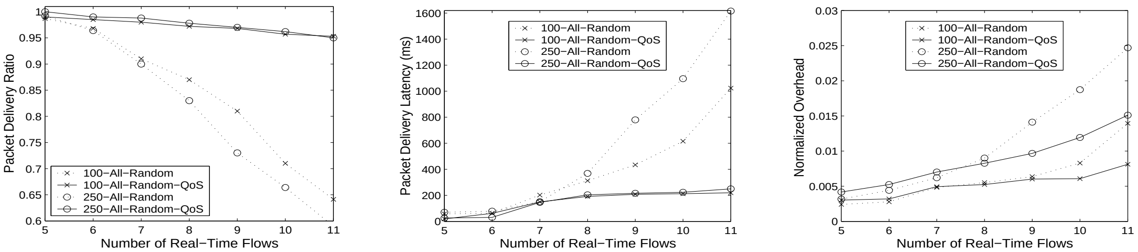
Number of Nodes	Network Size	Average Path Length
50	1000m × 1000m	3.3
100	1500m × 1500m	4.7
250	2400m × 2400m	8.3

We study networks of different sizes with randomly placed nodes. We increase the size of the network area as the number of nodes grows to keep an average node density of about seven neighbors per node [17]. The parameters of each network are shown in Table 1. Communication node pairs are randomly chosen among all the nodes. Generated traffic includes both real-time traffic and best-effort traffic, with characteristics as shown in Table 2. For real-time traffic, we model voice over IP data encoded with G.711. Each data point represents the average result of ten runs with different seeds and each run is executed for 300 seconds.

Table 2. Traffic Parameters

Traffic Type	Priority	Packet Size (bytes)	Data Rate (Kbps)
Real-Time (G.711 VoIP)	High	160	64
Best-Effort (CBR)	Low	500	80

The first set of simulations demonstrates the effectiveness of our QoS solution using resource prediction and call setup in a static network. For comparison, we also include performance results of an unmodified AODV routing protocol, i.e., no QoS provisioning is performed in this case. The second set of simulations examines the impact of mobility by introducing node movement. All nodes move according to random waypoint mobility model, except that during a real-time session between a randomly chosen source and destination pair, both the end nodes are static. This is a reasonable assumption because typically users do not move or move minimally when they are engaged in a multimedia session, such as typing an Instant Message, viewing a lecture, or watching



(a) Packet Delivery Ratio. (b) Average Delivery Latency. (c) Protocol Overhead.

Figure 8. Average received quality of real-time flows.

a multimedia stream [11]. Once the session ends, the users resume normal movement.

In the second simulation set, we first evaluate the received quality of real-time flows both with and without QoS provisioning using random mobility. We call the former scenario *All-Random* and the latter *All-Random-QoS*. In this case, all nodes are randomly placed in the network and they are moving according to the random mobility model. We then include fixed wireless routers in the network to serve as the relay nodes for real-time traffic. We call this scenario *Fixed-Random* when QoS is not used, and *Fixed-Random-QoS* when QoS is incorporated. For simplicity, fixed wireless routers are placed in a grid 200m apart. More advanced placement techniques [30] could also be used. In these two simulation sets, real-time traffic sessions are generated at intervals of 5 seconds. Each session lasts for 100 seconds.

The third set of simulations evaluates the performance gains of the prioritized scheduling algorithm when varying numbers of real-time and background traffic flows co-exist in large-scale mobile networks. Fixed wireless routers placed in a grid structure are also used in this simulation set.

4.2 Performance metrics

The efficiency of our proposed framework is evaluated through the following performance metrics:

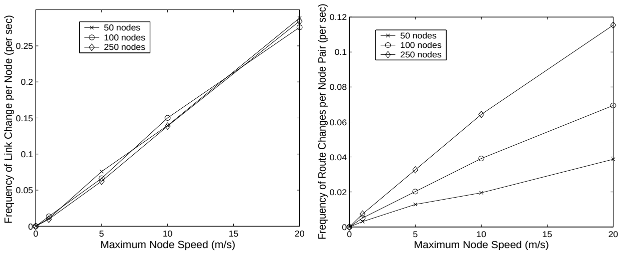
- **Packet Delivery Ratio**: the average fraction of transmitted data packets that are successfully delivered at the destination.
- **Average Transmission Delay**: the average end-to-end delivery latency from the source to the destination.
- **Jitter**: the delay variation between consecutive packets. This is an important metric for real-time traffic. A smaller jitter indicates a higher quality flow.
- **Protocol Overhead**: the number of routing control packet transmissions per data packet delivered in the network. Each hop-wise transmission of a routing packet is counted as one transmission.

4.3 Static network

This set of simulations evaluates the effectiveness of our resource prediction and QoS setup in a static network. No fixed wireless routers are introduced in the network.

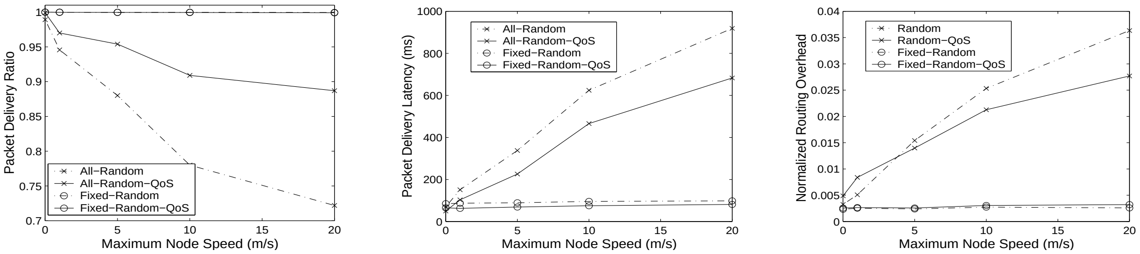
Figure 8 shows the performance received by the flows as the number of real-time flows increases in static networks with 100 and 250 nodes. Due to the lack of accurate quality prediction and admission control, the packet delivery ratio drops significantly as the traffic increases when quality of service is not implemented, as shown in Figure 8(a). With our modified QoS solution, the packet delivery ratio for admitted flows does not experience significant degradation. This is because the admission of new flows is delayed when the predicted flow quality does not meet the required quality, or when unsatisfactory quality for ongoing sessions is predicted. Figure 8(b) shows the average end-to-end delay of received packets. Our QoS approach reduces the delay as compared to the non-QoS solution. As shown in Figure 8(c), the normalized control overhead of our QoS solution is slightly higher than without QoS, due to the additional control packet exchanges during call setup, i.e., NREP packets. However, when traffic is heavy, our QoS approach has less control overhead. When no call admission is performed, the network is more congested, resulting in very lossy links. The frequent link breaks consequently trigger more route discoveries, thereby resulting in higher overhead.

In terms of network size, the results show that a larger network is not necessarily able to support more concurrent real-time flows. This is because in a larger network a flow needs to traverse longer paths, resulting in increased latency and decreased delivery ratio.



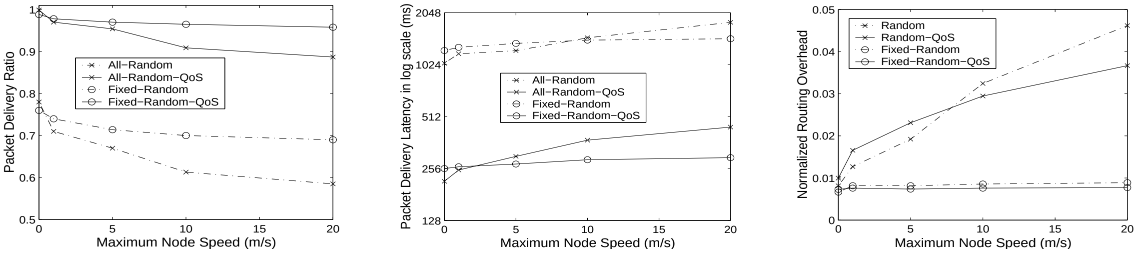
(a) Frequency of Link Changes. (b) Frequency of Route Changes.

Figure 9. Impact of Mobility.



(a) Packet Delivery Ratio. (b) Delay. (c) Protocol Overhead.

Figure 10. Average quality of flows with 5 sessions in a 250 node network.



(a) Packet Delivery Ratio. (b) Delay. (c) Protocol Overhead.

Figure 11. Average quality of flows with 10 sessions in a 250 node network.

4.4 Mobile network

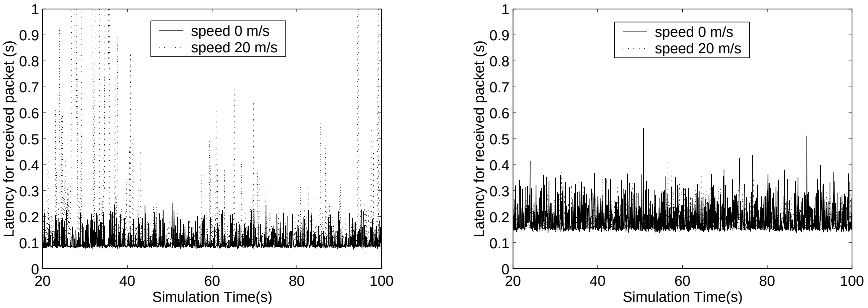
This set of simulations examines the impact of node mobility on our QoS provisioning solution and evaluates the benefits of fixed wireless routers.

Figure 9(a) shows the frequency of link breaks with varying node speed. The frequency of link breaks for all networks is similar, i.e., the average number of link breaks per second for a node is approximately constant, since the network density also remains constant. As the nodes move faster, the frequency of link changes increases accordingly. The average number of route changes, on the other hand, increases much faster as the network scales up, as shown in Figure 9(b). As node mobility increases, the frequency of route changes within larger networks grows dramatically faster than that of smaller ones. In large networks, the average path length is longer. A single link break between two nodes may lead to the breakage of an entire path. Hence, as the network size scales, the impact of mobility is more significant.

Figure 10 presents the performance results of a 250 node network with five real-time sessions. Here, the network load is relatively light. As we can see in the figures, our QoS approach achieves better performance than the non-QoS routing protocol due to the flow quality prediction. As mobility increases, the solutions using the mobile nodes as routing relays experience a significant drop in packet delivery ratio, as well as an increase in latency, due to the intermediate link breaks. Leveraging the fixed wireless routers, on the other hand, mitigates the impact of mobility by avoiding link breaks in the intermediate nodes. Similar to the static network, the control overhead of our QoS solution when random mobile nodes are used for routing is not higher than the non-QoS solution in general, as shown in Figure 10(c). This is due to its avoidance of network congestion. When fixed wireless routers are utilized, the overhead is further reduced because network-wide control packet flooding is avoided.

Figure 11 shows the results of 10 real-time sessions, i.e., the network load is heavier. QoS provisioning helps to improve the delivery ratio and decrease the packet latency by rejecting the admission of a new flow when the network is at full capacity. The rejected flows are serviced when the current flows end. However, the quality of all flows degrades when mobility increases. When the amount of traffic is high, reliance on fixed wireless nodes alone cannot provide satisfactory flow quality if no admission control is enabled. However, combining these two techniques, i.e., enabling admission control along the paths of fixed wireless routers, results in a high packet delivery ratio with low delivery latency. Again, the control overhead of the fixed-random approach is significantly lower than the random one. We also notice that when there is no mobility, the All-Random approach results in slightly lower packet delivery latency, as shown in Figure 11(b). This is because the hop-count diameter of $O(\sqrt{n})$ in grid-like structures is larger than the $O(\ln(n))$ diameter for randomly connected structures [5].

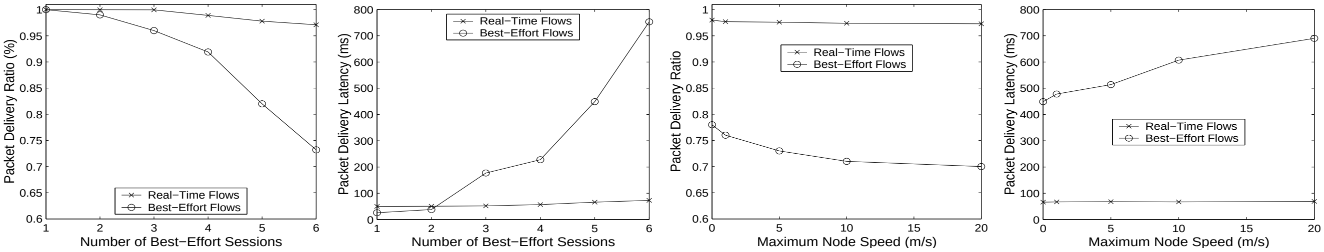
Figure 12 illustrates a sample latency trace of an admitted real-time flow of the QoS provisioning scheme both with



(a) All-Random-QoS.

(b) Fixed-Random-QoS.

Figure 12. Sample Latency Traces for 250 node networks.



(a) PDR with increasing number of best-effort traffic.

(b) Delay with increasing number of best-effort traffic.

(c) PDR with varying node speed.

(d) Delay with varying node speed.

Figure 13. Results for mixed traffic in a 100 node network.

and without fixed wireless routers. In a static network, the QoS protocol is effective in achieving a small delay variance, i.e., small jitter. However, in a mobile environment, the frequent broken paths and subsequent route discoveries result in a large delay variation, thereby causing high jitter. With the assistance of fixed wireless routers, the delay variance is much less than that of the total random scenario. When the nodes are not moving, the All-Random approach provides slightly lower packet delivery latency than the Fixed-Random approach. This can be explained again because the hop-count diameter of $O(\sqrt{n})$ in grid-like structures is larger than the $O(\ln(n))$ diameter for randomly connected structures.

4.5 Mixed traffic

This set of simulations evaluates the protocol performance for a mixed set of network traffic. The traffic set includes both real-time traffic and best-effort traffic. The number of real-time traffic sessions is five. Our QoS solution, including admission control, fixed wireless routers, and prioritized scheduling, is evaluated in the simulations.

Figures 13(a) and (b) show the quality of both flow types in a 100 node static network as the number of best-effort sessions increases. As we can see, the packet delivery ratio of the real-time traffic does not drop significantly when the number of best-effort flows increases and the network becomes more congested. Accordingly, the average delay of the real-time traffic remains roughly constant, as shown in Figure 13(b). On the other hand, the quality of the best-effort traffic degrades when the traffic level of the network is high. There is a significant drop in PDR and a surge in the delay. The results indicate the effectiveness of the prioritized medium access mechanism in providing higher priority to real-time traffic. This helps to meet the quality requirement of the real-time flows.

Figures 13(c) and (d) show the quality of both types of flows with varying node movement. There are five best-effort sessions. As we explained in Section 3.3, node movement can bring unregulated best-effort traffic into contention with the real-time traffic. The results show that the packet delivery ratio of the real-time traffic remains constant as nodes move, as does the average packet latency. This illustrates the effectiveness of our prioritized scheduling mechanism in reducing the interference from best-effort traffic, thereby providing satisfactory quality to real-time flows.

5 Conclusion

This paper proposes a QoS framework to provide real-time traffic support in large-scale mobile networks. Specifically, the framework first utilizes a call setup protocol at the IP layer to discover paths for real-time flows, as well as to perform admission control by accurate service quality prediction. We then use a prioritized MAC protocol to provide priority access for flows with real-time constraints to reduce interference from unregulated non-real-time traffic.

We foresee the utility of our proposed solution in large-scale ad hoc networks, such as campus or community-wide wireless networks. In these environments, fixed wireless routers may further be leveraged to achieve better service quality when node movement is significant.

References

[1] I. Ada and C. Castelluccia. Differentiation Mechanisms for IEEE 802.11. In *Proceedings of the IEEE Conference on Computer Communications (INFOCOM)*, Anchorage, Alaska, April 2001.

[2] G.-S. Ahn, A. Campbell, A. Veres, and L.-H. Sun. Supporting Service Differentiation for Real-Time and Best-Effort Traffic in Stateless Wireless Ad Hoc Networks (SWAN). *IEEE Transactions on Mobile Computing*, 1(3):192–207, July-Septemeber 2002.

[3] Bay area wireless users group. http://www.bawug.org.

[4] A. Banchs, X. Perez-Costa, and D. Qiao. Providing Throughput Guarantees in IEEE 802.11e Wireless LANs. In *Proceedings of the 18th International Teletraffic Congress (ITC-18)*, Berlin, Germany, September 2003.

[5] B. Bollobas. *Random Graphs*. Academic Press, 1985.

[6] F. Calí, M. Conti, and E. Gregori. Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit. *IEEE/ACM Transactions on Networking*, 8(6), December 2000.

[7] S. Chen and K. Nahrstedt. Distributed Quality-of-Service Routing in Ad-Hoc Networks. *IEEE Journal of Selected Areas in Communications*, 17(8), August 1999.

[8] T. Chen, M. Gerla, and J. Tsai. QoS Routing Performance in a Multi-hop, Wireless Network. In *Proceedings of the IEEE ICUPC'97*, San Diego, CA, October 1997.

[9] D.Grossman. New Terminology and Clarifications for Diffserv. Request For Comments (Draft Stardard) 3260, Internet Engineering Task Force, April 2002.

[10] K. Fall and K. Varadhan. ns Manual. *http://www.isi.edu/nsnam/ns/doc/*. The VINT Project.

[11] T. Henderson, D. Kotz, and I. Abyzov. The Changing Usage of a Mature Campus-wide Wireless Network. Technical Report TR2004-496, Dartmouth College, Computer Science, Hanover, NH, March 2004.

[12] T. S. Ho and K. C. Chen. Performance Evaluation and Enhancement of CSMA/CA MAC Protocol for 802.11 Wireless LANs. In *Proceedings of the IEEE PIMRC*, Taipei, Taiwan, October 1996.

[13] Mesh Networks Inc. Mesh Networks Technology Overview. http://www.meshnetworks.com.

[14] V. Kanodia, C. Li, A. Sabharwal, B. Sadeghi, and E. Knightly. Distributed Multi-Hop Scheduling and Medium Access with Delay and Throughput Constraints. In *Proceedings of the Seventh Annual International Conference on Mobile Computing and Networking (MobiCOM'01)*, Rome, Italy, July 2001.

[15] R. Karrer, A. Sabharwal, and E. Knightly. Enabling Large-scale Wireless Broadband: The Case for TAPs. In *Proceedings of HotNets 2003*, Cambridge, MA, November 2003.

[16] H. Kim and J. C. Hou. Improving Protocol Capacity with Model-based Frame Scheduling in IEEE 802.11-operated WLANs. In *Proceedings of the Ninth Annual International Conference on Mobile Computing and Networking (MobiCOM'03)*, pages 190–204, San Diego, CA, September 2003.

[17] L. Kleinrock and J. Silvester. Optimum Transmission Radii for Packet Radio Networks or Why Six is a Magic Number. In *Proceedings of the IEEE National Telecommunications Conference*, Birmingham, Alabama, December 1978.

[18] S. Lee, G.-S. Ahn, X. Zhang, and A. T. Campbell. INSIGNIA: An IP-Based Quality of Service Framework for Mobile Ad Hoc Networks. *Journal of Parallel and Distributed Computing, Special issue on Wireless and Mobile Computing and Communications*, 60:374–406, 2000.

[19] Nokia Networks. Nokia rooftop wireless routing. http://www.americasnetwork.com/americasnetwork/data/ articlebrief/americasnetwork/412003/34898/article.pdf, 2001.

[20] K. Nichols and B. Carpenter. Definition of Differentiated Services Per Domain Behaviors and Rules for Their Specification. Request For Comments (Draft Stardard) 3086, Internet Engineering Task Force, April 2001.

[21] C. E. Perkins and E. M. Royer. Ad-hoc On-Demand Distance Vector Routing. In *Proceedings of the* 2nd *IEEE Workshop on Mobile Computing Systems and Applications*, pages 90–100, New Orleans, LA, February 1999.

[22] MIT Roofnet. http://www.pdos.lcs.mit.edu/roofnet.

[23] R. Rozovsky and P. Kumar. "SEEDEX: A MAC Protocol for Ad Hoc Networks. In *Proceedings of the* 2nd *ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc'01)*, Long Beach, CA, October 2001.

[24] K. Sanzgiri and E. M. Belding-Royer. Leveraging Mobility to Improve Quality of Service in Mobile Networks. In *Proceedings of The First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services*, Boston, MA, August 2004.

[25] S. Shenker, C.Partridge, and R. Guerin. Specification of Guaranteed Quality of Service. Request For Comments (Draft Stardard) 2212, Internet Engineering Task Force, September 1997.

[26] P. Sinha, R. Sivakumar, and V. Bharghavan. CEDAR: A Core-Extraction Distributed Ad Hoc Routing Algorithm. In *Proceedings of the IEEE Conference on Computer Communications (INFOCOM)*, pages 202–209, New York, NY, 1999.

[27] IEEE Computer Society. IEEE Standard for Wireless LAN-Medium Access Control and Physical Layer Specification, Novermber 1999.

[28] Y. Sun, E. M. Belding-Royer, X. Gao and J. Kempf. A Priority-based Distributed Call Admission Protocol for Multi-hop Wireless Ad hoc Networks. Technical report 2004-20, Computer Science Department, UC Santa Barbara, August 2004.

[29] Y. Sun, X. Gao, E. M. Belding-Royer, and J. Kempf. Model-based Resource Prediction for Multi-hop Wireless Networks. In *Proceedings of the* 1st *IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS)*, Ft. Lauderdale, FL, October 2004.

[30] Wireless Valley. http://www.wirelessvalley.com.

[31] A. Veres, A. T. Campbell, M. Barry, and L.-H. Sun. Supporting Service Differentiation in Wireless Packet Networks Using Distributed Control. *IEEE Journal of Selected Areas in Communications*, 19(10), October 2001.