High Throughput Route Selection in Multi-rate Ad Hoc Wireless Networks

Baruch Awerbuch, David Holmer, and Herbert Rubens

Johns Hopkins University, Baltimore MD, USA
{baruch, dholmer, herb}@cs.jhu.edu

Abstract. Modern wireless devices, such as those that implement the 802.11b standard, utilize multiple transmission rates in order to accommodate a wide range of channel conditions. Traditional ad hoc routing protocols typically use minimum hop paths. These paths tend to contain long range links that have low effective throughput and reduced reliability in multi-rate networks. In this work, we present the *Medium Time Metric* (MTM), which is derived from a general theoretical model of the attainable throughput in multi-rate ad hoc wireless networks. MTM avoids using the long range links favored by shortest path routing in favor of shorter, higher throughput, more reliable links. We present NS2 simulations that show that using MTM yields an average total network throughput increase of 20% to 60%, depending on network density. In addition, by combining the MTM with a medium time fair MAC protocol, average total network throughput increases of 100% to 200% are obtained over traditional route selection and packet fairness techniques.

1 Introduction

Ad hoc wireless networks are self-organizing multi-hop wireless networks where all nodes take part in the process of forwarding packets. One of the current trends in wireless communication is to enable devices to operate using many different transmission rates. Many current and proposed wireless networking standards have this multi-rate capability. These include the 802.11b [1], 802.11a [2], 802.11g draft, and HiperLAN2 [3] standards. The reason for this multi-rate capability stems directly from some of the fundamental properties of wireless communication.

Due to the physical properties of communication channels, there is a direct relationship between the rate of communication and the quality of the channel required to support that communication reliably. Since distance is one of the primary factors that determines wireless channel quality, there is an inherent trade-off between high transmission rate and effective transmission range.

This range speed trade-off is what has driven the addition of multi-rate capability to wireless devices. Consumer demands for wireless devices always include both higher speed and longer range. Unfortunately a single rate represents a single trade-off point between these two conflicting goals. Since multi-rate devices support several rates, they provide a wide variety of trade-offs available for use.

This gives them a great deal of flexibility to meet the demands of consumers. This added flexibility is the primary driving force behind the adoption of multi-rate capability. It is also reasonable to assume that this type of capability will also be present in future wireless networking standards.

While multi-rate devices provide increased flexibility, they cannot change the inherent trade-off between speed and range. Both high speed and long range cannot be achieved simultaneously. Long range communication still must occur at low rates, and high-rate communication must occur at short range. This multi-rate capability merely provides a number of different trade-off points. Multi-rate devices must have protocols that select the appropriate rate for a given situation.

In infrastructure based networks, all communication takes place between nodes and access points. In this case, an additional protocol required to support multi-rate is necessary only at the medium access control (MAC) layer. Single rate nodes already have the ability to select the best access point based on the received signal strength. Thus the only additional task necessary is that of selecting the actual rate used to communicate. Since the distance between the user and the access point is dictated by the physical geometry of the network, the rate selection task must react to the existing channel conditions. In other words, the only option available to a wireless device is to select the fastest modulation scheme that works reliably.

However, this is no longer the case in ad hoc multi-hop wireless networks. In these networks, the routing protocol must select from the set of available links to form a path between the source and the destination. While in single-rate networks all links are equivalent, in multi-rate networks each available link may operate at a different rate. Thus the routing protocol is presented with a much more complex problem. Which set of trade-offs does it choose? Long distance links can cover the distance to the destination in few hops, but then the links would be forced to operate at a low speed. Short links can operate at high rates, but more hops are required to reach the destination. In addition, the path selected by the routing protocol will not only affect the packets moving along that path, but will affect the level of congestion at every node within the interference range of the path as well.

Our Contribution. We provide a general theoretical model of the attainable throughput in multi-rate ad hoc wireless networks. This model is derived from the properties of the physical and medium access control layers. The traditional technique used by most existing ad hoc routing protocols is to select minimum hop paths. These paths tend to contain long range links that have low effective throughput and reduced reliability. We present the *Medium Time Metric* (MTM) that selects higher throughput paths and tends to avoid long unreliable links. The MTM minimizes the total medium time consumed sending packets from a source to a destination. This results in an increase in total network throughput.

2 Related Work

Ad Hoc Routing Protocols. A large number of routing protocols have been proposed by the ad hoc wireless networking community. Typically these have adopted one of two major strategies: on-demand such as in AODV [4] and DSR [5], and proactive such as in DSDV [6] and OLSR [7]. The vast majority of these protocols where originally designed for single-rate networks, and thus have used a shortest path algorithm with a hop count metric (min hop) to select paths. While min hop is an excellent criteria in single-rate networks where all links are equivalent, it does not accurately capture the trade-offs present in the more complicated multi-rate networks. As ad hoc networks are likely to be deployed in multi-rate networks, it should be possible to enhance the network performance of almost any existing shortest path based protocol by adapting it to use our medium time metric.

Signal Stability Based Ad Hoc Routing Protocols. In [8] the authors show that the minimum hop path generally contains links which exhibit low reliability. In [9] and [10] the authors present routing protocols which are based on signal stability rather then just shortest path in order to provide increased path reliability. In our work, signal stability information is used not only to increase path reliability, but also to increase network throughput.

MAC Layer. Since our proposed solution is derived from properties of the MAC and physical layers, it is important to understand existing MAC layer techniques. The IEEE 802.11 standard [11] defines the most commonly used MAC protocol in ad hoc wireless networks. 802.11 based devices are used because of their widespread availability, low cost, and 802.11's ability to provide distributed medium access control when operated in "ad hoc" mode. This mode causes the stations to use the Distributed Coordination Function (DCF) protocol that operates using Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA).

The method of rate selection in multi-rate capable networks has been left unspecified by the 802.11 standards. As a result, several auto rate protocols have been proposed. The most commonly used protocol is Auto Rate Fallback (ARF). ARF was originally developed for Lucent's WaveLAN II devices [12], and was later enhanced for 802.11b devices [13]. ARF operates using the link level ACK frames specified by the 802.11 standard. Each node increases the rate it is using to communicate with its neighbor after a number of consecutively received acks, and decreases the rate after a number of consecutively missed acks. The advantage of this technique is that it is easy to implement because it is purely sender based, requires no modifications to the 802.11 standard.

As an alternative, the Receiver Based Auto Rate (RBAR) protocol was presented in [14]. RBAR allows the receiving node to select the rate. This is accomplished by using the SNR of the RTS packet to choose the most appropriate

rate and communicating that rate to the sender using the CTS packet. This allows much faster adaptation to the changing channel conditions than ARF, but requires some modifications to the 802.11 standard.

The Opportunistic Auto Rate (OAR) protocol, which is presented in [15], operates using the same receiver based approach, but allows high-rate multi-packet bursts to take advantage of the coherence times of good channel conditions. These bursts also dramatically reduce the overhead at high rates by amortizing the cost of the contention period and RTS CTS frames over several packets. By picking appropriate sized bursts, OAR also changes the fairness characteristic from each node sending an equal number of packets to each node getting an equal allocation of medium time. This produces a dramatic increase in overall throughput when links of multiple rates operate together in the same space. OAR also requires modifications to the 802.11 standard.

3 Network Model

Network Assumptions. This work relies on a few specific network assumptions. We assume that the ISO/OSI physical layer is capable of operating using multiple rates. We also assume that the ISO/OSI MAC layer is capable of selecting the rate used by the physical layer. In addition, we assume that the MAC layer is capable of providing information to the ISO/OSI network layer that indicates the selected rate. The network layer can then use this information to improve its routing decisions. This work stresses the importance of inter-layer communication in wireless networks.

Multi-Rate Model. The multi-rate model presented in this paper is based on the 802.11b standard [1]. The topics discussed here apply to other multi-rate standards, but all examples, ranges, and rates shown in this work are based on 802.11b.

Throughout the remainder of the paper we present the results of a number of NS2 [16] simulations. In order to simulate multi-rate 802.11b, we started with the ns-2.1b7a code base and the multi-rate extensions available from the Rice Networks Group [17] that contain implementations of the RBAR and OAR protocols. The 802.11 MAC and physical wireless parameters were further modified to match the published specifications of a Lucent ORiNOCO PC Card [18], a commonly used 802.11b wireless adapter (see Table 1). Since the carrier sense (CS) threshold specification is not published, we provide an estimate. This estimate was produced by setting the difference between the carrier sense threshold estimate and the 1.0 Mbps receive threshold equal to the difference between the NS2 default carrier sense threshold (-78 dBm) and default receive threshold (-64 dBm).

Table 2 shows the ranges resulting from these simulation parameters. Real world ranges are considerably smaller due to non-zero system loss, additional noise sources, obstructions, and propagation effects beyond the simple two ray

Table 1. NS2 Simulation Parameters

Parameter	Value
Frequency	2.4 GHz
Transmit Power	15 dBm
11.0 Mbps Receive Threshold	-82 dBm
5.5 Mbps Receive Threshold	-87 dBm
2.0 Mbps Receive Threshold	-91 dBm
1.0 Mbps Receive Threshold	-94 dBm
Carrier Sense Threshold	-108 dBm
Capture Threshold	10
Propagation Model	Two Ray Ground
System Loss	0 dBm

Table 2. 802.11b Ranges

Rate (Mbps)	Maximum Range
11.0	399 m
5.5	531 m
2.0	669 m
1.0	796 m
CS	1783 m

ground model. The results presented here should be valid for any set of ranges with similar proportions regardless of magnitude.

4 Minimum Hop Route Selection

Most existing ad hoc routing protocols have utilized hop count as their route selection criteria. This approach minimizes the total number of transmissions required to send a packet on the selected path. This metric is appropriate in single-rate wireless networks because every transmission consumes the same amount of resources. However, in multi-rate networks this technique has a tendency to pick paths with both low reliability and low effective throughput.

Throughput Loss. In multi-rate wireless networks, the selection of minimum hop paths typically results in paths where the links operate at low rates. This is because the shortest path contains the fewest number of nodes between the source and destination. Fewer intermediate nodes corresponds to longer links in order to cover the same distance. Since distance is one of the primary factors that determines channel quality, the long links have low quality, and thus operate at low rates. So given the opportunity, in an effort to minimize the number of hops, shortest path selection protocols will pick paths composed of links close to their maximum range that must operate at the minimum rate.

Not only do the low link rates produce a low effective path throughput, but as a result of the shared wireless medium, this path selection degrades the performance of other flows in the network. This occurs due to the large amount of medium time required to transmit a packet at a slow link speed. All nodes within interference range of the transmission must defer while it takes place. Thus, slow transmissions reduce the overall network throughput by consuming a large amount of medium time.

Reliability Loss. Multi-rate wireless devices are inherently designed to deal with changes in connectivity due to mobility and interference. The devices provide multiple link speeds to accommodate fluctuations in link quality. In 802.11b,

as two nodes move in opposite directions, the auto rate protocol will gracefully reduce their link speeds from 11 Mbps down to 1 Mbps before they are finally disconnected.

Minimum hop path route selection has a tendency to choose routes that utilize the lowest link speed, leaving the auto rate protocol no flexibility in dealing with channel quality fluctuations. As a result, routes are often established between nodes that are on the fringe of connectivity. This occurs when nodes are able to receive broadcast transmissions, but data/ack packets are unable to be successfully delivered. While routing broadcasts are typically extremely small in size, data packets typically occupy the full frame size, making them more susceptible to corruption at high bit error rates (BER). This tendency is even further exaggerated by the way 802.11 handles broadcast transmissions as opposed to unicast transmissions. While broadcasts are sent as a single frame, unicasts require a full RTS-CTS-DATA-ACK exchange for successful delivery, which is more likely to be disrupted by a low quality channel. The end result is that small broadcasts can often be delivered even when data communication is not possible.

5   General Model and Optimality Analysis

There is some ambiguity in the literature regarding what constitutes an optimal solution for the routing problem in multi-hop wireless networks. One of the main reasons for this is the inherent difficulty in modelling the complex environment of wireless multi-hop networks. We provide a model that captures many of the effects present in such a network.

5.1   General Model of Attainable Throughput

In this work, we ignore packet scheduling issues and consider a steady-state flow model. In this model, each network edge may be fractionally shared by several flows; however, the sum of shares cannot exceed 100%. Our model of the wireless network is defined by a *transmission graph* and *interference graph*.

The *transmission graph* is defined as $G(V, E, \rho)$. V is defined as the set of nodes in the network. A transmission edge $(u, v) \in E$ if node u is capable of transmitting to node v. ρ is a function that assigns a transmission rate to each transmission edge $\rho : E \to R^+$. $\rho(e) = \hat{\rho}$ where $\hat{\rho}$ is the maximum flow rate obtainable over edge e when no other traffic exists in the network. $\hat{\rho}$ should take into account any sources of overhead such as contention, headers, and multiple frame exchanges, and represents the "real" capacity of edge e. In this general definition, the transmission graph may be directed, and the transmission rate in the reverse direction of a bi-directional edge may be different than that in the forward direction. This is possible in real wireless networks because of different node configurations and asymmetric channel effects.

The *interference graph* is defined as $G(\tilde{V}, \tilde{E})$. We define the vertices of the interference graph to be the edges of the transmission graph, so $\tilde{V} = E$. An edge

in the interference graph represents the interaction between packets transmitted on nearby transmission edges. $((a, b), (c, d)) \in \tilde{E}$ if $(a, b), (c, d) \in E$ and if a transmission on (a, b) interferes with a transmission on (c, d).

In the general case, modelling the interference graph of an arbitrary network may be quite difficult due to complex propagation effects caused by obstacles and reflections. However, in the open space simulation configuration used in this, and many other papers, modelling the interference graph is much simpler. In this open space environment, the interference graph includes "edges" between each possible transmission edge, and all other transmission edges with an endpoint within carrier sense range of one of the transmission edge's endpoints. This roughly corresponds to everything within a two hop neighborhood of a transmitting node.

Given the interference graph, we can define the interference neighborhood of any given edge (u, v) as follows.

$$\chi(u, v) = \{(u, v)\} \cup ((x, y) : ((x, y), (u, v)) \in \tilde{E}) \tag{1}$$

Consider a set of i flows, where each flow ϕ_i originates from source s_i and is sinked by receiver r_i. Without loss of generality, we can represent each flow as a sum of path flows (indexed by j).

$$\phi_i = \sum_j \phi_{ij} \tag{2}$$

Each path flow ϕ_{ij} exists only on π_{ij}, where π_{ij} is a path from s_i to r_i in the transmission graph. In other words, $\phi_{ij}(x, y)$ equals the magnitude of the path flow $|\phi_{ij}|$ if the edge lies on its path, $(x, y) \in \pi_{ij}$, or zero otherwise. Thus we have effectively decomposed the general flow ϕ_i which may traverse multiple paths simultaneously into a set of flows ϕ_{ij} that each traverse only a single path, but sum to the original flow ϕ_i.

With this setup, we can now specify a flow constraint that captures the phenomena discussed above. For each edge (u, v) in the transmission graph, the sum of the fractional shares used by all flows in the interference neighborhood of (u, v) must be less than or equal to 100%. This is a more complicated version of the classic edge capacity flow constraint.

$$\sum_{(x,y) \in \chi(u,v)} \sum_{i,j} \left(\frac{\phi_{ij}(x, y)}{\rho(x, y)} \right) \leq 1 \tag{3}$$

In this general case, Linear Programming (LP) methods are required to achieve an optimal thruput solution. Opportunity-cost based approximations are possible in both the off-line case [19] (all connections are known ahead of time) and in the online case [20, 21]. Single path solutions are even harder to achieve as they require integer LP approaches.

5.2 Optimal Routing Assuming a Complete Interference Graph

Consider the special case of the general model where the interference graph is a clique (completely connected graph), i.e. each node can carrier sense each other node. In this special case, the constraint can be simplified since the interference neighborhood of any edge $\chi(u, v)$ is the same and consists of every edge in the transmission graph. In this case we wish to show the following theorem:

Theorem 1. *In the case of a complete interference graph in the stated multi-rate ad hoc wireless network model, a routing protocol that chooses a single path that minimizes the sum of the transmission times optimally minimizes network resource consumption, and optimally maximizes total flow capacity.*

Given the complete interference condition, we can rewrite the general flow constraint.

$$\sum_{(x,y)\in E} \sum_{i,j} \left(\frac{\phi_{ij}(x,y)}{\rho(x,y)} \right) \leq 1 \tag{4}$$

We can reverse the order of summation.

$$\sum_{i,j} \sum_{(x,y)\in E} \left(\frac{\phi_{ij}(x,y)}{\rho(x,y)} \right) \leq 1 \tag{5}$$

We can also decompose $\phi_{ij}(x, y)$ by moving its magnitude out of the inner sum, and changing the inner sum to include only non-zero terms.

$$\sum_{i,j} \left(|\phi_{ij}| \cdot \sum_{(x,y)\in \pi_{ij}} \left(\frac{1}{\rho(x,y)} \right) \right) \leq 1 \tag{6}$$

Since $\rho(x, y)$ was defined as the real capacity of transmission edge (x, y), we can define the transmission time used by a unit of flow on this edge to be the inverse of this capacity.

$$\tau(x,y) = \frac{1}{\rho(x,y)} \tag{7}$$

Thus the final constraint equation becomes

$$\sum_{i,j} \left(|\phi_{ij}| \cdot \sum_{(x,y)\in \pi_{ij}} \left(\tau(x,y) \right) \right) \leq 1 \tag{8}$$

In other words, the flow over each sub path consumes a certain fraction of the capacity. The sum of these fractions must be less than one. The fraction consumed by each sub path is equal to the amount of flow on that path times the sum of the transmission times along that path. The magnitude of flow on a sub path, $|\phi_{ij}|$, will be maximized when the sum of the transmission times along that path, $\sum_{(x,y)\in \pi_{ij}} \tau(x,y)$, is minimized. Therefore, a routing protocol that selects

paths that minimize the sum of the transmission times maximizes the flow along those paths. Also, it is only necessary for each flow to have a single sub path that minimizes the sum of the transmission times, because any other sub paths will be at best equivalent to the minimum, and thus offer no additional flow capacity. Even if a flow does not use its maximum available capacity, minimizing the path transmission time minimizes the flow's consumption of the common network resource and allows other flows to increase. Thus we have shown Theorem 1 to be true.

6 Medium Time Metric

We propose a *medium time metric* (MTM) that is designed to allow any shortest path routing protocol to find throughput optimal routes assuming full interference. The MTM assigns a weight to each link in the network that is proportional to the amount of medium time used by sending a packet on that link. The weight of any given path is thus a sum that is proportional to the total medium time consumed when a packet traverses the whole path. As a result, shortest path protocols that use the medium time metric find paths that minimize the total transmission time.

We have shown that the MTM is globally optimum in the case of complete interference. In real networks the interference graph is primarily determined by the carrier sense range. While the carrier sense range is not infinite, in 802.11b networks it is greater that twice the maximum transmission range. Therefore, full interference networks are limited to four or less maximum length hops. While these small networks are useful for some applications, many applications require larger networks.

Once we consider larger networks, we can no longer claim that the MTM is globally optimal because traffic patterns and congestion may shift the optimal routes, and in very large networks multiple "disjoint" paths may be used. However, the MTM still exhibits desirable characteristics in these larger networks. If we consider a reduced class of single flow (non-congestion sensitive) single path algorithms, we find that the MTM continues to perform optimally up to a much longer path length limit. This occurs because the sum of the medium times is still an accurate predictor of total path throughput, even for paths much longer than the complete interference case. The MTM will continue to be accurate as long as one link in the path is in interference range of all other links in the path. This occurs with paths of up to seven maximum length hops, and also corresponds to the length where pipelining begins to occur (see Figure 1). Once the paths are long enough to exhibit significant pipelining, the MTM begins to underestimate their throughput potential.

We have shown that the MTM performs completely optimally in small full interference networks, and that the MTM selects optimal non-congestion sensitive single path routes for lengths up to the pipelining distance. The reader should note that this second property places no restriction on the total size of the network, which may be extremely large, and only restricts the length of the
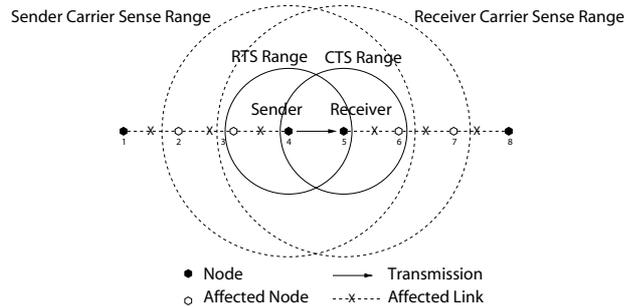
Fig. 1. Link Interference Range

actual communication paths. This is an important observation because prior research has shown that ad hoc networks scale to large sizes only if the traffic patterns remain local [22]. A local traffic pattern, such as when every node accesses the nearest Internet gateway, provides a natural path length limit allowing the MTM to operate in its optimal regions even in very large networks. The reason that non-local traffic patterns do not scale is that even with a globally optimal routing protocol, the attainable throughput at pipelining distance and beyond is extremely small, thus communicating over these distances consumes a large quantity of medium time with little gain.

6.1 Computing Link Weights

Our medium time metric states that paths that minimize the total consumed medium time should be selected. In order to accomplish this using existing shortest path protocols, we must assign a weight to each link that is directly proportional the medium time consumed by sending a packet across that link. The initial obvious solution is to use weights that are inversely proportional to the rate of the link. Using this scheme, if an 11 Mbps link was assigned a weight of 1, then a 1 Mbps link would be assigned a weight of 11 (see Figure 2).

However, we find that inverse rate weights do not accurately predict the amount of medium time consumed when sending a packet because they because they do not accurately represent an 802.11b packet transmission exchange. In 802.11b a packet is typically transmitted using an idle contention period and a four frame MAC level exchange (RTS, CTS, DATA, ACK). Much of this exchange takes place at the 1 Mbps base rate, so a large nearly constant amount of medium time is consumed by per packet MAC overhead regardless of the actual link rate. This overhead becomes a large fraction of the total consumed medium time at the higher rates, because the actual data payload transmission time becomes small (see Figure 2). This overhead is why two nodes never achieve anywhere close to 11 Mbps of real throughput over an 11 Mbps link. For example, inverse weights would select a path of ten 11 Mbps links over a single 1 Mbps link. However, a 1 Mbps link is faster (and therefore consumes less medium time) than ten 11 Mbps links.
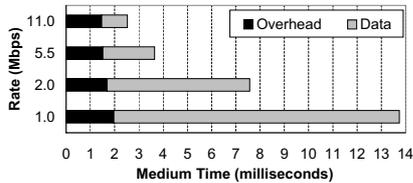
Link Rate	Inverse	1500 byte packet	
(Mbps)	Weights	μsec	MTM Weights
11.0	1.00	2542	1.00
5.5	2.00	3673	1.44
2.0	5.50	7634	3.00
1.0	11.00	13858	5.45

Fig. 2. Medium Times and Weights for 802.11b Transmissions

The almost fixed amount of medium time overhead caused by 802.11b introduces a dependency on packet size into our protocol. For example, the transmission time of a small packet will be dominated by the MAC overhead and will be almost the same regardless of the link rate. The implication of this phenomena is that the medium time metric would ideally use different link weights for each different packet size. This should be fairly easy to implement in link state protocols because they already have the topology information necessary to compute alternate routes using different sets of weights. However, this would be much more difficult for distance vector protocols, which require additional communication overhead for each additional set of weights. While it may be worth while in some networks to track more than one set of weights, usually the bulk of data transferred in a given network is of a single size.

An implementation of the MTM for a distance vector protocol should be tuned for the dominant packet size used by the network. This is accomplished by using link weights that are proportional to the medium time used by packets of the tuned size. These tuned weights represent the best trade-off point between short low-rate paths and long high-rate paths for packets of the tuned size. Packets that are much larger than the tuned size may have been better off traversing a longer path with even higher rate links. Similarly, packets much smaller than the tuned size may be better off taking paths that are shorter but with lower rate links. Performance of the MTM should not be significantly affected by transmissions of packets larger and smaller than the tuned size as long as those packets do not consume a large fraction of the total medium time.

In this work, the tuned packet size was chosen to correspond to a 1500 byte IP packet. This size is representative of the majority of the data transferred by the Internet [23] and corresponds to the standard Ethernet maximum transferable unit (MTU) [24]. This size was chosen over the larger native MTU of 802.11b (2314 bytes) because wireless networks today are mostly used to provide mobile access to LAN and Internet resources. In this environment, packets that flow over fixed links as well as wireless links would be limited to a 1500 byte path MTU. Purely peer to peer wireless networks would be free to use the native MTU, and could gain an additional measure of throughput due to the increased ratio of data to overhead in each packet.

Figure 2 shows the expected medium times, and corresponding proportional weights, for each rate computed according to the 802.11b standard specifications. These weights are significantly different then the inverse weights. The times are

calculated assuming a full RTS, CTS, DATA, ACK exchange. All information is sent at the base 1 Mbps rate except for the contents of the data and acknowledgement frames, which are sent at the chosen link rate. These computed times also include an estimate of the time spent backing off during contention. We used the value of half the minimum contention window size multiplied by the slot time (310 μsec). This estimate was derived from the average time spent in the single sender case, but should function sufficiently for multiple senders. When we have an increased number of senders contending, the average idle medium time should decrease dramatically because the time spent for any particular packet is the minimum of all the senders random back offs. However, the probability of a collision also increases so the average time wasted while performing contention should not change as much as we might expect.

Even though a large number of acknowledgement packets are present in the network when TCP is used, the time total consumed by these packets is small in comparison to the data. This is particularly true when the delayed acknowledgement option of TCP is used which effectively halves the number of acknowledgements. OAR further reduces the proportion of time consumed by acknowledgement packets sent at high-rate by amortizing much of the contention and control overhead over several packets.

Since the OAR protocol significantly changes the MAC layer packet exchange, the expected medium time consumed by a packet at a given rate changes significantly. Thus in networks where OAR, or a significantly different MAC exchange, is used, different MTM weights must be calculated to match the change in consumed medium time.

6.2 Advantages

The medium time metric has several advantages over other possible routing strategies. One of its primary advantages is its simplicity. As a shortest path metric, it can be incorporated into existing distance vector or link-state protocols. The majority of existing wireless ad hoc routing protocols fall into these categories (AODV, DSR, OLSR, DSDV). It would be much more difficult to incorporate the MTM into protocols that use routing strategies other than shortest path, such as TORA [25].

The medium time metric also sidesteps the most serious problems exhibited by the optimal solution under the general model. MTM protocols only need to track changes in link rates as opposed to changes in utilization. This results in drastically lower protocol overhead. Also, there is no danger of route oscillation because MTM routes do not depend on traffic patterns. Finally, there is no danger of disrupting higher level protocols such as TCP due to out of order packet delivery because the MTM selects a single path.

Another interesting property of MTM paths is that since they naturally avoid low-rate links, they exhibit some of the properties of signal stability based routing protocols. Nodes connected by a high-rate link must move a considerable distance before the link breaks. As the nodes move further apart, the auto rate protocol reduces the link speed. As a result, proactive routing protocols, which continually

update their paths based on the MTM will naturally avoid path failures by continuously switching to higher rate links.

6.3 Discussion

Increased Hop Count. Typically, the MTM selects paths that have a greater number of hops than the minimum. While these higher rate hops consume less total medium time than the minimum number of hops, the increased number of senders could cause other detrimental effects. For instance, the increased number of senders creates higher contention for the medium. If this higher degree of contention causes a significant degradation in the throughput of the underling MAC protocol, then the efficiency of the MTM will be degraded. The authors of [26] specifically explore this contention issue. Their paper shows that when RTS/CTS is used and the packet sizes are large (1000 bytes), the throughput of the 802.11 MAC is only reduced by 6% with 100 contending nodes. Furthermore, if the authors' proposed model based frame scheduling scheme is used, even the relatively small throughput reduction in this case is virtually eliminated. Therefore we would not expect the effect of increased contention to significantly affect the MTM when RTS/CTS is used.

An additional result of increased hop count is that there are more interface queue buffers along the path a packet must traverse. This increased amount of buffering may lead to an increase in end-to-end latency when the network is congested. While trading end-to-end latency for increased throughput is completely appropriate for bulk data transfer applications, this is not the case for delay sensitive traffic. Priority queues should be used on the intermediate nodes regardless of the routing metric used. This eliminates the need to wait in line at multiple buffers. It is also important to realize that although a min hop path may seem appropriate for delay sensitive traffic, it may actually take longer to deliver a packet over the min hop path than an MTM path. This is because it takes longer for a non-zero sized packet to be delivered across across a low-rate link as opposed to a high-rate link. Many types of delay sensitive traffic, such as Telnet, use relatively small packet sizes. Small delay sensitive packets would benefit from MTM routes tuned for small packet sizes, or an MTM implementation which tracks multiple packet sizes.

Effect of Density. Routing protocols that use the medium time metric choose paths that minimize the total consumed medium time. We have argued that these paths should yield significant throughput gains when compared with minimum hop paths. However, this assumes that a path exists that utilizes less medium time than the minimum hop path. This may not be the case. Whether a better MTM path exists depends solely on the current network topology. In general, the likelihood of there existing a smaller medium time path increases as the density of the network increases.

When the density of the network is low, the topology becomes sparsely connected. This yields few choices for routing protocols to select from. In this situation, MTM and min hop will tend to pick the same path. Conversely, as the
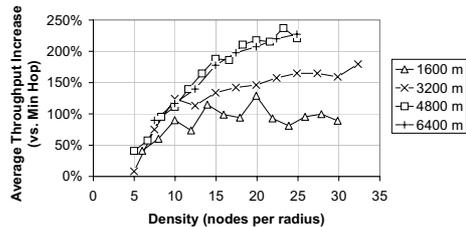
Fig. 3. Average throughput increase of MTM along a randomized straight line path.

network density increases, the abundance of nodes creates a dense, heavily interconnected topology. Routing protocols are provided with a multitude of paths from which to choose. This large number of choices allows the natural tendencies of each metric to be expressed fully.

We have constructed a simple experiment designed to illustrate the relationship between density and the performance of the MTM. A variable number of nodes are randomly placed along a straight line path of fixed length. A single UDP flow is setup between the source and destination, which are placed at opposite ends of the line. Figure 3 shows the relative throughput of the MTM and min hop routing protocols as the number of nodes and the line length are varied. The vertical axis shows the percent increase in achieved throughput over the min hop path when using the MTM. The horizontal axis shows the normalized density of the topology. We define the *normalized density* as the average number of nodes within the maximum transmission range of a given node.

The results show a clear relationship between node density and increased throughput. As expected, at low densities we see low increases as both the MTM and min hop metric pick nearly the same path. As the density increases, we see the full potential of the MTM revealed. The MTM path yields greater than three times (+200%) the throughput of the min hop path with the higher densities and longer path lengths. Longer paths yield more increased throughput than shorter paths because the MTM path utilizes the extra medium time available in long paths (from spatial reuse) much more efficiently than the min hop path.

7   Simulation Results

The purpose of this section is to evaluate the techniques proposed in this paper in a full simulated network environment. We explored the throughput gains provided by both our proposed medium time metric (MTM) and the temporally fair opportunistic auto rate (OAR) protocol over the traditional minimum hop (min hop) metric and the packet fair receiver based auto rate protocol (RBAR).

In order to implement the MTM we modified the DSDV routing protocol [6]. DSDV was selected because it is a simple example of a distance vector based proactive protocol. A distance vector based protocol is desirable in order to show that tuning for a single packet size works in a full network. A proactive

protocol was chosen in order to allow the MTM to fully extract the maximum throughput potential in a continually changing network environment. Typical on-demand protocols do not reroute until the path breaks, but a typical MTM path is constructed of short links and thus will allow considerable mobility before breaking. While this is in general a good feature, as the links lengthen due to mobility they will drop to lower rates and the throughput of the path will degrade. A proactive protocol will do a better job of preventing this degradation from occurring because it will not wait for the path to break, and thus will serve as a better platform to illustrate the potential of the MTM.

To enable the MTM two modifications to DSDV were required. The main modification was to change the computation of the routing metric, but an additional modification related to settling time was also required. In order to accomplish the metric change, we used signal information passed up from the MAC to predict the operational rate of links. Once the link rates are known, then integer weights 5, 7, 14, and 25 are used for link rates 11, 5.5, 2, and 1 respectively. These weights are directly proportional to the MTM weights discussed above, and allow all paths of up to 10 hops to be encoded in the single byte metric field used by DSDV.

Fixed Parameters. The wireless physical parameters given in Section 3 are used. In every simulation, a random way-point mobility model is used. Our simulations are setup for high mobility: the maximum speed is set to 20 meters per second and the pause time is set to zero seconds. In order to emulate a network under high load, we setup 20 flows of TCP traffic. We use the delayed acknowledgement option of TCP in order to reduce the medium time consumed by TCP acknowledgements. Each average gain result is computed from the gains in at least 25 random scenarios. Each scenario is created using a random number seed that generates the initial node placement and mobility pattern. The gains are computed by simulating each of the four protocol combinations (RBAR & min hop, RBAR & MTM, OAR & min hop, and OAR & MTM) in the exact same scenario, and then dividing the resulting total throughput by the base combination (RBAR & min hop). The base combination is representative of both the standard 802.11 MAC fairness model and the metric used by the majority of existing ad hoc routing protocols. This technique of computing gains prevents scenarios with high throughput from skewing the final average. Min hop results are obtained by using the standard DSDV protocol. MTM results are obtained using the modified DSDV. MTM link weights are tuned to match both the TCP traffic, which carries a 1460 byte payload in these simulations, and the selected auto rate protocol (RBAR or OAR).

Varying Parameters. The primary variable examined in this section is node density. The effect of node density on throughput gains was shown at the end of Section 6.3, but only in a simpler one dimensional line case. The central question this section hopes to answer, is how many nodes are required to reach the point where the MTM metric can increase throughput by selecting better
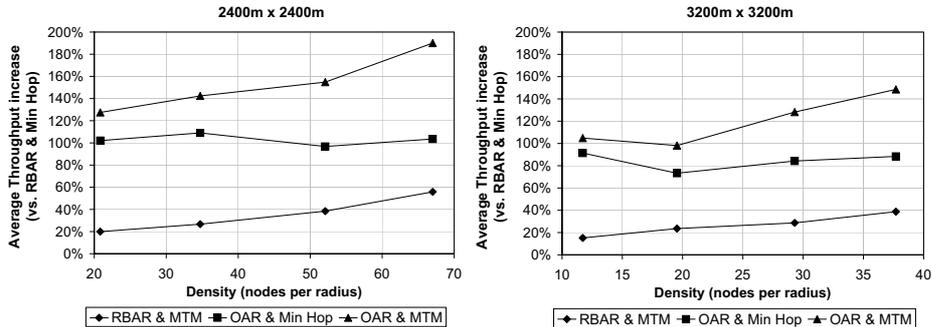
Fig. 4. Random Motion Average Throughput Gains

paths. It is clear that almost any reasonable routing metric will achieve similar performance when the network density is low because the number of available paths to choose from is limited. Since we have defined the *normalized density* as the average number of nodes within the maximum transmission range of a given node, the density in these scenarios is a function of both the number of nodes in the simulation and the total area of the simulation topology. We present simulation results for 60, 100, 150 and 193 nodes in 2400 meter by 2400 meter and 3200 meter by 3200 meter sized topologies.

Results. Figure 4 shows the average throughput gains with respect to the throughput of the RBAR & min hop combination. The average gains of the RBAR & MTM combination represent the throughput increase achieved by our proposed medium time metric under the standard packet fairness model of 802.11. As expected, we see a clear increasing trend in average gain as the density increases. Even at the lowest node density, MTM provides a modest 18% average increase. At the highest simulated density, we see a more substantial 56% average increase. The gains should continue to climb with even higher densities until a plateau is reached. Due to the increased degree of freedom in comparison to the line case, the plateau should not occur until high densities are reached.

For reference, the two sizes of simulations used in [27], a performance comparison of AODV and DSR, were 1500 meters by 300 meters with 50 nodes and 2200 meters by 600 meters with 100 nodes. Given the 250 meter nominal range used in this comparison, these simulations have the normalized densities of 21.8 and 14.9 respectively.

The average gains of the OAR & min hop combination show the throughput increase produced by the OAR protocol without changing the routing metric. As shown in the results, OAR provides quite a substantial boost in total network throughput. The gains provided by OAR come from two sources: increased overall network efficiency due to the increased proportion of time spent sending at high rates, and reduced MAC overhead due to amortization over a multiple packet burst. As a result, the OAR gains are relatively constant with respect to the node density. This experiment illustrates that the OAR protocol should

be used in high throughput multi-rate networks even if min hop is used as the routing metric.

Our analysis of the wide variety of phenomena that affect the throughput in multi-rate ad hoc wireless networks suggests that our proposed medium time metric and the OAR protocol should function well together. The MTM generally selects paths with higher rate links than the min hop, and thus gains an increased benefit from the reduced MAC overhead of high-rate links provided by OAR. Since MTM picks a greater number of high-rate links, it receives less benefit from the temporal fairness property of OAR, but is still helped in the case where paths with fast links are not available. The simulation results show that OAR and MTM do indeed function well together. The contribution of the MTM introduces the same kind of dependance on density that we saw in the pure MTM results. In the most dense simulated case, the total network throughput is almost tripled on average. These massive throughput gains lend support to the validity of both the analysis and solution presented in this paper.

8   Conclusion

In this work we have shown that minimum hop protocols tend to select paths with long slow links. As a result, these paths have low effective throughput and increase total network congestion. In addition, these paths are likely to contain long links that result in low reliability.

We have presented an improved technique for route selection in multi-rate ad hoc wireless networks. The medium time metric is proportional to the time it takes to transmit a packet on a given link. This metric selects paths that have the highest effective capacity. We have also shown the optimality of this technique under the full interference condition by presenting a formal theoretical model of the attainable throughput of multi-rate ad hoc wireless networks.

Our simulation results show an average throughput gain of 20% to 60%, depending on network density, over traditional minimum hop route selection in 802.11b networks. By combining the MTM with the Opportunistic Auto Rate (OAR) protocol, an increase of 100% to 200% is obtained over the traditional route and rate selection techniques. Our results demonstrate the importance of inter-layer communication in ad hoc routing protocol design.

References

1. *IEEE Std 802.11b-1999*, http://standards.ieee.org/.
2. *IEEE Std 802.11a-1999*, http://standards.ieee.org/.
3. *Draft ESTI EN 301 893 version 1.1.1: Broadband Radio Access Networks; HIPER-LAN Type 2*, http://www.etsi.org/.
4. Charles E. Perkins and Elizabeth M. Royer, *Ad hoc Networking*, chapter Ad hoc On-Demand Distance Vector Routing, Addison-Wesley, 2000.
5. David B. Johnson, David A. Maltz, and Josh Broch, *DSR: The Dynamic Source Routing Protocol for Multi-Hop Wireless Ad Hoc Networks. in Ad Hoc Networking*, chapter 5, pp. 139–172, Addison-Wesley, 2001.

6. Charles E. Perkins and Pravin Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers," in *ACM SIGCOMM'94 Conference on Communications Architectures, Protocols and Applications*, 1994.

7. A. Laouiti P. Muhlethaler a. Qayyum et L. Viennot T. Clausen, P. Jacquet, "Optimized link state routing protocol," in *IEEE INMIC*, Pakistan, 2001.

8. Douglas S. J. De Couto, Daniel Aguayo, Benjamin A. Chambers, and Robert Morris, "Performance of multihop wireless networks: Shortest path is not enough," in *Proceedings of the First Workshop on Hot Topics in Networks (HotNets-I)*, Princeton, New Jersey, October 2002, ACM SIGCOMM.

9. R. Dube, C. Rais, K. Wang, and S. Tripathi, "Signal stability based adaptive routing (ssa) for ad hoc mobile networks," February 1997.

10. Henrik Lundgren, Erik Nordstrom, and Christian Tschudin, "Coping with communication grey zones in ieee 802.11b based ad hoc networks," in *WoWMoM 2002*, September 2002.

11. *ANSI/IEEE Std 802.11, 1999 Edition*, http://standards.ieee.org/.

12. A. Kamerman and L. Monteban, "WaveLAN-II: A high-performance wireless lan for the unlicensed band," in *Bell Labs Technical Journal*, Summer 1997, pp. 118–133.

13. Anand R. Prasad and Henri Moelard, "WaveLAN-II system design note 225: Enhanced data rate control," March 1999.

14. Gavin Holland, Nitin H. Vaidya, and Paramvir Bahl, "A rate-adaptive MAC protocol for multi-hop wireless networks," in *Mobile Computing and Networking*, 2001, pp. 236–251.

15. B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly, "Opportunistic media access for multirate ad hoc networks," September 2002.

16. "The network simulator - ns2," http://www.isi.edu/nsnam/ns/.

17. "Rice networks group," http://www-ece.rice.edu/networks/.

18. "Orinoco wireless networks," http://www.orinocowireless.com/.

19. T. Leighton, F. Makedon, S. Plotkin, C. Stein, E. Tardos, and S. Tragoudas, "Fast approximation algorithms for multicommodity flow problem," in *Proc. 23rd ACM Symp. on Theory of Computing*, May 1991, pp. 101–111.

20. J. Aspnes, Y. Azar, A. Fiat, S. Plotkin, and O. Waarts, "On-line machine scheduling with applications to load balancing and virtual circuit routing," in *Proc. 25th ACM Symp. on Theory of Computing*, May 1993, pp. 623–631.

21. Baruch Awerbuch, Yossi Azar, and Serge Plotkin, "Throughput competitive online routing.," in *Proc. 34th IEEE Symp. on Found. of Comp. Science*. Nov. 1993, pp. 32–40, IEEE.

22. Jinyang Li, Charles Blake, Douglas S. J. De Couto, Hu Imm Lee, and Robert Morris, "Capacity of ad hoc wireless networks," in *Proceedings of the 7th ACM International Conference on Mobile Computing and Networking*, Rome, Italy, July 2001, pp. 61–69.

23. Cooperative Association for Internet Data Analysis (CAIDA), "Analysis of NASA Ames Internet Exchange Packet Length Distributions," http://www.caida.org/.

24. *IEEE Std 802.3-2002*, http://standards.ieee.org/.

25. Vincent D. Park and M. Scott Corson, "A highly adaptive distributed routing algorithm for mobile wireless networks," in *INFOCOM (3)*, 1997, pp. 1405–1413.

26. Hwangnam Kim and Jennifer 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*, 2003.

27. C. Perkins, E. Royer, S. Das, and M Marina, "Performance comparison of two on-demand routing protocols for ad hoc networks," in *IEEE INFOCOM*, 2000.