WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. (2016)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/wcm.2663

RESEARCH ARTICLE

Performance and scalability evaluation of IEEE 802.11v/aa multicast transport

Yousri Daldoul[1,2], Djamal-Eddine Meddour[1]*, Toufik Ahmed[2] and Raouf Boutaba[3]

[1] Orange Labs Network, Lannion, France
[2] University of Bordeaux / Bordeaux-INP, France
[3] David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada

ABSTRACT

IEEE 802.11v and 802.11aa are two recent amendments that define new functionalities in order to support a reliable multicast transport over wireless networks. The first amendment introduces directed multicast service (DMS). On the other hand, 802.11aa defines the groupcast with retries (GCR) service, which proposes two retransmission policies: block acknowledgement (GCR-BACK) and unsolicited retry (GCR-UR). In this paper, we evaluate the throughput and the scalability of these new proposals using both analytical and simulation approaches. We show that DMS has the lowest scalability, while GCR-BACK is not appropriate for groups with a large number of receivers. We conclude that GCR-UR is the most appropriate for large groups. However, increasing the number of transmission retries reduces significantly the achieved throughput of the unsolicited retry policy. Copyright © 2016 John Wiley & Sons, Ltd.

KEYWORDS
IEEE 802.11aa; IEEE 802.11v; GCR-block ack; GCR-unsolicited retries; DMS

***Correspondence**
Djamal-Eddine Meddour, Orange Labs Network, Lannion, France.
E-mail: djamal.meddour@orange.com

1. INTRODUCTION

Multicast transport is becoming popular nowadays because of the increasing deployment of multimedia services over Internet protocol (IP) networks. The IEEE 802.11 standard 1 defines the principal wireless access network. However, the legacy multicast transport of this standard does not ensure any reliability for the delivered traffic. Therefore, multicast packets may experience a significant loss rate, which degrades the quality of the offered services greatly.

Recently, two amendments to the standard have been proposed: 802.11v 2 and 802.11aa 3. The 802.11v defines the directed multicast service (DMS). On the other hand, the 802.11aa amendment introduces the groupcast with retries (GCR) service, which proposes two retransmission policies: block acknowledgement (GCR-BACK) and unsolicited retry (GCR-UR). In this paper, we briefly describe these new protocols, and we compare their performance. Specifically, we evaluate their throughput and scalability as a function of the multicast group size, and we measure their reliability. We compare these new proposals using analytical and simulation results. We show that DMS has the lowest scalability, while GCR-BACK is not appropriate for large multicast groups. We conclude that GCR-UR is the most appropriate for large groups. However, increasing the number of the transmission retries reduces significantly the achieved throughput of the unsolicited retry policy.

The remainder of this paper is organized as follows. In Section 2, we discuss the most relevant works having studied the 802.11v/aa amendments. We highlight the novelties of 802.11v and of 802.11aa in Sections 3 and 4, respectively. We devote Section 5 to present our analytical model. We present the analytical and the simulation results in Section 6. Finally, we conclude this paper in Section 7.

2. RELATED WORKS

IEEE 802.11v/aa have recently been approved but not implemented within real devices and in network simulators yet. Therefore, only few works have evaluated and compared the new multicast protocols. In 4, the authors provide a general presentation of GCR and DMS. They introduce the new buffering architecture of 802.11aa. However, they do not evaluate any of the following parameters:

throughput, scalability, delays, and loss rate. In other words, the paper only provides a global overview of the aforementioned amendments without evaluating and comparing any of the key parameters of the new retry policies.

In 5, the authors evaluate and compare the new amendments using OPNET simulator 6. They compare the scalability, delays, efficiency, and reliability of the legacy multicast, DMS, GCR-BACK, and GCR-UR. However, we notice the following two limitations. First, the authors consider a partial evaluation of GCR-BACK. Specifically, they consider that only one member is allowed to send a feedback. This simple scenario requires few modifications to the unicast BACK mechanism already implemented within the network simulator. However, it cannot evaluate the scalability of GCR-BACK because it does not allow the evaluation of the impact of the multicast group size on the network throughput and on the delivery delays. Besides, this scenario is not accurate in measuring the reliability of GCR-BACK. This is because the obtained loss statistics are mainly caused by the absence of feedbacks from the other members. Therefore, they do not reflect the protocol reliability. The second limitation of this work is that there is no analytical validation of the obtained results. We believe that an analytical model is required, not only to corroborate the simulation results but also to provide an easy performance estimation for any particular network configuration, such as the expected throughput for a given group size and a specific transmission rate.

In 7, the authors introduce the operating mode of 802.11aa/v. They evaluate the reliability and the overhead of the different multicast protocols as a function of the packet error rate (PER). Their results show that none of the different proposals is reliable. This is because the authors consider all the loss factors together, including the queue drops. However, this is typical for network congestion and is not specifically related to the multicast protocol. We noted that the authors do not evaluate the throughput; they only measure the overhead ratio. Moreover, they do not evaluate the delays, and they do not show the impact of the group size on the protocol scalability. Furthermore, the authors do not provide any analytical model to determine the throughput under various network configurations. As such, it is not possible to deduce the efficiency of transmission rates other than the used one (i.e., 54 Mbps).

In 8, the authors evaluate the impact of collisions on the throughput of 802.11aa/v. But they do not evaluate the maximum achievable throughput of the different multicast protocols when the medium is not shared with other flows. Moreover, their solution did not provide any protection mechanism against the simultaneous access to the medium. Such a protection is required by 802.11aa and would enhance the reliability and the throughput of GCR-UR and GCR-BACK significantly. Hence, it is necessary to revise the proposed analytical model in order to consider a collision prevention feature. According to the obtained results, none of the protocols is reliable. This is because the authors consider a particular scenario of a highly loaded network in addition to a high-throughput video. Therefore, an important number of packets are rejected because of queue overflow. Thus, we believe that the major limitation of this work is that it does not show when the aforementioned protocols are reliable and when they are not. Furthermore, they do not evaluate the incurred transmission delays.

We note that none of all the aforementioned studies has introduced any of the following new features: GCR service period (GCR-SP) and GCR active (GCR-A) defined by 802.11aa, and flexible multicast service (FMS) introduced by 802.11v. These new mechanisms tackle the power management when a multicast traffic is being delivered. Furthermore, a data rate selection procedure is added for the first time into IEEE 802.11. This procedure is defined by 802.11v for multicast transmissions but has not been documented or studied in existing research literature. This is hopefully covered by this paper.

3. OVERVIEW OF IEEE 802.11V

3.1. Directed multicast services

The IEEE 802.11v amendment 2 defines the DMS in order to resolve the unreliability issue of multicast transport. This service allows the members to receive the multicast traffic as individually addressed packets (i.e., using unicast transport). Therefore, DMS guarantees the same unicast reliability degree to multicast transport at the expense of bandwidth. Hence, it can be used to stream a standard traffic to a limited group size but does not scale well for high-throughput streams like high-definition television. Moreover, DMS does not allow the use of block transfer. We note that 802.11v defines the required procedures to establish DMS sessions but does not define any functionality to manage multicast groups and to detect their members. According to the amendment, the method used by an access point (AP) to determine that the multicast group members is outside the scope of IEEE 802.11 but is typically performed by snooping IP packets.

3.2. Data rate adaptation

IEEE 802.11v defines a rate adaptation procedure for multicast transmissions. The data rate is selected according to the following steps. In the beginning, the AP sends a multicast diagnostic request in a radio measurement request frame indicating the measurement duration. This packet includes the address of one or many multicast sessions to be considered. It is transmitted in multicast if all the associated stations support the diagnosis capability. Otherwise, the packet is individually addressed to the compliant nodes. This request can be ignored by receivers, which did not join any of the mentioned sessions. On the other hand, the group members start counting the number of received packets from the specified sessions and record the maximum observed data rate to receive these packets. These values (i.e., the number of the received packets and the maximum used data rate per joined multicast address) are then returned to the AP in a Radio Measurement Report.

Triggered reports are also defined and allow the AP to gather statistics from the group members without having to send diagnostic requests periodically. These reports are transmitted subject to the following two conditions: (i) the multicast trigger condition occurs and (ii) the specified delay between two successive reports, called re-activation delay, expires. The trigger condition occurs when no multicast packet is received within a specific time period. But this period may be unspecified. In this case, the trigger condition is always true, and the reports are sent every re-activation period. This period should be greater or equal to a minimum threshold in order to limit the number of generated reports. These reports indicate the number of received packets and the highest data rate per joined session, because the previous diagnostic request. Therefore, it is necessary that the AP sends a new request each time the data rate is modified. This allows the members to report the recently observed highest data rate.

We notice many performance issues regarding this procedure. First, the diagnostic request may be lost when it is transmitted in multicast. Therefore, some solicited reports may be missing, and this may reduce the accuracy of the selected data rate. Besides, probing high transmission rates may cause the loss of all the delivered packets during the measurement periods. Therefore, the quality of the multicast services may be disrupted frequently. Moreover, this selection procedure is not appropriate for large multicast groups because an important collision rate may occur between the different reports themselves and between these reports and any other traffic (particularly multicast packets). This reduces the network throughput and further increases the loss rate of the multicast traffic.

3.3. Flexible multicast service

The FMS intends to reduce the power consumption of multicast receivers. It allows a group member in the power save mode to request an alternate delivery traffic indication map (DTIM) interval for the multicast traffic. FMS defines eight alternate intervals from 1 to 8. The first one corresponds to two DTIM intervals, and each successive interval is one DTIM interval larger. The multicast packets are buffered then sent in one block following the selected alternate interval. Recall that a DTIM is a variable number of beacon intervals and that the legacy power saving mode delivers multicast packets in one block each DTIM. Therefore, FMS allows the AP to schedule the transmission of the multicast packets at longer intervals. This allows the stations to spend more time in the doze state and allows a significant energy saving compared with the legacy procedure.

We note that FMS is accurate for low bit rate and delay-tolerant multicast streams. This is because packets may be buffered for more than 1 s at the AP. During this time, a limited number of packets may be stored without queue overflow. Similar to the legacy power saving procedure, FMS is not appropriate for high-throughput applications such as video services. Moreover, many packets may arrive and cause queue overflow. Because FMS delivers the multicast traffic in one single burst, the channel may be used for very long transmission durations. This may be an impediment for time-sensitive flows sharing the medium.

Furthermore, FMS enables group members to request a specific data rate. Therefore, the AP selects any transmission rate equal or lower than the lowest rate value provided by the group members.

4. OVERVIEW OF IEEE 802.11AA

4.1. Stream classification

The 802.11aa amendment 3 defines medium access control (MAC) enhancements for robust audio and video streaming. Hence, it defines additional transmit queues called Alternate VO (A_VO) and Alternate VI (A_VI). Incoming packets from the upper layer are mapped similarly to 802.11e except that flows with user priority (UP) of seven and four are buffered in A_VO and A_VI, respectively. These new queues are optional and may be disabled. In this case, only the four principal queues (i.e., VO, VI, BE, and BK) are used. We note that the standard does not define new enhanced distributed channel access functions for A_VO and A_VI. Therefore, they share the same access functions as VO and VI as illustrated in Figure 1. Packets are selected from the primary and alternate queues in such a way that packets with higher UP are selected with higher priority. However, 802.11aa does not define a specific scheduling function but requires that the default algorithm selects the transmission queue according to the selection procedure of IEEE 802.1Q 9. Therefore, the packets are selected by default based on the strict priority algorithm. This algorithm selects packets from the highest priority queue (i.e., in our case A_VO and VI) whenever this queue is not empty. Otherwise, the other queue is processed.

As can be observed from Figure 1, one alternate queue (i.e., A_VO) has more priority than a primary queue (i.e., VO). Then another primary queue (i.e., VI) has more priority than a second alternate queue (i.e., A_VI). Hence, a primary queue does not always stand for a queue with the highest priority. We believe that this is confusing and is not justified. The main reason to have two different queues for voice and video streams is to give more priority for time sensitive applications. Therefore, flows with strict delay requirements (i.e., less than 200 ms), such as VoIP, are delivered first, while delay-tolerant streams (i.e., about 1 s), such as IPTV, still have a high transmission priority compared with other flows.

Although the new queues are defined for both unicast and multicast flows, the principal novelty of 802.11aa is

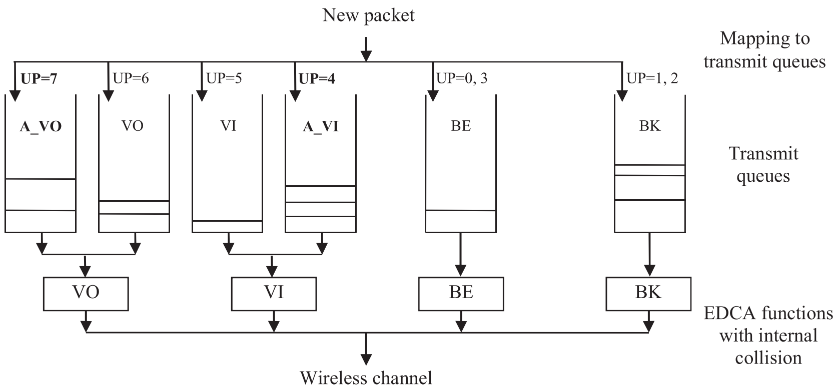


Figure 1. Primary and alternate queue of 802.11aa.

the inclusion of new retransmission policies and delivery methods for the multicast mode.

4.2. Group membership management

IEEE 802.11aa proposes one simple method to discover the members of different multicast sessions. This method relies on the awareness of the receiver itself about the groups to which this receiver belongs. Thus, the AP sends a packet, called Group Membership Request, individually to every associated station in order to request the addresses of the joined sessions. Each receiver replies to this request using a Group Membership Response packet, which contains the list of all the joined groups. Every time this list changes, the station sends an unsolicited Group Membership Response with the updated list to the AP. This allows the multicast source to be aware of the different members of the available multicast sessions. It is worth noting that the standard does not define any mechanism allowing the immediate detection of an unexpected departure of a group member. Therefore, if a disconnection failure occurs, the multicast source cannot take the required actions in time.

In addition to the aforementioned procedure, the amendment allows the use of any other group membership detection method such as Internet group management protocol snooping 10. But the definition of such a procedure is beyond the scope of the current standard.

4.2.1. Groupcast with retries

Groupcast with retries service is defined to improve the reliability of multicast transmissions. GCR defines two additional retransmission policies for multicast flows: GCR-UR and GCR-BACK. These policies allow the retransmission of multicast packets in order to reduce the loss rate. GCR defines two new delivery methods as well: GCR-SP and GCR-A. These methods define the way the multicast packets should be transmitted under different power states and particularly when the members are in the power save mode.

4.2.2. GCR service period

Unlike FMS, the GCR-SP method transmits the multicast packets at intervals, called service intervals, which may be smaller than the beacon interval. The AP informs the group members about the SP intervals, then sends the packets assigned to this group at the SP periods. We note that the SP duration is not limited and may reach an important number of transmission opportunities (TXOPs). Moreover, two or more SPs may be linked. During all this time, only the multicast stream is delivered. Therefore, GCR-SP should not be used with a high throughput service in order to allow a fair sharing of the channel with other flows.

4.2.3. GCR active

The second delivery method of GCR is GCR-A. It allows the AP to send multicast packets at any time, regardless of the power state of the group members. When this method is used, multicast members, which are in the power save mode, should enter the awake state in order to be ready to receive any multicast packet. These receivers should remain awake indefinitely until the delivery method is modified or the GCR agreement is canceled (e.g., the receiver leaves the group or the multicast session is over). Therefore, this is the most appropriate method to deliver high-throughput streams without impacting time sensitive flows sharing the same network.

4.2.4. GCR unsolicited retry

When the GCR-UR policy is used, the multicast source defines a retry limit, say “N”, and transmits each multicast packet “N” times without waiting for any feedback after each transmission. The retry limit is not defined by the standard and is implementation dependent. Additionally, GCR-UR allows the transmission of multicast packets in blocks separated by SIFS. However, a packet and its retransmission should not occur in the same block. On the other hand, the standard requires the use of a collision protection mechanism such as RTS/CTS or CTS-to-Self, in order to reduce the collision probability. The retransmission of the same packet several times allows the sender to increase the probability of the successful delivery. The main advantage of this policy is its scalability. Therefore, it is appropriate for large multicast groups. However, its reliability depends on the retry limit and on the accuracy of the selected transmission rate. We note that 802.11aa does not define any rate adaptation scheme for GCR-UR.

The main issue of GCR-UR is that this policy does not allow the detection of packet losses. Consequently, the multicast sender is not able to adapt the transmission rate when the losses are caused by signal attenuation. Therefore, the appropriate operation of GCR-UR relies on the use of the lowest transmission rate. In this way, the multicast packets are reliably delivered to any member in the network even those located at the limit of the coverage area. However, the use of the lowest rate significantly limits the overall network throughput. On the other hand, if a high transmission rate is used instead of the lowest one, all the members located beyond the coverage area of the used data rate (but are still connected at a lower transmission rate) cannot receive the multicast packets. We note that GCR-UR may use the rate selection procedure of 802.11v. However, the limited scalability of this procedure limits that of the multicast protocol.

It is worth highlighting that transmitting every packet several times increases the protocol reliability but reduces the network throughput significantly. Thus, losses over the wireless channel may be avoided, but rejections due to the queue overflow may occur more frequently. This may limit the reliability of the multicast traffic even though the loss rate over the wireless link is low.

4.2.5. GCR block ack

The GCR-BACK feedback policy is similar to the basic block transfer of unicast. It allows the AP to establish a block ack agreement with one or many of the group members at the beginning of the multicast session. Then the sender transmits a block of multicast packets followed by multiple exchanges of block ack requests (BAR) and BACKs. A member is allowed to reply only upon the reception of an explicit request. The received feedbacks allow the AP to detect any transmission failure. Furthermore, missing packets are retransmitted until their lifetime limit is reached. Therefore, the GCR-BACK guarantees the same reliability degree of unicast transport. New agreements may be established, and existing ones may be deleted during the streaming duration. If all the agreements are deleted, the AP switches to another delivery policy such as the legacy multicast or GCR-UR till all the members leave the group.

Similar to GCR-UR, the GCR-BACK policy requires the use of a protection mechanism (like RTS/CTS with one member, CTS-to-Self) to avoid collisions. CTS, data packets, BARs, and BACKs within a block transfer are separated by a SIFS period. If the medium remains idle within a period of PIFS (i.e., SIFS plus one SlotTime) after the transmission end of a BAR, the AP concludes the reception failure of the last BAR and sends it again immediately. The AP retransmits a BAR in this way until it detects a transmission before the PIFS expiry or the lifetime of all the multicast packets expires. If the AP detects the BACK transmission but does not receive it correctly, then the AP retransmits the BAR following channel contention. Besides, packets transmitted within a block are subject to the TXOP duration. Therefore, if this duration is not enough to gather all the feedbacks, the AP should interrupt the block and contends for the channel again in order to request the remaining BACKs during a new TXOP. The GCR-BACK procedure using the CTS-to-Self protection is illustrated in Figure 2. We note that GCR-BACK can also be used to deliver aggregated packets, that is, A-MPDU and A-MSDU.

The GCR-BACK policy allows the AP to adapt the transmission rate according to loss statistics. Accordingly, it enables the selection of the most appropriate data rate. If a member in the group does not establish a BACK agreement with the AP, this member does not receive any BAR and is not allowed to send any feedback. In this case, the selected rate may be inappropriate for this receiver. Therefore, the AP needs the feedbacks of all the members in order to deliver the multicast stream reliably using the most appropriate data rate. However, this policy requires an important number of feedbacks. This number depends on the

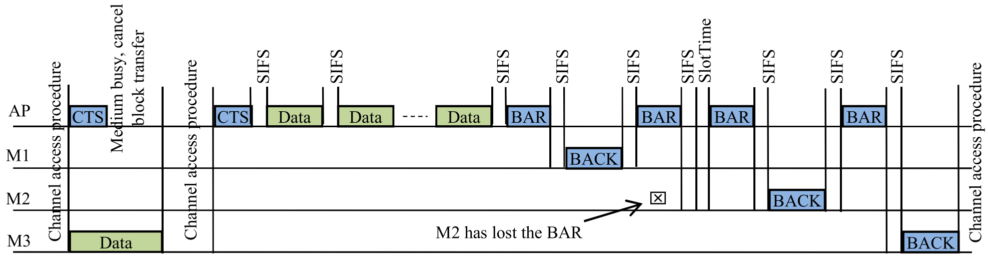


Figure 2. Typical frame exchange scenario with groupcast with retry block acknowledgement (GCR-BACK) policy.

DOI: 10.1002/wcm

multicast group size. Thus GCR-BACK has a limited scalability and is not appropriate for large groups.

5. MODEL DESCRIPTION

In this section, we define an analytical model to evaluate the throughput and the scalability of DMS, GCR-UR, and GCR-BACK. We consider multicast User Datagram Packet (UDP)/IP packets with the maximum transmission unit size of 1500 octets. Thus, the MAC packet length is 1538 Bytes. We consider that ACKs, BNRs, and BNAKs are transmitted at the lowest data rate of 6 Mbps and that they are always delivered successfully. This is a reasonable assumption because these packets have a small size and use the most robust rate.

It is worth noting that losses in wireless networks are classified into two principal categories: (i) losses caused by collisions: these losses are correlated between the different receivers. This is because a collision is experienced by the different members and leads to the distortion of the received data. (ii) Losses caused by the path-loss: the loss probability depends on the signal-to-noise ratio (SNR), and packet losses are not correlated between the different receivers (for the same SNR value, a member may receiver a packet correctly and another member may lose the same packet).

To define our analytical model, we only consider uncorrelated losses for the following reason. GCR-BACK and GCR-UR use a protection feature (CTS-to-Self in our case) to avoid the collisions. So reception failures are only related to path loss. This will be illustrated further in Figure 9 using simulation results. Besides, DMS delivers a multicast stream using multiple unicast sessions. So any collision will affect one single (unicast) transmission. Therefore, the only losses that may affect the modeled protocols (i.e., GCR-BACK, GCR-UR, and DMS) are not correlated.

Let $\boldsymbol{G}$ be the multicast group size. Each member in the group experiences a PER of p_i for $i = 1,\ldots, G$. As previously explained, we consider that losses are not correlated between different receivers. We fix the transmission limit to seven for DMS. However, for GCR-BACK we set the limit to 100. We choose this value because the retransmission of a packet is subject to lifetime limit. Thus we fix a transmission limit by excess. We note that the probability to reach high retransmission stages is negligible when the PER is limited. Let $\boldsymbol{N}$ be the block size and $Nr(k)$ be the number of packets transmitted for the k-th time within a block. $Nr(1)$ is the number of packets transmitted for the first time. Every block is composed of $\sum_{k=1}^{100} Nr(k)$ packets. $Nr(k)$, for $k = 1,\ldots, 100$, depends on the PER of the network.

Table I presents the used variables and their values at different transmission rates. We consider that the CTS-to-Self is always delivered at the highest data rate of 54 Mbps in order to have the shortest length. Thus, it allows the efficient detection of simultaneous transmissions (as depicted in Figure 2).

We define $\boldsymbol{X}$ as the number of transmission attempts. The probability for a given member $\boldsymbol{M_i}$, for $i = 1,\ldots, G$, to receive correctly a packet in any of the k first transmissions is given by equation (1).

$$P^i(X \le k) = 1 - p_i^k \tag{1}$$

We derive the probability to serve all the $\boldsymbol{G}$ receivers in any of the k first transmissions in equation (2).

$$P_G\,(X \le k) = \prod_{i=1}^{G}\left(1 - p_i^k\right) \tag{2}$$

We obtain $Nr(1)$ and $Nr(k)$, $k = 2,\ldots, 100$, in equations (3) and (4), respectively.

$$Nr(1) = N - \sum_{k=2}^{100} Nr(k) \tag{3}$$

$$Nr(k) = Nr(1).(1 - P_G(X \le k-1)),\ k = 2, \ldots, 100 \tag{4}$$

We resolve equations (3) and (4), and we obtain $Nr(k)$, for $k = 1,\ldots 100$, in equation (5). It is obvious that the probability to receive a packet correctly in $X = 0$ attempts is nil; hence, $P(X = 0) = P_G(X = 0) = 0$.

$$Nr(k) = \frac{N.(1 - P_G(X \le k-1))}{\sum_{k0=1}^{100}(1 - P_G(X \le k0-1))},\ k = 1, \ldots, 100 \tag{5}$$

We express the average packet transmission time using GCR-BACK, in equation (6). This time takes into account

Table I. Parameters description and values.

Variables	Values
Network	IEEE 802.11a
T_{PPDU_Data}: PHY packet duration, 1538 B.	252 μs (at 54 Mbps)
T_{PPDU_BAR}: PHY BAR duration, 30 B.	64 μs (at 6 Mbps)
T_{PPDU_BACK}: PHY BACK duration, 38 B.	76 μs (at 6 Mbps)
PROTECTION_DURATION, CTS	40 μs (at 54 Mbps + SIFS)
SlotTime	9 μs
SIFS	16 μs
DIFS (SIFS + 2 SlotTime)	34 μs
CWmin: contention window min	15

DOI: 10.1002/wcm

the minimum waiting time between two successive transmission opportunities, that is, one DIFS plus the average backoff time. Besides, we consider the duration of any protection mechanism against the collisions; for the case of CTS-to-Self transmitted at 54 Mbps, the protection duration is 40 µs. It corresponds to the required time to send CTS plus one SIFS. The average time to send a block depends on the block size. We note that the packets within a block are separated by SIFS. Following the transmission end of the data packets, the AP exchanges BAR/BACKs with every group member. Then, we divide the total time by the number of packets transmitted for the first time within a block. This allows us to obtain the average time to send one multicast packet. It is worth noting that retransmitted packets are considered as overhead, and they increase the average time per packet.

$$T_{BACK}(N,G) = \Big(DIFS + \frac{CW\min}{2} \times SlotTime + PROTECTION\_DURATION + (T_{PPDU\_Data} + SIFS) \times N - SIFS + G \times (SIFS + T_{PPDU\_BAR} + SIFS + T_{PPDU\_BACK})\Big)/Nr(1) \quad (6)$$

The average packet transmission time using GCR-UR does not depend on the group size. Instead, this time depends on how many times every packet is transmitted. We define $\boldsymbol{U}$ as the transmission number, and we derive the average transmission time of GCR-UR in equation (7). Similarly to GCR-BACK, we consider the waiting time, the average backoff time, and the duration of the used protection mechanism. Then, we add the transmission duration of the multicast packets. We multiply this duration by the retry count, and we divide by the number of data packets within a block to obtain the average transmission time per packet.

$$T_{UR}(N,U) = \Big(DIFS + \frac{CW\min}{2} \times SlotTime + PROTECTION\_DURATION + (T_{PPDU\_Data} + SIFS) \times N - SIFS\Big) \times U/N \quad (7)$$

We measure the PER of GCR-UR as experienced by member M_i, for $i = 1, \ldots, G$, in equation (8).

$$P^i_{UR}(U) = p_i^U \quad (8)$$

The average transmission time of DMS is a function of the group size. We express this average time in equation (9). This time depends on the group size. On the other hand, the average backoff time of DMS depends on the retry count. Besides, we do not use any protection mechanism as this is not required by 802.11v.

$$T_{DMS}(G) = \sum_{i=1}^{G}\left[\sum_{k=1}^{7}\left(\left(DIFS + \frac{CW(k)}{2} \times SlotTime + T_{PPDU\_Data} + SIFS + T_{PPDU\_ACK}\right) \cdot \left(1 - P^i(X \le k-1)\right)\right)\right] \quad (9)$$

where $CW(k)$ is the contention window of the k-th transmission. Equations 10–12 illustrate the packet transmission rate of GCR-BACK, GCR-UR, and DMS, respectively.

$$Throughput_{BACK} = 1/T_{BACK}(N,G) \quad (10)$$

$$Throughput_{UR} = 1/T_{UR}(N,U) \quad (11)$$

$$Throughput_{DMS} = 1/T_{DMS}(G) \quad (12)$$

6. PERFORMANCE EVALUATION

We use NS-3 11 to evaluate DMS, GCR-UR, and GCR-BACK and to validate our analytical model. We build an IEEE 802.11a infrastructure network, and we consider the simulator configuration of Table II. In the remainder of this paper, we consider multicast packets of 1538 Bytes (including the MAC header) transmitted at the highest rate of 54 Mbps. The CTS-to-Self is continuously sent at 54 Mbps, while the other control packets (i.e., ACK, BAR, and BACK) are always delivered at the lowest rate of 6 Mbps.

6.1. Model validation

To validate our analytical model, we consider that all the group members have the same PER. We compare the analytical and the simulation results of GCR-BACK, GCR-UR, and DMS in Figure 3(a)–(c), respectively. We consider a multicast group of 10 members for all the protocols. In Figure 3(b), we illustrate the obtained results using three different transmission limits: (i) GCR-UR1 transmits each packet one single time, (ii) GCR-UR2 sends each packet two times, and (iii) GCR-UR3 allows each packet to be delivered three times. We observe a very good

Table II. Simulator configuration.

Parameters	Values
Simulator version	Ns-3.13
Transmission power	40 mW (16.02 dBm)
Transmission gain	1 dB
Reception gain	1 dB
Reception noise figure	7 dB
Propagation loss model	Log distance
Path loss exponent	3
Reference distance	1 m
Reference loss (at 1 m)	46.677 dB
Propagation delay model	Constant speed propagation
Speed	3.108 m/s
Error rate model	Nist
Energy detection threshold	−96 dBm
Network	IEEE 802.11a
Beacon interval	1 s
Packet lifetime limit	60 ms
Queue size	20 packets
CWmin	15
CWmax	31

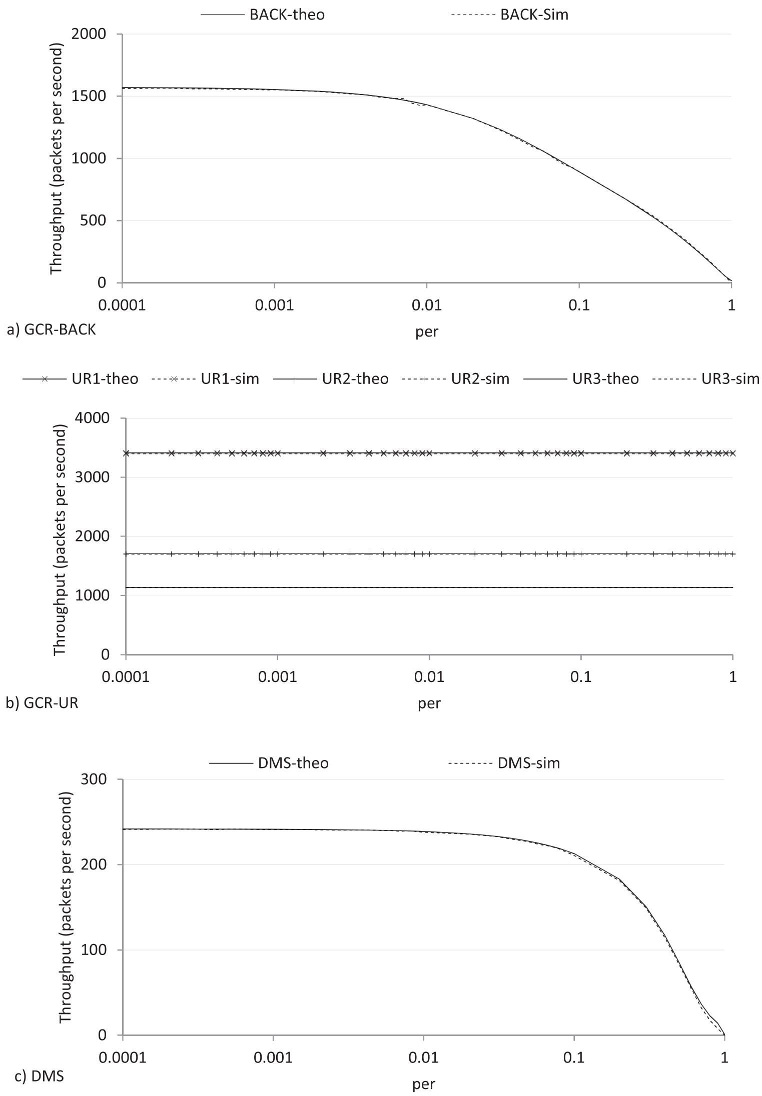


Figure 3. Throughput estimation of (a) groupcast with retry block acknowledgement (GCR-BACK), (b) groupcast with retry unsolicited retry (GCR-UR), and (c) directed multicast service (DMS), for a group of 10 members and a data rate of 54 Mbps.

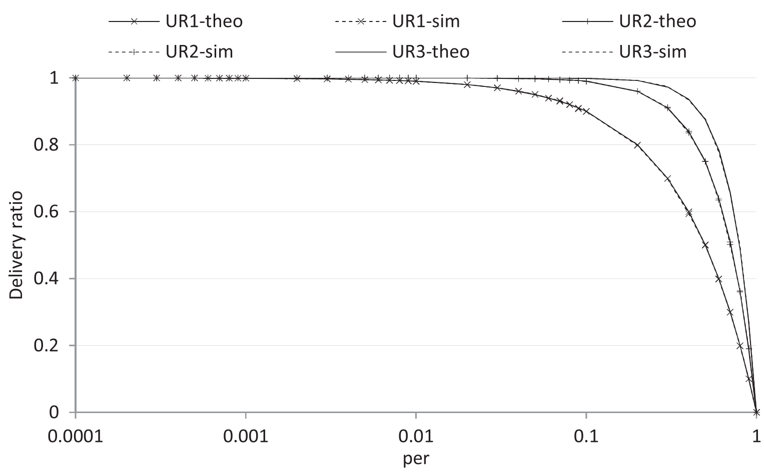


Figure 4. Packet delivery ratio of groupcast with retry unsolicited retry.

DOI: 10.1002/wcm

accuracy of our model for all the protocols and regardless of the loss rate.

Furthermore, we validate the accuracy of our mathematical estimation of the reliability of GCR-UR. Therefore, we compare the simulation and the analytical results for three different transmission limits: GCR-UR1, GCR-UR2, and GCR-UR3 (one, two, and three transmissions for each packet, respectively). We depict the obtained results in Figure 4. We conclude that our analytical model is accurate in estimating the delivery ratio of GCR-UR as a function of the PER. We highlight that these results do not depend on the group size because GCR-UR does not require any feedback from the multicast receivers.

6.2. Simulation results

In the remainder of this section, we compare DMS, GCR-UR, GCR-BACK, and the legacy multicast procedure using simulation. In Figure 5, we show the throughput of the different protocols for a variable group size. We set all the receivers at a distance of 10 m from the AP. This distance is suitable for the used data rate of 54 Mbps. Therefore, the incurred loss rate due to signal attenuation is very limited and is almost nil. Besides, GCR-UR and GCR-BACK use CTS-to-Self. We set the block size limit to 5. Thus, GCR-UR and GCR-BACK may send up to five packets within a block.

We note that the main difference between GCR-UR1 and the legacy multicast procedure is that the former policy uses a protection feature against collisions and is allowed to deliver more than one packet within a transmission opportunity. Therefore, GCR-UR1 has an enhanced throughput and is less vulnerable to collisions compared with the legacy procedure.

We notice that GCR-UR with one single transmission per packet (i.e., UR1) has the highest efficiency and delivers more than 3300 packets per second (pps), regardless of the group size. Furthermore, we notice that increasing the number of transmissions per packet reduces significantly the achieved throughput of the unsolicited retry policy; we observe that the efficiency of GCR-UR2 is less than 50% of that of GCR-UR1. Moreover, GCR-UR3 achieves only 1125 pps. This is about 30% of the throughput of GCR-UR1. However, the major advantage of the unsolicited retry policy is that this protocol is scalable and the achieved throughput does not depend on the group size. Thus, under the assumption that the appropriate data rate is carefully selected using a scalable rate adaptation algorithm, GCR-UR becomes suitable for large multicast groups.

We observe in Figure 5 that the throughput of GCR-BACK decreases with the increase of group size. Therefore, the protocol is able to deliver about 3000 pps when there is one single member in the group. This throughput decreases by 50% when there are 10 members and falls to less than 270 pps for a group of 100 receivers. Therefore, this policy is appropriate for groups with a few members but is not efficient for large groups.

As expected, DMS has the lowest scalability. The efficiency of this protocol falls significantly when a second member joins the group. Moreover, the highest throughput is limited to 236 pps when 10 members are present and is limited to 23 pps when 100 receivers join the group. Therefore, DMS is appropriate for small groups of two or three members.

We conclude that DMS has the lowest scalability, while the throughput of GCR-BACK is significantly impacted by the group size. Furthermore, the efficiency of GCR-UR decreases significantly when increasing the number of transmissions per packet. However, the scalability of this policy does not depend on the group size. To maintain such scalability, the protocol needs a scalable data rate selection procedure, which is currently missing from the standard.

Now, we evaluate the impact of the distance on the throughput. We note that the distance is the principal parameter to determine the signal attenuation and to vary the bit error rate. We measure the packet delivery ratio in order to evaluate the reliability of the different protocols.

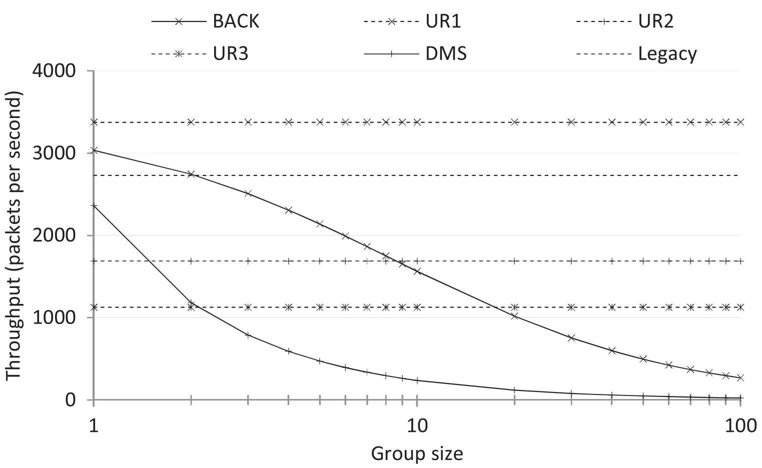


Figure 5. Throughput versus group size. BACK, block acknowledgement; DMS, directed multicast service; UR, unsolicited retry.

We build a group of 10 members, and we set them all at the same distance from the AP. We vary this distance, and we measure the throughput and the reliability. We first assume a collision-free environment in Figures 6 and 7, and then we introduce the collision factor in Figures 8 and 9, respectively.

As stated earlier, in Figures 6 and 7, the results are obtained *in the absence of collisions*. Thus, the AP is the only sender. This scenario allows us to compute the maximum achievable throughput.

In Figure 6, we observe that the throughput of GCR-BACK and of DMS depends on the PER. Thus, the loss rate experienced by one single member can reduce the throughput of the entire multicast session. On the other hand, the throughput of GCR-UR and of the legacy multicast procedure does not depend on the loss rate.

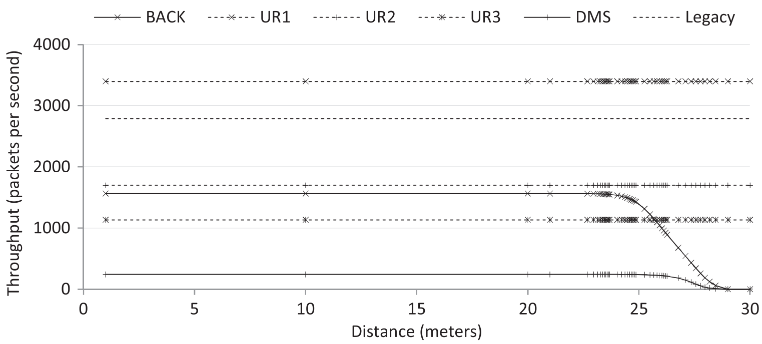


Figure 6. Throughput in the absence of collisions for a group of 10 members. BACK, block acknowledgement; DMS, directed multicast service; UR, unsolicited retry.

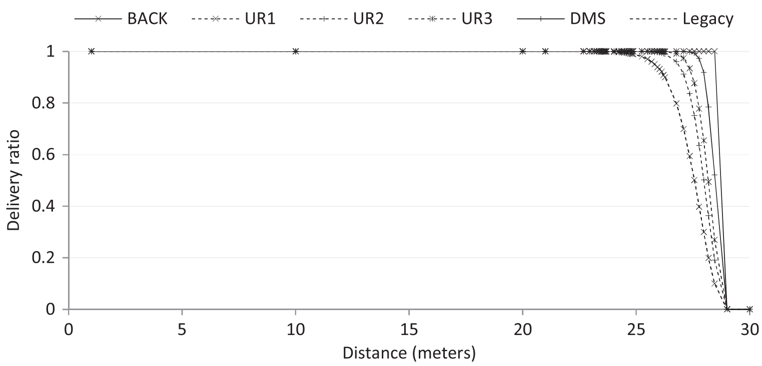


Figure 7. Delivery ratio in the absence of collisions for a group of 10 members. BACK, block acknowledgement; DMS, directed multicast service; UR, unsolicited retry.

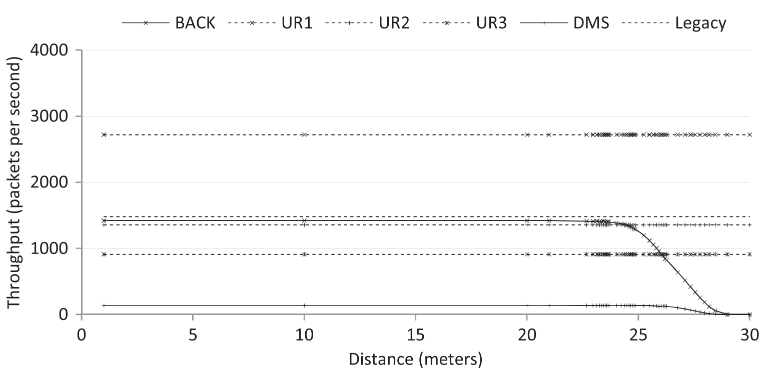


Figure 8. Throughput in the presence of collisions for a group of 10 members. BACK, block acknowledgement; DMS, directed multicast service; UR, unsolicited retry.

DOI: 10.1002/wcm

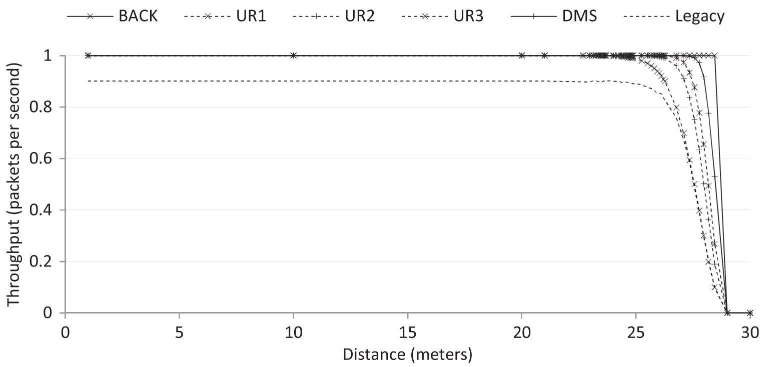


Figure 9. Delivery ratio in the presence of collisions for a group of 10 members. BACK, block acknowledgement; DMS, directed multicast service; UR, unsolicited retry.

According to Figure 7, we notice that the reliability of all the protocols depends on the distance. Thus, when the loss rate increases significantly, none of the protocols is reliable. However, we notice that GCR-UR1 and the legacy multicast have the lowest reliability when the receiver is located at a distance of 24–29 m from the sender. This reliability improves slightly for UR2 and UR3. The reliability of DMS is better than that of GCR-UR at these locations. This is because DMS is allowed to retransmit a packet up to seven times according to our configuration. On the other hand, GCR-BACK provides the highest delivery ratio. This is because this protocol is allowed to retransmit a packet till the expiry of its lifetime limit. In our configuration, we set this limit to 60 ms. Thus, a packet may be transmitted even more than 100 times.

We evaluate the throughput and the reliability of the different protocols in Figures 8 and 9, respectively, *in the presence of collision* this time. We configure a station to send unicast packets to the AP and to be in the saturation condition (i.e., the transmission queue of this station is never empty). The main goal of this configuration is to evaluate the impact of collisions on the loss rate of multicast packets.

In Figure 8, we observe that the throughput of the different protocols decreases compared with the first scenario of Figure 6. This can be explained by the fact that the remaining time to send multicast packets decreases because the medium is shared with another traffic. However, we notice that the throughput of the legacy multicast and of DMS is more impacted than that of the other protocols. This is because these two protocols do not take advantage of block transfer.

In Figure 9, we observe that the legacy multicast experiences an important loss rate even when the receiver is located near to the sender. This is caused by collisions. On the other hand, we notice that the loss rate of GCR-UR is not impacted by the unicast traffic. This is because this protocol uses a protection mechanism to avoid collisions.

In the following, we evaluate the transmission delays as a function of the distance (i.e., bit error rate), the throughput, and the group size. We consider that only the multicast packets are transmitted and that the channel is not shared with any other traffic. In the first scenario, we build a group of 10 members located at the same distance from the AP. We vary this distance progressively till reaching a loss rate of 100%. We send a multicast traffic with a rate of 1 pps. The main advantage to use a very low throughput is to avoid the buffering delays. Thus, the transmission of each multicast packet starts immediately upon the arrival of that packet to the MAC layer, and the obtained results are limited to the delays incurred by the multicast protocol. We illustrate the obtained results in Figure 10.

For the case of DMS, we measure the average transmission delays as experienced by the first and last (i.e., the 10th) members. For the other protocols, all the members experience the same average delays. We notice that the delays experienced by member 10 using DMS (i.e., DMS10) are significantly more important than the average delays of the first member (i.e., DMS1). These delays increase at important distances because of the need to retransmit the missing packets. Furthermore, we observe that the delays of GCR-BACK increase with the increase of the loss rate. But the delays of the unsolicited retry policy are the lowest because a packet is retransmitted up to three times.

In Figure 11, we depict the transmission delays under variable network load, and we consider a constant bit rate multicast traffic. We obtain these results when all the group members are located at a distance of 10 m from the sender. We observe that the transmission delays increase significantly when the throughput exceeds the maximum capacity of the used protocol. This is because the buffering delays will be added. Furthermore, a packet is rejected when it exceeds the lifetime limit. Therefore, the highest delays are limited to this limit, which is 60 ms (as shown in Table II). Similar to the previous scenario, we observe that member DMS1 experiences lower delays than DMS10 when the throughput is up to 200 pps. For the case of GCR-BACK, we notice that the delays are very limited when the packet rate is lower than 500 pps. This is because a packet is immediately transmitted when it arrives to the MAC layer. Then the delays increase slightly for data rates from 500 to 1500 pps. This is because a packet may arrive while the protocol is in the feedback phase, that is, the AP is exchanging BAR/BACK with the members. In this case, the

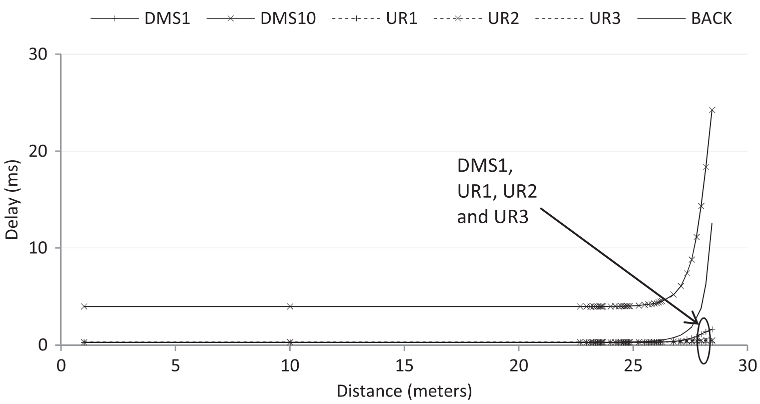


Figure 10. Delivery delays as a function of the distance for a group of 10 members. DMS, directed multicast service; UR, unsolicited retry.

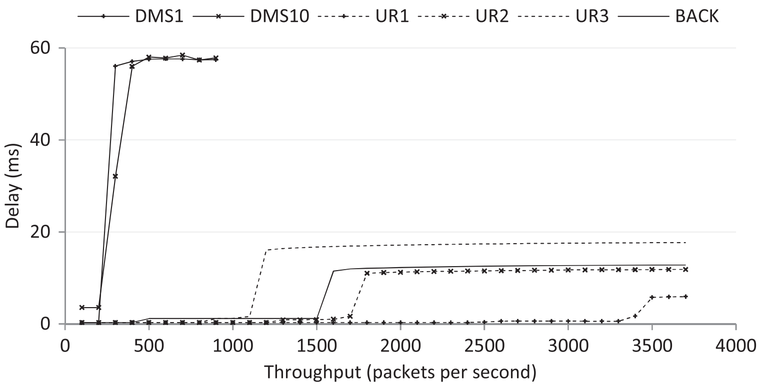


Figure 11. Delivery delays as a function of the throughput for a group of 10 members. BACK, block acknowledgement; DMS, directed multicast service; UR, unsolicited retry.

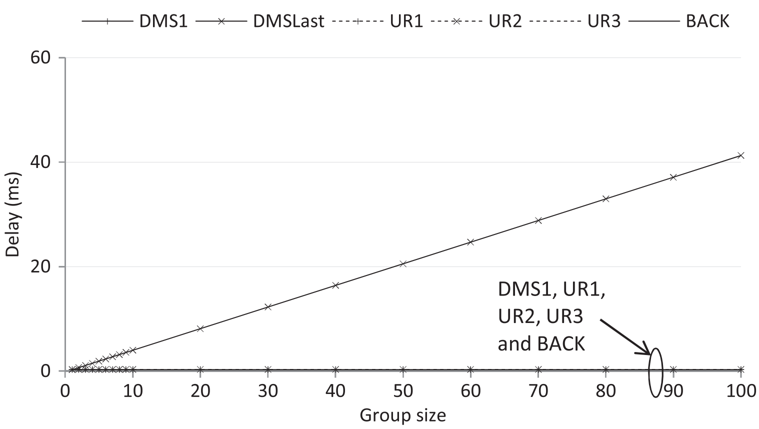


Figure 12. Delivery delays as a function of the group size. BACK, block acknowledgement; DMS, directed multicast service; UR, unsolicited retry.

new packets wait the end of the exchange and the channel contention before being transmitted. When the throughput exceeds 1600 pps, the delays of GCR-BACK increase significantly because of the buffering delays. This is because the highest supported throughput without queue overflow is limited to 1564 pps, according to Figure 5. However, we notice that the highest delays of GCR-BACK are much lower than the lifetime limit of 60 ms. This is because these delays depend on the queue size; the packets are rejected if they arrive when the queue is full. But when a packet is in the queue, it should wait the transmission end of the older packets. In our case, the queue size is 20. In a saturated network, the delays of GCR-BACK are bounded by the maximum delay to send 20 packets, whenever this delay is lower than 60 ms. We observe the same curve behavior for GCR-UR1, UR2, and UR3. However, GCR-UR3 reaches the saturation condition first. Moreover, the maximum delays of UR3 are higher than those of UR1 in a saturated network. This is because the average service time for a packet under UR3 is more important than that required by UR1 and UR2.

In Figure 12, we measure the average delays for variable group sizes. We use a very low throughput of 1 pps, and we set all the members at a distance of 10 m from the AP. We observe that all the protocols, except DMS, incur very limited delays. It is worth noting that the curve of GCR-BACK does not illustrate the required time to send the feedbacks. This is because all the packets are delivered correctly at the first transmission attempt. Therefore, the feedback duration is not added to the transmission latency because the feedback step occurs following the packet transmission. For the case of DMS, however, the incurred delays depend on the receiver rank. Thus, the first member experiences very low delays, while the last receiver (i.e., DMSLast) encounters the worst latencies. These delays increase linearly with the group size.

7. CONCLUSION

This paper presented the most important protocol amendments for multicast transport introduced as part of 802.11v and 802.11aa. Specifically, we discussed DMS, FMS, and the data rate selection procedure of 802.11v. We also discussed the GCR service of 802.11aa, particularly describing GCR-SP, GCR-A, GCR-UR, and GCR-BACK. We described the FMS approach designed to ensure very low power consumption but incurs important buffering delays, which is not appropriate for high-throughput and time-sensitive multicast flows. We also explained how GCR-SP is appropriate for low throughput applications. We concluded that GCR-A is the most appropriate method to deliver high bit rate flows to receivers in the power save mode. Furthermore, we defined an analytical model to determine the throughput of DMS, GCR-UR, and GCR-BACK for different values of group size, transmission rate, PER, and packet size. The proposed model can also be used to determine the delivery ratio of GCR-UR and was validated using simulations. In addition, we have evaluated, in this paper, the scalability of different protocols, and we found that DMS has the lowest efficiency. We found that GCR-BACK is not appropriate for large group members. Moreover, we showed that the throughput of GCR-UR does not depend on the group size but is significantly impacted by the increasing number of transmissions per packet. Finally, we measured the incurred delays and found that they increase significantly in a saturated network. We also found that the delays experienced by the last DMS member increase linearly with the group size increase.

REFERENCES

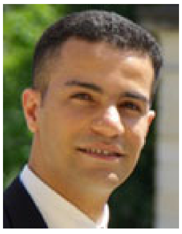
1. Wireless LAN medium access control (MAC) and physical layer (PHY) specifications. IEEE std 802.11, 2012.
2. Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: IEEE 802.11 Wireless Network Management. IEEE std 802.11v, February 2011.
3. Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: MAC enhancements for robust audio video streaming. IEEE std 802.11aa, May 2012.
4. Maraslis K, Chatzimisios P, Boucouvalas A. IEEE 802.11aa: improvements on video transmission over wireless LANs. in IEEE International Conference on Communications (ICC) 2012, Ottawa, Canada.
5. Santos MA, Villalon J, Orozco-Barbosa L. Evaluation of the IEEE 802.11aa group addressed service for robust audio–video streaming. in IEEE International Conference on Communications (ICC) 2012, Ottawa, Canada.
6. OPNET. Home page, http://www.opnet.com/
7. Oliva ADL, Serrano P, Salvador P, Banchs A. Performance evaluation of the IEEE 802.11aa multicast mechanisms for video streaming. in the 14th IEEE International Symposium and Workshops on a World of Wireless, Mobile and Multimedia Networks (WoWMoM) 2013, Madrid, Spain.
8. Barranco LE. Performance evaluation of IEEE 802.11aa MAC enhancements for robust audio video streaming. University Carlos III of Madrid, Department of Telematics Engineering, Master of Science Thesis, 2011.
9. Media access control (MAC) bridges and virtual bridge local area networks. IEEE Std 802.1Q, 2011.
10. Christensen M, Kimball K, Solensky F. Considerations for Internet group management protocol (IGMP) and multicast listener discovery (MLD) snooping switches. Request for Comments 4541, May 2006.
11. The NS-3 Simulator. Home page, http://www.nsnam.org/

AUTHORS' BIOGRAPHIES

Yousri Daldoul received the Engineering degree in Telecommunications from the Higher School of Communication of Tunis (SUP'COM), Tunisia, in 2008. He received the MSc and the PhD degrees in Computer Science from Bordeaux University, France, in 2009 and 2013, respectively. From

2009 to 2013, he was with Orange Labs networks in Lannion, France, where he extensively investigated reliability issue related to unicast and multicast stream transport in WiFi and mobile networks. His research field includes Wireless Networks, QoS of multimedia applications, and protocols optimization.

Djamal-Eddine Meddour received his computer engineering degree with honors from the INI (Institut National d'Informatique), Algiers, Algeria, in 2000, his master's degree in computer science from the University of Versailles, France, in 2001, and his Ph.D. in Computer Science from University of Paris VI in 2004. He is currently a senior researcher with Orange Labs Networks (formerly France Telecom R&D), Lannion, France. His main research activities concern the design and management for light wireless and mobile infrastructure networking, management, and interoperability in new-generation wireless networks. He served as a guest editor and TPC member for numerous international conferences and journals. He is the co-author of several international journal articles and book chapters, and holds many patents.



Toufik Ahmed is a Professor at IPB (Institut Polytechnique de Bordeaux) in the ENSEIRB-MATMECA school of engineering. He is doing his research activities in CNRS LaBRI Lab, UMR 5800 at University of Bordeaux. He received the BSc in Computer Engineering, with high honors, from the I.N.I (National Institute of Computer Science) Algiers, Algeria, in 1999, an MSc and a PhD degree in Computer Science from the University of Versailles, France, in 2000 and 2003, respectively. In November 2008, he obtained the HDR degree (Habilitation Diriger des Recherches) at University of Bordeaux-1 on adaptive streaming and control of video quality of service over wired/wireless IP networks and P2P architectures. His main research activities concern quality of service for multimedia wired and wireless networks, end-to-end signaling protocols, P2P network, and wireless sensors network.

Raouf Boutaba received the MSc and PhD degree in computer science from the University Pierre & Marie Curie, Paris, in 1990 and 1994, respectively. He is currently a professor of computer science at the University of Waterloo (Canada). His research interests include resource and service management in networks and distributed systems. He is the founding editor in chief of the IEEE Transactions on Network and Service Management (2007–2010) and on the editorial boards of other journals. He received several best paper awards and recognitions including the Premiers Research Excellence Award, the IEEE ComSoc Hal Sobol, Fred W. Ellersick, Joe LociCero, Dan Stokesbury, Salah Aidarous Awards, and the IEEE Canada McNaughton Gold Medal. He is a fellow of the IEEE, the Engineering Institute of Canada, and the Canadian Academy of Engineering.

DOI: 10.1002/wcm