Performance Issues with IEEE 802.11 in Ad Hoc Networking

Claude Chaudet, ENST
Dominique Dhoutaut, Université de Franche-Comté
Isabelle Guérin Lassous, INRIA/CITI

ABSTRACT

Although 802.11 was initially designed for infrastructure-based networks, the distributed coordination function allows mobiles to access the radio medium without the need for a base station. Thus, studies in wireless multihop networks, also called ad hoc networks, often rely on the use of IEEE 802.11 for the physical and MAC layers. However, for a couple of years the use of 802.11 in ad hoc networks has been discussed. Different scenarios show serious performance issues. The performance offered by 802.11 is often low and directly impacts the performance of higher-layer protocols. In this article we provide a summary of the different performance issues extracted so far. We classify basic situations according to the main effects. We then present a quick survey of the possible solutions to these problems. This classification is intended to help design appropriate MAC protocols dedicated to ad hoc networks.

INTRODUCTION

Mobile ad hoc networks are wireless and infrastructureless networks. They are formed solely by the gathering of wireless terminals and do not require any base station to operate. Communications between out-of-range peers is ensured by a dedicated routing protocol. The terminals are mobile, and these networks are totally self-organized. The randomness of the topology coupled with mobility make these systems difficult to analyze.

IEEE 802.11 [1] has become the standard for wireless networking. This protocol defines both physical and medium access control (MAC) layers. Initially designed for base station managed networks, it includes a totally distributed medium access protocol, the distributed coordination function (DCF). With the DCF, ad hoc networks can be built on top of IEEE 802.11 as soon as a routing protocol is provided.

Most of the protocols designed for ad hoc networks assume that IEEE 802.11 is used for lowest-layer communications. Some work has highlighted a certain number of performance issues arising with IEEE 802.11 used in an ad hoc context. Each study evaluates one specific scenario; there is no work that summarizes all encountered performance issues. Moreover, some studies evaluate 802.11 via specific applications, and it is difficult to isolate the performance of 802.11 from that of higher layers.

This article presents an overview of all known performance problems with the 802.11 DCF that can appear in these networks. We present the DCF mode. We compile a classified list of the elementary issues and present some simulation results. We discuss possible and proposed solutions to these issues.

IEEE 802.11 DCF

The IEEE 802.11 DCF is part of the carrier sense multiple access with collision avoidance (CSMA/CA) family. To prevent collisions, emitters have to wait for the channel to become free before sending a frame. When a frame is ready to be emitted, if the medium is free when the frame is dequeued, it is emitted after a fixed time interval called the distributed interframe space (DIFS) during which the medium shall stay idle. If the medium is or becomes busy during this interval, the sender draws a random number called backoff in an interval called contention window. When the medium becomes idle again, the mobile waits for 1 DIFS before starting to decrement its backoff slot by slot. When the medium becomes busy, the process is stopped and will be resumed later after a new DIFS with the remaining number of backoff slots. As soon as the backoff reaches 0, the frame is emitted.

When a collision occurs, the contention window size of involved emitters is doubled, and the same frame is re-emitted by the process described above. If another collision occurs, the contention window size is doubled again until it reaches its maximum value defined by the standard. After a fixed number of retransmissions, the frame is dropped and the contention window size is reset. A successful transmission also resets the contention window size.
IEEE 802.11 protocol operation.

As collision detection is not possible due to the half-duplex characteristic of the wireless interfaces, each unicast frame has to be acknowledged. When a receiver successfully receives a frame, it waits for a short interframe space (SIFS) and then emits the acknowledgment. The SIFS is much shorter than the DIFS in order to give priority to acknowledgments over data frames. The emitter treats the lack of reception of an acknowledgment as a collision.

When an erroneous frame is detected at the MAC level (due to collisions or transmission errors), emitters will defer their transmission of an extended interframe space (EIFS) interval times instead of a DIFS, when the medium becomes idle. This timing is much larger than the DIFS and sufficient to protect acknowledgment transmission. The reception of an error-free frame during the EIFS wait cancels this long wait mechanism.

To prevent a hidden node situation in which two independent emitters simultaneously send a frame to the same receiver, an optional request to send (RTS)CLEAR to send (CTS) exchange can be used. Before transmitting a frame, the emitter asks the receiver if the medium is free in its vicinity by emitting an RTS frame. If no interfering transmission is present, the receiver answers by a CTS frame and the transmission can begin. Neighbors of both emitter and receiver overhear these frames and consider the medium reserved for the duration of the transmission, acting as if it was busy for this whole time. This mechanism is called virtual carrier sense.

Figure 1 summarizes part of these operations. This figure represents the MAC operations in a network of three nodes in mutual communication range. At the beginning, node A has a frame to send. As the medium is free, it can emit its frame after a DIFS, which is followed by the acknowledgment from B. A then wishes to emit another frame. If the new frame has been enqueued while the previous exchange was taking place, A has to draw a backoff. In this example, a backoff of 2 slots is drawn. In parallel, C also wishes to send a frame to B. As the medium was sensed free, it emits as soon as DIFS is elapsed. Unfortunately, A and C choose the exact same moment to start their emission, and a collision happens at B. As no acknowledgment is sent back, both emitters re-emit their frame, drawing a backoff in a doubled contention window. This time, C wins the contention, and its emission stops the backoff decrementation of A two slots before it finishes. After the data-acknowledgment exchange is finished, A restarts the backoff process with its two remaining slots.

**PERFORMANCE ISSUES**

Performance study of this standard has been the subject of many contributions. Simulation-based studies are the most frequent and can help to understand the behavior of the MAC layer even though the physical layer is often oversimplified. Analytical modeling is also an active field, but the designed models usually study single-hop networks or at least fully synchronized networks. Finally, experimentations represent the most reliable method of performance analysis, but experimental conditions are difficult to control and results are often hard to analyze. Most work that presents performance issues with 802.11 in the ad hoc context gives simulation results. The presented performance issues can be classified in three main categories:

- Configurations that lead to long-term fairness issues in which some flows suffer from starvation while other flows capture the whole channel bandwidth
- Configurations that result in short-term fairness issues where the frames are emitted in bursts
- Configurations that result in overall throughput decrease, where a part of the network capacity is not used and thus wasted

Some configurations fall into several categories. Usually, performance issues do not appear until the medium capacity is overloaded. But with greedy applications such as file or video transfer, situations where capacity is overloaded can frequently arise in ad hoc networks since the radio medium bandwidth is scarce and shared. The 802.11 DCF mechanisms that prevent an emitter from accessing the radio medium in normal conditions (i.e., reduce its throughput) can be classified in three categories:

- The emitter cannot transmit because the medium is busy for a long time.
- Its random wait phase (given by the back-off) is long due to its contention window increase after repeated collisions.
Its fixed wait phase is longer than in usual conditions since it uses an EIFS in place of a DIFS due to the presence of distant transmissions.

Figure 2 groups most of the known elementary configurations that show performance issues. Plain lines represent the existence of a communication link between two nodes. Depending on the attenuation on the path, the link rate can be used for 1, 2, 5.5, or 11 Mb/s communications. Each of these data rates uses different physical encodings, and high rates require a better signal-over-noise ratio. Dashed lines represent the “links” triggering an EIFS; that is, on such a link, the emission of one station forces the other node to use an EIFS for the next transmission. This happens when the physical header of the frame is decoded but the MAC frame reception is incorrect. Note that the physical header is sent at 1 Mb/s (or possibly 2 Mb/s with 802.11b).

Some performance issues depicted later on require some emitters to be totally independent. It means that two emitters are not in mutual carrier sensing range. This kind of situation can arise either in the presence of obstacles or when emitters are distant enough. In [2] a study shows that the carrier sensing range is around 300 m regardless of the used transmission rates. Some scenarios requiring mobiles to be independent can therefore only happen at low data rates, using the first IEEE 802.11 flavor, or in the presence of obstacles. In Fig. 2, an obstacle creates the absence of relationship when needed.

To analyze the depicted scenarios, we use the NS-2 network simulator with modified parameters of the 802.11 DCF in order to match the rate and the parameters of the 802.11b wireless cards. The modifications concern only physical parameters and modulation and multi-rate are not implemented. The simulation results therefore correspond to wireless hosts using AVAYA cards at the maximum throughput. In order to isolate the performance of 802.11 from that of higher layers, we use ad hoc on demand vector routing (AODV) as the routing protocol with increased timeouts in order to get the behavior of static routing without route expirations. Moreover, the emitters send UDP packets since UDP does not add any complexity in analysis of 802.11 performance compared to TCP.

**THE LONG-TERM FAIRNESS ISSUE**

In [3] the scenario represented in Fig. 2a, involving three emitter-receiver pairs, is depicted. Three flows try to occupy the entire channel capacity. Let us first consider that the links between emitters are communication links, not as shown in Fig. 2a. Emitters A and E are totally independent and evolve asynchronously. This can happen either when flows are emitted at a low throughput, or if obstacles block the line of sight between A and E. Moreover, receivers are near enough to their associated emitters that collisions do not happen. Whenever two packets are emitted simultaneously, both transmissions are supposed to be successful. This way, the sole medium access problem can be isolated. When emitter C wishes to transmit a frame, it has to wait for both A and E to leave the medium free simultaneously. Since A and E are independent, their silence periods almost never coincide, and C sees the medium as busy more often than the two others. This basic imbalance in medium access results in a severe fairness issue as the (C, D) flow can expect no more than 15 percent of the medium capacity, whereas the two others can use 75 percent of the capacity, as shown in [4] by both analytical modeling and simulation. In this configuration the long-term unfairness comes from the fact that some emitters are independent and unaware of each other, and prevent one emitter from accessing the medium. The share of the radio medium is represented in Fig. 3 as

---

**Figure 2. Configurations leading to performance issues: a) three pairs; b) colliding flows; c) large EIFS; d) hidden nodes; e) small EIFS; f) chain of mobiles.**
the achieved throughput of each emitter according to the packet size. The throughput of the exterior transmitters increases with packet size, whereas the throughput of the central emitter remains constant. Modifying packet size on all three emitters improves the overall performance of the network at the cost of an unfairness increase. Now let us consider that the emitters are no longer in communication range of each other but at a distance that triggers an EIFS between each other, as depicted on Fig. 2a. The situation is worsened by the use of EIFS as the central pair has to wait not only for both exterior pairs to be silent simultaneously, but also for them to be silent simultaneously for a time longer than an EIFS in order to be able to decrement a single backoff slot. This leads, as pointed out in [3, 4], to a medium share of about 5 percent vs. 95 percent. Compared to the previous situation, unfairness is worsened by a longer fixed wait phase of the DCF.

In [5] the authors describe another scenario that presents long-term fairness issues, shown in Fig. 2b. In this topology, A and C are independent but can both communicate with B. This can happen either when flows are emitted at a low throughput and emitters are distant enough, or if obstacles block the line of sight between A and C. Node A always senses the medium free and therefore alternates backoff decrementation periods and emissions. Transmission from node C to node D always succeeds if it manages to start, as collisions cannot happen at D. C sees the medium free unless virtual carrier sense blocks it or B transmits an acknowledgment to A. Therefore, if RTS-CTS is not active, the chance for A to successfully transmit a frame depends on the frame size, as it should fit in a backoff interval of C that never increases its contention window. With RTS-CTS, the chance for A to be able to reserve the medium access is increased; nevertheless, if the CTS is emitted during the time C sends a RTS, C will not understand the CTS and will soon emit another frame. This results in great unfairness, shown in Fig. 4. Without RTS-CTS, the unfairness is high and gets worse when the packet size increases. With RTS-CTS, this share remains constant and corresponds to 10 percent of the used bandwidth for successful transmissions of A and to 90 percent for C. In that case, the unfairness comes from the fact that two emitters are independent and one receiver experiences many more collisions than the other receiver.

In [6] the authors present a scenario, called the large EIFS problem, leading to unfair behavior. If we consider two simultaneous flows as in Fig. 2c, acknowledgment from C triggers an EIFS at node A. The situation is unbalanced between the two contending emitters, as one has to wait for a DIFS while the other needs to wait an EIFS before decrementing any backoff slot. Simulations show that the throughput of each emitter increases with packet size, but the ratio between the two emitters’ throughputs remains the same: B sends around seven times more packets than A.

In [7] the authors point out that the presence of slow terminals in a single-hop network slows down every other terminal: during the transmission of a slow terminal, the medium is busy for a longer period than during the transmission of a fast terminal. As 802.11 provides per-packet fairness in one-hop networks, each emitter statistically sends one frame in turn. As soon as multiple rates are present in an ad hoc network, a loss of performance can appear for some terminals due to the existence of slow transmissions. Note that with actual wireless cards, the control frames (RTS, CTS, etc.) and broadcast frames are transmitted at a slower rate than unicast frames. We classify this anomaly in the long-term fairness issues insofar as some flows achieve the desired rate and others do not; moreover, it can also be seen as a global throughput decrease. In this case the performance issue comes from DCF’s per packet fairness.
THE SHORT-TERM FAIRNESS ISSUE

In [8] the authors point out that even though RTS-CTS exchange can solve hidden node situations and results in fair medium access in the long term, the behavior of the medium access is unfair in the short term. This can happen either when flows are emitted at a low throughput and emitters are distant enough, or if obstacles block the line of sight between the emitters. Figure 2d represents the well-known hidden node scenario when RTS-CTS exchange is active. RTS from A and B will collide at C until both emitters’ backoffs are large enough to allow successful transmission of an RTS. When one emitter succeeds, the CTS that follows blocks the other emitter from decreasing its backoff, which is statistically large as contention window sizes increased just before. On the other hand, the successful emitter resets its contention window and draws a small backoff at the next round. Thus, it is likely to access the medium again and so on until collisions happen again. This results in bursty behavior for both emitters. The simulations we have carried out show that the average number of packets consecutively sent by the emitter that has won the contention is nine, regardless of packet size. In that configuration, the short-term unfairness comes from resetting the contention window of the station that wins the medium access.

Now we only consider two of the three pairs of Fig. 2a, say (A, B) and (C, D), and assume that an EIFS is triggered on one emitter when the receiver of the opposite pair sends an acknowledgment. This can happen with the different communication rates of 802.11 as soon as the two pairs are well located. Long-term fairness is preserved, but we can observe short-term unfairness. When a station accesses the medium it has to wait for a DIFS when the medium becomes free, whereas the competing emitter has to wait for an EIFS. The simulations we have carried out show that the average number of packets consecutively sent by the emitter that has won the contention is around four regardless of packet size. In this case, the short-term unfairness comes from a longer fixed wait interval for an emitter.

LOW AGGREGATE THROUGHPUT

Let us now consider the configuration depicted in Fig. 2c. This situation, presented in [6] and called the small EIFS problem, exhibits a significant performance decrease in terms of overall used bandwidth. This situation may arise with low-speed communication links or high-speed communication links with obstacles making A and D independent. CTS and acknowledgments from one receiver trigger the EIFS on the other emitter. Collisions happen since this EIFS is not large enough to protect the associated data frame. The resulting contention window increase represents additional performance loss. This problem is also depicted in [9]. This study shows the inefficiency of RTS-CTS in these situations. The simulations we have carried out show that the overall performance loss increases with packet size and ranges from 9 to 25 percent of the capacity. This scenario also exhibits a short-term fairness issue as the frames are emitted in bursts. This kind of situation arises whenever two emitters are independent and their emissions collide at the receivers.

In [10, 11], the authors study the behavior of a chain of mobiles, as represented on Fig. 2f. This configuration happens as soon as two mobiles not in communication range wish to communicate and routing is involved. According to the used rate, the dependency links are not the same. With high rates, each mobile may detect the activity of its two- or three-hop neighbors. With lower rates, more mobiles are independent. The combination of multiple problems leads to a global performance problem. They show, with NS-2 simulations, that achievable throughput stabilizes around 250–300 kb/s in a six-hop chain for a network capacity of 2 Mb/s where 400 kb/s could be expected. This fact may be explained by the multiple collisions that occur on the second and third mobiles on the chain and the triggering of EIFS due to distant emissions. Thus, very few packets reach the fourth
terminal of the chain. Packets transmitted by this mobile are very likely to reach the destination, though, since the medium is not overloaded in this part of the chain.

Figure 5a compares throughputs achieved with different chain lengths with RTS-CTS. The throughput decreases as the number of hops increases. Increasing the number of hops from two to three results in a throughput loss of 50 percent. The greatest gap occurs when increasing to four hops and results in a further decrease of 80 percent due to the carrier sense dependencies. Then the situation stabilizes if we consider the carrier sense range to be twice the communication range.

Figure 5b compares by simulation the throughput with and without RTS-CTS exchange. The throughput decreases when RTS-CTS is activated. In that case, the spatial reuse possibility is reduced: without RTS-CTS, parallel communications between the first and second terminals and the fourth and fifth terminals are possible, whereas with RTS-CTS, if the fourth terminal is already sending packets to the fifth mobile, the second terminal cannot answer the RTS of the source since it is blocked by the transmission of the fourth mobile. The spatial reuse possibility is reduced, which leads to a noticeable decrease in throughput.

In [10] the authors show that the same inefficiencies of 802.11 apply on regular topologies like mesh and on random topologies with random traffic.

**POSSIBLE SOLUTIONS**

The problems depicted in the previous sections are essentially due to the impact of particular topologies on the MAC protocol operation. As topology is, in most cases, impossible to control, a solution consists of modifying the MAC protocol. Nevertheless, the freedom in such an approach is quite low, provided that the CSMA/CA approach is preserved. Performance or fairness issues are usually caused by the emission of nodes unaware of the fact that the traffic they generate can cause problems somewhere in the network. Several solutions can be found to particular problems, but only two types of general solutions seem possible: either explicitly informing the nodes that they are the cause of some medium access problems, or restraining the nodes from emitting in order to avoid medium saturation.

Bringing the nodes knowledge of the congestion level of neighbor emitters can be difficult because it requires explicit signaling, which would be difficult to transmit on an already overloaded medium. Using a separate channel for these particular messages could be a solution, but apart from specific countries’ regulations that would prevent this operation, it would not solve situations like the three pairs configurations in which communication between emitters is impossible.

Statically limiting the nodes’ emissions is inefficient because the congestion level is highly dependent on the number of emitters. Many authors propose to modify the medium access method statistically by modifying the backoff algorithm or the contention window increase method at each node according to its perception of the medium load. Nevertheless, it is not easy to find a good compromise between global performance reduction and fairness increase. Indeed, an emitter cannot, by its sole perception, determine in every situation whether it is alone in the network or if it prevents another node from transmitting.

In [8] the authors propose to modify the algorithm that manages the contention window size in order not to favor the terminal that has just won the contention for the subsequent transmission, thus reducing short-term unfairness. In [5] the authors propose a fair MAC protocol based on the DCF: by listening to packets sent on the medium, each terminal is able to determine if it accesses the medium more often than its neighbors or vice versa. With this knowledge, the terminals can adapt their contention window size in order to limit or increase their throughput. The difficulty is to define the thresholds and modification algorithms of contention window size.

In [12] the authors propose to modify each node’s medium access method according to the perception of the occupation of the medium due to its emissions and other nodes’ transmission. According to these measurements, nodes compute a probability of transmission. When a backoff timer expires, nodes only transmit according to this probability, which decreases as the medium load due to the node’s emissions increases. Although originally designed for single-hop networks, it seems suitable for multihop networks with minor modifications.

To solve EIFS-related problems, the authors of [6] propose to measure the length of the frames detected but not decoded, and to base deferment on these measures. Such a solution requires using different lengths for the different frame types.

In [9] the authors allow a CTS reply only if the RTS is received with a power larger than a given threshold. Thus, they decrease the possible communication range but ensure that the RTS informs all potentially interfering nodes.

Many solutions are specific to some configurations and cannot solve the whole range of performance issues encountered. Work still remains to design a global solution that proves to be efficient in every situation. The approaches described in this section have to face the challenge of finding a good trade-off between overall performance of the system and the bandwidth equity between mobiles. To overcome this challenge, the use of multiple channels in terms of time, frequency, or code is also proposed. But the problem in ad hoc networks is to allocate the channels since the system is fully distributed and mobile. Due to dynamism and lack of synchronization, their implementation is not straightforward, and much work remains to achieve a usable and efficient solution. Moreover, it is still unclear that the proposed solutions have a significant gain in overall performance compared to the overhead required for allocation and control of the channels. Other directions are to use directional antennas or power control in order to reduce interference and increase spatial...
reuse. These solutions nevertheless introduce new problems that still require careful attention.

**CONCLUSION**

In this article we summarize the performance issues that arise with the IEEE 802.11 distribution coordinator function when used in ad hoc networks. We present the main known elementary topologies that lead to performance issues and classify them into three categories: long-term and short-term unfairness situations, and overall throughput decrease. Simulation results are presented for some issues. Finally, we discuss some possible solutions to overcome these problems.

Work still remains to design a global solution that proves to be efficient in every situation. The right compromise between global performance and access fairness is not easy to find and even harder to enforce. Identifying the problems is a first step toward this goal.

Nevertheless, due to the complexity of these networks, an isolated MAC layer might not be the right choice; information from lower and higher layers could help the MAC layer to efficiently adapt to problems encountered. Cross-layer protocol design is probably one of the most promising areas of research in this area, although layer independence will still be preserved.

**REFERENCES**


**BIographies**

**CLAUDE CHAUDET** (Claude.Chaudet@enst.fr) has been an assistant professor at Ecole Normale Supérieure de Télécommunications (ENST), Paris, France, since 2005. He received his Master’s degree in computer science from INSA de Lyon in 2001. He received his Ph.D. thesis in computer science from INSA de Lyon in 2003. His research interests are in the areas of wireless networks (ad hoc, sensor, and hybrid networks) and QoS protocols.

**DOMINIQUE DHOUTAUT** (Dominique.Dhoutaut@pu-pm.univ-fcomte.fr) has been an assistant professor at the University of Franche-Comté, France, since 2004. He received his Master’s degree in computer science from the University of Franche-Comté in 2000. He received his Ph.D. thesis in computer science from INSA de Lyon in 2003. His research interests are in the areas of wireless networks and experimental evaluation.

**ISABELLE GUÉRIN LASSOUS** (Isabelle.Guerin-Lassous@inrialpes.fr) has been a research officer at INRIA, France, since 2000. She received her Ph.D. degree in computer science from the University of Franche-Comté in 1999. She is currently in the Architecture of Networks of Services (ARES) project of INRIA Rhône-Alpes, located in the CIT laboratory at INSA de Lyon. Her research interests are in the area of local wireless networks.