Limitations of the IEEE 802.11e EDCA Protocol when Supporting Real-Time Communication

Ricardo Moraes[⊕◇], Paulo Portugal[⋆], Francisco Vasques[⋆], José Alberto Fonseca[♣]
[⊕] *LabSEC, Federal University of Santa Catarina, Brazil*
[◇] *IT-Instituto de Telecomunicações, University of Aveiro, Portugal*
[⋆] *Faculdade de Engenharia, University of Porto, Portugal*
[♣] *Department of Electronics and Telecommunications, University of Aveiro, Portugal*
rmoraes@inf.ufsc.br, {pportugal, vasques}@fe.up.pt, jaf@det.ua.pt

Abstract

In this paper, we analyze the timing behavior of the EDCA communication mechanism defined in the IEEE 802.11e standard, when it is used to support real-time (RT) traffic. In the context of this paper, RT traffic means small sized packets generated in periodic intervals, which must be delivered before the end of the message stream period. Otherwise, the message is considered to be delayed and a deadline loss occurs. The target of this paper is to understand the limitations of the highest priority level of the EDCA mechanism (voice category) when supporting RT communication. We have assessed this mechanism considering an open communication environment, where there are RT and non-RT stations operating in the same frequency band. Furthermore, a realistic error-prone model channel was used to measure the impact of interferences against an error-free channel. We show that in the most cases evaluated, both the number of packet losses and the average packet delays forecast an unacceptable number of deadline losses for the RT message streams, even for intermediate load cases. As a conclusion of this paper, we present some potential future directions toward improved QoS in wireless networks.

1 Introduction

In the last few years, there has been a growth in the use of wireless technologies in several application areas that requires a trustworthy Quality of Service (QoS). Driving examples range from voice over IP (VoIP) to Networked Control Systems (NCS). For such type of application domains, the support of reliable communications is one of the major requirements. For instance, in automation systems, real-time (RT) control data must be periodically transferred between sensors, controllers and actuators according to strict transfer deadlines.

The IEEE 802.11e [1] amendment is a published standard, intending to provide differentiated levels of QoS to the supported applications. A number of studies have evaluated this standard considering typical multimedia traffic requirements. That is, requirements usually applied for transferring voice and video streams together with background traffic. However, when the communication services are used to support RT applications, specific communication requirements must also be considered, including specific RT and reliability constraints [2].

Besides, a relevant aspect that must be taken into account, when addressing wireless networks, is that all stations share the access to the same radio channel, as the medium is an *open communication environment*. Thus, any new participant can try to access the medium at any instant according to the medium access control (MAC) rules and establish its own communication channels. Furthermore, the wireless communication environment is susceptible to interferences, not only from devices using the same communication technology, but also from other technologies working in the same frequency band [3]. Therefore, the bit error rate (BER) cannot be considered negligible in wireless networks. As a consequence, the system load cannot be predicted at system setup time, nor can it be effectively controlled during the system run-time.

Most published works on the performance of IEEE 802.11 networks have focused on typical metrics used in multimedia domain. Besides, the impact of varying networks conditions, which is one of the main challenges to ensure QoS support in IEEE 802.11 networks identified in [4], are usually not considered. Relevant exception is [5], which experimentally assesses the impact of background traffic upon RT traffic.

In this paper, we evaluate the timing behavior of the EDCA mechanism of the IEEE 802.11e amendment by simulation. Basically, we assess the impact of the unconstrained external traffic load upon the behavior of the voice category, when this access category is used to transfer small sized packets, generated in periodic intervals. Besides, we consider both the case of error-free and error-prone channels. That is, the effects of interferences typical of industrial environments (*e.g.* EMI) are taken into account and compared to the ideal environment behavior.

The rest of this paper is organized as follows. In

978-1-4244-2350-7/08/$25.00 ©2008 IEEE.

Section 2 we describe the basics of the IEEE 802.11e MAC mechanism. In Section 3 we briefly refer the used Stochastic Petri Nets (SPN) simulation model. Afterwards, in Sections 4 and 5 we carefully analyze relevant RT simulation scenarios. These scenarios target the timing behavior of the EDCA mechanism supporting RT traffic. In Section 6 a summary and future directions for the support of RT communication in open communication environments are drawn. Finally, the paper is concluded in Section 7.

2 IEEE 802.11 Medium Access Mechanisms

The original IEEE 802.11 MAC protocol defined two MAC mechanisms: the Distributed Control Function (DCF) and the Point Coordination Function (PCF). The DCF mechanism uses the well-known Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol to control the medium access. The PCF mechanism uses a centralized polling scheme upon the DCF mechanism to support synchronous data transmissions, through an Access Point (AP) station.

Recently, the IEEE 802.11e standard was published as an amendment to the original standard, intended to provide QoS in IEEE 802.11 networks. The IEEE 802.11e amendment incorporates an additional coordination function called Hybrid Coordination Function (HCF) that is only used in QoS network configurations. The HCF uses both a contention-based channel access method, called the *Enhanced Distributed Channel Access* (EDCA) and a controlled channel access, referred as the HCF *Controlled Channel Access* (HCCA) [1] (Figure 1).

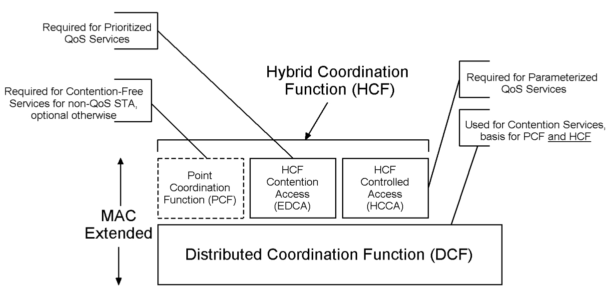

Figure 1. IEEE 802.11e MAC architecture.

The HCF mechanism schedules the access to the channel by allocating transmission opportunities (TXOP) to each of the stations. Each TXOP is defined by a starting time and a maximum length and may be obtained through one of two access mechanisms specified by the HCF function (HCCA and EDCA).

The HCCA mechanism has been proposed as an improvement to the PCF mechanism. It was proposed to guarantee bounded delay requirements. However, some preliminary studies [6, 7] have already shown that HCCA may not be able to guarantee the expected RT communication requirements. Furthermore, it is still not clear if the HCCA mechanism will be implemented in next generation WLAN networks cards, solving the unavailability problem of the PCF mechanism [8]. Therefore, it is expected that multiple applications will use the EDCA mode to support RT communications, due to the widespread use of this access mode.

Therefore, this paper will be focused on the analysis of the timing behavior of the EDCA mode, when it is used to support RT communications.

2.1 IEEE 802.11e EDCA

The EDCA provides a distributed mechanism for the channel access under the control of the HCF coordination function. It can be used as the only access mechanism or it can be used during the contention period (CP) of the HCCA mechanism. The EDCA is designed to provide differentiated transmission services, with 4 different priorities. It enhances the DCF scheme, as each frame arriving at the MAC layer with a defined priority will be mapped into one of 4 access categories (AC): voice (VO), video (VI), best-effort (BE) and background (BK). These access categories are based on the 8 priority levels defined by the IEEE 802.1D standard, as follows: priorities 1 and 2 for BK traffic; priority 0 and 3 for BE traffic; priorities 4 and 5 for VI traffic; and, finally, priorities 6 and 7 are mapped for VO traffic that is the highest priority level.

Different levels of service are provided to each of the AC traffics, based on three independent mechanisms: the Arbitration Interframe Space (AIFS), the TXOP time interval and the Contention Window size (CW). Firstly, for a station operating under the EDCA mode, each frame will wait during an *AIFS[AC]* interval, instead of waiting during a DIFS interval (as it was the case for DCF in IEEE 802.11). Only after the channel remaining idle during an *AIFS[AC]* interval (Equation 1), the station will start to transmit the frame.

$$AIFS[AC] = AIFSN[AC] \times aSlotTime + aSIFSTime \quad (1)$$

where the *AIFSN[AC]* must be greater than or equal to 2 for all stations, except for the QoS Access Points where it shall be greater than or equal to 1. Figure 2 shows the relationships between the multiple *AIFSs* in the EDCA scheme.

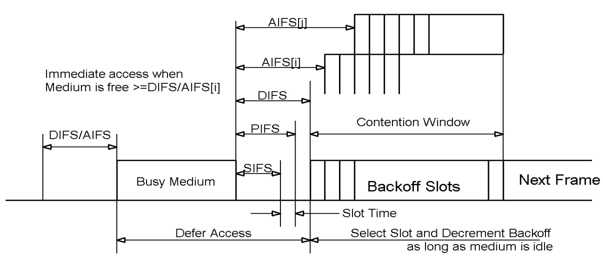

Figure 2. Interframe spaces in the EDCA mechanism.

Secondly, the EDCA mechanism introduces the TXOP concept, i.e., a time interval during which the station keeps the medium access rights. Consequently, multiple frames may be transmitted during an acquired TXOP, if there is more than one frame pending to be transferred in the AC for which the channel has been acquired.

Finally, if a station wants to transmit a frame while the channel is busy, or becomes busy before the expiration of the *AIFS[AC]*, the backoff procedure is invoked (third traffic differentiation mechanism). The contention windows are defined by the aCW_{min}*[AC]* and aCW_{max}*[AC]* attributes, in contrast to the legacy DCF where the initial values are randomly selected from the [0, CW] interval defined by the physical layer. In the EDCA function, the backoff procedure selects a random number, in the range [*0, CW*], where the *CW* size is initialized at aCW_{min}*[AC]*. When a transmission fails, *CW* is increased as $newCW[AC] = [(oldCW[AC]+1) \times 2 - 1]$. On the other hand, the backoff counter decreases the backoff interval whenever the medium is detected to be idle for *AIFS[AC]*. In contrast to the DCF mode, where the station would try to transmit as soon as the backoff timer reaches zero, in EDCA the station initiate the transmission in the first slot boundary after the backoff counter has reached zero. Nevertheless, DCF counters decrement at the end of idle backoff slots, while EDCA backoff counters decrement on *backoff slot boundaries*. Consequently, there is no difference on the initial transmission time between DCF and EDCA, considering the same number of selected slots to backoff. Additional information about IEEE 802.11 MAC protocol can be found at [9].

3 The Simulation Analysis

Traditionally, the analysis of the RT communication behavior should be made for worst-case scenarios. A common specification for the delay requirement is through the use of a deterministic delay bound, where $D_i <= D_{max}\ \forall i$, being D the delay of each message i, and D_{max} the upper bounded delay specified by the client. Usually, this upper bounded delay is also called deadline. However, worst-case scenarios in probabilistic medium access networks, as it is the case of IEEE 802.11 networks, address rarely occurring cases and, those cases may only be relevant for safety-critical applications.

This paper has been focused on the timing analysis of the average communication behavior. The main reason for this option is that the target applications are usually loss tolerant in what concerns the lost of some message deadlines. For instance, the transfer of a video stream may be specified to tolerate a maximum of 10% deadline loss rate, if the loss frames are "adequately" spaced. Other examples of relevant loss tolerant applications are NCSs scheduled according to the (m,k)-firm model [10], and the support of VoIP applications, where an average packet delay below $150ms$ and an average jitter $< 50ms$ are acceptable for most user applications [11]. However, for VoIP applications, only packet loss rates up to 3% are generally acceptable.

In the case of loss tolerant applications, the RT behavior can be specified through the use of statistical metrics. For example, the statistical delay bound can be represented by $Prob(D_i \leq D_{max}) \geq Z_{min}$, where D_i and D_{max} are defined as above, and Z_{min} is the lower bound of the probability of successful and timely delivery. Adequate values for Z_{min} within the industrial environment can be defined among 0.95 to 0.98, when dealing with soft RT applications.

The simulation model was built using a *Stochastic Petri Net* (SPN) tool [12] and it describes the dynamics of the Contention-Based Channel Access function (EDCA) of the Hybrid Coordination Function (HCF) of IEEE 802.11e. For all the simulations, it has been used a Semi-Markov error model, where the channel is always in one of two states: Good or Bad. The state holding times are according to a log-normal distribution. This model assume that bit errors are independent, with a fixed error rate. For the parametrization of the Semi-Markov model, it has been used the values defined in [13], which are realistic values for wireless transmission in an industrial environments.

Therefore, for all channels the mean duration of good state is $65ms$, the mean duration of bad state is $10ms$ and, the coefficient of variation (CoV) for the bad state holding times has been set to 10 and for the good state to 20. These mean burst lengths lead to a rather bad channel, where the steady-state probability for finding the channel in bad state is approximately 13.3%. Two sets of simulations are assessed, differing in their respective mean bit error rate (BER). The first set defines a mean BER of 10^{-4}, while the second set defines that no bit errors occur. Therefore, during the bad channel states for the first set, the BER is about *7.5E-3* and, for the good state no errors will occur.

4 Simulation Scenarios

The proposed scenarios analyze the behavior of the highest access category (*voice*), when this category is used to transmit RT data (small sized packets in periodic intervals) from RT stations. The communication environment is shared with unconstrained traffic sources that are out of the sphere-of-control of the RT architecture. These RT periodic data exchanges are intended to model both sensor messages sent to plant controllers, and output messages sent from plant controllers to the actuators.

Therefore, a simulation model was built, considering multiple ST and RT stations operating in the same frequency band. The RT stations only transfer RT traffic, using the default set of parameters defined by the EDCA function for the voice (VO) access category. On the other hand, ST stations transmit three types of traffic: voice (VO), video (VI) and background (BK), also using the default set of parameters defined by the EDCA mechanism.

Basically two simulation cases are analyzed. The first scenario (*small population case*) considers 10 ST stations operating in the same frequency band together with 10, 20, 30 or 40 RT stations. The second scenario (*large population case*) extends the number of ST stations to 40.

The physical parameters used in the simulations are based on the IEEE 802.11a PHY mode. Each station operates at OFDM (orthogonal frequency division multiplexing) PHY mode, control frames are transmitted at a basic rate equal to 1 Mbps, while the MSDU (MAC service data units) are transmitted at 36 Mbps. The maximum number of transmission attempts is set to 4. The MAC queue size is set to 50 positions. All other relevant simulation parameters are shown in Table 1, where it becomes clear that RT traffic is transferred at the same priority level as VO (voice) traffic. The only distinguishing difference is the length of transferred packets.

Table 1. Simulation data.

Parameters	RT stations	ST stations		
		VO	VI	BK
CW_{min}	7	7	15	31
CW_{max}	15	15	31	1023
AIFSN	2	2	3	7
TXOP (ms)	1.504	1.504	3.008	0
Packet Size - bytes	45	160	1280	1600
Message stream (ms)	2, 10, 20	Variable	Variable	Variable

The generated data frames have a constant size. The RT traffic is characterized by periodic traffic sources with a small amount of imposed jitter. To model this behavior a normal distribution with $\sigma/\mu \leq 1\%$ (σ is the standard deviation and μ is the average expected value) is used. It is also guaranteed that the RT traffic is not correlated among RT stations. Each RT station generates packets with fixed message stream periods (MSP) of $2ms$, $10ms$ or $20ms$ with 45 bytes for data payload, which is equivalent to generate 500, 100 or 50 packets/s. Therefore, each RT station imposes a constant network load of 180, 36 or 18 kbits/s that represents about 0.5%, 0.1% or 0.05% of the total network load (without considering the MAC and PHY headers).

The ST stations have Poisson traffic sources, where the imposed network load ranges from 10% to 90%. In order to impose these network loads, each ST station generates $\{14.80; 29.60; 44.40; 59.21; 74.01; 88.82; 103.62; 118.42$ or $133.22\}$ VO, VI and BK packets/s for the small population scenario. Similarly, each ST station generates $\{3.70; 7.40; 11.10; 14.80; 18.50; 22.20; 25.90; 29.60$ or $33.31\}$ VO, VI and BK packets/s for the large population scenario.

5 Simulation Results

All the simulation results have been obtained with a 95% confidence interval with a half-width interval of 5%. The performance metrics analyzed in this paper include: average delay, average queue size and packet loss rate. The *average delay* is the average delay required to transfer a packet, measured from the start of its generation at the application layer to the end of the packet transfer at the receiving station. The *average queue size* represents the average output buffer occupancy. The *packet loss* metric represents the percentage of packets that are lost for each traffic stream.

As a first step, a set of simulations was performed to characterize the network behavior, when just RT stations are transferring messages in the communication medium. This set of results will be used for comparison purposes. Then, a full set of simulations is made to analyze the timing behavior of the RT stations, when the communication medium is shared with a set of timing unconstrained devices.

5.1 Simulation Results: RT Traffic

This set of results consider 10 to 40 RT stations generating RT packets (45 bytes for data payload) with fixed message stream periods (MSP) of $2ms$, $10ms$ or $20ms$. Therefore, it describes the network behavior when there are no ST stations trying to transfer its own messages, i.e., it represents an unrealistic *closed* communication environment.

The results presented in Figures 3, 4 and 5 represent the obtained values for average queue size, packet loss rate and the average packet delay, respectively.

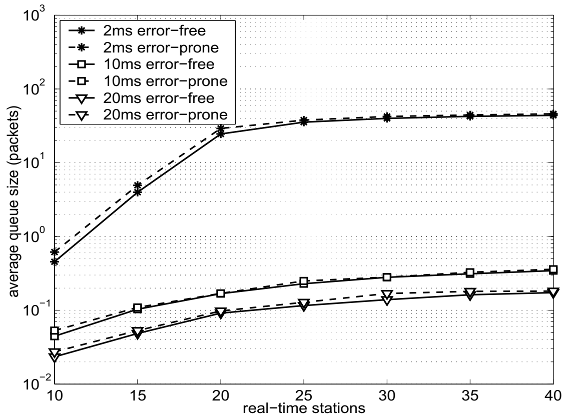

Figure 3. Average queue size: undisturbed scenario.

From such results, it is clear that the EDCA mechanism is not able to provide a RT communication service for more than 20 RT stations generating packets with MSP of $10ms$ or $20ms$. This conclusion derives from the following observations: on the one hand, the average queue size is smaller than 1 packet, which is an indication that, in average, all the deadlines for the RT traffic are accomplished (considering that, in the most usual case, the deadline of a message stream is equal to its generation period) and; the average packet delay is smaller than the period of the related message stream ($10ms$ or $20ms$). However, on the other hand, the packet loss rate is larger than 10% for a number of stations larger than 20. This means that the RT

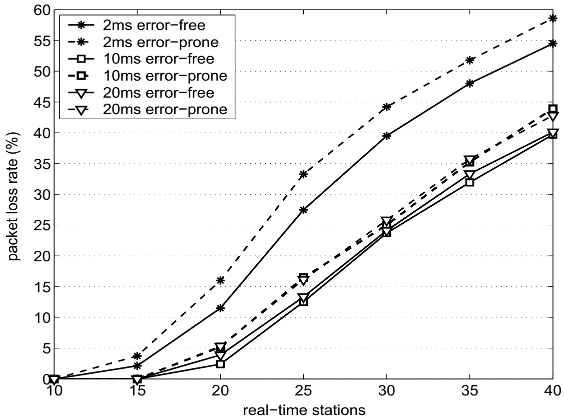

Figure 4. packet loss: undisturbed scenario.

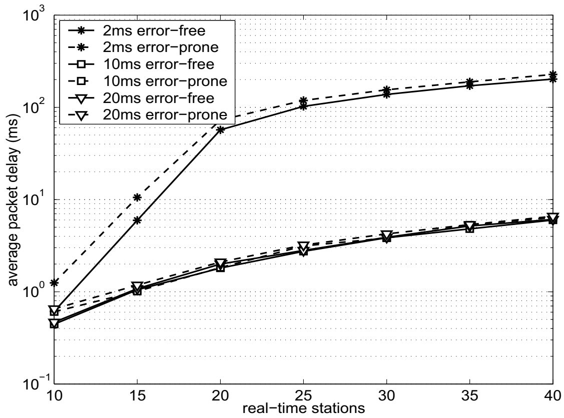

Figure 5. Average packet delay: undisturbed scenario.

message streams will loose more than 10% of its deadlines and, despite many RT applications being loss-tolerant, this value cannot be well accepted by the majority of them. Loss tolerance levels are different for different type of applications. As mentioned before, a video stream may be specified to tolerate a maximum of 10% deadline loss rate, only when the lost frames are "adequately" spaced.

Applying a similar reasoning, it can be also observed that for RT message stream periods of $2ms$, the EDCA mechanism is not able to provide an acceptable RT service when the number of RT stations exceed 10 stations. It can be easily verified through the average queue size results (Figure 3), where the average number of packets in the queue is already approaching 1 packet even for a number of RT stations as low as 10 stations. These values are consistent with the average packet delay values (Figure 5), where for $MSP = 2ms$, the average packet delay is already approaching the message stream deadline of $2ms$.

Therefore, from the analysis of this subsection, we conclude that the relevant scenarios for the assessment of the RT characteristics of the EDCA mechanism are restricted to, at most, 10/20 RT stations generating packets with MSP of $10ms$ or $20ms$ and; up to a maximum of 10 RT stations generating packets with MSP of $2ms$. Simulation scenarios above these thresholds are no longer relevant, as the EDCA mechanism is not able to support a RT communication service even in the case of a closed communication environment.

5.2 Simulation Results: ST-RT Traffics

In this subsection, we assess the behavior of both the small and the large population scenarios, where ST stations are joined with RT stations, both transferring its own messages in the same frequency band. For the sake of simplicity, only the values for RT traffic are plotted in the following figures, as the target of this study is to illustrate the impact of the ST timing unconstrained traffic upon the RT traffic, for each RT configuration (10 RT stations with a fixed MSP of $2ms$; 10/20 RT stations with a fixed MSP of $10ms$ or $20ms$). Furthermore, Ni et al. [14] have already demonstrated that the EDCA mechanism improves the performance behavior for high priority traffic by downgrading the service of the low-priority one. Thus, it is not relevant to further discuss the timing behavior of the ST traffic.

5.3 The impact of timing unconstrained ST traffic upon the average queue size of RT stations

A first simulation analysis concerns the assessment of the average queue size. Figure 6 shows the average queue size for message stream periods of $2ms$, considering the case of 10 RT stations operating in both the small and large population scenarios (10ST - 40ST).

When comparing the undisturbed scenario with the case of 10 RT stations operating together with 10/40 ST stations ($MSP = 2ms$), it becomes clear the impact that the ST timing unconstrained traffic has upon the average queue size of RT message streams. Figure 6 also illustrates the impact of the increasing number of stations contending for the medium access, which shows a clear degradation of the QoS for the large population scenarios. This is a clear consequence of the effect of multiple collisions occurring among the different contending stations.

Summing up, the EDCA mechanism is not able to provide RT communication service for MSPs of $2ms$ in open communication environments, since the average queue size of the RT stations is already larger than 1 packet when the network load imposed by external stations is only slightly above 10%.

Another important result, that is in contradiction with a common belief when dealing with wireless communications, is that the major source of perturbation upon the RT communication is caused by the increasing external network load and not by the error-prone characteristics of the wireless medium. It is interesting to note that the average queue size increases more than one order of magni-

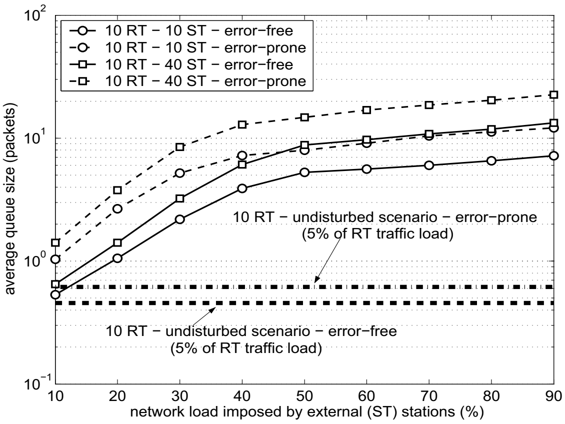

Figure 6. Average queue size (small and large pop.): error-free *vs.* error-prone - $MSP = 2ms$.

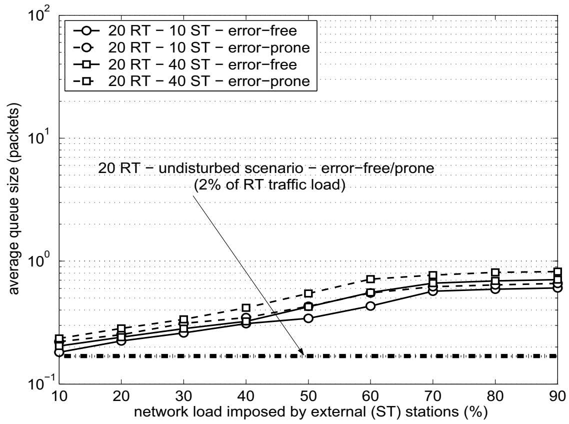

Figure 7. Average queue size (small and large pop.): error-free *vs.* error-prone - $MSP = 10ms$.

tude when the network load increases from 10% to 50%, whereas the average queue size for an error-prone channel is only slightly larger than the average queue size for an error-free channel.

We discuss now the case of RT stations with MSP of $10ms$. From Figure 7, it can be observed the impact of the timing unconstrained ST traffic upon the average number of packets waiting to be transmitted from 20 RT stations. Whatever the external network load, the average number of queued packets is always kept under 1 packet. This indicates that the EDCA mechanism can be suitable to support RT traffic, when the message stream periods are of $10ms$. Similar results were obtained for MSP of $20ms$, and in the case of 10 RT stations sending messages with MSP of $10ms$ or $20ms$. Therefore, those results can be generalized, stating that, in what concerns the average queue size of RT messages, the EDCA mechanism can be suitable to support RT traffic in disturbed communication environments.

5.4 The impact of uncontrolled external traffic upon the packet loss rate

A second simulation analysis concerns the assessment of the percentage of packet loss when the communication medium is shared with timing unconstrained traffic. Figures 8 and 9 show the packet loss results for message stream periods of $10ms$ and $20ms$, in both the small and large population scenarios.

These Figures show that, for a number of RT stations above 10, the impact of the timing unconstrained ST traffic upon the packet loss rate of RT messages becomes clearly undesirable. It is clear that the EDCA mechanism is not able to provide any acceptable RT guarantee for a number of RT stations above 10, as the percentage of packet losses is above 10%, when the network load imposed by external stations increases from just 10% to 30%.

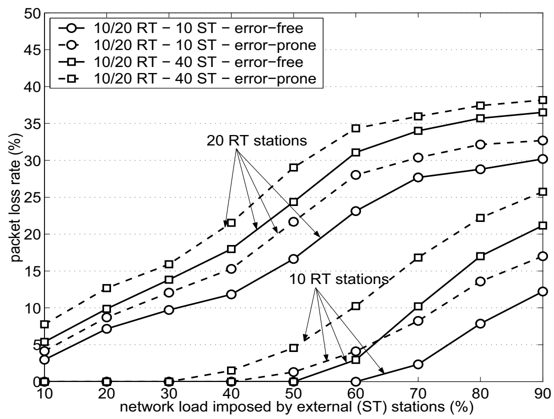

Figure 8. Packet loss rate (small and large pop.): error-free *vs.* error-prone - $MSP = 10ms$.

5.5 The impact of timing unconstrained ST traffic upon the packet delay of RT stations

The average packet delay for transferring a RT packet in small and large population scenarios is finally assessed. We have been previously concluded through the average queue size and packet loss rate analysis that the EDCA mechanism is not able to provide a RT communication service for more than 10 RT stations, even for MSPs of $10ms$ or $20ms$. Therefore, the average packet delay will be plotted only in the case of 10 RT stations with MSPs of $10ms$ and $20ms$, which are represented in Figures 10 and 11, respectively.

From this set of simulations (Figures 10 and 11), it can be concluded that the EDCA mechanism can be able to support the RT traffic with MSPs of $10ms$ or $20ms$ generated by up to 10 RT stations. The upper bound for ex-

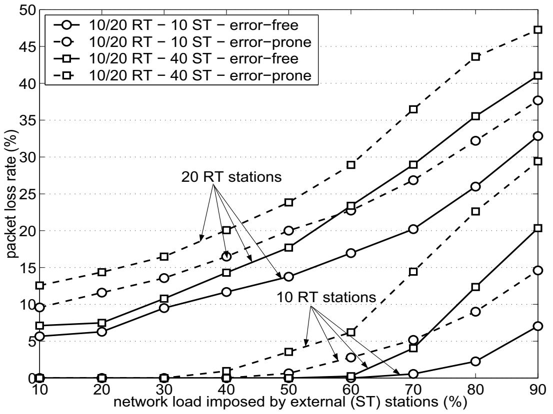

Figure 9. Packet loss rate (small and large pop.): error-free *vs.* **error-prone -** $MSP = 20ms$**.**

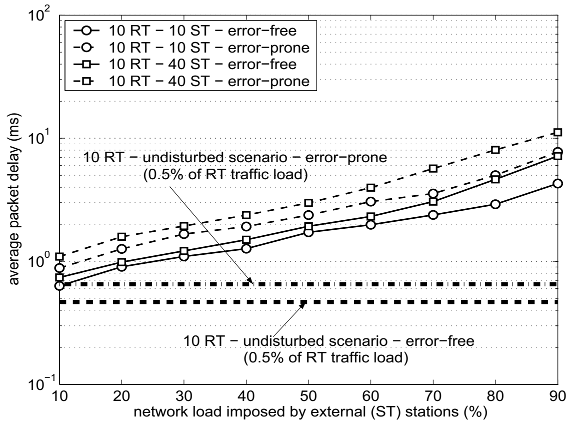

Figure 11. Average delay (small and large pop.): error-free *vs.* **error-prone -** $MSP = 20ms$**.**

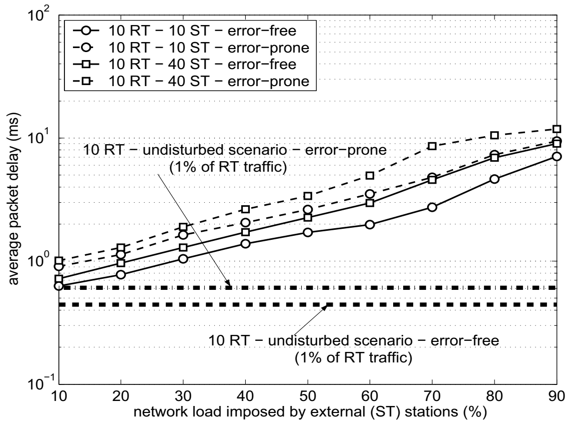

Figure 10. Average delay (small and large pop.): error-free *vs.* **error-prone -** $MSP = 10ms$**.**

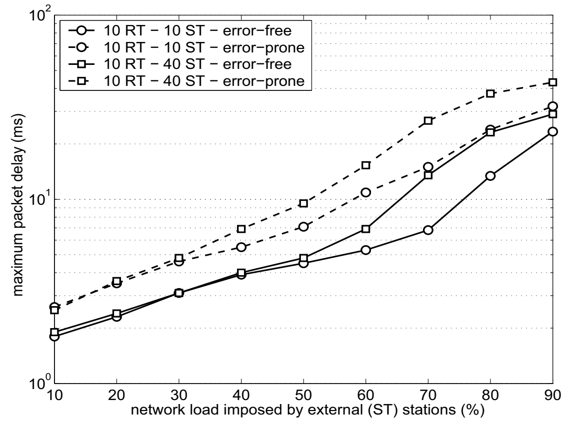

Figure 12. Maximum delay (small and large pop.) for $Z_{min} = 95\%$**: error-free** *vs.* **error-prone -** $MSP = 10ms$**.**

ternal disturbances may be as high as 70% or 90% in the case of message stream periods of $10ms$ or $20ms$, respectively. Additionally, and most interestingly, the effect of the error-prone channel becomes more relevant when the network load increases. However, when comparing the percentage of packet loss in error-prone *vs.* error-free channels, it can be observed that the difference is always close to 5%. Conversely, this difference is much larger when comparing the increase of the network load imposed by external ST stations. This means that the effect of collision upon the timing behavior of the RT communication is much more relevant than the effect of communication errors. Remark the level of communication errors has been set to a realistic value when dealing with wireless transmission in industrial environments [13].

Finally, as mentioned in Section III, an important measurement when dealing with support of RT communication is the maximum transmission delay. Therefore, we have also analyzed the maximum value of the packet delay, when considering 95% of the successful message transmissions (Figures 12 and 13). This value is relevant in the case of loss tolerant applications, where up to 5% of deadline misses can be encompassed. The previous simulation analysis (average packet delays) are only indicative of the average communication behavior, but may be valuable enough to provide an indication of the QoS that can be supported by the EDCA mechanism.

Considering that the deadline of a message stream is equal to its generation period, it can be concluded that EDCA mechanism is able to support MSPs of $10ms - 20ms$ only if the network load imposed by external stations is kept under about 50%-65%, respectively. For network load above these thresholds more than 5% of the deadlines will not be accomplished, which is usually un-

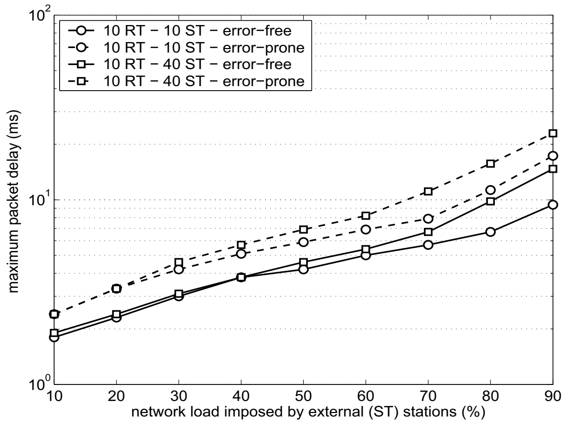

Figure 13. Maximum delay (small and large pop.) $Z_{min} = 95\%$: error-free $vs.$ error-prone - $MSP = 20ms$.

acceptable even for the typical loss-tolerant applications.

6 Summary and Future Directions

In this paper, we have analyzed the timing behavior of the EDCA mechanism when supporting RT communication in an error-prone channel. In summary, we concluded that the EDCA mechanism is not able to provide an acceptable QoS level for RT communication scenarios, unless the message stream periods are larger than $10ms$ and the external network load is kept smaller than 50%.

An interesting conclusion is that such misbehavior is mainly due to the influence of the external network load imposed by non-RT stations instead of the error-prone channel characteristics. This behavior can be confirmed in all results, where we note that when the network load imposed by external stations increases just from 10% to 50%, the EDCA analyzed metrics (average packet delay, average queue size and packet loss) significantly increases, while the difference between error-prone and error-free channels only slightly increases.

Traditionally, when supporting RT communication in CSMA wired networks, the timing behavior has been guaranteed through the tight control of every communicating device. The coexistence of RT controlled stations with non-RT stations has been made possible by constraining the traffic behavior of the latter. Unfortunately, this approach cannot be enforced in wireless environments, since any stations will share the access to the same radio channel. Therefore, it is not possible to impose any traffic smoothing strategy upon stations that are out of the sphere-of-control of the RT architecture.

Thus, we entirely agree with Bianchi et al. [9] that in wireless architectures, the service differentiation mechanism must be compulsory introduced as medium access control (MAC) layer extension. Furthermore, we argue that the performance analysis of the upcoming wireless solutions must always consider the impact of external traffic load upon the assessed communication scenarios. Otherwise, the simulation results will become useless.

Several approaches and techniques have been developed to provide RT behavior to IEEE 802.11 wireless-supported applications. However, few of those techniques allow timing unconstrained devices to coexist with RT stations, i.e., the majority of the proposed solutions only offer RT guarantees under the unrealistic assumption that is possible to have a closed communication environment.

Relevant exceptions are those based on the black-burst (BB) scheme [15]. This kind of solution consists in *forcing* the collision resolution in favor of the RT stations.

The BB is a distributed MAC scheme applied to *ad hoc* CSMA wireless networks. It requires the shut-down of the random retransmission scheme. Real-time stations implementing the BB scheme contend for the channel access after a medium interframe spacing t_{med}, rather than after a long interframe spacing used by standard stations implementing CSMA protocols. Thus, RT stations have priority over standard stations. When a RT station wants to transmit, it sorts its access rights by jamming the channel with energy pulses (BB's) immediately after the channel becomes idle during t_{med}. The length of the BB transmitted by a RT node is an increasing function of the contention delay experienced by the node.

In [16], it is presented a similar scheme, where voice nodes (RT stations) use energy-burst (EB) or black burst (BB) periods to prioritize RT packets over data packets. The access point (AP) can transmit its VoIP packets after PIFS without backoff or contending. On the other hand, each voice station has its own address (ID), referred as VID. The VID can be assigned at the traffic stream (TS) setup procedure provided in IEEE 802.11e. The station with the highest VID wins the contention and the VID is expressed as a binary value based on the fixed total bits which are determined by the voice packet resolution period (VPRP).

The main disadvantage of the BB scheme is that it compels the modification of the MAC layer and possibly also of parts of the PHY layer (*e.g.* radio ICs), which impairs the use od COTS hardware.

In [17], we have proposed a new RT-communication approach (VTP-CSMA) that is based on similar traffic separation mechanisms. Such mechanisms are able to prioritize RT-traffic over other traffic, without directly controlling the latter. The proposed architecture is based on a Virtual Token Passing procedure that circulates a virtual token among a number of RT devices. This virtual token is complemented by an underlying traffic separation mechanism that prioritizes the RT traffic over the non-RT traffic. Although VTP-CSMA mechanism also needs to modify parts of the MAC layer, it could be easily implemented in COTS hardware (*e.g.*, FPGA) upon standard 802.11e hardware.

For next generation communication environments, allowing the coexistence of both RT and non-RT stations in the same communication domain is likely to become a major requirement. Furthermore, solutions that enable the prioritization of RT traffic only by controlling every communicating device will be unacceptable.

Specifically in the case of wireless networks, the underlying wireless communication protocol must be able to guarantee the timing constraints of the RT traffic in a communication environment shared with timing unconstrained devices. Therefore, the most promising solutions to provide RT communication in IEEE 802.11 are those that force the collision resolution in favor of the RT stations, compelling all the other contending stations to temporally abandon the medium access.

More precisely, we consider that a 2-tier architecture (Figure 14) is an adequate architecture to support RT communication in wireless communication domains, providing that: (a) in the lower layer (MAC layer), a forcing collision resolution (FCR) mechanism enforces a high priority level to RT stations; (b) in the upper layer, a coordination mechanism ensures collision-free access among the set of RT stations. This coordination mechanism can be based on a token passing scheme as proposed in [17] or on a time division multiple access (TDMA) mechanism. It must serialize the medium access of only RT stations, as the FCR layer ensures that non-RT stations will always loose the medium contention when contending for the medium access with a RT-station.

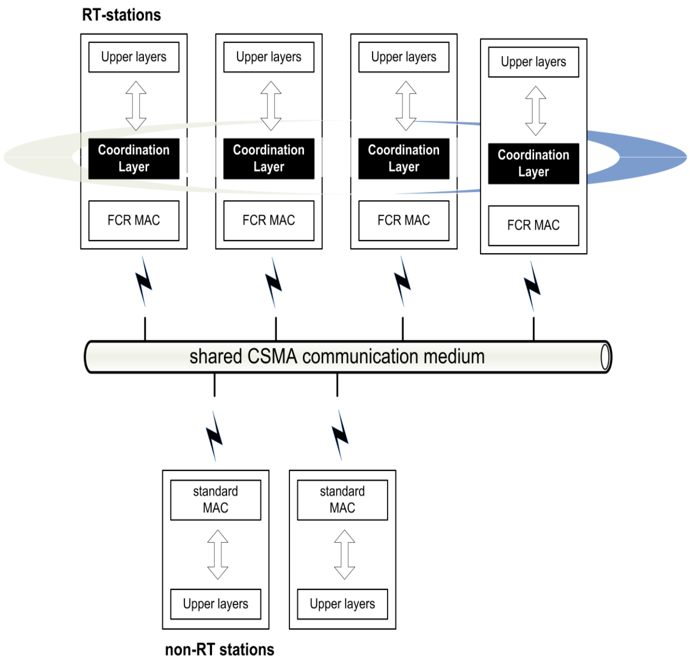

Figure 14. A 2-tier architecture to support real-time communication in wireless networks.

This contrasts with the traditional RT communication approaches that usually consider just the coordination layer among the set of RT stations. It is our opinion that one of the fundamental assumptions that must be considered when supporting RT communication is that the wireless physical medium is an open communication environment. Thus, state-of-the-art approaches that do not allow the coexistence of both RT and non-RT stations will not be able to handle next generation communication scenarios. Therefore, whatever the RT communication solution, it is mandatory to define a lower communication layer that enables the separation of the RT and the non-RT traffic in any shared communication environment.

7 Conclusion

We have studied the timing behavior of the EDCA function when supporting RT traffic, similar to those usually found in industrial environments. The simulation scenarios consider a set of RT stations operating in the same frequency band together with a set of ST stations. Basically, we have assessed the impact of timing unconstrained traffic (generated by ST stations) upon the behavior of the highest access category (voice), when this access category is used to transfer small sized packets in periodic intervals. Moreover, both error-free and error-prone channels are considered. The simulation analysis shows that:

- Both increasing the number of stations and the network load strongly influence the RT traffic, by means of higher packet delays, higher percentage of packet loss, higher average queue sizes;
- The packet loss rate for transferring RT packets (45 bytes) is highly influenced in an error-prone channel, mainly for high rates of network load imposed by external (ST) stations;

The main conclusion of the simulated scenarios is that default parameter values of the EDCA mechanism are just able to guarantee industrial communication (timing) requirements for a number of RT stations smaller than 10 stations with message stream periods above $10ms$. Therefore, new communication approaches must be devised in order to use IEEE 802.11e networks in the factory floor, where the RT traffic may be disturbed by timing unconstrained traffic from generic stations sharing the same communication medium.

References

[1] "IEEE Standard for Information Technology - Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements", IEEE Std 802.11e-2005, 2005.

[2] A. Willig, K. Matheus, and A. Wolisz, "Wireless technology in industrial networks", *Proceedings of the IEEE*, vol. 93, no. 6, pp. 1130 – 1151, 2005.

[3] D. Miorandi, E. Uhlemann, S. Vitturi, and A. Willig, "Guest Editorial: Special Section on Wireless Technologies in Factory and Industrial Automation, Part I", *IEEE*

Transactions on Industrial Informatics, vol. 3, no. 2, pp. 95–98, 2007.

[4] N. Ramos, D. Panigrahi, and S. Dey, "Quality of service provisioning in 802.11e networks: Challenges, approaches, and future directions", *IEEE Network*, vol. 19, no. 4, pp. 14 – 20, 2005.

[5] G. Cena, I. Bertolotti, A. Valenzano, and C. Zunino, "Evaluation of Response Times in Industrial WLANs", *IEEE Transactions on Industrial Informatics*, vol. 3, no. 3, pp. 191–201, Aug. 2007.

[6] C. Casetti, C.-F. Chiasserini, M. Fiore, and M. Garetto, "Notes on the Inefficiency on 802.11e HCCA", Dipartimento de Elettronica, Politecnico di Torino, 2005.

[7] P. Garg, R. Doshi, R. Greene, M. Baker, M. Malek, and X. Cheng, "Using IEEE 802.11e MAC for QoS over wireless", *Conference Proceedings of the 2003 IEEE International Performance, Computing, and Communications Conference*, , pp. 537 – 42, 2003.

[8] D. Miorandi and S. Vitturi, "Analysis of master-slave protocols for real-time industrial communications over IEEE 802.11 WLANs", *2nd IEEE International Conference on Industrial Informatics*, , pp. 143 – 148, 2004.

[9] G. Bianchi, I. Tinnirello, and L. Scalia, "Understanding 802.11e Contention-based Prioritization Mechanisms and their Coexistence with Legacy 802.11 Stations", *IEEE Network*, vol. 19, no. 4, pp. 28 – 34, 2005.

[10] M. Hamdaoui and P. Ramanathan, "A Dynamic Priority assignment technique for streams with (m,k)-firm deadlines", *IEEE Transactions on Computers*, vol. 44, no. 12, pp. 1443 – 51, 1995.

[11] B. Goode, "Voice over Internet Protocol (VoIP)", *Proceedings of the IEEE*, vol. 90, no. 9, pp. 1495 – 1517, 2002.

[12] R. Moraes, P. Portugal, and F. Vasques, "A Stochastic Petri Net Model for the Simulation Analysis of the IEEE 802.11e EDCA Communication Protocol", *10th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, , pp. 38 – 45, 2006.

[13] A. Willig and A. Wolisz, "Ring stability of the PROFIBUS token-passing protocol over error-prone links", *IEEE Transactions on Industrial Electronics*, vol. 48, no. 5, pp. 1025 – 1033, 2001.

[14] Q. Ni, "Performance analysis and enhancements for IEEE 802.11e wireless networks", *IEEE Network*, vol. 19, no. 4, pp. 21 – 27, 2005.

[15] J. Sobrinho and A. Krishnakumar, "Quality-of-service in ad hoc carrier sense multiple access wireless networks", *IEEE J. Sel. Areas Commun*, vol. 17, no. 8, pp. 1353 – 68, 1999.

[16] G.-H. Hwang and D.-H. Cho, "New access scheme for VoIP packets in IEEE 802.11e wireless LANs", *IEEE Communications Letters*, vol. 9, no. 7, pp. 667 – 669, 2005.

[17] R. Moraes, F. Vasques, P. Portugal, and J. A. Fonseca, "VTP-CSMA: A Virtual Token Passing Approach for Real-Time Communication in IEEE 802.11 Wireless Networks", *IEEE Transactions on Industrial Informatics*, vol. 3, no. 3, pp. 215–224, August 2007.