Differentiation mechanisms for heterogeneous traffic integration in IEEE 802.11 networks

Luca Scalia, Ilenia Tinnirello
Universita' di Palermo, Dipartimento di Ingegneria Elettrica
Viale delle Scienze, 90128, Palermo, Italy
luca.scalia@tti.unipa.it, ilenia.tinnirello@tti.unipa.it

Abstract

Recently a great amount of work in the research community has been focused on the support of multimedia traffic over wireless networks, which are known not to be specifically designed to support time sensitive applications. This study's primary objective is to analyze the behaviour of 802.11e differentiation mechanisms in heterogeneous traffic scenarios, when data and multimedia traffic sources share the access to the wireless medium. Secondarily, our aim is to provide some specific feedbacks in terms of MAC parameter settings in order to both ensure the QoS needs of multimedia stations and simultaneously optimize the overall system performance. To do that, an object oriented C++ simulator, which simulates the IEEE 802.11e MAC behaviour, has been developed. Simulation results show that it is possible to find specific MAC parameter configurations for different traffic patterns both to probabilistically provide delay guarantees for time-sensitive applications and simultaneously to maximize performance for backlogged (and non-prioritized) data traffic[1].

1. Introduction

Nowadays the IEEE 802.11 [1] represents the most exploited protocol for the deployment of Wireless LANs (WLANs). The reason is that the 802.11 Wireless technology is able to provide cheap, mobile and easy to deploy network connectivity. However, the growth in the demand of multimedia applications such as real-time voice, audio and video, has evidenced the limitations of the 802.11 protocol in terms of QoS support. Time-sensitive applications require strict QoS guarantees in terms of delay, jitter and bandwidth which basically the DCF 802.11 mechanism isn't able to provide. The 802.11 MAC in fact was not originally designed to support service differentiation. Nonetheless, the capability of supporting multimedia applications within 802.11 wireless environments will represent the key for the success of such technology in the next future. To overcome and then fill the lack of QoS support in the 802.11 MAC, the IEEE 802.11 working group has formed the task group E with the aim to introduce prioritized access mechanisms within the 802.11 context. The final version of the 802.11e standard [2] is expected for the end of the 2004/beginning of 2005. The 802.11e proposal introduces a new function called HCF (Hybrid Coordination Function) which basically extends the mandatory Distributed Coordination Function (DCF) and the optional Point Coordination Function (PCF) of the 802.11 MAC layer, while maintaining backward compatibility with the basic DCF/PCF access schemes. Based on them, the HCF respectively provides both a contention-based channel access mechanism and a controlled channel access one. The former is known as Enhanced Distributed Channel Access (EDCA) and is based on the introduction of unfairness in the channel access probability.

AC	CW_{min}	CW_{max}	$AIFSN$
AC_BK	aCW_{min}	aCW_{max}	7
AC_BE	aCW_{min}	aCW_{max}	3
AC_VI	$aCW_{min}/2$	aCW_{min}	2
AC_VO	$aCW_{min}/4$	$aCW_{min}/2$	2

Table 1. EDCA default settings

The latter is known as HCF Controlled Channel Access (HCCA) and enhances the 802.11 PCF mechanism. It provides a deterministic support of the QoS by opportunely scheduling the prioritized traffic needs of the requiring QoS stations in both the Contention Periods (CPs) and Contetion Free Periods (CFPs).

In this work we exclusively deal with the EDCA mechanism, referring to an heterogeneous data/real-time traffic scenario. We face the problem of resource reparti-

---

[1] The work reported in this paper was partially supported by the Italian Research Program PRIMO.

tion among stations with different requirements in terms of bandwidth and delays. We show that the best way to obtain service differentiation and contemporaneously maximize the overall system performance is represented by the joined utilization of the backoff and inter-frame spacing MAC parameters. The rest of the paper is organized as follows: in section 2 we discuss some significant research works about the 802.11e service differentiation mechanisms; in section 3 a brief overview of the EDCA mechanism is provided; in section 4 performance results about coexistence among high priority and low priority stations as well as EDCA MAC parameter tuning issues are discussed. Finally, in section 5 some conclusion are drawn.

2. Related Work

A large amount of work on service differentiation via distributed mechanisms has been carried out starting from 2000. Two different point of views have been considered for the analysis of the EDCA QoS support: the global perspective of the application layer and the partial perspective of the MAC layer itself. About the first issue, the most comprehensive overview of the various complementary as well as mutually alternative mechanisms for service differentiation in 802.11 appears to be in the technical report [5], which also addresses the performance evaluation of a subset of the considered mechanisms. Another comparative performance evaluation, although for a limited number of proposals can be found in [6]. Schemes based on the differentiation of one or more of the CW_{min}, CW_{max} and IFS parameters have been considered in [7] (which focus on differentiated CWs) and in [8, 9, 10] (which - additionally - focus on differentiated IFS). Specific work on performance evaluation of EDCA, modeled with all the details defined in the 802.11e draft standard, can be found in [11, 12, 13, 14] and in [15], which analyze the performance effectiveness of the 802.11e mechanisms in various traffic patterns. Finally, regarding the integration of the service differentiation mechanisms with transport layers, it is worth to mention [21], which tackles the problem of the reduction of the service differentiation due to the absence of priority in the TCP-ACKs. Pioneering work dealing with the usage of service differentiation mechanisms to provide coexistence of overlapping BSS can be found in [22, 23].

About the second issue, adaptive setting of CW to achieve better MAC layer performance is considered in [20] and [16], which proposes an adaptive algorithm to dynamically re-calculate the CW_{min} value accordingly with the specific traffic class and the changes of the network load. [17] also discusses the problem of supporting a distributed admission control rule on top of an enhanced version of DCF (they use different CWs), via the definition of a Virtual MAC algorithm, that passively monitors the radio channel and estimates the service levels available, plus a Virtual Source algorithm that used the above mentioned algorithm to adapt application parameters to the radio channel conditions. Work related to distributed and adaptive traffic scheduling can be found in [18, 19]. Both papers deal with distributed scheduling, which is a mean to provide service differentiation as well as overall network efficiency advantages.

We focus on this second aspect, i.e. on the MAC layer performance analysis for different EDCA parameters, by roughly modeling higher layers as simple MAC Service Data Unit (MSDU) sources. Previous papers faced this problem in the case of saturated traffic sources, i.e. assuming that all the stations are always in the contending state. Indeed, we considered a particular traffic scenario, in which saturated stations share the bandwidth with low-rate not saturated stations. As we detail in the next station, this scenario can model the interaction of data and real-time traffic sources.

3. Enhanced Distributed Channel Access

The IEEE 802.11 channel access protocol is fair in terms of access probability. This means that each station has the same probability to win the contention and, in long term, obtains the same number of channel access (i.e., packet transmissions). This feature is often referred as "throughput-fairness", where the throughput here is defined in terms of packets/s. Several mechanisms are currently debated in order to alter the throughput-fairness property of the DCF. The EDCA proposal of the IEEE 802.11e Task Group is devised to differentiate the channel access probability among different traffic sources. The mechanism is designed to manage 8 different traffic priorities. The packet priority represents the quality of service defined at higher layers. Packets arriving to the MAC (Mac Service Data Units MSDUs) are then mapped into four different access categories (ACs), which represent four different levels of service for the contention to the shared medium. Basically, each AC contends to the medium with the same rules of standard DCF, i.e., wait until the channel is idle for a given amount of silence time, and then access/retry following exponential backoff rules. The access probability differentiation is provided by means of different silence monitoring times (namely, Arbitration Inter-Frame Space AIFS) instead of the constant DIFS, and different values for the minimum/maximum contention windows for the backoff extractions. Then, each AC is specified by the values $AIFS[AC]$, $CW_{min}[AC]$, and $CW_{max}[AC]$. The $AIFS[AC]$ values differ each one for an integer number of backoff slots. In particular, $AIFS[AC] = AIFSN[AC] \cdot aSlotTime + aSIFSTime$, where $AIFSN[AC]$ is an integer greater than 1 for normal stations and greater than 0 for APs. Table

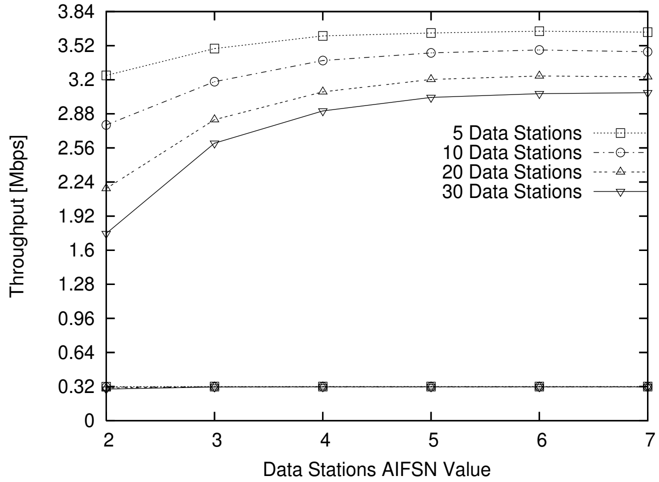

Figure 1. Throughput vs. different Data Stations AIFSN settings

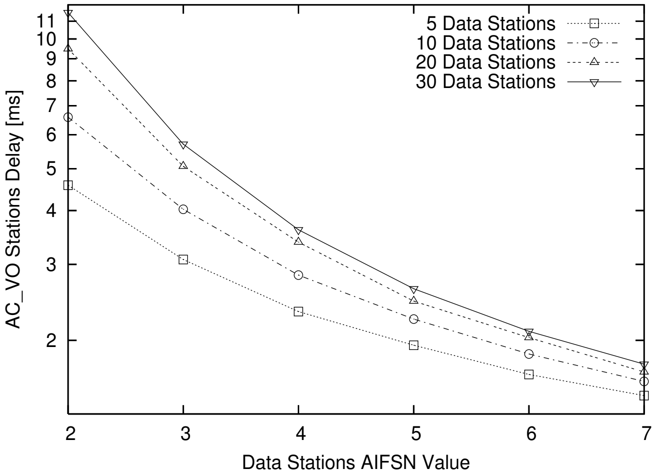

Figure 2. Real-Time stations Delay vs. different Data Stations AIFSN settings

1 shows the default values of the channel access parameters defined in EDCA for each AC. In each beacon frame, the AP broadcasts the values of these parameters chosen for each AC. In fact, they can be dynamically adapted according to the network conditions. Obviously, the smaller the $AIFSN[AC]$ and the $CW_{min}[AC]$, the higher is the probability to win the contention with the other ACs. Separate queues are maintained in each station for different ACs and each one behaves as a single enhanced DCF contending entity. When more than one AC of the same station expires its backoff counter, a virtual collision occurs and highest priority packet among the colliding ones is selected for actual transmission on the radio channel.

Once a station wins the contention and starts its transmission grant, EDCA also specifies new channel utilization operations based on the concept of transmission opportunity (TXOP), which represents a time interval in which the station is authorized to hold the channel. Since this mechanism can also be disenabled, in the following we do not consider specifically the TXOP feature.

4. Differentiation Mechanisms Analysis with Heterogeneous Sources

The EDCA access mechanism faces the problem of channel access differentiation in a completely distributed manner. This implies that no form of QoS guarantee can be provided. Here differentiated access priority means that a station is given a greater probability to gain a transmission opportunity respect to other stations contending the access to the shared channel. As we said before, the MAC parameter differentiations produces the loss of access fairness to the channel, and introduces differentiated bandwidth repartition and service delays among the contending stations. However, it is to be remarked that just differentiating the MAC payload size of the contending stations allows to easily manage the bandwidth repartition among the traffic classes, as it is shown in [16]. Nonetheless such mechanism does not account for the service delay differentiation, thus showing that further differentiation mechanisms have to be necessarily considered in order to meet the QoS needs of certain traffic patterns.

In this section we analyze the effectiveness of the proposed EDCA differentiation mechanisms in order to discover their capabilities to meet the bandwidth and delay requirements of heterogeneous applications. Our results are based on simulations. We developed an object-oriented event-driven C++ simulator, where we implemented all the described EDCA access features. Such simulator has been recently cross-validated with other 802.11 simulation programs developed on top of the NS-2 simulator [3] and analytical models [4] available in the scientific community. We have chosen a physical layer parameters configuration compliant with the 802.11b 11 Mbps operation mode. This assumption is not limitative, and the final considerations we have drawn are obviously applicable to other PHY configurations (802.11a and 802.11g).

4.1. Traffic Scenario

We assume that in each station a single application is running. This corresponds to have only one MAC layer queue and one AC for each station. Two different types of MSDU sources are considered: long packet saturated

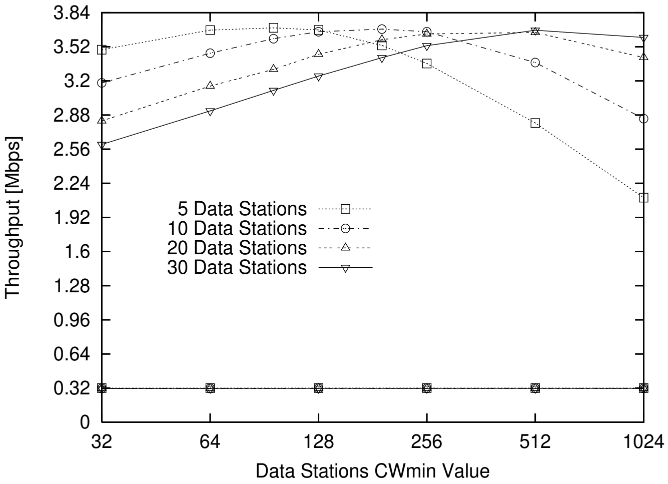

Figure 3. Throughput vs. different Data Stations CWmin settings

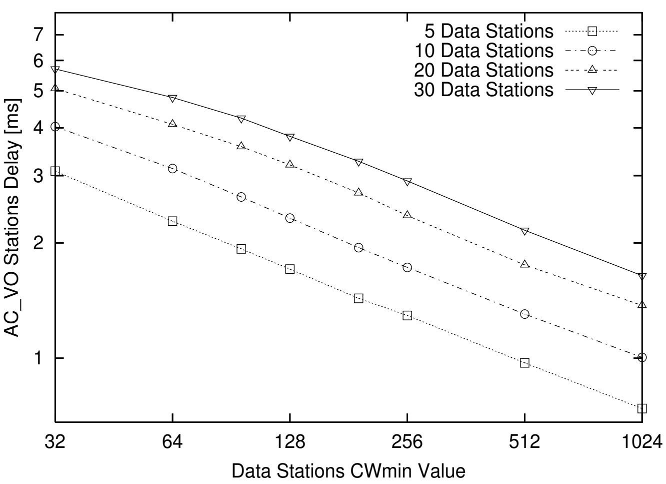

Figure 4. Real-Time stations Delay vs. different Data Stations CWmin settings

sources, (i.e. sources which make the MAC layer queue never empty) and short packet low-rate sources (i.e. sources with the packet arrival rate lower than the packet service rate). In the following, a station employing the saturated source is called data station, since this type of source corresponds to a data transfer application model. A station employing the low-rate source is called real-time station, since this type of source corresponds to a voice-like application model. Two difference performance figures result significant for the two different traffic sources. For the data application, the requirement is the minimization of the overall data transfer time and then the maximization of the granted bandwidth; for the real-time application, the requirement is a constraint on the maximum tolerable packet delivery delay. These two different requirements imply two different channel access Operations, which correspond to two different opportune EDCA parameter settings.

In order to give priority to delay sensitive traffic (real-time) over the data traffic, we exclusively consider $AIFS$-based and CW_{min}-based differentiation schemes. Specifically, we give higher inter frame times and contention window values to the data stations, thus allowing very quick contention resolutions for real-time active stations. We only focus on these two mechanisms because of their major effectiveness in providing service differentiation with respect to other differentiation mechanisms such as the CW_{max}, Persistence Factor (PF) (which has not been anymore considered since the version 5.0 of the 802.11e draft). In particular, it is intuitive to comprehend that CW_{max}-based differentiation mechanisms, together with PF-based ones, are able to provide an acceptable level of service differentiation only when the collision probability results very high [16].

In the following, we describe three different simulation scenarios, which respectively analyze the $AIFSN$-only based differentiation mechanism, the CW_{min}-only based differentiation mechanism and finally the joined utilization of both of the two differentiation schemes.

4.2. AIFSN-based differentiation

We consider a fixed number of 10 real-time stations. The real-time sources generate equal length short packets (80 bytes) at the rate of 32 Kbps. For this settings, at the data rate of 11 Mbps, the system capacity is much higher than the real-time offered load and, even for very short $[CW_{min}, CW_{max}]$ ranges, collisions are very rare. Thus, we assume that the real-time traffic is mapped to the default AC_VO access parameters (see table 1). A varying number of N data stations, employing the maximum 11 Mbps data rate, share the medium with the real-time stations. The data sources generate equal length long packets (1500 bytes) and work in saturation conditions. Data traffic is mapped to the AC_BE access category. The corresponding CW_min value is set to 32, while the $AIFSN$ settings is considered as a simulation parameter.

Fig. 1 reports the throughput of AC_VO stations (curves in the bottom of the figure) and the throughput of data stations (the curves on top of the figure). The curves are all parameterized according to the number of data stations. From the figure, we notice that the AC_VO MAC parameters configuration for delay-sensitive stations allows for them to drain all the offered traffic, independently from the number of contending data stations. Also we notice that the throughput curves of data stations are almost parallel, and their behaviour is substantially constant starting from

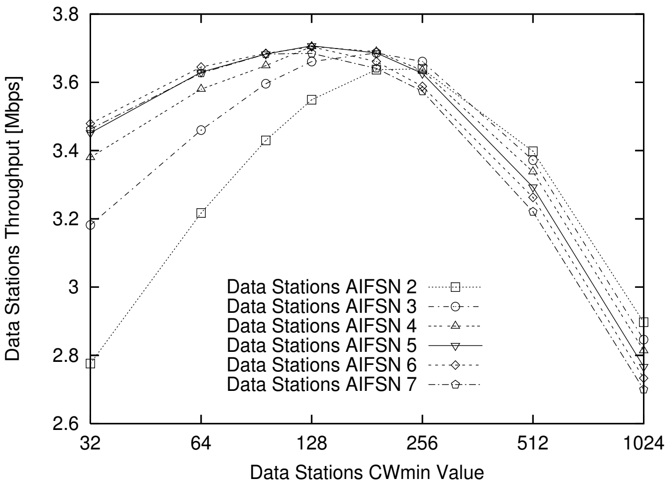

Figure 5. Impact of CWmin on data performance, in the case of 10 data stations, for different AIFSN values

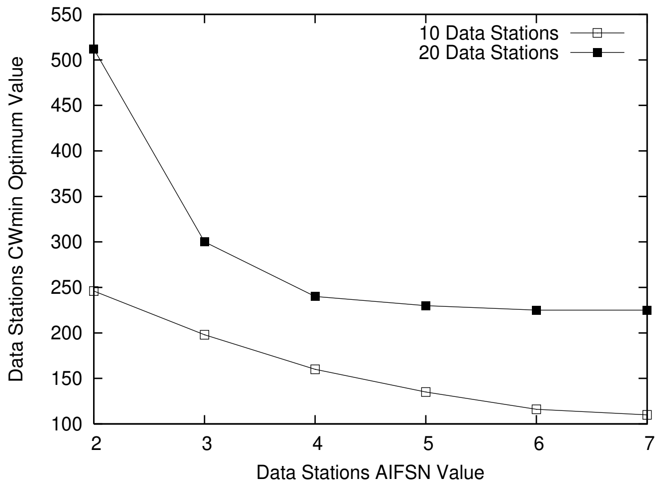

Figure 6. Optimum CWmin values for Data Station throughput maximumization

$AIFSN$ values greater than 3. Looking at the AC_VO delay performance in Fig. 2, we see that just one slot of differentiation (i.e. $AIFSN$ equal to 3), drastically reduces the delays to few milliseconds. Finally, we notice how, as the $AIFSN$ differentiation increases, such curves tends to be more and more clustered towards a limit value.

We can conclude that the AC_BE $AIFSN$ setting can be a control parameter for the real-time delays. A single time slot of difference in the inter-frame spaces is sufficient to reduce such delays under 10 ms in all the considered load cases and, as the $AIFSN$ setting of the data stations increases, delay performance is less and less dependent on the data congestion state. Moreover, the $AIFSN$ increment has no remarkable effect on the data throughput.

4.3. CWmin-based differentiation

We consider the same network scenario described in the previous section (i.e., 10 real-time stations which share the medium with N data stations) and the same source settings (i.e., 80 bytes packets at 32 kbps and 1500 bytes packets in saturation). In this case, we assume that the AC_BE $AIFSN$ setting is fixed to 3 and

we vary the CW_{min} values. Fig. 3 shows that, for each considered CW_{min} value, the throughput of the real-time traffic results equal to the offered load. From Fig. 4 we see that also the CW_{min} parameter works as a control parameter on the AC_VO delays. In fact, as the CW_{min} grows, the delivery delay for real-time stations decreases as a result of the reduction of the channel access probability experienced by data stations. Note that, the number of data stations affects the real-time delay performance even in the case of very large CW_{min} setting. What substantially differentiates the CW_{min}-based control from the $AIFSN - based$ control, is the effect on the data throughput. In fact, from Fig. 3 we also see that data throughput is strongly affected by the CW_{min} setting and, in particular, it results maximized (almost 3.65 Mbps) for a given CW_{min} value, which depends on the number of competing stations N. This implies that, according to the specific real-time delay requirements, the data stations throughput can be significantly degraded. For example, in the case of 5 data stations, a delay constraint of almost 1 ms, corresponds to an AC_BE CW_{min} value equal to 512, which in turns reduced to 2.7 Mbps the data throughput.

4.4. Joined AIFSN/CWmin differentiation

Previous considerations allow to conclude that the best possible AC_BE settings should be based on an opportune tuning of both the parameters $AIFSN$ and CW_{min}. If the $AIFSN$ setting satisfies the real-time delay requirement for the case $CW_{min} = 32$, it is clear that higher values of CW_{min} can only improve the real-time performance (since the data station access probability is further reduced). Thus, the CW_{min} value can be considered as an optimization parameter for the data throughput. Fig. 5 shows the data throughput performance, in the case of 10 data stations and 10 real-time stations, as the CW_{min} value changes, for different $AIFSN$ values. In the figure it is evident that the optimum AC_BE CW_{min} value results function of the AC_BE $AIFSN$ value. For example, in the case $AIFSN = 2$ the optimum corresponds to $CW_{min} = 256$,

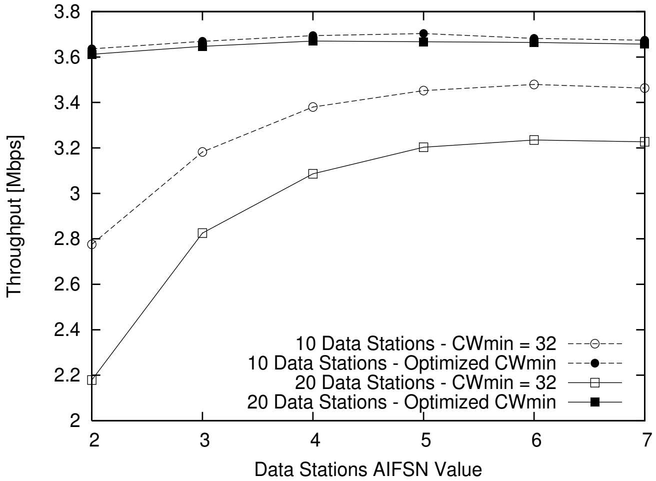

Figure 7. Data Stations throughput vs. Data Stations AIFSN settings with and without optimized CWmin values

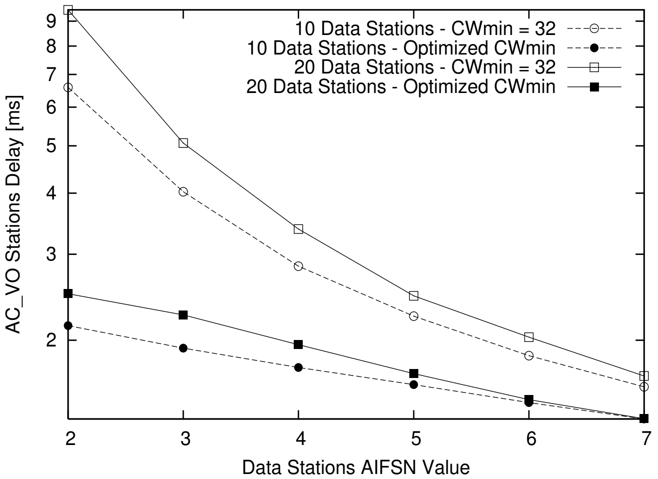

Figure 8. Real-Time Stations delay vs. Data Stations AIFSN settings with and without optimized CWmin values

while in the case $AIFS = 7$ the optimum corresponds to $CW_{min} = 128$. This phenomenon is due to the fact that, as the $AIFSN$ increases, the collision number among data and real-time packets gets lower and lower, since most part of real-time transmissions are originated before the AC_BE $AIFS$ expiration. Since the throughput optimum is given by the tradeoff among the minimization of the data collision probability and the resource wastes due to collisions and idle slots, it is clear that as the real-time stations reduce their impact on the data collision probability, such tradeoff is reached for lower CW_{min} values.

Fig. 6 shows the AC_BE CW_{min} optimum values as a function of the AC_BE $AIFSN$ settings, for the two cases $N = 10$ and $N = 20$. The optimum values are more accurate than in the previous figures, since we considered shorter CW_{min} increments to find the optimum. Clearly, the optimum value is also function of the number of data competing stations N.

Finally, Fig. 7 and 8 shows, respectively, the data throughput and the real-time delays, for the scenarios $N = 10$ and $N = 20$, as a function of the AC_BE $AIFSN$ value, in the two cases of $CW_{min} = 32$ and optimum CW_{min} settings. We can see from these figures, that the joined tuning of these two AC_BE parameters allows to optimize the performance for both data and real-time stations.

5. Conclusions

In this paper we have analyzed the influence of the $AIFS$ and CW_{min} parameters settings (introduced by the 802.11e proposal) in the case of heterogeneous traffic sources, focusing on the possibility to cope with the optimization of network resources. From simulation results we have concluded that a joined utilization of the above differentiation 802.11e parameters results in a optimal and effective solution for optimizing the overall system performance, thus allowing to meet the voice delay constraints while maximizing the throughput for data stations.

References

[1] IEEE 802.11 WG, IEEE Std 802.11, 1999 edition. International standard for Information Technology. Telecommunications and information exchange between systems - Local and metropolitan area networks. Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, 1999.

[2] IEEE 802.11 WG, IEEE Std 802.11e/D8.0, Draft Supplement to IEEE standard for Telecommunications and Information exchange between systems. Local and metropolitan area networks. Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, February 2004.

[3] Sven Wiertholter, Christian Hoene, "Design and Verification of an IEEE 802.11e EDCF simulation model in ns-2.26", November 2003

[4] G. Bianchi, I. Tinnirello, "Analysis of Priority Mechanisms based on Differentiated Inter-Frame Spaces in

CSMA/CA", in Proc. IEEE VTC 2003, Orlando (FL), October 2003

[5] Qiang Ni, Lamia Rhomdani, Thierry Turletti and Imad Aad, QoS Issues and Enhancements for IEEE 802.11 WLAN, INRIA technical report RR4612, November 2002

[6] Anders Lindgren, Andreas Almquist, Olov Schelen, Evaluation of Quality of Service Schemes for IEEE 802.11 Wireless LAN, LCN 2001,

[7] A. Veres, Campbell, A.T, Barry, M. and L-H. Sun, Supporting Service Differentiation in Wireless Packet Networks Using Distributed Control, IEEE Journal of Selected Areas in Communications (JSAC), Special Issue on Mobility and Resource Manegement in Next-Generation Wireless Systems, Vol. 19, No. 10, pp. 2094-2104, October 2001.

[8] I. Aad, C. Castelluccia, Differentiation mechanisms for IEEE 802.11, Proc. of IEEE INFOCOM 2001, pp. 209-218.

[9] I. Aad, C. Castelluccia, Priorities in WLANs, Computer Networks (Elsevier), vol. 41, 2003, pp. 505-526.

[10] X. Pallot, L. E. Miller, Implementing message priority policies over an 802.11 based mobile ad hoc network, IEEE Military Communications Conference (MILCOM), 2001, Vol. 2, pp. 860-864.

[11] S. Mangold, S. Choi, P. May, O. Klein, G. Hiertz, L. Stibor, IEEE 802.11e Wireless LAN for Quality of Service, In Proceedings of the European Wireless, Vol. 1, pp. 32-39, Florence, Italy, February 2002.

[12] Sunghyun Choi, IEEE 802.11e MAC-Level FEC Performance Evaluation and Enhancement, in Proc. IEEE GLOBECOM 02, Taipei, Taiwan, November 17-21 2002.

[13] Sunghyun Choi, Javier del Prado, Stefan Mangold, and Sai Shankar, IEEE 802.11e Contention-Based Channel Access (EDCF) Performance Evaluation, Proc. IEEE ICC 03, Anchorage, Alaska, USA, May 2003.

[14] H.L Truong, G.Vannuccini, The IEEE 802.11e MAC for Quality of Service in Wireless LANs, SSGRR 2003, L'Aquila, January 2003

[15] H. Dajiang, Q. Charles, Q. Shen, Simulation Study of IEEE 802.11e EDCF

[16] V.A. Siris, M. Kavouridou, Achieving Service Differentiation and High Utilization in IEEE 802.11

[17] M. Barry, A. T. Campbell, A. Veres, Distributed Control Algorithms for Service Differentiation in wireless Packet Networks, Proc. of IEEE INFOCOM 2001, pp. 582-590

[18] V. Kanodia, C. Li, A. Sabharwal, B. Sadeghi, and E. Knightly, Distributed Priority Scheduling and Medium Access in Ad Hoc Networks, ACM Wireless Networks Journal (WINET), September 2002

[19] N. H. Vaidya, P. Bahl, and S. Gupta, Distributed Fair Scheduling in a Wireless LAN, Proceedings of ACM International Conference on Mobile Computing and Networking (MobiCom 2000), Boston, Massachusetts, USA, August 2000

[20] Lamia Rhomdani, Qiang N, Thierry Turletti, AEDCF: Enhanced Service Differentiation for IEEE 802.11 Wireless Ad-Hoc Networks, INRIA technical report RR4544,September 2002

[21] I. Aad, C. Castelluccia, Remarks on Per-Flow Differentiation in IEEE 802.11. In Proceedings of European Wireless 2002

[22] S.Mangold, "IEEE 802.11e: Coexistence of Overlapping Basic Service Sets", In Proceedings of the Mobile Venue 02, pp. 131-135, Athens, Greece, May 2002.

[23] S. Mangold, S. Choi, P. May, G. Hiertz, "IEEE 802.11e - Tair Scheduling between Overlapping Basic Service Sets", Proc. of IEEE PIMRC 2002, pp.166-171.