In Proc. IEEE Wireless Communications and Networking Conference (WCNC), Atlanta, Georgia, USA, March 2004

Queue-driven Cut-through Medium Access in Wireless Ad Hoc Networks

David Raguin, Martin Kubisch, Holger Karl, Adam Wolisz
Telecommunication Networks Group, Technische Universität Berlin
Sekr. FT 5-2, Einsteinufer 25, 10587 Berlin, Germany
raguin|kubisch|karl|wolisz@tkn.tu-berlin.de

***Abstract*—In multi-hop ad hoc networks the IEEE 802.11 MAC protocol, in its distributed version, dramatically degrades the performance in terms of throughput and delay. This is due to a protocol property which tries to provide equal probability of channel access to all nodes. But in multi-hop ad hoc network a more frequent channel access might be necessary at certain nodes that are intermediate hops and forward other nodes' data. We propose a modification to the protocol that leads to a higher throughput and lower delay by combining the ACK for one packet with the channel access procedure for another, queued-to-be-sent packet. Hence, nodes receiving many packets, probably forwarder, have a higher chance of accessing the channel. Additionally, these changes provide a scheme of packet forwarding more robust to different layouts than original IEEE 802.11, where performance can heavily deteriorate in certain configurations.**

Keywords: ad hoc networks, relaying, medium access control (QoS, fairness)

I. Introduction

The most widely used wireless communication protocol for computer networks is the group of IEEE 802.11 standards. This standard has different modes of operation; for ad hoc networks, the distributed coordination function (DCF) along with its request to send (RTS)/clear to send (CTS) [1] handshake is the mode of choice. With this handshake the collisions due to the hidden terminal effect can be reduced, but it requires additional bandwidth for the transmission of the RTS and CTS frames, which are solely used for the handshake.

As this mechanism was developed with one-hop data transfers in mind, it was a later finding that MAC protocols utilizing this mechanism perform poorly in ad hoc multi-hop networks. In simple scenarios such as chain or lattice networks, the achieved throughput is about 40 % below [2] the theoretically achievable one [3].

To improve throughput and latency for a chain of hops in the DCF, Acharya et. al. proposed the Data-driven Cut-through Medium Access protocol (DCMA) [4], which needs additional labels in each packet to mark the flow it belongs too. This protocol merges the acknowledgment (ACK) for a packet reception for one hop with an RTS for the same packet's transmission over the next hop. Thus, the implicit right to send an ACK is exploited to reserve the channel for the next hop, combining two channel accesses into one single "cut-through" access.

To identify a forwarding packet of a flow, a special label is added. A network interface card (NIC) can then quickly lookup the next hop address, necessary for the creation of the ACK/RTS frame — a quick operation is necessary, otherwise the timing requirements for ACK packets would be violated. The drawback is that it requires additional information from the network layer to be present in the NIC, which requires a heavily modified NIC architecture, and that the needed memory in the NIC grows with the size of the network.

But the principle of cut-through access is not limited to flow oriented packet forwarding. Cut-through access can be used to generally speed up the packet send off of a node. Hence, we propose another cut-through protocol, the Queue-driven Cut-through Medium Access protocol (QCMA): Instead of confining the cut-through process to flows of packets, QCMA creates the necessary cut-through ACK/RTS frame with information from the next packet in its outgoing queue. This allows a simpler hardware structure (relaxed processing requirements in the NIC) and the potential for supporting more sophisticated packet scheduling algorithms, e.g., such as picking the packet with the highest priority from multiple queues like in IEEE 802.11e [5] — an option not foreseen for DCMA.

Additionally, as Archarya et. al. showed improvements [6] of DCMA for only regular spaced topologies (equally spaced chain of nodes and regular grid), they did not evaluate it for other configurations. Hence, our paper provides, apart from our new cut-through protocol, also the first performance evaluation of DCMA in more complex setups.

The remainder of the paper is organized as follows: In Section II we present the related work, followed by the description of the QCMA protocol in Section III. Then, we describe the simulation scenarios used in Section IV followed by the presentation of the results in Section V. Finally, we draw conclusions in Section VII.

II. Related work

The disadvantages of the IEEE 802.11 MAC protocol in multi-hop networks has been discussed by Xu and Saadawi [7]. They presented stability and fairness problems which the IEEE 802.11 MAC protocol experiences in such networks. In [2], Li et al. showed that in simple setups such as a chain, IEEE 802.11 uses only $\frac{1}{7}$ of the raw channel bandwidth. Whereas the ideal capacity, as a result of collisions between packets of the same flow, could be as high as $\frac{1}{4}$. Additionally, as there are certain scenarios, which would justify using the Gupta and

Kumar [3] results to assess the maximum capacity, we also show how DCMA and QCMA compare to this bound.

There have been efforts to improve throughput and reduce delay in multi-hop networks, but so far it resulted in quite complex reservation mechanisms. Differentiations between data packets, which are transmitted via the IEEE 802.11 DCF, and time bounded packets, which are transmitted in a dedicated time slot [8], [9] are made. In order to provide quality of service using the IEEE 802.11 MAC, Vaidya et al. proposed to prioritize packets by assigning different weights to different flows of data and change the backoff mechanism to determine the first transmitter [10]. The performance of prioritizing packets is discussed in [11] and [12] and will influence the definition of the upcoming IEEE 802.11e standard.

But reservation and priority mechanisms are complex to implement and do not address the main problem of the IEEE 802.11 standard: it was not designed for multi-hop communication and does not provide the right mechanism to support an efficient packet forwarding operation. In the following we show a forwarding protocol, requiring only minimal changes to the IEEE 802.11 MAC protocol.

III. Proposed Algorithm

A. Design

To overcome the inefficiency of the IEEE 802.11 MAC protocol in intermediate, forwarding nodes of an ad hoc network, we propose the Queue-driven Cut-through Medium Access (QCMA): After receiving a packet and when there is at least one queued-to-be-sent packet present in the NIC, QCMA combines the necessary transmission of an ACK packet for the received packet to perform an immediate medium access procedure for another queued-to-be-sent packet. Due to time constraints of the IEEE 802.11 MAC protocol and the current technique of performing a routing decision outside of the NIC the received packet can not be used for this forwarding operation.

The queued-to-be-sent packet is simply the first packet in the forwarding queue of the NIC when one unique packet queue is used (IEEE 802.11a/b/g), or the first packet of the queue with the highest priority when several prioritized queues are used (IEEE 802.11e). This packet can be from the intermediate node ora packet which must be forwarded on behalf of another node — a decision made by the routing layer — for simplicity, we will refer to it uniformly as a packet “to be forwarded”. An advantage of using the next queued-to-be-sent packet is that QCMA kicks in when necessary — under high load when the queues are full.

More detailed, QCMA works as follows:

1) When the outgoing queue(s) of the NIC are not empty, than after the reception of a DATA packet, QCMA creates an ACK/RTS packet that acknowledges the received packet and contains the MAC address of the next packet to be sent, acting simultaneously as both an ACK and a RTS packet.
2) The creation of the ACK/RTS is not allowed when the destination of the next packet to be sent is the same as the source of the packet to be acknowledged — a bounce back; it would lead to high unfairness as two directly communicating nodes could grab the entire bandwidth and push aside all other nodes.

It is worth mentioning that no modifications is needed as the network allocation vector (NAV) is concerned. As in the IEEE 802.11 MAC protocol, every node which overhears the ACK/RTS exchange and is not addressed, increments its NAV by the time interval included in the ACK/RTS frame. Also no modifications are necessary at the backoff mechanism. The cut-through access is considered as another attempt to transmit the packet. When it fails, the contention window is doubled and when it succeeds, the contention window is initialized to its minimal value, as done in the IEEE 802.11 MAC protocol.

Figure 1 illustrates the cut-through communication with QCMA and one packet queue. Upon receiving a DATA packet

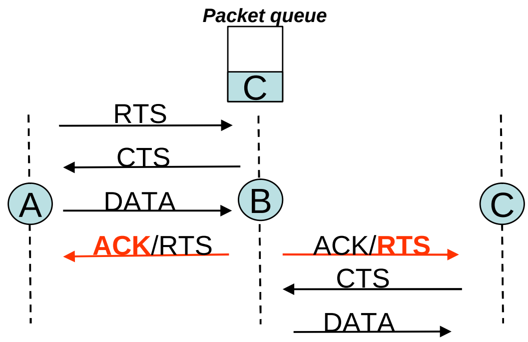

Fig. 1. QCMA cut-through

from node A, node B sends an ACK/RTS frame to node A and simultanueously to C; acknowledging the packet reception from node A and requesting another packet transfer of the first packet in its queue to C.

B. Comparison of QCMA and DCMA

1) Hardware modifications: Since DCMA uses the ACK/RTS frame to forward packets belonging to a certain flow, a label switching table is required in the NIC to resolve the identity of the next hop and avoid the time consuming routing lookup in the host kernel. In contrast, QCMA creates the ACK/RTS frame with the next packet in the queue and no routing lookup in the NIC is needed, thus the hardware modifications should be minimal and the stringent processing requirements (the label switching of DCMA has to be done before a packet is acknowledged) are not necessary; address resolution and routing lookups can be kept in only one place — the routing protocol in the host’s kernel.

2) Compability with the IEEE 802.11e standard: The upcoming IEEE 802.11e standard proposes an Enhanced Distributed Coordination Function (EDCF) that supports up to eight traffic classes with independent transmission queues and Quality of Service (QoS) parameters that determine their priorities. It is obvious that DCMA contradicts this idea to differentiate traffic streams or sources in several queues, because the forwarded packets are not queued but directly send thanks to the cut-through access. This contradiction does not exist with QCMA, given that the ACK/RTS frame is created with the first queued-to-be-sent packet.

3) Frequency of ACK/RTS: While DCMA performs a cut-through access for each packet of a flow, the ACK/RTS usage in QCMA depends on the status of the receiving node's queue. In a system with empty queues QCMA behaves like the IEEE 802.11 MAC protocol does in absence of congestion. When congestion occurs, i.e., the queues of a nodes NIC are going to be populated, QCMA kick is an performs cut-through communication.

In some specific topologies such as a chain of nodes, QCMA may use the ACK/RTS less often than DCMA because of the interdiction to send the ACK and the RTS to the same pair of nodes (we will illustrate this case in Section V). But, in some cases QCMA uses the ACK/RTS where DCMA does not, for example, when a node is the final destination of a packet, but nevertheless has packets to be sent.

IV. Simulation setup

A. Simulation environment

The simulation parameters are chosen to be comparable with [6], i.e., a 550 meters interference range, a 250 meters effective transmission range, a 2 Mbps data rate and a 50 packets queue size in each NIC. A collision occured each time the SNIR of a packet was smaller than 4 dB. Each simulation was running for 3600 second.

All aspects of the IEEE 802.11b standard regarding the DSSS PHY and the DCF are implemented (Inter Frame Spaces, CWmin, CWmax, RTS, CTS, ACK length, etc.). The RTS/CTS exchange was always used. When a node can not perform a cut-through access (the queue(s) are empty or a bounce back), it sends an ACK/RTS frame with an empty RTS address, i.e., it is an conventional ACK with a small overhead. But this small overhead simplifies the packet processing, since at the beginning of a RTS/CTS/DATA/ACK handshake a node does not know whether a cut-through access can be performed or not.

In order to find the shortest route between any two nodes in the network, we chose to use an off-line computation by applying Dijkstra's algorithm [13] at the beginning of the simulation, thus having a forwarding table in every node. Another option would be the use of an ad hoc routing protocol, but given that all nodes remained static this simplification is immaterial to the evaluation of QCMA.

To evaluate a scheme like the one presented here, we considered a fairly detailed modeling of the physical layer with a proper interference representation as necessary. The NS/2-based models did not satisfy our requirements, and a modification appeared cumbersome. Hence, we have implemented the IEEE 802.11b standard with the DSSS PHY and the Distributed Coordination Function (DCF), as well as the DCMA and QCMA enhancements using the OMNeT++ [14] simulation system. This model is available for download at *http://www.tkn.tu-berlin.de/∼kubisch/QCMA*.

B. Scenarios

We used two different simulation scenarios, a pipe scenario and a random scenario. The pipe scenario as seen in Figure 2 consists of a three-hop chain (a "pipe") between two clusters each with ten nodes. The random scenario was a large random network containing 80 nodes, uniformly randomly placed on an area of 1000 m $\times$ 1000 m.

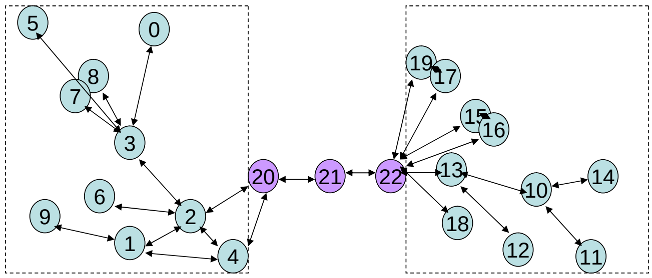

Fig. 2. Example pipe scenario

Every 30 seconds, a node randomly selects a destination node (in the opposite cluster in the case of the pipe scenario or among all other nodes in the case of the random scenario), and sends unidirectional CBR traffic to this destination. The packets have a size of 800 bytes and the packet generation time is used to vary the load. The three nodes in the pipe do not generate own traffic.

We simulated three MAC protocols (plain IEEE 802.11b, with DCMA enhancements, and with QCMA enhancements). For each protocol and scenario 20 random placements of nodes are used. And a confidence interval of 95 % is applied in all Figures. Eight different loads are simulated for the pipe scenario and six for the random scenario.

C. Metrics

For each simulation the performance of the protocol was evaluated by measuring the goodput and the mean delay per hop. The goodput was measured by counting the number of packets correctly received at the final destination during the simulation and for every node. For every correctly received packet the time elapsed between the creation of the packet and its reception was measured. Having this times the mean delay per hop was computed by dividing the elapsed time by the number of hops a packet traversed and finally by taking the average of all these values.

We also considered fairness between users in the network using the Herfindahl index [15]: We counted the packets created at each node and how many of these packets are correctly received at the final destination. At the end of a simulation with N nodes, the ratio of received to send packets

U_k was calculated for each node k as:

$$U_k = \frac{\sum_{j=0}^{N-1} R_{k,j}}{\sum_{j=0}^{N-1} S_{k,j}}$$

where $R_{i,j}$ is the number of packets received at node j from node i and $S_{i,j}$ is the number of packets sent from node i to j. Using U_k, the fairness index I_f is defined as:

$$I_f = \frac{\left(\sum_{k=0}^{N-1} U_k\right)^2}{N \times \sum_{k=0}^{N-1} U_k^2}$$

with

$$\frac{1}{N} \text{ (worst case)} \leq I_f \leq 1 \text{ (optimal).}$$

A fairness index of $\frac{1}{N}$ denotes that only one of the N nodes received all of its packets correctly ($Uk = 1$), but all packets from the other nodes are lost. A fairness index of 1 reflects that all N nodes had the same ratio between the number of send and received packets.

V. Simulation results

A. Pipe scenario

Figure 3 shows the total goodput in the pipe network when the total load offered is varied. All protocols saturate with a load starting at 325 Kbps. At 325 Kbps, a peak goodput of 275 Kbps is obtained for DCMA, whereas a peak goodput of 250 Kbps is obtained for IEEE 802.11b and QCMA. QCMA does not gain over IEEE 802.11 at this point, because QCMA uses cut-through communication only when there is a packet in the queue of the NIC. But as higher load increases the probability of having a packet in the queue, QCMA starts using cut-through communication. At low load QCMA behaves like IEEE 802.11 and does not use cut-through communication. At 512 Kbps offered load, the goodput improvement over IEEE 802.11b is 64 % for DCMA and 27 % for QCMA.

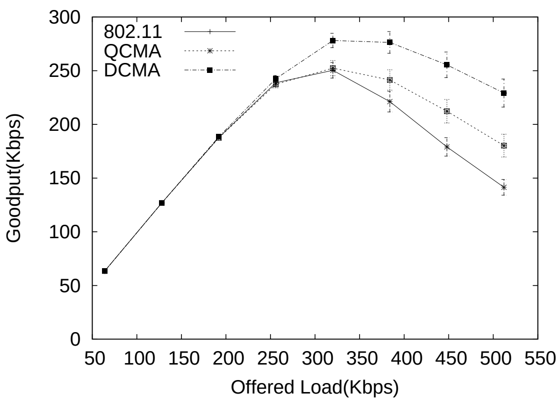

Fig. 3. Goodput vs. load in pipe scenario

Figure 4 shows the mean delay per hop when the total offered load is varied. The mean delay per hop starts to increase at 250 Kbps for IEEE 802.11b, at 325 Kbps for QCMA, and at 400 Kbps for DCMA. At 512 Kbps, the mean delay per hop improvement over IEEE 802.11b is 31 % for DCMA and 13 % for QCMA.

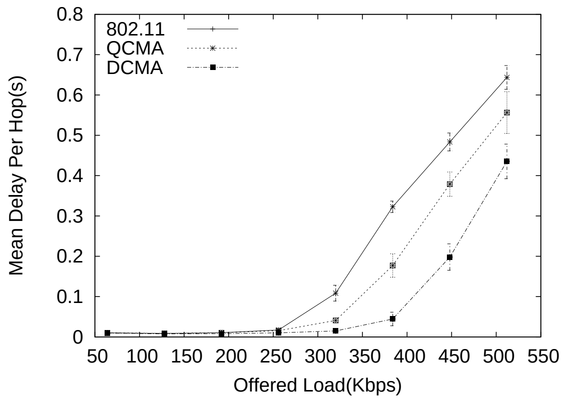

Fig. 4. Mean end-to-end delay/hop vs. load in pipe scenario

It is important to note that both QCMA and DCMA are improving goodput and end-to-end delay over IEEE 802.11b. As shown in [2], in a long chain of nodes, IEEE 802.11b suffers from collisions between packets of the same flow due to its channel access mechanism. The performance of the IEEE 802.11b standard in our pipe scenario is limited by this poor scheduling at the MAC level in the three hop chain. But in contrast, the use of cut-through communication in the pipe reduces the number of collisions, thus increases the goodput capacity of the network and also reduces the end-to-end delay. This reduction of end-to-end delay comes by the fact that forwarding nodes in the pipe do not suffer from queue build-ups of packets to be sent. Performance improvement over IEEE 802.11b is smaller with QCMA than with DCMA, because QCMA uses the ACK/RTS frame less often than DCMA.

As shown in Figure 5, both QCMA and DCMA are deteriorating fairness while improving throughput and delay. But as the fairness index for our simulation is between 1 (optimal case) and $\frac{1}{80}$ (worst case for 80 nodes) the difference between IEEE 802.11b (0.96 at 512 Kbps) and the cut-through protocols (0.92 at 512 Kbps) is only a minimal claim.

B. Random scenario

As shown in Figure 6, the total goodput in the random scenario is nearly the same for all protocols.

All protocols have a peak rate of 425 Kbps when the offered load is 625 Kbps, and for higher loads than 625 Kbps, goodput is slowly decreasing, to end up at 375 Kbps when the offered load is 1024 Kbps. Thus, the first important conclusion we can draw is that cut-through protocols do not improve goodput in a large random network. But this result comes as no surprise, as [2] already showed that IEEE 802.11 approaches the maximum capacity of such wireless

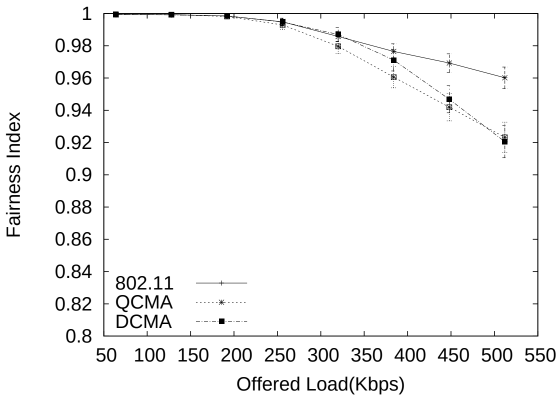

Fig. 5. Fairness vs. load in pipe scenario

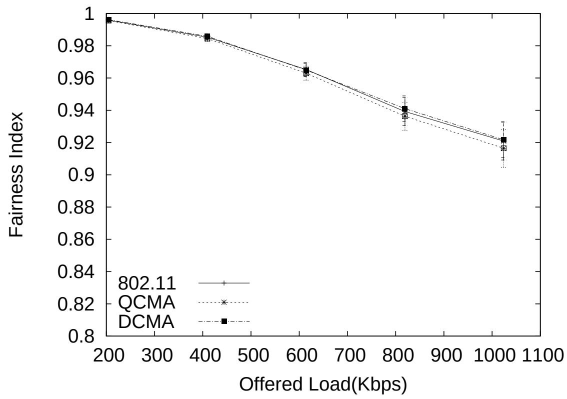

Fig. 7. Fairness vs. load in random scenario

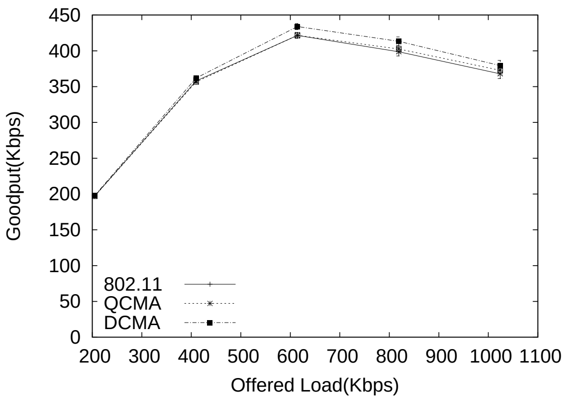

Fig. 6. Goodput vs. load in random scenario

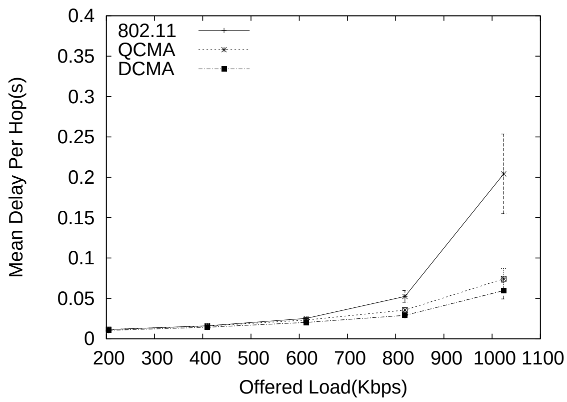

Fig. 8. Mean end-to-end delay/hop vs. load in random scenario

networks [3]. Additionally, all protocols have practically the same fairness as shown in Figure 7.

But in contrast to the other performance metrics, Figure 8 shows significant differences in end-to-end delay between IEEE 802.11b and both cut-through protocols. While the mean delay per hop of IEEE 802.11b increases a lot at high load, both QCMA and DCMA perform much better. At 1024 Kbps, the mean delay per hop improvement over IEEE 802.11b is 71 % for DCMA and 64 % for QCMA.

VI. Main benefit

This reduction in end-to-end delay in large ad hoc networks is a very interesting property of QCMA and DCMA. A look at Figure 9, which shows the number of queue drops in the network when the total load varies, helps us to understand how it is possible to reduce the end-to-end delay while the goodput remains the same.

As can be seen, a relation between Figure 9 and the mean delay per hop (Figure 8) exists: the mean delay per hop and the number of queue drops in the network start to increase in a similar manner when load reaches 800 Kbps. This shows that the high delay, which IEEE 802.11b experiences at saturation, is due to queue build-ups at certain nodes in the network. This way long end-to-end delays are introduced. The large confidence intervals show that these queue build-ups strongly dependent on the topology of the network, thus the end-to-end delay of IEEE 802.11b networks strongly depends on the node layout.

In contrast, queue drops almost disappear when a cut-through protocol is used, i.e., they have smaller confidence intervals and are more robust with respect to node layout. This queue drop diminution is not significant as far as the total amount of transported traffic is concerned (these drops correspond to about 1.5 % of the total traffic injected into the network) but it has a large impact on end-to-end delays. This is caused by the queuing behavior of some strategic nodes of the network. With IEEE 802.11b the queues of these nodes are full, but empty when a cut-through protocol is used. Thus, the queuing times at these important forwarding nodes are much smaller than with IEEE 802.11b, i.e., not adding so much time to the end-to-end delay.

In fact, with QCMA and DCMA, each time a forwarding node receives a packet, it can often remove a packet from its send queue. And this almost at the same time (if the creation of an ACK/RTS is possible and if the RTS/CTS/DATA/ACK is correctly accomplished): it leads to a better equilibrium in the queues of the nodes with a lot of forwarding activity.

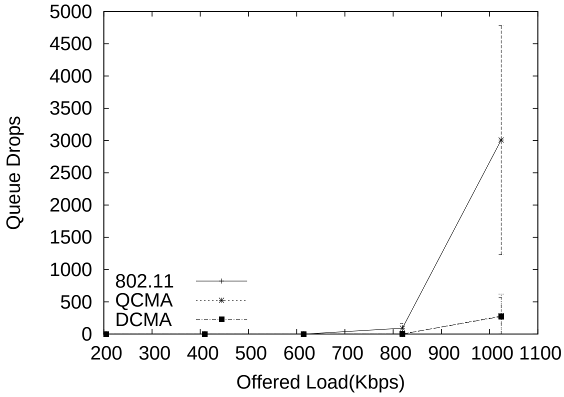

Fig. 9. Queue drops vs. load in random scenario

VII. Conclusion

As our results show, cut-through communication can decrease the end-to-end delay in multi-hop networks compared with the IEEE 802.11b standard. As the delay is reduced due to smaller queuing delays in intermediate hops, it can also lead to increased network throughput. Similar results can be expected for the IEEE 802.11a and 802.11g standard, as cut-through access is not affected by the differences between these standards.

While our proposed QCMA protocol does not achieve quite as big a performance increase as DCMA (both in goodput and in delay) because we considered cut-through on MAC level without routing interaction, it is still a considerable improvement over plain IEEE 802.11. What is more, it should be much simpler to implement, possibly even on existing hardware with a firmware update, and is much more amenable to the support of priority-based QoS schemes like the incipient IEEE 802.11e scheme. Hence, we believe that we struck a good compromise between complexity of the scheme and practicality considerations.

Consequently, evaluating how QCMA would integrate with IEEE 802.11e is the most important next step for future work. Additionally, as our main focus is the comparison of 802.11b, DCMA, and QCMA, we use the same set of parameters as in [6].

As one of these parameters is the length of the packet queue in every node, this queue length can have a great impact on end-to-end delay in a multi-hop ad hoc network. It is also interesting how changes to the QCMA protocol disburse, e.g., permitting the creation of an ACK/RTS for the same pair of communicating nodes or looking for a proper packet in the outgoing queue. As this could possible lead to unfairness, we will investigate these possibilities in a further step.

References

[1] P. Karn, "Maca - a new channel access method for packet radio," in *9th ARRL Computer Networking Conference*. London, Ontario, Canada: ARRL, Sept. 1990.

[2] J. Li, C. Blake, D. S. J. D. Couto, H. I. Lee, and R. Morris, "Capacity of Ad Hoc Wireless Networks," in *ACM International Conference on Mobile Computing and Networking*, August 2001.

[3] P. Gupta and P. R. Kumar, "The Capacity of Wireless Networks," *Transactions on Information Theory*, pp. 388–404, March 2000.

[4] A. Acharya, A. Misra, and S. Bansal, "Challenges in high performance data forwarding in multi-hop wireless networks," IBM, Tech. Rep., 2002.

[5] I. C. Society, *IEEE Std 802.11e/D2.0. Part 11: Wireless Medium Access Control (MAC) and physical layer (PHY) specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS)*. IEEE, November 2001.

[6] A. Acharya, A. Misra, and S. Bansal, "A Label-switching Packet Forwarding Architecture for Multi-hop Wireless LANs," IBM, Tech. Rep., 2002.

[7] S. Xu and T. Saadawi, "Does the IEEE 802.11 MAC Protocol Work Well in Multihop Wirelles Ad Hoc Networks?" *IEEE Communications Magazine*, June 2001.

[8] C. Lin and M. Gerla, "Asynchronous multimedia multihop wireless networks," in *Proc. Joint Conf. of the IEEE Computer and Communications Societies (INFOCOM)*, Apr. 1997.

[9] J. Sobrinho and A. Krishnakumar, "Real-time traffic over the ieee802.11 medium access control layer," in *Bell Labs Technical Journal*, 1996, pp. 172–187.

[10] N. Vaidya, P. Bahl, and S.Gupta, "Distributed fair scheduling in wireless lan," in *Proc. 6th Intl. Conf. on Mobile Computing and Networking*, Aug. 2000.

[11] S. Mangold, S. Choi, P. May, O. Klein, G. Hiertz, and L. Stibor, "Ieee 802.11e wireless lan for quality of service (invited paper)," in *Proceedings of the European Wireless*, vol. 1, Florence, Italy, Feb. 2002, pp. 32–39.

[12] P. Garg, R. Doshi, R. Greene, M. Baker, M. Malek, and M. Cheng, "Using ieee 802.11e mac for qos over wireless," in *Proceedings of the 22nd IEEE International Performance Computing and Communications Conference (IPCCC 2003)*, Phoenix, Arizona, Apr. 2003.

[13] E. Dijkstra, "A Note on Two Problems in Connection with Graphs," *Numerical Mathematics*, vol. 1, pp. 269–271, 1959.

[14] A. Varga, *OMNeT++: Discrete Event Simulation System*, Technical University of Budapest, Faculty of Electrical Engineering and Informatics, Mar. 2001, http://www.hit.bme.hu/phd/vargaa/omnetpp.htm.

[15] A. O. Hirschman, *The Paternity of an Index*. AER, Sept 1964.