Channel-Adaptive Relaying in Mobile Ad Hoc Networks with Fading

Michael R. Souryal* and Nader Moayeri
Wireless Communication Technologies Group
National Institute of Standards and Technology
Gaithersburg, Maryland
Email: {souryal,moayeri}@nist.gov

***Abstract*— This paper describes an approach for relaying in multihop networks that adapts to the time-varying channel and exploits spatial diversity to mitigate multipath fading. Ignored in some performance analyses, fading arises from multipath propagation and causes fluctuations in the signal strength in mobile networks, adversely affecting communication performance. Our approach uses limited cross-layer interactions between the physical, link and routing layers to provide adaptivity to both large and small-scale channel effects and to achieve spatial diversity gain without the use of multiple antennas. The routing layer uses long-term measurements of link quality in the form of the average signal-to-noise ratio (SNR) to opportunistically select next-hop relays on a hop-by-hop basis. Small-scale variations are overcome at the MAC layer through efficient multicast polling of multiple next-hop candidate relays prior to data transmission. A performance analysis for networks employing geographic routing and an IEEE 802.11-based MAC (i) demonstrates significant improvements in network capacity and end-to-end delay achieved with these channel-adaptive techniques in Rayleigh and Ricean fading environments, (ii) shows that most of the small-scale diversity gain is obtained through the use of only two next-hop relay choices, and (iii) assesses the practical limit of the short-term adaptive component in terms of maximum node velocity.**

I. Introduction

Among the channel propagation effects present in most terrestrial mobile networks in practice is multipath fading. Caused by the constructive and destructive summing of multipath signal components of differing phase arising from reflections, diffraction and scattering, fading can result in large fluctuations in signal strength across distances on the order of the carrier wavelength. Short-term variations in the signal strength of 10-20 dB due to multipath fading are not atypical and can cause a link to experience intermittent behavior. Most ad hoc network routing protocols rely on the consistent and stable performance of individual links, and therefore intermittent links can result in high packet loss rates and control overhead [1].

A common means of mitigating the adverse effects of fading is to use some form of diversity in which the signal is transmitted through multiple channels, each of which experiences uncorrelated fading, ideally. The use of diversity typically involves a cost, such as additional bandwidth or a lower data rate in frequency or temporal diversity, or additional antennas in conventional spatial diversity. However, multihop ad hoc networks possess an inherent source of spatial diversity that is often overlooked: the availability of multiple next-hop relay alternatives at any given forwarding node in a packet's path to its destination. In this paper, we describe and analyze a cross-layer approach to relaying in ad hoc networks that adapts to the time-varying channel and exploits the inherent spatial diversity in these networks without the use of multiple antennas per node. This approach utilizes existing multipath routing protocols along with measurements of the channel state made by the physical layer with appropriate metrics to opportunistically select one of a choice of next-hop relays to which to forward a packet. Whereas multipath routing as described in [2] sends a partitioned packet simultaneously over multiple routes, packets in our approach follow a *single*, albeit potentially changing route end-to-end.

Other routing approaches incorporate channel state information of some form. For example, signal stability-based adaptive routing is a source-initiated on-demand routing protocol that favors the selection of stronger, more stable links in route discovery [3]. Adaptive multipath routing in [4] forms multiple paths during route discovery and uses periodic control packets to monitor the quality of each path and select the best path. A similar approach is used in receiver-initiated [5] and bandwidth-guarded channel-adaptive routing [6]. These protocols differ from our approach in that they use channel state information to construct or select end-to-end routes as a whole, whereas we use channel-adaptive relaying on a hop-by-hop basis with the ability to adapt to short-term channel variations that occur on a shorter timescale than can be captured using end-to-end metrics.

Previous analytical and numerical performance studies have shown a substantial potential for performance improvement using channel-adaptive relaying [7], [8]. Here, we present and analyze an implementation of channel-adaptive relaying for a geographic routing protocol operating over an IEEE 802.11-based MAC layer. An implementation for a multipath version of AODV routing is described in [9], where performance is evaluated for a linear topology with a slow fading channel. Our approach uses channel measurements to adapt to long-term channel variations due to path loss and shadowing on one timescale and to adapt to small-scale fading on a shorter

*This research was performed while the author held a National Research Council Research Associateship Award at the National Institute of Standards and Technology.

0-7803-9012-1/05/$20.00 (C) 2005 IEEE

timescale. End-to-end performance is evaluated for both fixed and mobile networks for a range of channel speeds, with an eye towards finding the point at which the short-term adaptive protocol is unable to keep up with the channel variability.

II. Fading and Diversity in Ad Hoc Networks

A. Effect of Fading

The impact of multipath propagation that leads to fading is sometimes modeled and predicted for specific environments using ray-tracing, which requires detailed knowledge of the paths of reflection a signal follows and the material characteristics of the reflectors. For more general environments, however, a stochastic model is more often applied. For example, in the case of frequency non-selective (flat) fading, when the differential delays of the multipath components are small compared to the symbol duration, the effect of fading can be described by a single, multiplicative complex random process. Furthermore, when the number of multipath components is large and the scattering is isotropic, the fading process is modeled as Gaussian with independent and identically distributed in-phase and quadrature components, giving rise to Rayleigh fading on non-line-of-sight links and Ricean fading on line-of-sight links. While the Rayleigh and Ricean distributions describe the instantaneous statistics of the envelope of the fading process, the autocorrelation function of the envelope describes its statistics over time. Considering two samples in time of a given fading process, the magnitude of their correlation generally decreases with the product of the time between the samples and the maximum Doppler frequency, the latter being given by $f_m = v/\lambda$, where v is the relative velocity of the transmitter and receiver, and λ is the carrier wavelength. For example, the correlation of two samples of a Rayleigh fading envelope taken at t and $t+\tau$ is given by

$$\rho(\tau) = J_0^2(2\pi f_m \tau) \tag{1}$$

where $J_0(x)$ is the zero-order Bessel function of the first kind [10].

To illustrate the dynamic nature of fading in time, Fig. 1 shows the squared envelope as a function of time of three uncorrelated Rayleigh fading waveforms generated numerically using the modified Jakes model [11]. The maximum Doppler of each waveform is 40 Hz, which would correspond, for example, to a 2.4 GHz carrier and a 5 m/s node velocity. One observes that during the 20 ms window of this example, the gain of a given waveform varies by as much 18 dB, and that the peak-to-valley transition occurs in as little as 8 ms. Depending on the long-term average power of a link, a short-term change in signal strength of 18 dB can significantly impact the reliability of a link.

To provide some perspective on the timescale of these variations with respect to practical transmission times, consider the duration of a single packet transmission attempt over an IEEE 802.11 link. Including the request-to-send (RTS)/clear-to-send (CTS) handshake and subsequent acknowledgement, the transmission of a 120-byte data unit would require approximately 1.6 ms at the 2 Mbps data rate. While the signal strength would be relatively stable during a single transmission attempt, varying by a few decibels at most in this example, it could change dramatically over just a few packet transmissions, especially when contention delays are present. Thus, channel measurements that must propagate over multiple hops before they are acted upon for adaptive routing can easily become outdated and useless for adapting to multipath fading.

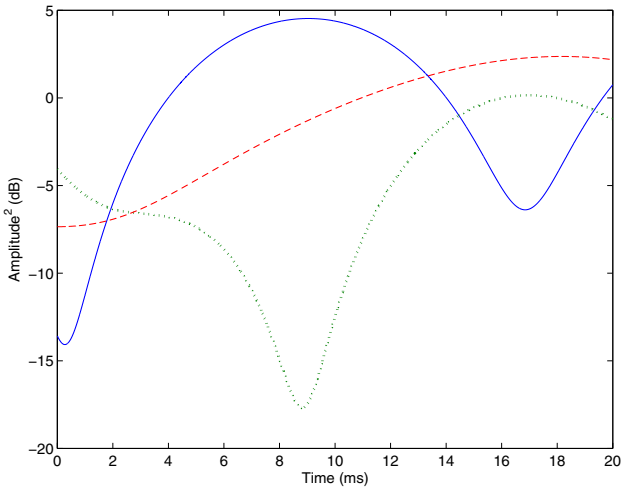


Fig. 1. Sample uncorrelated Rayleigh fading waveforms for a 2.4 GHz carrier with a relative tx-rx velocity of 5 m/s

B. Diversity as a Countermeasure

From wireless communication theory, we know that diversity is an effective means for mitigating the adverse effects of fading. In general, diversity is the protection provided against fading by transmission of the signal through multiple channels, each of which experiences uncorrelated or minimally correlated fading, ideally. In that case, the likelihood of transmission failure—namely, that all the channels are in a poor state simultaneously—is greatly reduced. Diversity is typically obtained in one of three forms: temporal, spatial and frequency, referring to whether the channelization is in time (e.g., repeated transmissions[1]), space (e.g., multiple antennas), or frequency (e.g., multiple carriers).

At the receiver, the multiple received replicas of the signal are combined to yield a decision statistic. Common combining techniques include (i) equal gain combining, in which the replicas are coherently summed at the physical layer, (ii) maximal ratio combining, in which the signals are weighted by their channel amplitudes before being summed, and (iii) selection combining, in which the strongest signal is selected while the others are ignored. The benefits of diversity in terms of increased link reliability are significant. For example, the probability of symbol error in Rayleigh fading with L independent channels decreases inversely with the Lth power of the SNR [12].

[1]Channel coding is another means for obtaining temporal diversity in non-static fading channels, and can be viewed as a generalization of symbol repetition using more sophisticated redundancy.

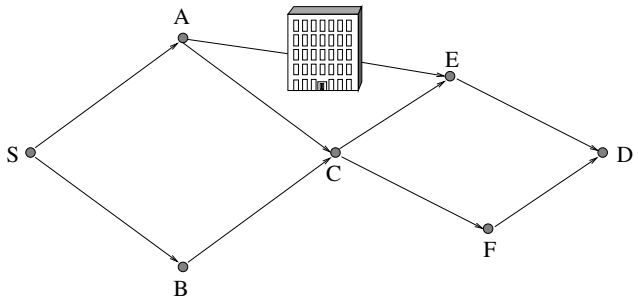


Fig. 2. Sample topology with multiple next-hop relays

C. Multiple Next-Hop Spatial Diversity

As with any other wireless communication system, diversity can be built into the physical layer of ad hoc networks to combat fading. Typically, though, the implementation of diversity incurs the cost of additional resources needed for the provision of multiple channels, whether it be additional bandwidth for frequency diversity, additional time slots (or, equivalently, a lower data rate) for temporal diversity, or additional antennas (or, a larger form factor) for spatial diversity. However, multihop ad hoc networks possess an inherent source of diversity against fading that is often overlooked and does not incur these traditional resource costs.

In a multihop ad hoc network, there typically are multiple route alternatives from a source to a destination. Some of these routes utilize different neighbors of the source as the first-hop (e.g., links S-A and S-B in Fig. 2). Viewing the second node in the chosen route as a new source to the same destination, it too may have multiple neighbors from which to choose for route selection, and so on, until a next-to-last node in the route has the destination as one of its neighbors. The availability of multiple next-hop choices at a node, whether it be the source node or an intermediate node along the route, presents a natural, built-in mechanism for obtaining diversity against fading. In effect, the multiple next-hop relays can serve as the multiple channels of a spatial diversity system.

We describe and analyze a system for achieving multiple next-hop spatial diversity in which a node forwards a packet to *only one* of its next-hop relays. The choice of next-hop relay is made as a function of the channel states of the next-hop links. A major challenge in using next-hop relays as a source of diversity is the reaction time of relay selection relative to the time variability of the channel. To improve responsiveness to small-scale channel variations, the choice is not fixed over multiple packet transmissions but is made independently for each packet forwarded by that node. A design objective is that any overhead associated with this selection be kept minimal.

The potential impact of next-hop spatial diversity on communication reliability can be significant. Referring again to the Rayleigh fading waveforms of Fig. 1, consider a scenario in which the three waveforms represent the fading processes of three next-hop link choices for a transmitting node. Though all three waveforms have the same average power over a long period of time, at any given time their instantaneous powers differ by as much as 22 dB. Enabling the transmitting node to select from these links at any given time can easily make the difference between a failed transmission and a successful one. In addition, though not addressed in this paper, adaptive modulation and coding can be used on better links to increase spectral efficiency for greater throughput and/or lower delay.

The next-hop diversity scheme described here is analogous to a conventional spatial diversity scheme consisting of a single transmitting antenna, multiple receiving antennas, and selection combining (i.e., strongest signal selection). However, instead of having the receiver equipped with multiple antennas, here the candidate next-hop relays represent a distributed antenna array. Next-hop spatial diversity differs from conventional spatial diversity in another important way. While conventional spatial diversity with multiple antennas combats small-scale fading, next-hop diversity protects against both small-scale fading and larger-scale, more slowly varying channel effects such as shadowing and path loss with distance. For example, save for the higher energy collected by the extra antennas, a multiple antenna transceiver provides no benefit in the face of severe signal attenuation due to a large obstruction (e.g., a building or mountain in the propagation path). However, next-hop diversity can potentially provide an alternate path averting the obstruction using the fortuitous locations of one or more neighbors (e.g., path A-C-E in Fig. 2).

The next two sections describe channel-adaptive relaying with multiple next-hop relays more fully, describing how a next-hop relay selection is made and detailing a specific implementation based on existing protocols.

III. Channel-Adaptive Relaying

This section first discusses the type of routing protocols that permit channel-adaptive relaying and several metrics for selecting a relay. It then describes channel-adaptive relaying at two levels. The first, large-scale adaptive relaying, adapts to the slowly varying large-scale channel attenuation effects due to path loss with distance as well as shadowing. The second, small-scale adaptive relaying, attempts to adapt to the multipath fading variability of the channel which can occur on a much smaller timescale.

A. Suitable Routing Protocols

Early on-demand routing protocols determine a single end-to-end route which packets follow for as long as is necessary and the route is available. Channel-adaptive relaying as described here, however, relies on a new class of routing protocols that have the following two characteristics:

1) Any node that forwards a packet may have a choice of relays from which to select the next-hop recipient of that packet.
2) The choice of next-hop relay can change from one packet to the next routed by that node.

Many *geographic*, or position-based, routing protocols [13] inherently possess these two characteristics and are therefore suitable for channel-adaptive relaying. Geographic routing protocols require that each node knows its own geographic

coordinates as well as those of its neighbors. Furthermore, the source must know the position of the final destination, obtained through a required location service. In the greedy mode of a geographic routing protocol, a forwarding node compares the coordinates of the packet's destination, contained in the packet header, with those of its neighbors and transmits the packet to a neighbor that is closer to the final destination. In case no neighbor is available closer than itself to the destination, a loop-free fall-back protocol is used to route around voids in the network. The evaluation of potential relays and forwarding selection is performed for every packet. One example of a geographic routing protocol that operates in this way is Greedy Perimeter Stateless Routing (GPSR) [14].

While geographic routing is a natural candidate for channel-adaptive relaying for multiple next-hop diversity, there is a class of non-geographic, on-demand routing protocols that are also amenable. Generally referred to as *multipath on-demand* protocols, they are variations of existing on-demand protocols that provide alternate paths for faster, more efficient recovery from route failures as well as for load balancing. Of particular interest, here, are protocols that form multiple paths at intermediate nodes as well as at the source. One example, Ad Hoc On-demand Multipath Distance Vector (AOMDV) routing [15], is an extension of AODV that permits nodes to accept multiple route advertisements (RREQ, RREP) and record a routing table entry for each one, subject to loop-free constraints. When forwarding a packet, a node can have multiple next-hop links from which to choose.

In this paper, we concentrate on an implementation of channel-adaptive relaying for geographic routing. The concept can be extended to multipath on-demand routing in analogous fashion but with the absence of position information.

B. Maximizing Expected Progress

As noted earlier, geographic routing algorithms select a next-hop relay based on position information. For example, in its greedy forwarding mode, GPSR chooses the neighbor that is closest to the packet's destination. From a set of neighbor nodes, $\mathcal{N}$, that are closer to the destination than the forwarding node, this criterion selects the relay j as follows:

$$j = \arg\min_{i \in \mathcal{N}} \{d_{i,D}\}$$

where $d_{i,D}$ is the Euclidean distance between neighbor i and the destination, D.[2] Expressed slightly differently, this selection criterion is equivalent to maximizing the geographic *progress* toward the destination:

$$j = \arg\max_{i \in \mathcal{N}} \{d_{S,D} - d_{i,D}\} \quad (2)$$

where $d_{S,D}$ is the distance between the forwarding node and the destination.

The maximum progress metric (2) favors the selection of next-hop relays that reside at the edge of a node's transmission range. However, because signal strength generally decreases with distance, these next-hop links also tend to be among the least reliable links available, on average.

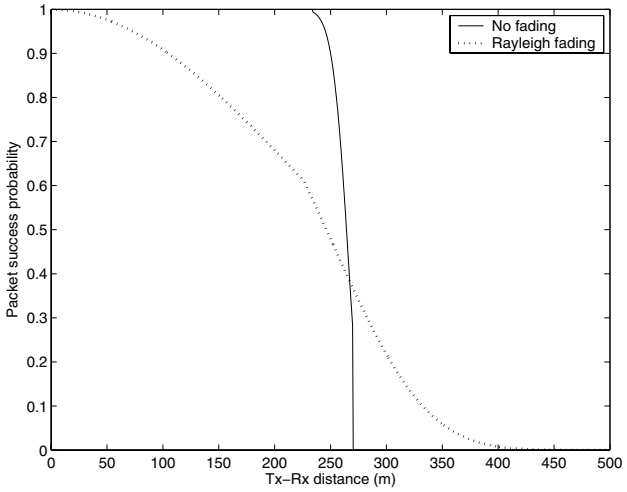


Fig. 3. Packet success probability versus distance for a 512-byte packet on a 2 Mbps IEEE 802.11 link

To quantify link reliability, let us consider the probability of a successful packet transmission for a given system. Fig. 3 plots the probability of a successful packet transmission versus the transmitter-receiver distance for a 512-byte packet transmitted on a 2 Mbps IEEE 802.11 link experiencing path loss and additive white Gaussian noise, with and without Rayleigh fading. The non-fading curve was generated using the bit error probability for differential quadrature phase shift keying (DQPSK) given in [12], and the fading curve is based on the packet success probability derived in the appendix. In this example, transmission power is 4 dBm, noise power is -102 dBm, the receiver sensitivity is -89 dBm, and the large-scale path loss model is two-ray ground reflection lower-bounded by free space path loss at short distances.

We observe from Fig. 3 that, without fading, packet success probability drops rapidly at a certain transmission range (around 260 meters in this case), while with Rayleigh fading it decays gradually. Therefore, in realistic environments with multipath fading, link quality is not a hard quantity, being either on or off, but rather diminishes gradually with distance. Intermediate link quality has also been observed in practice in various field trials [1], [16], [17].

The fact that link reliability is a soft quantity suggests that selecting a relay with maximum progress may be a poor choice, leading to retransmissions which increase delay and consume network capacity.[3] This problem is exacerbated by the fact that potential relays can be discovered on the basis of `HELLO` packets that, by virtue of their shorter length relative to data packets, have a higher probability of reception. Thus, a number of a node's "neighbors," especially those that are farther, may not be good candidates for receiving a data packet. On the other hand, consistently choosing a relay that is close to

[2]If no neighbor closer to the destination is found (i.e., $\mathcal{N}$ is an empty set), loop-free perimeter forwarding is used to route around the void [14].

[3]For non-geographic routing protocols, the analogous metric to maximum progress would be the minimum hop metric.

the transmitting node can lead to an excessive and unnecessary number of hops.

The relationship between the progress of a hop and the link throughput on that hop is a well known tradeoff in wireless multihop networking. Previous studies have sought optimum network parameters (such as transmission power) that maximize the *expected progress* per hop [18]–[22]. An expected progress metric takes into account the packet success probability; maximizing this metric balances the objective of reducing the number of hops (maximizing progress) with that of choosing reliable links (maximizing packet success probability).

Since its early use in optimization analysis, the expected progress metric has recently been proposed for use in making geographic routing decisions [7], [23], [24]. In the context of relay selection, the expected progress of a candidate next-hop relay is simply the progress offered by that relay multiplied by the probability that it successfully receives the packet. In terms of notation used above, the maximum expected progress (MEP) rule selects the relay j that satisfies

$$j = \arg\max_{i \in \mathcal{N}} \{(d_{S,D} - d_{i,D})\, P_{s,i}\} \tag{3}$$

where $P_{s,i}$ is the packet success probability of the link to relay i.

Maximizing the expected progress per hop can be related to minimizing the average number of transmissions. If packet success can be modeled as independent Bernoulli trials, then the number of transmissions until a packet success on a link is a geometric random variable with mean equal to the inverse of $P_{s,i}$. Furthermore, the number of links in a path is inversely related to the progress of each hop. For example, in a simple linear topology with uniform hop distance and uniform packet success probability, the average number of transmissions end-to-end is simply $d_{S,D}/\left(\Delta d \cdot P_s\right)$, where Δd is the hop distance. For a given source-destination pair, minimizing this ratio is equivalent to maximizing the denominator, or the expected progress per hop. In more general networks, maximizing the expected progress at a hop can be viewed as a greedy approach to minimizing that hop's contribution to the total number of transmissions of that packet end-to-end. Benefits of reducing the number of transmissions per packet include improving energy efficiency [23], reducing end-to-end delay, and increasing network capacity.

Recent applications of metric (3) have been studied in the context of non-fading, contention-free networks [23] and in networks with fixed noise and random packet errors [24]. Below, we apply the MEP metric to geographic routing with channel-adaptive relaying where the channel model includes realistic time-varying correlated fading and multiple access interference. Channel adaptivity is incorporated through the evaluation of the probability of success factor, $P_{s,i}$, using real-time channel measurements and taking into account the impact of dynamic fading and interference.

C. Large-Scale Adaptivity

The purpose of large-scale channel-adaptive relaying is to select links in such a way as to adapt to the slowly varying components of the channel which are typically associated with distance and shadowing. In low mobility or stationary environments, it may also adapt to slow multipath fading.

For large-scale adaptivity, selection rule (3) is employed where $P_{s,i}$ is the packet success probability averaged over small-scale fading. For example, in quasi-static Rayleigh fading, the appendix shows that for certain receiver sensitivities the packet success probability as a function of the average SNR on the link to relay i, μ_i, can be approximated as

$$P_s\left(\mu_i\right) \simeq e^{-\mu^*/\mu_i} \tag{4}$$

where μ^* is the receiver sensitivity-to-noise ratio.

To utilize (4) for evaluating the packet success probability in the MEP metric (3), a node needs estimates of the average SNR to its potential next-hop relays. One way to obtain these estimates would be for each node to maintain a running average of the SNR from each of its neighbors. Many routing protocols already rely on periodic `HELLO` packets that inform nodes of their neighbors. Sample means of the SNR on each neighbor link could be computed from measurements made by the physical layer on receipt of these packets. Provided the transmission power and average noise power are the same at both ends of the link, the measured average SNR on the reverse link should reflect the average SNR on the forward link and be useful for estimating its long-term packet success probability.

To isolate the relative benefit of position information, it will be helpful to compare the performance of the maximum expected progress metric with that obtained by removing the distance factor from (3):

$$\begin{aligned} j &= \arg\max_{i \in \mathcal{N}} \{P_s\left(\mu_i\right)\} \\ &= \arg\max_{i \in \mathcal{N}} \{\mu_i\}. \end{aligned} \tag{5}$$

The second line follows from the monotonicity of $P_s\left(\mu_i\right)$ with μ_i. Referred to as the maximum SNR metric, this metric tends to select relays that are close to the transmitter, trading off progress for link reliability.

D. Small-Scale Adaptivity

Channel-adaptive relay selection based on average SNR estimates is unable to adapt to short-term fluctuations due to multipath fading and interference. The transmission interval of `HELLO` packets on which these estimates are based is typically on the order of one to three seconds, much longer than the coherence time of a fading channel, even at pedestrian speeds. Hence, to obtain diversity against short-term variability of the channel, a faster form of relay selection is needed.

Small-scale channel-adaptive relaying incorporates the fast polling of a small number of candidate next-hop relays prior to the forwarding of each packet. The objective of the polling is to obtain current signal-to-interference-and-noise ratio (SINR) measurements at both ends of each link. With

these measurements, a forwarding node chooses the relay that maximizes the expected progress as in (3), where now the packet success probability $P_{s,i}$ is a function of these current SINR measurements.

The relationship of the current SINR to packet success probability depends on the modulation-coding scheme in use at the physical layer. For example, with the IEEE 802.11b physical layer, which does not utilize forward error correction, a frame is received correctly if no bit errors occur. In this case, the probability of frame success is simply $[1 - P_b(\nu)]^N$, where N is the frame length in bits, ν is the current SINR (assumed to be fixed for the duration of the packet), and $P_b(\nu)$ is the bit error probability for the given modulation scheme. For instance, with differential binary phase-shift keying (DBPSK), which is used at the 1 Mbps data rate, the bit error probability in Gaussian noise and interference is given by

$$P_b(\nu) = \frac{1}{2}e^{-\nu} \quad ; \quad \nu > \mu^*.$$

Thus, for small-scale channel-adaptive relaying, knowledge is needed of the modulation-coding scheme or rate mode in use, an implicit cross-layer dependency.

IV. Protocol Implementation

This section describes an implementation of channel-adaptive relaying for GPSR routing and the IEEE 802.11 MAC distributed coordination function (DCF) with modest modifications to these protocols as needed. Using SNR measurements provided by the physical layer and position information contained in the headers of received `HELLO` and other packets, the routing layer selects one or more candidate next-hop relays providing large-scale adaptivity. The maximum number of candidate relays sought, L, is a system parameter. The routing layer selects the top L relays in terms of expected progress using (3) and (4) and passes their addresses to the MAC layer. The average SNR to a relay, μ_i in (4), is estimated by the sample mean of the SNR measurements of the previous five `HELLO` (or piggybacked data) packets received from that relay; if a packet is not received in a given beacon interval, a zero value is recorded for that measurement. Restricting the candidate relays to those with positive expected progress, and using GPSR's standard perimeter forwarding when greedy forwarding fails, provides for loop-free routing.

In general, the routing layer will find and pass $M \leq L$ candidate relays to the MAC layer. If $M > 1$, the MAC layer then polls the M candidate relays for current SINR and position measurements and forwards the packet to the relay that maximizes the expected progress for small-scale adaptivity. The polling of candidate relays is performed using a form of MAC-layer *anycasting* [25], [26]. Anycasting is applied to the Request-to-Send/Clear-to-Send (RTS/CTS) channel reservation mechanism of the IEEE 802.11 MAC protocol; a conceptual view of relay polling with anycasting is depicted in Fig. 4. Instead of unicasting the RTS to a single next-hop relay, the forwarding node multicasts the RTS to the M candidate relays. The multicast RTS (MRTS) message

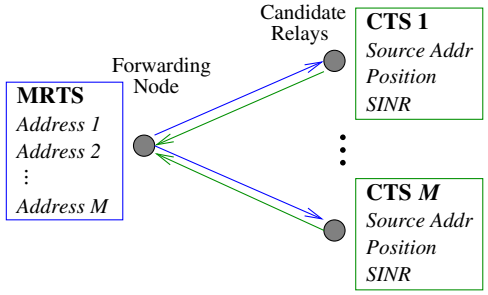


Fig. 4. Polling for current position and channel state information

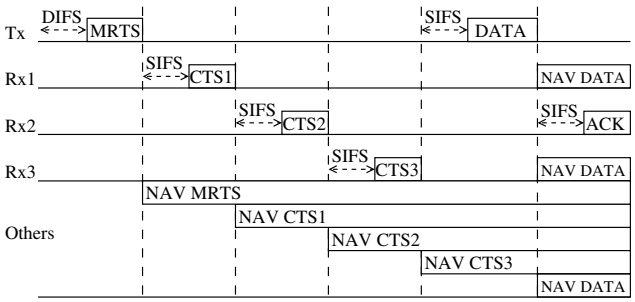


Fig. 5. Timeline of polling protocol for three next-hop nodes

contains the list of addresses of the relays. Each relay that successfully receives the MRTS, and that is available to receive the data packet, replies with a CTS. To avoid collisions, the relays reply in the order specified by the MRTS address list. The CTS frame includes three new fields: the current position coordinates of the relay, the current channel state of the link (the SINR) detected upon receipt of the MRTS, and the source address of the relay. Furthermore, the forwarding node measures and records the channel state for each CTS it receives. Thus, at the conclusion of the MRTS/CTS exchange, the forwarding node has up-to-date measurements of the SINR at both ends of each potential relay link as well as up-to-date position information.[4] Using this current information, the forwarding node transmits the data frame to the relay that maximizes the expected progress. If the data frame is successfully received, the relay replies with an acknowledgment (ACK). Fig. 5 illustrates the timeline of a sample exchange where the MRTS is multicast to three nodes, each one replies with a CTS, and the forwarding node selects the second relay. The timeline also shows the network allocation vector (NAV) states of nearby nodes as a result of the exchange.

Subsequent to successful receipt of a CTS frame, a successful data exchange requires error-free reception of both the data and acknowledgment frames. Thus, in choosing the relay that maximizes the expected progress for small-scale adaptivity, the forwarding node evaluates the packet success probability for (3) as

$$P_s(\nu_r, \nu_l) = [1 - P_b(\nu_r)]^{N_{\mathrm{DATA}}} [1 - P_b(\nu_l)]^{N_{\mathrm{ACK}}}$$

[4]In this version of anycasting, the forwarding node waits until all M relays have had an opportunity to respond.

where ν_r and ν_l are the SINR measurements of the remote and local ends of the link, respectively, and N_{DATA} and N_{ACK} are the number of bits in the DATA and ACK frames, respectively. This evaluation assumes, of course, that the SINRs are fixed and that bit errors are independent; it serves as a first-order approximation of the true probability.

In terms of overhead relative to the standard RTS/CTS exchange, the anycast exchange allots time for at most an additional $L-1$ CTS replies. Furthermore, the MRTS message requires an additional $6(L-1)$ bytes to carry the L addresses, and each CTS requires an additional 22 bytes to carry position information (12 bytes), the SINR (4 bytes) and a source address (6 bytes). In terms of duration at the 2 Mbps data rate, for which a standard CTS including short interframe space lasts 258 μs and each additional byte lasts 4 μs, this overhead translates to an additional $370(L-1)+88$ μs compared to the standard exchange.

A second source of overhead less straightforward to evaluate is the impact of increasing spatial reservation with L. Each candidate relay responding with a CTS causes its respective neighbors to set their NAVs, thereby silencing them from transmitting or relaying other traffic. A larger L implies a larger geographic area of nodes that are blocked from sending and relaying for the duration of the NAV. The extent to which this blocking overhead of anycast polling affects performance will be observed and discussed in the subsequent section.

The ability to adapt to small-scale channel variations is clearly related to the channel correlation between the time the channel measurements are taken and the completion of the acknowledgment transmission. The higher the Doppler (faster channel) or the longer the time interval is, the less correlated is the fading over that interval, and therefore the less accurate the SINR measurements will be. We can use (1) to find the correlation of a Rayleigh fading envelope at the start and finish of an anycast exchange for some number of candidate relays, L. For example, for a node velocity of 5 m/s and a data payload of 120 bytes transmitted at 2 Mbps on a 2.4 GHz carrier, the correlation is equal to 0.87 for $L=2$, and decreases to 0.82 and 0.77 for $L=3$ and 4, respectively. Thus, there is a tradeoff between the order of the diversity and the accuracy of the channel measurements upon which the small-scale channel-adaptive relay selection is based. This tradeoff, among others, is examined in the performance analysis in the next section.

V. Performance Analysis

A. Simulation Model

The protocols for channel-adaptive relaying described in Sections III and IV were implemented and evaluated in the QualNet 3.7 simulation environment. Routing is based on GPSR with average beacon interval of 1.5 seconds. Medium access is based on the IEEE 802.11 DCF with the extensions described in Section IV. The physical layer is that of 802.11 operating at the 2 Mbps data rate. Antennas are omnidirectional, and channel propagation is modeled as a combination of two-ray path loss and time-varying correlated fading. Transmission power is 4.145 dBm, and the receiver sensitivity is -89 dBm, resulting in a 1024-byte packet success rate of 0.9 at a transmitter-receiver separation of 250 m in additive white Gaussian noise (i.e., if fading and interference were absent).

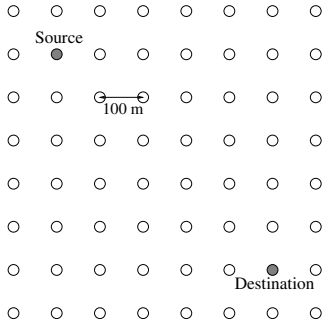


Fig. 6. 8×8 grid topology

The detection of duplicate MAC frames in the QualNet 3.7 implementation of 802.11 is based on the use of sequence numbers that are unique to each transmitter-receiver pair (one-to-one sequence numbers). In channel-adaptive relaying, it is not unlikely that a failed transmission attempt to one relay due to a lost ACK is followed by a successful attempt to a different relay. To prevent the false detection of a duplicate frame the next time a new frame is transmitted to the first relay, we replaced one-to-one sequence numbers with one-to-many so that the transmitter increments a single sequence number for every new frame, and the receiver detects a duplicate only if the frame's sequence number is less than its cached sequence number for that transmitter.

B. Grid Topology

The following results were obtained for a fixed topology consisting of an 8×8 grid of nodes spaced 100 m apart. A single file transfer protocol (FTP) connection transmits 512-byte packets from the source to destination nodes indicated in Fig. 6. Although the nodes are fixed, fading channels for a range of mobility levels are simulated to test the responsiveness of the channel-adaptive protocols to time-varying channels. Each data point is averaged over six independent channel realizations, with each realization lasting 900 seconds. File transfer starts at 6 seconds and ends at 880 seconds.

Fig. 7 plots the average FTP throughput as a function of the channel Doppler in terms of the maximum velocity, where the fading is Rayleigh. Shown is the performance for the maximum progress (MP), maximum SNR (MS), and the maximum expected progress (MEP) metrics. The latter is also shown for small-scale diversity with $L=2$, 3 and 4 next-hop relays sought for the channel-adaptive MAC layer. The MEP metric with $L=1$ (large-scale adaptivity) yields higher throughput than the MP and MS metrics for the full range of channel speeds studied. The benefit of small-scale diversity is especially pronounced at low channel speeds, as expected, since the SINR measurements of the MRTS/CTS exchange are good indicators of the channel state during the DATA

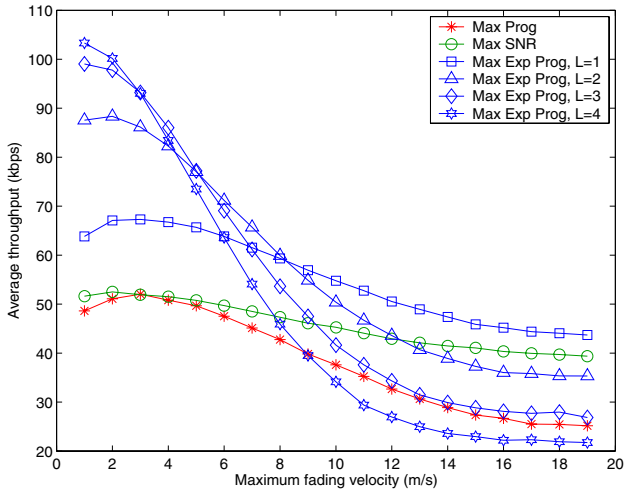


Fig. 7. Average FTP throughput vs. maximum fading velocity for 8×8 grid topology with Rayleigh fading

L	1	2	3	4
FTP throughput (kbps)	108	140	133	126

TABLE I

AVERAGE FTP THROUGHPUT FOR 8×8 GRID TOPOLOGY WITH MEP METRIC IN NON-FADING CHANNELS

and ACK transmissions. At pedestrian speeds (2 m/s), small-scale diversity increases throughput by up to 50% relative to large-scale channel-adaptive relaying with the MEP metric and nearly doubles the throughput achieved by the non-adaptive MP metric. Furthermore, at these channel speeds we observe the diminishing marginal benefit of small-scale diversity with L, a general characteristic of diversity schemes.

However, as channel speed increases, the small-scale diversity gain decreases and throughput levels eventually dip below that achieved with only large-scale adaptivity. At higher speeds, relay selection is based on increasingly outdated channel measurements, leading to a higher risk of making a poor selection. As expected, the impact is greater for larger L, for which the delay between channel measurement and channel use is longer. The MEP curves in Fig. 7 illustrate the tradeoff between small-scale diversity and the timeliness of channel measurements over a range of channel speeds.

Also embedded in these results, but not as clearly seen, is the impact of the increasing spatial reservation with L that results from anycasting to more relays. First, we consider the effect in *non-fading* environments. Table I shows the average throughput for the MEP metrics in the same scenario as above but without fading. Second-order relay diversity increases throughput by offering an alternative to relays made busy by other traffic (e.g., higher layer acknowledgments returned by the destination). However, throughput then *decreases* for $L = 3$ and $L = 4$ due to the spatial blocking and delay overhead of anycasting. Nevertheless, in fading environments (Fig. 7) we see that the benefit of the additional diversity offered by $L > 2$ can outweigh the costs of anycasting, especially at the lower channel speeds. This observation is also made in the random, mobile scenarios discussed below.

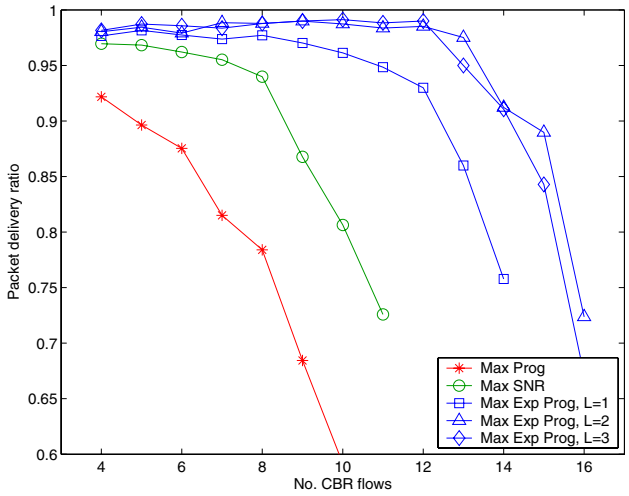
C. Random, Mobile Topologies

The following results were obtained for mobile networks of 200 nodes with random topology. The initial locations of the nodes are randomly, uniformly distributed in a rectangular region of size 3000 m $\times$ 600 m. Node mobility follows the random waypoint model with zero pause time and speeds selected from the interval $(0, v_{\max})$, where $v_{\max}$ is the maximum node velocity. Traffic is generated from multiple constant bit rate (CBR) flows, each generating 64-byte packets at 4 packets/sec. Source-destination pairs are mutually exclusive, so that a node is the source or destination of no more than one CBR flow. Each data point is averaged over six independent realizations of a topology/mobility pattern, with each realization lasting 900 seconds. The start times of the CBR flows are randomly, uniformly distributed on the interval $[6, 180)$ seconds, and they all end at 880 seconds.

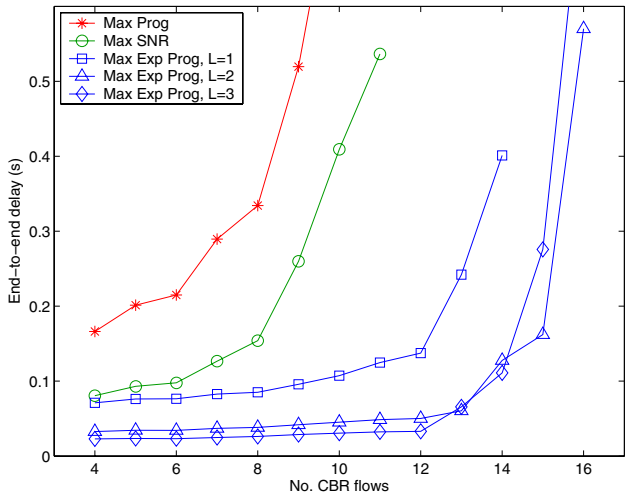
The first set of results are for nodes moving with a maximum velocity of 2 m/s in a Rayleigh fading environment. Fig. 8(a) plots the packet delivery ratio (PDR) as a function of the number of CBR flows in the network for the various metrics. In each case, the PDR eventually decreases with increasing traffic as increasing levels of congestion are encountered. However, the point at which the PDR decreases markedly varies by metric. For example, the MP metric can only sustain up to five CBR flows with an overall PDR of 90% or greater. The MEP metric with second-order small-scale diversity can sustain up to 14 flows at 90% PDR, or a factor of 2.8 increase in traffic capacity over the MP metric. Relative to large-scale adaptivity, the benefit of small-scale adaptivity with $L = 2$ is a modest 17% increase in traffic handling capacity, and the additional benefit for $L = 3$ in terms of PDR is marginal at best.

Figs. 8(b) and 8(c) show corresponding results for average end-to-end delay and average jitter. Again, the MEP metric performs better than the MP and MS metrics, maintaining lower delay and jitter over a broader range of traffic flows. Furthermore, the incremental benefit of small-scale diversity is more apparent here, with second-order diversity reducing both delay and jitter by 50% compared with large-scale adaptivity. These figures also highlight the diminishing marginal benefit with increasing L, showing again that most of the available gain from small-scale adaptivity is achieved with $L = 2$.

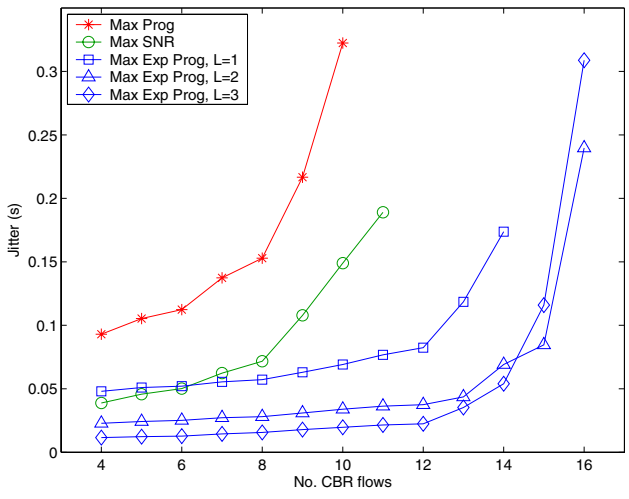
Results for Ricean fading are shown in Fig. 9 for a Rice factor of $K = 5$ dB. Since the Ricean distribution is more concentrated around a specular component resulting in less variation of the signal strength due to fading, performance is better, in general, relative to the Rayleigh fading results. Furthermore, the incremental benefit in PDR of small-scale diversity relative to large-scale channel-adaptive relaying is somewhat less since the fading is less severe. Nevertheless, second-order diversity at the MAC layer still achieves a 50%



(a) Packet delivery ratio

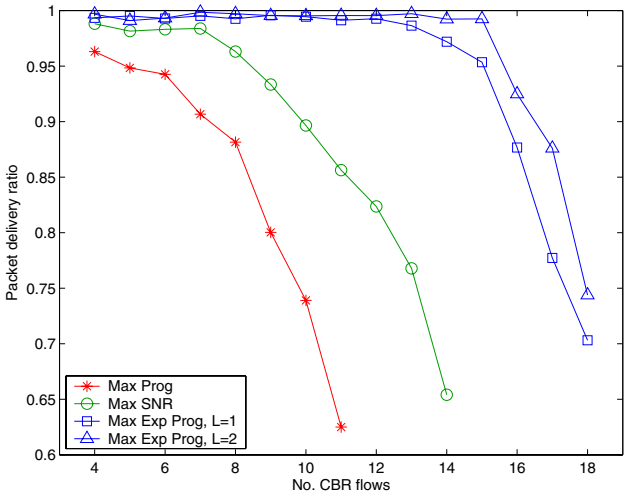


(b) End-to-end delay

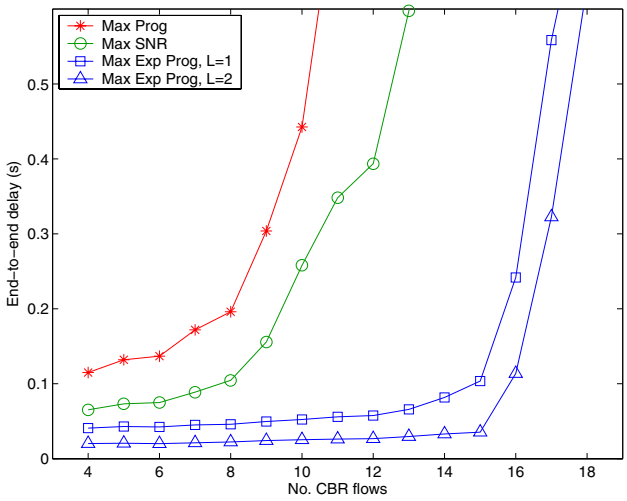


(c) Jitter

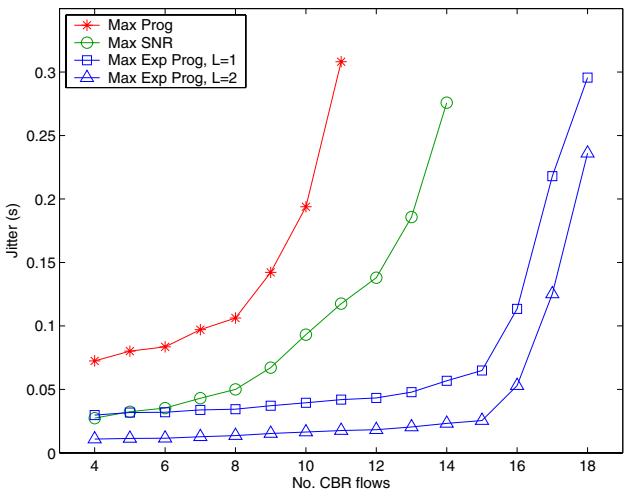
Fig. 8. Performance vs. number of CBR flows for random, mobile topology with Rayleigh fading, $v_{\max} = 2$ m/s



(a) Packet delivery ratio



(b) End-to-end delay



(c) Jitter

Fig. 9. Performance vs. number of CBR flows for random, mobile topology with Ricean fading ($K = 5$ dB), $v_{\max} = 2$ m/s

or more reduction in end-to-end delay and jitter relative to MEP with $L = 1$.

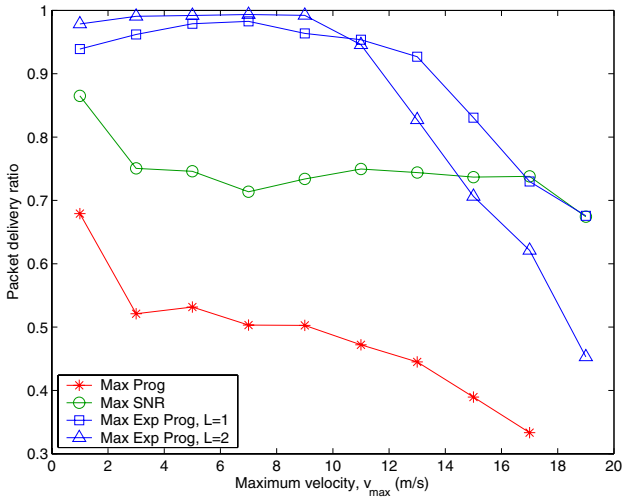
The decrease in delay with L in fading channels (for the values studied) points again to the fact that the small-scale diversity gain from channel-adaptive relaying with $L > 1$ outweighs the costs of the MAC layer polling of these additional relay alternatives. Indeed, results for non-fading channels (not shown) reveal that end-to-end delay *increases* with L, as one might expect since a larger L implies longer anycast exchanges and greater spatial blocking. In fading environments, however, the benefits of diversity overcome this overhead. A similar observation is made in terms of the PDR: whereas without fading the PDR with $L = 2$ drops at fewer CBR flows than with $L = 1$, the converse is true when fading is taken into account.

While channel-adaptive relaying with the MEP metric clearly provides a benefit at pedestrian speeds, it is important to investigate its behavior in higher mobility environments. Fig. 10 shows plots of PDR and delay for a fixed traffic level of 10 CBR flows as a function of maximum node velocity, $v_{\max}$, in Rayleigh fading. (Jitter plots follow similar trends as those in delay.) The MEP metric exhibits a clear advantage for node velocities up to 16 m/s. Furthermore, the break-even point between small and large-scale channel-adaptivity using MEP occurs at $v_{\max} = 11$ m/s, or about 25 mph. This is the point at which the time variability of the fading channel exceeds the ability of the small-scale adaptive protocol to make beneficial selections of the next-hop relay. These results illustrate, though, that while small-scale channel-adaptive relaying is certainly beneficial in low mobility environments, it can also be beneficial at the vehicular speeds that may be typical of some urban environments.

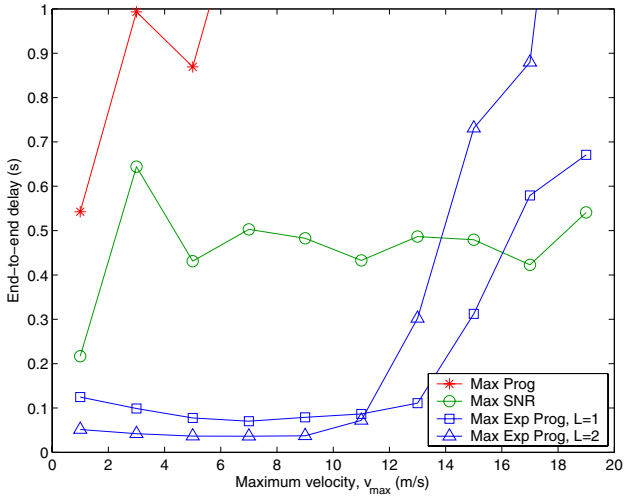
VI. Conclusion

We have described a cross-layer approach for relaying messages in multihop ad hoc networks that adapts to time-varying channel effects and exploits spatial diversity by opportunistically selecting an appropriate relay from a set of candidate relays at each hop. This approach utilizes existing multipath routing protocols that can provide multiple next-hop alternatives to any forwarding node. Using physical layer measurements of the SINR, the routing and MAC layers jointly adapt to both long-term and short-term changes in the channel. Network simulations of an implementation based on geographic routing and the IEEE 802.11 MAC demonstrate increased throughput of delay-insensitive traffic, as well as an increase in capacity and decrease in delay and jitter of delay-sensitive traffic in time-varying Rayleigh and Ricean fading channels. Short-term adaptivity to multipath fading was found to be most effective in slower fading channels and to be useful even at vehicular speeds that are typical in an urban environment. Furthermore, a choice of two next-hop relays was found to provide most of the small-scale diversity gain.

Future work will extend these concepts to include joint adaptive routing and adaptive modulation (or fast rate adaptation) so that high quality links are utilized with greater spectral efficiency. Other possible extensions include the application to non-geographic routing and the use of channel-prediction to enhance short-term adaptivity in faster fading channels.



(a) Packet delivery ratio



(b) End-to-end delay

Fig. 10. Performance vs. maximum node velocity for random, mobile topology with Rayleigh fading, 10 CBR flows

Appendix
Packet Success Probability with Fading

For simplicity of bit error rate expressions, we first consider the 1 Mbps IEEE 802.11 physical layer, which uses differential binary phase shift keying (DBPSK) and no forward error correction. Furthermore, we assume a quasi-static fast fading channel, where the fading amplitude is fixed for the duration of a packet and is modeled as a sequence of independent, unit-mean-square random variables from one packet to the next. The probability of bit error for DBPSK, conditioned on the fading attenuation Γ, is given by [12]

$$P_b\left(\mu|\Gamma=\gamma\right) = \begin{cases} \frac{1}{2}e^{-\gamma\mu} & ; \quad \gamma\mu \geq \mu^* \\ 1 & ; \quad \gamma\mu < \mu^* \end{cases} \tag{6}$$

where μ is the average SNR (averaged over the fading) and μ^* is the receiver sensitivity-to-noise ratio.

For a packet of length N bits, the conditional packet success probability is

$$P_s\left(\mu|\Gamma=\gamma\right)=\left[1-P_b\left(\mu|\gamma\right)\right]^N$$

and the average (unconditioned) packet success probability is

$$\begin{aligned} P_s\left(\mu\right) &= \mathrm{E}_\Gamma\left\{\left[1-P_b\left(\mu|\gamma\right)\right]^N\right\} \\ &= \int_0^\infty \left[1-P_b\left(\mu|\gamma\right)\right]^N f_\Gamma\left(\gamma\right)\, d\gamma \end{aligned} \tag{7}$$

where $\mathrm{E}_\Gamma\left[\cdot\right]$ is the statistical expectation with respect to the random fading attenuation Γ (square of the fading amplitude) and $f_\Gamma\left(\gamma\right)$ is the probability density function of Γ. Substituting (6) in (7) yields

$$P_s\left(\mu\right)=\int_{\mu^*/\mu}^\infty \left(1-\frac{1}{2}e^{-\gamma\mu}\right)^N f_\Gamma\left(\gamma\right)\, d\gamma\,. \tag{8}$$

We may approximate (8) for the condition that $P_b\left(\mu|\gamma\right)\ll 1$ for $\gamma\geq\mu^*/\mu$ such that $\left[1-P_b\left(\mu|\gamma\right)\right]^N\simeq 1$. Under this condition,

$$P_s\left(\mu\right)\simeq\int_{\mu^*/\mu}^\infty f_\Gamma\left(\gamma\right)\, d\gamma = 1-F_\Gamma\left(\mu^*/\mu\right) \tag{9}$$

where $F_\Gamma\left(\gamma\right)$ is the cumulative distribution function of Γ. For Rayleigh fading, Γ is exponentially distributed, so that $F_\Gamma\left(\gamma\right)=1-e^{-\gamma}$, $\gamma>0$. More generally, for Ricean fading with Rice factor K, $F_\Gamma\left(\gamma\right)=1-\mathrm{Q}\left(\sqrt{2K},\sqrt{2(K+1)\gamma}\right)$, $\gamma>0$, where $\mathrm{Q}\left(a,b\right)$ is the Marcum Q-function [10].

Note that approximation (9) applies beyond the case of uncoded DBPSK to any system for which the conditional packet success probability is a steep function of the SNR at some threshold value. For example, in a system with strong forward error correction, the threshold μ^* would be a function of the specific modulation-coding scheme in use.

ACKNOWLEDGMENT

The authors would like to thank Marc Heissenbüttel of the Institute of Computer Science and Applied Mathematics, University of Bern, Switzerland, for providing the basic GPSR source code which we extended for this work.

REFERENCES

[1] H. Lundgren, E. Nordström, and C. Tschudin, "Coping with communication gray zones in IEEE 802.11b based ad hoc networks," in *Proc. 5th ACM Int. Workshop on Wireless Mobile Multimedia*, 2002, pp. 49–55.

[2] A. Tsirigos and Z. J. Haas, "Multipath routing in the presence of frequent topological changes," *IEEE Commun. Mag.*, vol. 39, no. 11, pp. 132–138, Nov. 2001.

[3] 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, Feb. 1997.

[4] P. Sambasivam, A. Murthy, and E. M. Belding-Royer, "Dynamically adaptive multipath routing based on AODV," in *Proc. 3rd Annual Mediterranean Ad Hoc Networking Workshop*, June 2004.

[5] X.-H. Lin, Y.-K. Kwok, and V. K. N. Lau, "RICA: A receiver-initiated approach for channel-adaptive on-demand routing in ad hoc mobile computing networks," in *Proc. 22nd Int. Conf. on Distributed Computing Systems (ICDCS)*, 2002, pp. 84–91.

[6] ——, "BGCA: Bandwidth guarded channel adaptive routing for ad hoc networks," in *Proc. Wireless Commun. and Networking Conf. (WCNC)*, vol. 1, 2002, pp. 433–439.

[7] M. R. Souryal, B. R. Vojcic, and R. L. Pickholtz, "Information efficiency of multihop packet radio networks with channel-adaptive routing," *IEEE J. Select. Areas Commun.*, vol. 23, no. 1, pp. 40–50, Jan. 2005.

[8] P. Larsson, "Selection diversity forwarding in a multihop packet radio network with fading channel and capture," in *Proc. ACM MobiHoc*, 2001, pp. 279–282.

[9] J. Wang, H. Zhai, W. Liu, and Y. Fang, "Reliable and efficient packet forwarding by utilizing path diversity in wireless ad hoc networks," in *Proc. MILCOM*, 2004.

[10] G. L. Stüber, *Principles of Mobile Communication*. 2nd ed., Boston: Kluwer Academic Publishers, 2001.

[11] P. Dent, G. E. Bottomley, and T. Croft, "Jakes fading model revisited," *Electronics Letters*, vol. 29, no. 13, pp. 1162–1163, June 24, 1993.

[12] J. G. Proakis, *Digital Communications*. 4th ed., New York: McGraw-Hill, 2001.

[13] I. Stojmenovic, "Position-based routing in ad hoc networks," *IEEE Commun. Mag.*, vol. 40, no. 7, pp. 128–134, July 2002.

[14] B. Karp and H. T. Kung, "GPSR: Greedy perimeter stateless routing for wireless networks," in *Proc. MobiCom*, 2000, pp. 243–254.

[15] M. K. Marina and S. R. Das, "On-demand multipath distance vector routing in ad hoc networks," in *Proc. IEEE Int. Conf. for Network Protocols*, 2001, pp. 14–23.

[16] D. Aguayo, J. Bicket, S. Biswas, G. Judd, and R. Morris, "Link-level measurements from an 802.11b mesh network," in *Proc. Special Interest Group on Data Communication (SIGCOMM)*, 2004, pp. 121–132.

[17] D. S. J. De Couto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multi-hop wireless routing," in *Proc. MobiCom*, 2003, pp. 134–146.

[18] L. Kleinrock and J. Silvester, "Optimum transmission radii for packet radio networks or why six is a magic number," in *Proc. Nat. Telecommun. Conf.*, Birmingham, Alabama, Dec. 1978, pp. 4.3.1–4.3.5.

[19] H. Takagi and L. Kleinrock, "Optimal transmission ranges for randomly distributed packet radio terminals," *IEEE Trans. Commun.*, vol. 32, no. 3, pp. 246–257, Mar. 1984.

[20] R. Nelson and L. Kleinrock, "The spatial capacity of a slotted ALOHA multihop packet radio network with capture," *IEEE Trans. Commun.*, vol. 32, no. 6, pp. 684–694, June 1984.

[21] T.-C. Hou and V. O. K. Li, "Transmission range control in multihop packet radio networks," *IEEE Trans. Commun.*, vol. 34, no. 1, pp. 38–44, Jan. 1986.

[22] E. S. Sousa and J. A. Silvester, "Optimum transmission ranges in a direct-sequence spread-spectrum multihop packet radio network," *IEEE J. Select. Areas Commun.*, vol. 8, no. 5, pp. 762–771, June 1990.

[23] K. Seada, M. Zuniga, A. Helmy, and B. Krishnamachari, "Energy-efficient forwarding strategies for geographic routing in lossy wireless sensor networks," in *Proc. 2nd Int. Conf. on Embedded Networked Sensor Systems (SenSys)*, Nov. 2004, pp. 108–121.

[24] S. Lee, B. Bhattacharjee, and S. Banerjee, "Efficient geographic routing in multihop wireless networks," in *Proc. 6th ACM Int. Symp. on Mobile Ad Hoc Networking and Computing (MobiHoc)*, May 2005, pp. 230–241.

[25] R. R. Choudhury and N. H. Vaidya, "MAC-layer anycasting in ad hoc networks," in *Proc. 2nd workshop on Hot Topics in Networks (HotNets II)*, 2003.

[26] S. Jain and S. R. Das, "Exploiting path diversity in the link layer in wireless ad hoc networks," in *Proc. 6th IEEE Int. Symp. on a World of Wireless Mobile and Multimedia Networks (WoWMoM)*, June 2005, pp. 22–30.