Congestion-Aware Routing Protocol for Mobile Ad Hoc Networks

Xiaoqin Chen[1]   Haley M. Jones[2]   A .D .S. Jayalath[1]

e-mail: {Sandra.Chen, Haley.Jones, Dhammika.Jayalath}@anu.edu.au

[1]Department of Information Engineering, CECS   [2]Department of Engineering, CECS

The Australian National University   The Australian National University

Canberra ACT 0200, Australia   Canberra ACT 0200, Australia

Abstract— **Congestion leads to transmission delay, packet loss, and wastes time and energy for congestion recovery. Routing protocol adaptive to mobile ad hoc networks congestion status can greatly improve network performance. This paper proposes a congestion-aware routing protocol for mobile ad hoc networks (CARM) which uses a metric incorporating data-rate, channel delay, buffer delay, and retransmission count to combat congestion and improve network utilization. This metric is used, together with the avoidance of mismatched link data-rate routes, to make ad hoc networks robust and adaptive to congestion.**

Index Terms—**ad hoc network, routing, congestion-aware.**

I. INTRODUCTION

Congestion occurs in ad hoc networks with limited resources. In such a network, packet transmission frequently suffers from collision, interference, and fading, due to shared radio and dynamic topology. Transmission errors burden the network load. Recently, there is an increasing demand on supporting multimedia communications in ad hoc networks. The massive real-time traffics are in bursts, bandwidth intensive, and congestion liable. Congestion in ad hoc networks leads to packet loss, bandwidth degradation, and wastes time and energy for congestion recovery. Congestion-aware routing protocol can preemptively settle a congestion through bypassing the congested links.

Wireless standards, such as IEEE 802.11a/b, support adaptive transmission in ad hoc networks to accommodate time-varying radios. However in [1], the authors point out that, when operating under heavy traffic conditions (every node always has packets to transmit), 802.11 DCF provides long term per packet fairness in single-hop networks, which incurs network performance anomaly: in an one-hop network, the active low data-rate nodes decrease the throughput of other high data-rate nodes. One of the solutions to the throughput decrease in multi-rate networks is to use multiple channels. However, another solution, which is feasible in multi-hop ad hoc networks, is to employ routing protocol to make the low data-rate nodes "asleep" as much as possible. The routing protocol selects higher data-rate links to build a route, to reduce the usage of the low data-rate nodes, or not use them as intermediate nodes at all. Because low data-rate nodes have a lower probability to attend a packet relay, each one-hop network operates with high data-rate nodes mostly, to improve the whole network throughput.

Choosing higher data-rate links, as the medium time metric (MTM) suggested in [2], will generally mean that they are short and more links are needed in any given route. This is an advantage in that we have higher data rates while the packets are in transmission along these links. However, more links in a route also means more access contention, potentially increasing congestion. Then in [3], the use of channel access delay was proposed as an enhancement to the MTM, providing awareness of congestion to help avoid routing through bottleneck regions. When applied in a reactive routing protocol, [3] is equivalent to the shortest-delay routing metric used in Dynamic Source Routing (DSR) [4]. However, accurately measurement of the link congestion level should combine channel occupation, packet drop rate, and buffer load [5]. Literatures of congestion measurement can be found in [5], [6].

Further, in multi-rate ad hoc networks, different data-rates will almost certainly lead to some routes having different links with quite different data-rates. If lower data-rate links follow higher data-rate links, packets will build up at the node heading the lower data-rate link, leading to long queueing delays. A further cause of congestion is link reliability. If links break, congestion is increased due to packet salvage.

In this paper, we first propose a congestion-aware routing metric which employs the retransmission count weighted channel delay and buffer queuing delay, with preference of less congested high throughput links to improve channel utilization. Then, we propose the Congestion Aware Routing protocol for mobile ad hoc networks (CARM). CARM applies a link data-rate categorization approach to prevent routes with mismatched link data-rates. In this paper, CARM is only discussed and simulated in relation IEEE 802.11b networks, however, it can be applied to any multi-rate ad hoc network.

The layout of this paper is as follows. In Section II we describe issues with congestion in multi-rate ad hoc networks. In Section III and Section IV we describe the proposed metric and congestion-aware routing protocol. Simulation results are presented in Section V and in Section VI we draw our conclusions.

II. CONGESTION IN MULTI-RATE AD HOC NETWORKS

A. Mismatched Link Data-Rate Routes

In multi-rate ad hoc networks, throughput via a given route is limited by the minimum data-rate over all of its constituent links. Consider a route with significantly different data-rates over each of its links (e.g., $A \rightarrow B \rightarrow D \rightarrow F \rightarrow H$ in Fig. 1). Let us call such a route a *mismatched data-rate*

route (MDRR). When large-scale traffics, such as multimedia
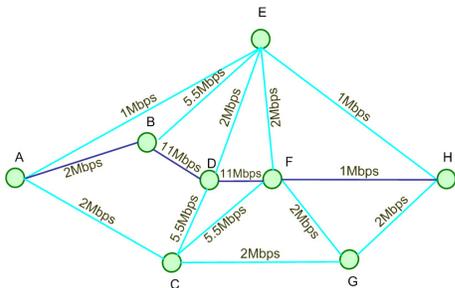

Fig. 1.   An example of an 802.11b multi-rate ad hoc network.

streams, are transmitted in such a route, the benefits of having multi-rate links can be compromised. There is potential for congestion at any node which heads a link with a slower data-rate than previous links, in a MDRR route, (e.g., node F in the example path), due to earlier high data-rate nodes forwarding more traffic into low data-rate nodes than they can handle. Long queuing delays may occur on such paths, dominating the end-to-end delay. Clearly, avoiding, or at least lessening the mismatch in, MDRRs is important in combatting congestion.

B. MAC Overhead in Congestion

In this paper, we consider the network with 802.11 [7] MAC with the distributed coordination function (DCF). In such networks, the standard packet sequence is: request-to-send (RTS), clear-to-send (CTS), data, acknowledge (ACK). The amount of time between the receipt of one packet and the transmission of the next is called a short interframe space (SIFS). So, the minimum channel occupation due to MAC overhead is

$$T_{\mathrm{MACmin}} = T_{\mathrm{RTS}} + T_{\mathrm{CTS}} + 3T_{\mathrm{SIFS}} \qquad (1)$$

where T_{RTS} and T_{CTS} are the time consumed on RTS and CTS, respectively, and T_{SIFS} is the SIFS period. Here we have not included the ACK, for the purpose of using the overhead later. If we include the time taken due to contention for the channel, the channel occupation due to MAC overhead is

$$T_{\mathrm{MACall}} = T_{\mathrm{MACmin}} + T_{\mathrm{cgs}} \qquad (2)$$

where T_{cgs} is the time taken due to access contention (including NAV waiting and back-off intervals) and retransmissions.

Let the channel delay for the link ($\ell_{i,j}$) between nodes i and j be defined as the interval between the start of the RTS transmission at node i and the time the data packet is correctly received at node j. The channel delay is given by

$$\tau_{i,j} = T_{\mathrm{MACall}} + T_{\mathrm{data}} \qquad (3)$$

where $T_{\mathrm{data}} = L_{\mathrm{data}}/R_{i,j}$ is the data transmission time, L_{data} is the length of the data, and $R_{i,j}$ is the data-rate of the link.

The amount of MAC overhead, T_{MACall}, is dependent upon the access contention for the shared medium, and the number of packet collisions. That is, T_{MACall} is strongly related to the congestion around a given node. With little or no access contention, the end-to-end delay is effectively a constant, given

by $T_{\mathrm{MACmin}} + T_{\mathrm{data}}$ from (1) and (3). When medium access contention or packet retransmissions are included, T_{MACmin} is replaced with T_{MACall} which is variable and can become relatively large if congestion is incurred and not controlled.

MAC overhead can dramatically decrease the capacity of the congested link. For example in Fig. 2, if only the physical bit rate is applied, the link in scenario II (11Mbps) would be said to have a higher capacity than the link in scenario I (5.5Mbps). However, when the MAC delay is included, the links in the two scenarios turn out to have identical overall channel delays, giving them the same *real* channel capacities. Therefore, in the design of a congestion aware metric for multi-rate ad hoc networks, the data-rate and the MAC delay should be jointly considered to more accurately indicate channel capacity.
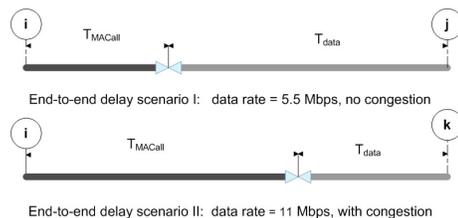

Fig. 2.   Two scenarios with the same overall delay but different MAC and transmission delays due to different data-rates and congestion levels.

C. Channel Reliability

Packet transmission in ad hoc networks is affected by the reliability of the radio due to factors such as interference and fading. Not only will congestion deteriorate performance with respect to packet losses, but increased packet losses will lead to more congestion due to higher packet retransmission. Thus, packet loss indicates channel congestion and reliability. Again using the example in Fig. 2, it can be seen that $\ell_{i,k}$ in scenario II has a longer T_{MACall} than $\ell_{i,j}$, in scenario I. This indicates that $\ell_{i,k}$ may have higher channel utilization or severer congestion. Now, if we used just the channel delay for link selection, $\ell_{i,k}$ and $\ell_{i,j}$ would be equally likely to be selected. However, the higher congestion levels in $\ell_{i,k}$ mean that it has greater potential for performance degradation, in terms of higher levels of corruption or retransmission. So, if we include some measure of packet loss into link selection, we reduce the chances of selecting a congested node. In the 802.11 DCF, packet is dropped after several failed transmission reties. So, we can use the MAC retransmission count to represent the channel reliability.

III. CONGESTION-AWARE ROUTING METRIC

Congestion-aware routing metric for multi-rate ad hoc networks should incorporate transmission capability and congestion around a link. In the previous section we saw that congestion in ad hoc networks is related to access contention, channel occupation, and channel reliability. Channel delay from (3) is a good measurement of congestion with combination of access contention, channel occupation, and data-rate, and it is easy to implement with low overhead. In addition to channel delay, queuing delay in interface queue is also an useful measure of

congestion. Another factor affecting congestion is channel reliability, which can be measured by retransmission count.

We now introduce the weighted channel delay (WCD) which assigns a cost to each link in the network using the aforementioned physical layer and MAC layer information. The WCD utilizes the channel delay $\tau_{i,j}$ from (3), as well as the queuing delay in the interface queue of an intermediate node, to select maximum throughput paths, avoiding the most congested links.

For an intermediate node i with established data packet transmission with several of its neighbours, the WCD for the link from node i to a particular neighbouring node j, is given by,

$$\text{WCD}_{i,j} = a \sum_{k \in \vartheta_i} \tau_{i,k} Q_{i,k} + \tau_{i,j}(1 + n_{i,j}b)$$
$$= Q_i + \tau_{i,j}(1 + n_{i,j}b) = Q_i + D_{i,j} + \tau_{i,j} \quad (4)$$

where ϑ_i is the set of all nodes neighbouring node i, $Q_{i,k}$ is the number of packets buffered in the node i queue bound for node k, $n_{i,j}$ is the retransmission count for $\ell_{i,j}$ and Q_i is the total queuing delay for node i. The constants a and b are network-specific parameters with values between 0 and 1. The WCD seeks to capture as many effects and indicators of congestion as possible, so that the network is aware of local congestion. A smaller WCD for a link is more favourable, meaning that the link is more likely to be selected in any given route.

IV. CONGESTION-AWARE ROUTING PROTOCOL

CARM is an on-demand routing protocol that aims to create congestion-free routes by making use of information gathered from the MAC and physical layer. The CARM route discovery packet is similar to that in DSR [4] where every packet carries the entire route node sequence. CARM employs the WCD metric in (4) to account for the congestion level. In addition, CARM adopts a route effective data-rate category scheme to combat the MDRR problem discussed in Section II-A. The combination of these two mechanisms enables CARM to ameliorate the effects of congestion in multi-rate networks. CARM uses the same route maintenance approach as that in DSR.

A. Addressing Mismatched Data-Rate Routes

Because the effective bandwidth of a link can be dramatically degraded by congestion, regardless of its specified physical bit rate, we introduce the effective link data-rate for $\ell_{i,j}$ as

$$\mathcal{D}_{i,j}^{\text{eff}} = \frac{L_{\text{data}}}{\tau_{i,j}} \quad (5)$$

where $\tau_{i,j}$ is defined in (3). We next introduce the effective link data-rate category (ELDC) scheme, where each link is marked by its ELDC type which is determined by its effective link data-rate range. For example, in an IEEE 802.11b network with data-rates ranging from 1Mbps to 11Mbps, we might choose the following two categories:

$$\text{ELBC I} : \mathcal{D}_{i,j}^{\text{eff}} < 6\text{Mbps}; \quad \text{ELBC II} : \mathcal{D}_{i,j}^{\text{eff}} \geq 6\text{Mbps}.$$

For a given route, the route ELDC is taken as that for the link directly connected to the source and is included in the route request (RREQ) packet. During route discovery, an intermediate node only forwards a RREQ if the ELDC type of the link preceding the current node is higher than or equal to that of the route. That is, for two ELDCs, if the route ELDC is I then all paths are possible. However, if the route ELDC is II, only links with ELDC II may be chosen, eliminating low data-rate links and lessening the chances of congestion. This lessens the occurrence of very slow initial links being teamed with very fast links in the same route.

While using the ELDC scheme helps to alleviate the MDRR problem, in extreme cases, the limiting of choice of links in a route could lead to route discovery failure. To counter this situation, we include a field, ELDCF, in the RREQ packet to flag whether or not the ELDC scheme is in operation. It is utilised in the following way. On an initial route discovery attempt, the ELDCF field is set to 1, indicating that the ELDC scheme is in use, such that nodes should only forward the RREQ under the ELDC rules given above. If this route discovery process is unsuccessful, another is initiated, this time with the ELDCF field set to 0. In this way, all RREQs are forwarded as in DSR.

B. Route Discovery

Now we describe the route discovery process in CARM.

1) RREQ Initiation: A source node, i, wishing to transmit data to a given destination node, generates a RREQ which it transmits to the neighbours. The RREQ packet to neighbour j contains the following fields: <source ID, source sequence number, destination ID, transmission start-time at i, Q_i, $n_{i,j}$, ELDC, ELDCF, record of route hop sequence>. The ELDC field is assigned appropriately at the first intermediate node. For the first route search cycle the ELDCF field is set to 1, indicating that the ELDC scheme is in use. If this route discovery is unsuccessful, ELDCF is set to 0 and a second cycle is initiated.

2) Processing a RREQ: Each intermediate node maintains a local forwarding list of the triples <source ID, source sequence number, ELDC> to record and keep track of the RREQs that it has received. Upon receiving a RREQ packet, an intermediate node compares the appropriate fields in the RREQ and its local list to avoid propagating duplicate RREQs. The ELDCF field is also checked. If ELDCF = 1, the ELDC of the preceding link is determined and compared with that in the RREQ. If the link ELDC is lower than the route ELDC, the RREQ is discarded.

Note that in DSR intermediate nodes drop any RREQ with the same source ID and lower or identical source sequence number to those in any RREQs they have already seen. So, in DSR each node only forwards a RREQ, during a given route discovery process, once. In CARM, as the ELDC is also taken into account, any node may forward a RREQ during a route discovery up to the number of ELDC types. This means that more routing information is required to establish feasible routes because more copies of the same RREQ are propagated around the network. This causes a slight increase in overhead during the route discovery phase of CARM over DSR.

3) Prioritizing RREQ with WCD: In the interface queue routing packets have higher priority over data packets, such that they are forwarded immediately, without queueing. Because of this, the congestion level information inherent in queueing delays is lost in DSR. This is addressed in CARM via the WCD described in Section III.

Having determined to keep a RREQ from node i, node j calculates $(Q_i + D_{i,j})$ from the $\text{WCD}_{i,j}$ in (4) according to the information recorded in the RREQ. Then, node j delays forwarding of the RREQ by this amount, so that the total time that a RREQ is delayed, from the time it is sent by node i, to the time that node j is ready to forward it is equal to $\text{WCD}_{i,j}$. This ensures that each node forwards RREQs on a priority basis related to congestion level as encompassed by the WCD. So, RREQs for routes with higher throughput and lower congestion will reach the destination first and, because the intermediate nodes will drop later arriving duplicate RREQs, congested links are much less likely to be included in any established routes.

4) Route Reply: As part of CARM, intermediate nodes are prohibited from generating route reply (RREP) packets. That is, only the destination node may generate and send RREPs, to avoid stale information at intermediate nodes. The destination responds to RREQs by sending a RREP packet back to the source along the route via which it came. The first RREP to reach the source establishes the route. Routes indicated in any subsequent RREPs are cached at the source in case of failure of the established route.

V. SIMULATIONS

In the simulations, we compare performance of DSR with two slightly different version of CARM protocol. In the first, only the WCD metric is taken into account in DSR, which we call CARMdelay. In the second, both the WCD and the ELDC scheme are taken into account, which we call CARM. The simulations were carried out using the network simulator, ns-2.29 [8] with adaptive auto-rate feedback [9] multi-rate extension. DSR works by building routes based on the shortest end-to-end delay, lessening control overhead by allowing intermediate nodes to issue RREPs using cached routing information.

The simulations assumed an IEEE802.11b network, configured with 80 nodes uniformly distributed over a 1500m × 1500m area, moving according to a random waypoint model [10] with a maximum speed of 5m/s and a pause time of 10 seconds. Ten nodes were randomly chosen to be constant bit rate (CBR) sources, generating 512 byte data packets to be sent to randomly chosen destinations. The network traffic was increased from 10 to 80 packets per second, with each simulation running for 300s. The MAC layer was based on IEEE 802.11 DCF with a control packet transmission rate of 1Mbps. The interface queue at the MAC layer had a length of 50 packets while the routing buffer at the network layer had a length of 64 data packets. The transmission power was fixed at 13dBm. The simulation receiving threshold powers for various data-rates and their respective transmission ranges (based on the two-ray ground propagation model) are shown in Table I. Note that for calculating the WCD component in CARMdelay and CARM, from (4), we chose $a = 0.25$, $b = 0.75$, based on an comparison of results for a range of values for a and b.

In developing the simulations, we considered the following properties to assess the performance of routing protocols: (1) **Packet delivery ratio** (PDR): the ratio of the number of data packets successfully received at the destinations to the number of data packets generated by the sources; (2) **Average end-to-end delay**: the average time taken to transfer a data packet from

TABLE I
SIMULATION PARAMETER FOR MULTI-RATE

data rate	receiving threshold	transmission range
carrier sense	-101dB	1071m
1Mbps	-91dB	597m
2Mbps	-89dB	532m
5.5Mbps	-87dB	475m
11Mbps	-82dB	356m

a source to the destination; (3) **Normalized routing control overhead**: the ratio of the number of control packets to the number of data packets received at the destinations.

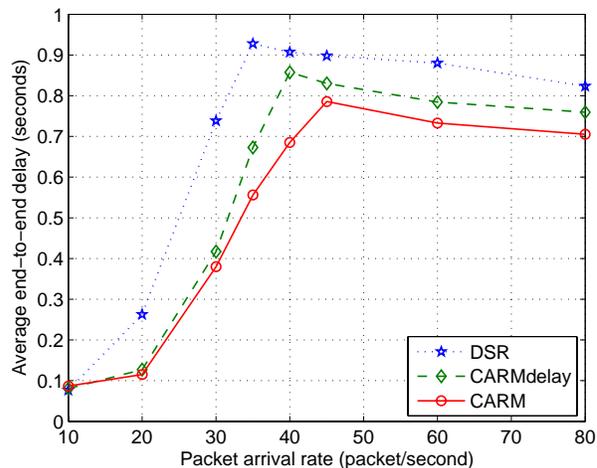### A. End-to-End Delay versus Packet Arrival Rate

Fig. 3. Comparison of end-to-end delay with increasing packet arrival rates from 10 to 80 per second for CARM, CARMdelay and DSR.

Fig. 3 illustrates the average end-to-end delay for CARM, CARMdelay and DSR with an increase in the packet arrival rate. It can be seen that both CARM and CARMdelay outperform DSR with respect to the effect of traffic level on end-to-end delay. In DSR, if a route becomes disconnected, the source then attempts to make use of cached routes in either the source node itself or intermediate nodes before initiating another route discovery. However, the DSR link error notification mechanism means that not all nodes necessarily find out about a breakage until they next use that link, so many cached routes may be stale. Unwittingly attempting to forward data through such routes uses up transmission time. The use of the WCD in effectively delaying RREQs which have come through congested links means that such links are unlikely to be included in established routes in CARMdelay and CARM. The additional use of the ELDC scheme in CARM further combats congestion.

B. Packet Delivery Ratio versus Packet Arrival Rate

Fig. 4 illustrates the trend of the packet delivery ratio with increasing packet arrival rate. It can be seen that CARM and CARMdelay outperform DSR, particularly for higher traffic
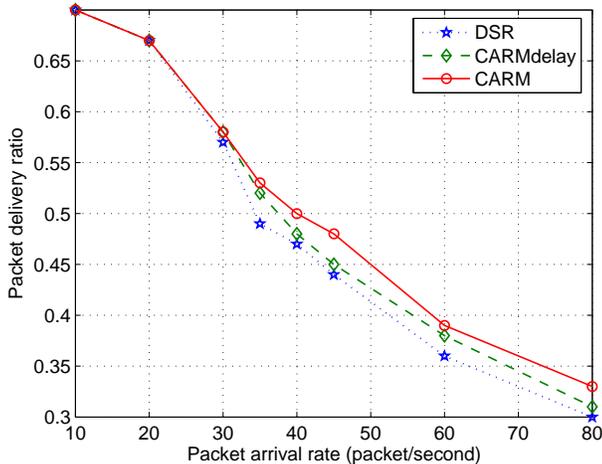
Fig. 4. Comparison of packet delivery ratio (PDR) with increasing packet arrival rates from 10 to 80 per second for CARM , CARMdelay and DSR.

loads. At higher traffic loads, in general, links face a higher probability of being unusable due to node congestion and the packet drop rate increases due to collisions or buffer overload. DSR uses cached routes to re-establish the connection when the current route becomes unusable. The cached routes may be stale, and no longer optimal for the current network topology. Higher traffic levels make the cached routes more prone to congestion, leading to more dropped packets. From the results, we can see that the usage of the WCD metric can increase the number of packets delivered in DSR by up to 10%. The employment in CARM of the WCD, combined with the avoidance of mismatched link data-rate routes, aids in the selection of routes more robust to congestion. Through these mechanisms, CARM is able to deliver more packets than DSR.

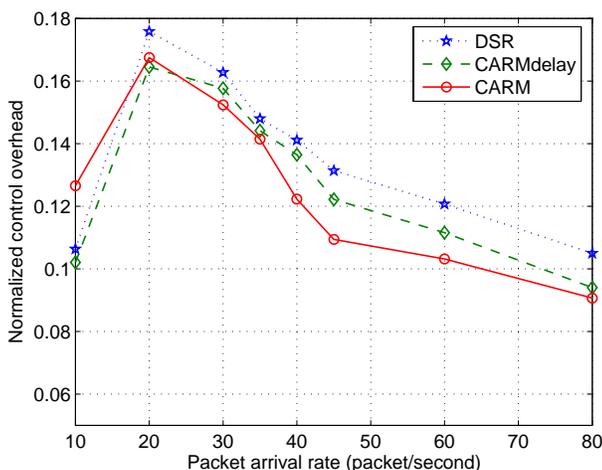### C. Control Overhead versus Packet Arrival Rate

Fig. 5. Comparison of normalized control overhead with increasing packet arrival rates from 10 to 80 per second for CARM, CARMdelay and DSR.

Finally, Fig. 5 shows the trend of the normalized control overhead with increasing traffic, for CARM, CARMdelay and

DSR. It can be seen that CARMdelay and CARM generally outperform DSR with respect to normalised control overhead, except under light traffic where DSR outperforms than CARM. When the traffic load is light, the route discovery packets dominate the control overhead. CARM requires more routing information to establish routes because of the ELDC mechanism. However, because CARM only allows RREQs to be broadcast if the ELDC is appropriate, thereby somewhat lessening the impact of RREQ propagation. At high traffic loads, the network is more prone to congestion, most likely increasing the stale cached routes in DSR and increasing the likelihood of the need for a new route discovery. It has been noted [11], that in DSR the routing load is dominated by RREP packets. However, in CARM this is not the case, due to the suppression of RREPs at intermediate nodes. CARMdelay and CARM work to exclude congested links via the use of the WCD. In CARM, ELDCs also contribute to congestion control. So, while DSR yields lower overhead due to route discovery, it requires route discovery more often due to congestion. In CARMdelay and CARM, the reduced number of heavily congested links in established routes contribute to better performance in high traffic loads.

VI. CONCLUSION

We have proposed a congestion-aware routing protocol for mobile ad hoc networks (CARM). CARM utilizes two mechanisms to improve the routing protocol adaptability to congestion. Firstly, the weighted channel delay (WCD) metric is used to select high throughput routes with low congestion. The second mechanism that CARM employs is the avoidance of mismatched link data-rate routes via the use of effective link data-rate categories (ELDCs). In short, the protocol tackles congestion via several approaches, taking into account causes, indicators and effects. The decisions made by CARM are performed locally. Our simulation results demonstrate that CARM outperforms DSR due to its adaptability and robustness to congestion.

REFERENCES

[1] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda, "Performance anomaly of 802.11b," *Proc. of INFOCOM*, vol. 2, pp. 836–843, 2003.
[2] B. Awerbuch, D. Holer, and H. Rubens, "High throughput route selection in multi-rate ad hoc wireless networks," *Proc. of the First working conference on wireless on-demand network systems*, pp. 251–269, 2004.
[3] S. Zhao, Z. Wu, A. Acharya, and D. Raychaudhuri, "PARMA: a PHY / MAC aware routing metric for ad-hoc wireless networks with multi-rate radios," *Proc. of WoWMoM 2005*, pp. 286–292, Jun. 2005.
[4] D. Johnson and D. Maltz, *Dynamic source routing in ad hoc wireless networks*. Mobile Computing, Kluwer Academic Publishers, 1996, ch. 5, pp. 153–181.
[5] J. Kang, Y. Zhang, and B. Nath, "Accurate and energy-efficient congestion level measurement in ad hoc networks," *Proc. of WCNC*, vol. 4, pp. 2258–2263, Mar. 2005.
[6] Y. J. Lee and G. F. Riley, "A workload-based adaptive loadbalancing technique for mobile ad hoc networks," *Proc. of WCNC*, vol. 4, pp. 2002–2007, Mar. 2005.
[7] "IEEE Std 802.11 Wireless LAN Medium Acess Control (MAC) and Physical Layer (PHY) Specifications," 1999.
[8] *http://www.isi.edu/nsnam/ns*.
[9] M. Lacage, M. H. Manshaei, and T. Turletti, "IEEE 802.11 rate adaptation: A practical approach," *Proc. of MSWiM'04*, pp. 126–134, Oct. 2004.
[10] J. Broch et al, "A performance comparison of multi-hop wireless ad-hoc network routing protocols," *Proc. of MOBICOM'98*, pp. 85–97, Oct. 1998.
[11] C. E. Perkins, E. M. Royer, S. R. Das, and M. K. Marina, "Performance comparison of two on-demand routing protocols for ad hoc networks," *IEEE Personal Communications*, vol. 8, pp. 16–28, Feb. 2001.