Improving the QoS of IEEE 802.11e networks through imprecise computation

Anna Lina Ruscelli* and Gabriele Cecchetti

TeCIP Institute,
Scuola Superiore Sant'Anna,
Via G. Moruzzi 1, Pisa 56124, Italy
Email: a.ruscelli@sssup.it
Email: g.cecchetti@sssup.it
*Corresponding author

Abstract: IEEE 802.11e HCCA reference scheduler is based on fixed value parameters that do not adapt to traffic changes, thus quality of service (QoS) for multimedia applications is a challenge, especially in the case of variable bit rate (VBR) streams, that requires dynamic resource assignment. This paper is focused on *immediate dynamic TXOP HCCA* (IDTH) scheduling algorithm and its new evolution *immediate dynamic TXOP HCCA plus* (IDTH+). Their reclaiming mechanisms, refined by the monitoring of transmission duration, aim at overcoming the limits of fixed preallocation of resources by varying the stations transmission time and avoiding waste of resources. Simulations and theoretical analysis based on the imprecise computation model show that the integration of IDTH and IDTH+ can achieve improved network performance in terms of transmission queues length, mean access delay and packets drop rate, and to efficiently manage bursty traffic. Moreover, the performance improvements of IDTH+ with respect to IDTH are highlighted.

Keywords: QoS; quality of service; scheduling algorithms; WLAN; wireless local area networks.

Reference to this paper should be made as follows: Ruscelli, A.L. and Cecchetti, G. (2016) 'Improving the QoS of IEEE 802.11e networks through imprecise computation', *Int. J. Ad Hoc and Ubiquitous Computing*, Vol. 23, Nos. 3/4, pp.152–167.

Biographical notes: Anna Lina Ruscelli is a Research Fellow at Institute of Communication, Information and Perception Technologies (TeCIP) of Scuola Superiore Sant'Anna, Pisa. Her research interests are in quality of service support over heterogeneous networks, IEEE 802.11e networks, scheduling algorithms for resource reservation, ERMT-ETCS railway systems, railways signalling systems and communications. Her works have been published in international journals and in proceeding of international conferences. She has been a Lecturer for the Courses of Real-Time Operation Systems, Theory and Techniques for Digital Systems, Computer Architecture and Contract Professor of Computer Architecture course at University of Siena.

Gabriele Cecchetti is a Researcher at Institute of Communication, Information and Perception Technologies (TeCIP) of Scuola Superiore Sant'Anna, Pisa. His current research activities span several areas, including railways signalling systems and communications, design and performance evaluation of scheduling algorithms for QoS provisioning in wireless networks with soft real-time constraints, architectures for QoS in heterogeneous wireless networks, and wireless link emulation. His works have been published in international journals and in proceeding of international conferences. He has been a Professor of Computer Architecture, and Theory and Techniques for Digital Systems at University of Siena.

This paper is a revised and expanded version of a paper entitled 'Providing variable TXOP for IEEE 802.11e HCCA real-time networks' presented at *2012 IEEE Wireless Communications and Networking Conference (WCNC 2012)*, Paris, 1–4 April, 2012.

1 Introduction

Quality of service (QoS) support over IEEE 802.11e networks, part of the IEEE802.11 (2007) standard devoted to the QoS provisioning, is a challenge owing to the increasing number of users, the spreading diffusion and the coexistence of different multimedia applications. Dealing with multimedia traffic generated by applications such as high digital TV, video chatting and conferences, VoIP, high-definition multiplayer video games etc., implies the provisioning of a tailored service differentiation that impacts on resource management. However IEEE 802.11e *hybrid coordination function controlled channel access* (HCCA) *reference* scheduler, suggested by the standard, deals on a fixed and guaranteed preallocation of capacity during the admission control phase, and *transmission opportunity*

($TXOP$) transmission time and *service interval* (SI) polling period are computed when the stations initially negotiate their service with the *QoS access point* (QAP) and resources are reserved. Thus, the value of these protocol parameters, used during the subsequent polling phases, is based only on the initial stations conditions, and will be changed only after a further activation of the admission control when a new *traffic stream* (TS) is accepted. Moreover, their computation considers restrictive conditions to meet the most strict requirements from those of all admitted stations and that of the more exigent TS. As a result of this aggregation of information, all different TSs are polled with the same period and those of a QSTA are served with the same transmission time, without any type of dynamic allocation of resources to reflect traffic variability. Thus, despite the service differentiation with a parameterised QoS of the reference scheduler, the granularity of the provided service is rough with respect to the heterogeneity of the TSs of a QSTA. Consequently it results tailored for *constant bit rate* (CBR) streams, whereas *variable bit rate* (VBR) ones put a strain on its performance owing to the variability of their profile, as highlighted by numerous studies (Cowling and Selvakennedy, 2004; Grilo and Nunes, 2002; Mangold et al., 2002; Tsao, 2000) on QoS performance and also on *Quality of Experience* (QoE) Pastrav et al. (2012). Moreover the issue of ensuring QoS to multimedia applications limiting the interference between different IEEE 802.11e systems through spectrum sharing techniques has been addressed by some studies, such as Siddique et al. (2010).

Many scheduling algorithms have been proposed to improve the provided QoS and to introduce a dynamic resource assignment, see Cecchetti and Ruscelli (2011), Cecchetti et al. (2012b), Siris and Courcoubetis (2006) and Skyrianoglou et al. (2006) and the references therein. Different methods are used to make variable the protocol parameters, such as queues length model (Ansel et al., 2006) or feedback mechanism (Boggia et al., 2007)), the combination of HCCA and *enhanced distributed channel access* (EDCA), the other IEEE 801.11e function, to increase the amount of available resources (Lai et al., 2009; Ruscelli et al., 2012; Siris and Courcoubetis, 2006). The concept of deadline, that sets the time until the transmission has to be finished, is suitable to consider temporal requirements (Cecchetti and Ruscelli, 2011; Cicconetti et al., 2007a; Fan and Huang, 2005; Grilo et al., 2003; Inanc et al., 2006; Skyrianoglou et al., 2006), whereas bandwidth reclaiming recovers unused resources to reduce the delay (Cecchetti et al., 2012a; Larcheri and LoCigno, 2006; LoCigno et al., 2007; Ruscelli et al., 2011). However, many of the proposed algorithms are underperforming in the case of VBR TSs as investigated in Cecchetti and Ruscelli (2011), for instance schedulers based on *earliest deadline first* (EDF) algorithm (Liu and Layland, 1973) owing to its postponing deadlines mechanism, leaving space to refined enhancements.

In this paper, differently than trying to improve the provided service through a new algorithm, the proposed solution, based on imprecise computation model introduced by Lin et al. (1987), Liu et al. (1987, 1991), suggests that the service offered by an existent algorithm can be refined by integrating a ‘patch’ suitable to add what is missing. In particular, since one of the drawbacks of many HCCA schedulers in the case of VBR TSs is the lack of flexibility owing to static preallocation of transmission time, the idea is integrating an additional scheduling module providing a dynamic resource allocation based on instantaneous needs of stations and avoiding waste of resources. The imprecise computation provides the theoretical model to formalise this integration: a huge scheduling activity, such as providing the negotiated QoS, can be split into a *mandatory* part, that deals with the necessary computation with strict requirements, carried out by the centralised scheduler, and an *optional* part that can tolerate an adjustable service. Their combination generates a global scheduler suitable to provide an ‘advanced’ service. In this context *immediate dynamic TXOP HCCA* (IDTH), a QoS scheduling algorithm with a dynamic mechanism for resource reclaiming and estimation of transmission duration recently presented by Cecchetti et al. (2012b), is deeply investigated extending and integrating the previous results. Moreover, a new algorithm, *immediate dynamic TXOP HCCA plus* (IDTH+) is presented. Inspired to the concept of statistical multiplexing, IDTH and IDTH+ can be considered as the optional scheduling part. They aim at filling in the gaps of algorithms based on deterministic and fixed capacity preallocation, such as the reference one, and at improving the instantaneous behaviour of dynamic schedulers when data rate varies, especially in the case of VBR traffic with QoS requirements. Thus, the global scheduler obtained through their integration can be considered as an advanced QoS scheduler. More precisely, the centralised scheduler continues assigning $TXOP$, whereas IDTH and IDTH+ use the resources residue from the sending session of a QSTA, when data rate drops down and its $TXOP$ is not exhausted, to instantly increase the transmission time of the subsequent polled station, deducing its instantaneous needs by estimating the transmission time used during its previous polling. These additional resources are useful to absorb traffic variability in the presence of bursts and when $TXOP$ assigned during the admission control is not sufficient to dispatch enqueued packets.

IDTH+ aims at limiting the hysteresis of the reaction of IDTH when data rate decreases with respect to its mean value and to guarantee a minimum resource threshold to cope with the changes of data rate, especially in the case of a sudden increase after a continuous decreasing. It will be shown that it is suitable to further improve the performance obtained by IDTH, without any additional computational load.

Theoretical analysis focused on computational complexity, real-time scheduling (time loss and scheduling error), impact on queues lengths and service rate when VBR and bursty traffic are scheduled, deeply illustrates the usefulness of IDTH and IDTH+, with the benefit of imprecise computation approach but without its drawbacks. These considerations will be corroborated by simulation results using VBR and CBR TSs to take into account different traffic categories, showing that the analysed algorithms are able to improve the performance of the centralised schedulers in terms of transmission queues length, mean access delay and packets drop rate.

In the following a summary of existing QoS HCCA scheduling algorithms is presented in Section 2, IDTH and IDTH+ are described in Section 3, and are deeply investigated analytically in Section 4, and by simulation in Section 5. Final considerations are drawn in Section 6.

2 Related works

In this section a summary of some HCCA scheduling algorithms, representative of different techniques used to ensure QoS, is presented along with the description of the basics of the HCCA *reference* scheduler. Their properties and issues are compared to highlight their pros and cons and to provide, at the best of our knowledge, a panorama of different strategies for making variable the computation of scheduler parameters with particular attention to $TXOP$.

2.1 HCCA reference scheduler

IEEE 802.11e standard proposes an HCCA *reference* scheduler as a guideline for the design of MAC scheduling algorithms. It sets a unique SI for all TSs, respecting their transmission period constraints and guaranteeing at least one polling for each station during the beacon interval. Instead $TXOP$ of each QSTA$_i$ is computed as the maximum time to transmit at the minimum physical rate Γ_i the total amount of bits that can arrive during SI:

$$TXOP_i = \max\left(\frac{N_i \cdot L_i}{\Gamma_i}, \frac{M_i}{\Gamma_i}\right) + O,$$

where M_i is the maximum *MAC service data unit* (MSDU) size (2304 bytes), L_i is its nominal size, R_i is the mean data rate, *O* is the transmission overhead owing to interframe spaces and to the sending of `ACK` and `CF-Poll` frames, and $N_i = \left\lceil \frac{SI \cdot R_i}{L_i} \right\rceil$.

Finally, the admission control test used for deciding whether to admit a new stream, taking into account the available resource threshold, is:

$$\frac{TXOP_{k+1}}{SI} + \sum_{i=0}^{k} \frac{TXOP_i}{SI} \leq \frac{T - T_{CP}}{T} \leq 1. \quad (1)$$

where k is the number of admitted streams, $k+1$ indexes the newly admitted one, T is the beacon interval and T_{CP} is the EDCA duration.

2.2 HCCA scheduling algorithms

Since the reference scheduler considers *worst case* conditions for the computation of SI and, then, of $TXOP$, the admission control test is conservative admitting less TSs than possible and the efficiency of the resource management is limited. Moreover, as summarised in Section 1, numerous studies, such as Cowling and Selvakennedy (2004), Grilo and Nunes (2002), Mangold et al. (2002), Tsao (2000) and Pastrav et al. (2012), shown that the reference scheduler is suitable to serve only CBR traffic. Instead it is unable to efficiently manage VBR TSs, since all different TSs are polled with the same period and those of a station are served with the same computation time without service differentiation. Hence new algorithms proposed possible QoS improvement, see for instance Cecchetti and Ruscelli (2011), Fattah and Leung (2002), Grilo et al. (2001), Lu et al. (1999), Tsao (2000), Lu et al. (1999). In the following, at the best of our knowledge, we summarise some significant QoS schedulers based on different techniques that have the common goal of making variable $TXOP$ and SI to reflect traffic variability. This can be performed, for instance, acting on the parameters used for their computation, improving their estimation or modelling, tuning or correct the computation taking into account expected and actual performance.

One of the meaningful parameters impacting on the network performance is the transmission queues length. Indeed the number of enqueued packets and the time they wait to be transmitted directly affect end-to-end delay and packets drops rate, furthermore it is dependent from transmission interval duration and, at its turn, it can be considered when $TXOP$ is set. As a consequence the information about the length of queues can play an important role in the QoS assurance; it can be obtained using different methods and can be used in divers approaches.

Fair HCF (FHCF), proposed in Ansel et al. (2006), and composed by two scheduling components, computes a variable $TXOP$ for each station by means of a mathematical model of the uplink TSs queues length. This model estimates the global packet delay, distinguishing between the packet queuing delay, and the waiting time delay until the next polling interval. This estimation is used by the centralised scheduler of FHCF to provide an initial value of $TXOP$ which is refined by the local scheduler of FHCF that runs in each node and that can exploit the actual value of the corresponding transmission queue length. The drawback of this articulated calculation is the computational overload imposed to the system.

Instead of modelling the queues, scheduler parameters can be tuned acting through different techniques, like feedback control and/or collecting the available information on queues. For instance, *feedback based dynamic scheduler* (FBDS) (Boggia et al., 2007) aims to bound the maximum delay, whereas $TXOPs$ are dynamically calculated considering the initial queues length and corrected by a proportional controller that uses the corresponding actual value sent by each station to the QAP. $TXOP$ is assigned a value suitable to deliver, at the mean data rate, the packets arrived during a CAP. SI, computed during the admission control, remains unchanged. With respect to Ansel et al. (2006), this computation uses a feedback information about the enqueued packets instead of trying to forecast the queue length by means of a mathematical model. However the computation is yet limited to mean value parameters.

Another method based on knowing the exact queues length is presented in Al-Maqri et al. (2013), where $TXOP$ is dynamically computed at each polling taking into account the length of the next MSDU received from the application layer; this feedback information is included in the MAC header *QS size* field of uplink data packets sent to the QAP by each station. On the basis of this information the mean size of MSDU, that is part of $TSPEC$ is updated and, applying the same $TXOP$ computation of the reference scheduler, the transmission duration is modified. With respect to Boggia et al. (2007), it assumes that only one packet is generated by the station during every SI and that the scheduler will try to dispatch assigning a tailored $TXOP$. This solution does not consider the presence of more enqueued packets but only the information about the next video frame. These two last approaches are on the opposite fronts: the first one considers the enqueued packets waiting to be transmitted and residue from the previous polling, whereas the second one is focused exclusively on the next packet that will be generated. The exact value of transmission queues length is used also by Lee et al. (2011) in the *explicit traffic aware scheduling with explicit Length notification* (ETA-EQN) to compute $TXOP$ with the difference that this information is sent during the EDCA function to save HCCA resources. To provide a transmission duration suitable to empty saturated transmission queues and reduce the transmission delay the standard parameter *TXOPLimit*, that sets the upper bound of $TXOP$, is dynamically increased.

Ju and Chung (2013) propose a dynamic computation of $TXOP$ considering an estimation of the next packets to be transmitted instead of the queues length. It acts on the transmission time computation by updating the N_i standard parameter, taking into account the number of burst packets of a multimedia application and introducing a *variation* factor V that resumes the fluctuations of the actual data rate with respect to the mean data rate. As a consequence $TXOP$ is assigned a value variable between a maximum and a minimum in dependency of the burst size and of data rate variations. However this algorithm uses a theoretical estimation without any validation based on real traffic profile.

A completely different method is presented in Arora et al. (2010), where the link adaptation is used to vary the $TXOP$ duration in dependency of the channel state and on the physical rate of the stations: the transmission time is decreased when bad channel conditions could penalise transmission rate and increased when the state improves. The transmission rate is monitored and a lead-leg counter ensures long-term and short-term fairness to the stations.

In Saheb et al. (2012) the estimation of queues length and data rate is based on the consideration of the *tolerable loss limit* (TLL) added to $TSPEC$, that allows to take into account the packet loss. In particular, starting from the computation of the *bit error rate*, the queue length at the end of $TXOP$ is estimated along with the packets loss. Considering the history of these parameters the data rate for the next polling is computed and $TXOP$ estimated along with the expected delay to reach the destination. Finally the packets with the minimum difference between the estimated and the actual delay are scheduled first. This algorithm differentiates from the previous ones from two different points of view: first of all the packet loss is directly involved in the $TXOP$ computation, secondly a polling reorder is dynamically modulated.

A completely different method consists in crossing the frontiers between HCCA and EDCA, exploiting the last one to speed up the emptying queues process. When resources allocated to EDCA are used to integrate HCCA, a channel model (Kuan and Dimyati, 2007) is suitable to show the behaviour of EDCA and HCCA networks in saturation conditions.

Adaptively Tuned HCF (AT-HCF), proposed by Lai et al. (2009), varies their durations until the optimal value, considering the supported applications and the global throughput.

In Ruscelli et al. (2012) a cooperative scheduling is proposed where the centralised scheduler retains the control of the main scheduling activities and the new *Overboost* scheduler, local to each station, aims at reducing the delay by enqueuing the unsent HCCA packets to the higher priority traffic class of EDCA. Overboost is conceived as a local node scheduler able to integrate any type of centralised algorithm. Priority is done to traffic with strict QoS requirements, with the goal to dispatch that using available resources, firstly from HCCA, then from EDCA.

In Hayajneh and Al-Mashaqbeh (2014) a new IEEE 802.11e MAC scheme is proposed based on the use of both HCCA and EDCA functions to improve the performance of video traffic. The polling mechanism is slightly changed reassigning the HCCA slots time of station that notified the QAP of no high priority traffic to send at the pollable stations with high level QoS streams. In this case it is the joint use of both HCCA and EDCA suitable to reach an enhanced QoS, without any preference among these two functions.

Users requirements can be taken into account by means of different methods. *scheduling estimated transmission time – earliest due date* (SETT-EDD) (Grilo et al., 2003) computes variable $TXOPs$ between its minimum and maximum values taking into account the stations expectations by means of a token bucket of time units; SI is computed considering traffic profile in the interval [$SI_{\min}$, $SI_{\max}$], and the polling order is set by *delay earliest due date* (delay-EDD) (Ferrari and Verma, 1990). In this scheduler the concept of deadline is used, which expresses the timing constraints of each TS. Thus, the diverse policies used to manage the polling list by scheduling the TSs deadlines are tools to impose and handling the respect of temporal requirements.

In Fan and Huang (2005) a timer-based scheduler computes the transmissions deadlines as the smallest between those of downlink and uplink TSs and sets accordingly a variable SI for each QSTA, whereas EDF handles streams polling. In this algorithm $TXOP$ is untouched.

Cicconetti et al. (2007a) presented *real-time HCCA* (RTH) that considers $TXOPs$ as critical sections, according to *stack resource policy* (SRP) algorithm of Baker (1991), distinguishing offline task that deals with admission control, transmission parameters computation, and online task limited to the transmissions scheduling. $TXOP$ and SI are computed as fixed value but the use of EDF to order the polling list introduces variability in the polling time. Moreover SRP implies a modified admission control formula that results to be more efficient whereas the two tier scheduler reduces the system overhead.

Adaptive resource reservation over WLANs (ARROW) (Skyrianoglou et al., 2006) computes dynamic $TXOP$ considering initial enqueued TSs at each polling. In particular,

$TXOP$ for each polling is obtained as the sum of the maximum time to transmit each TS. This allows to easily serve VBR and bursty traffic at the cost of allocated resources. An upper bound of MSI avoids deadline miss and guarantees delay constraints, whereas *earliest due date* (EDD) (Jackson, 1955) algorithm manages the polling order.

In Inanc et al. (2006) the uplink and downlink schedulers of *application-aware adaptive HCCA scheduler* assign respectively each QSTA and QAP a minimum and a maximum SI, considering application and network states and enqueued packets by monitoring transmission starting time and ending time plus packets size for each TS; the polling list is set by EDF.

A diverse scheduling scheme is that of Cicconetti et al. (2007b) where *wireless timed token protocol* (WTTP) schedules stations through *timed token protocol* (TTP), a MAC protocol based on token passing and introduced for ring networks by Grow (1982). In this case the polling list is intended as a ring network and a station is polled when reached by the token with periodicity set by the *target token revolution time* (TTRT) that plays the role of SI, equal for all the stations. In particular, TTRT is set as $TTRT = 1/2 \cdot \min_i D_i$, where D_I is the delay bound, thus the polling period computation is not dynamic and is conservative, ensuring the respect of all deadlines at the cost of polling stations more frequently than needed. The resource allocation is performed only for synchronous nodes with strict QoS requirements, that are assigned a *sejour* time ($TXOP$ computed during the admission control), whereas the asynchronous ones can transmit only if the token arrives earlier that expected, i.e., if the previous QoS stations do not exhaust their accorded transmission time. The method result to be extremely conservative, trying to efficiently serve only QoS stations and do not taking care of the remaining ones. Moreover, the use of a single SI does not provide any flexibility.

A different approach is to correct the resource assignment through reclaiming schemes. In this case the action is performed after the initial $TXOP$ computation and polling of a station by recovering the exceeding resources and allocating that to the next one in the polling list without modifying the $TXOP$ computation rules. In that it differs from the feedback approach.

First of all *wireless capacity based scheduler* (WCBS) (Cecchetti and Ruscelli, 2008), Cecchetti and Ruscelli (2011) schedules TSs using static and dynamic parameters, whereas EDF handles polling list. In WCBS, that will be used in the study of IDTH and IDTH+, the admission control assigns each TS_i a static budget Q_i, i.e., the maximum transmission time, and a fixed polling period P_i, considering TSs parameters. The dynamic scheduling uses the following parameters: the instantaneous remaining time c_i of the foregoing polling reserved for the next one, the absolute deadline d_i i.e., the due date of Q_i, the next polling time p_i (set considering P_i, d_i, EDF order of deadlines and the current polling time), and the stream $state$. These parameters are the 'dynamic' version of the static ones and are involved in the dynamic computation. Moreover, as it will be explained below, they are suitable to perform a embryonic form of resource reclaiming. In particular, when a QSTA is polled at time p_i, it is assigned the current budget c_i, if present, or Q_i. This transmission has to be exhausted until d_i, and the spare time, if any, is reserved if it is greater than the minimum time to send an MSDU. Thus, WCBS allows an elementary recovery of unused resources, that are, however, reserved for the same station and not for the subsequent ones with strict QoS requirements. Moreover this recovery does not take into account stations needs but simply avoiding waste of resource. Finally, WCBS offers a dynamic polling and a dynamic transmission duration, but EDF penalises the temporal performance, as shown in Cecchetti and Ruscelli (2011).

In Larcheri and LoCigno (2006) a bandwidth reclaiming mechanism, based on a proportional controller where a weighted function assigns the resources proportionally to traffic classes and queues length, uses optimisation methods, in particular two *max-min fairness* algorithms, to bound the delay using respectively a fixed and a variable SI. Instead in LoCigno et al. (2007) the same approach is used to schedule different applications trying to increase throughput and to reduce packets loss.

In Lee et al. (2007) a bandwidth reclaiming scheme for the IEEE 802.11 PCF function with *weighted round robin* (WRR) scheduling policy is proposed. It determines when the unused transmission time can be either used to advance the next polling opportunity or assigned to the contention period. Moreover the proposed algorithm modifies the WRR polling list to put the stations with higher probability of generating unused time at the end of the list. This rearrangement aims to reduce the number of reclaimed stations but it makes the solution not extensible to HCCA function, where the polling order is strictly related to real-time guarantees.

Unused time shifting scheduler (UTSS), presented and studied in Ruscelli et al. (2011); Ruscelli and Cecchetti (2014), integrates a greedy reclaiming algorithm with a centralised scheduler without traffic analysis or stations differentiation. It performs a recovery of resources remaining from the current transmission and simply assigns that to the next polled station in a greedy manner. No analysis or estimation of stations requirements are adopted thus the recovered capacity is assigned independently the next station in the polling list needs that or not. This implies an unnecessary addition of resources that will be yet recovered at the next polling. Considering more refined algorithms based on the use of queues length, UTSS plays in advance, trying to provides a $TXOP$ tailored to the traffic profile, before the packets are enqueued. With respect to schedulers using traffic model to predict the future stations needs, it has a practical approach, strictly focused on the next step scheduling, but with the advance of no additional scheduling load. However this is yet an evolved version of reclaiming compared to the idea behind WCBS where there is no propagation of recovered capacity to other stations in the polling list. Instead *dynamic TXOP HCCA* (DTH) (Cecchetti et al., 2012b) refines the assignment of recovered resources introducing an estimation of the future resource expectation. It adds a mechanism based on time series forecasting for the estimation of the actual needed transmission interval. Using a moving window that sets the monitoring time interval, DTH takes memory of the effective duration of the previous transmissions of a station and uses this information to deduce the required transmission time for the next polling. This statistical study of the resources used in the past history of the station allows a resource allocation more tailored to the station traffic profile with respect to a simply greedy approach. With respect to the approach based on the use of queues length, such as that of Boggia et al. (2007), the consideration of the past history of the data source to forecast its corresponding future behaviour allows to find a tradeoff between knowing the traffic profile and predict its future and knowing the number of enqueued packets, which is an effect of the traffic behaviour, balancing past and future history of the data source. Moreover the traffic profile estimation based on time series forecasting provides a solution to the problem of traffic profile prediction in the case of non-deterministic sources.

The above summary is just a spotlight to the different techniques proposed in the evolution of IEEE 802.11e HCCA scheduler to improve its QoS. At the basis of their proposal is the consideration of the lack of flexibility of its transmission time and polling period computation mechanisms. To highlight this issue in Table 1 the mentioned algorithms are listed distinguishing those that provide a dynamic $TXOP$, or a dynamic SI or both of them.

Table 1 HCCA scheduling algorithms: variable TXOP and SI

Scheduler	*TXOP*	*SI*	*Scheduler*	*TXOP*	*SI*
FHCF	X		*Ju2013*	X	
FBDS	X		*Hayajneh2014*	X	X
ETA-EQN	X		*Saheb2012*	X	X
SETT-EDD	X	X	*WCBS*	X	X
Overboost	X		*Larcheri2006*	X	
Al-Maqri2013	X		*LoCigno2007*	X	
RTH		X	*UTSS*	X	
ARROW	X	X	*DTH*	X	
Inanc2006	X	X	*Fan2005*		X

As shown, all the mentioned algorithms succeed in ensuring a dynamic $TXOP$ and some of them are also able to use a variable SI. However the common aim is to improve the dynamicity of the scheduling engine meeting different QoS expectations, especially when multimedia applications and, in general, VBR traffic are considered.

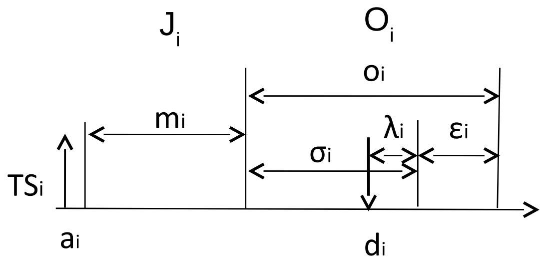
Each technique is focused on a particular aspect emerging from a deep analysis of the packets transmission methodology: understanding the precise mechanism of the transmission of the packets is suitable to highlight the involved variables, the mutual dependencies and the failing points where performance degradations are derived. This can directly and clearly reveal what are the efficient action points where focusing the interventions aiming to correct the scheduling behaviour and improve the offered service. In particular, considering the transmission time that, as demonstrated by the literature is a key point for the QoS, variable transmission duration can be derived taking into account the stations parameters whose variations impact on that. In this context a more accurate representation of the information can be useful. This information can be derived by the introduction of a tailored model suitable to represent their temporal evolution and their dependency from other variables. Otherwise an estimation of future values, based on statistical studies of the history of these parameters, can substitute the models or, finally, the exact monitoring and sampling of the considered parameters offers a real evaluation directly used in the computation or in correction mechanism for the computed scheduler parameters. Then this information can be used differently in dependency of the chosen scheduling techniques: feedback mechanisms can use that to perform correction of the MAC parameters, otherwise they can be directly inserted in the computation as set by the reference scheduler at the place of mean value parameters to take into account real traffic and network behaviour. EDCA can be involved to help HCCA to dispatch TSs with strict requirements. Deadlines are used to model temporal expectations providing a tailored scheduling method. The recovery of unexhausted resources can repair to inefficient resource assignment and can be declined from the rough greedy approach to refined techniques based on optimisation or statistical studies. The choice of the method can to be guided taking into account which performance (end-to-end delay, throughput, packet loss, jitter etc.) is of interest and the load imposed by the improving method.

3 QoS scheduling through imprecise computation

HCCA schedulers based on static preallocation of resources are unable to efficiently manage VBR traffic. Indeed, since at each polling a general deterministic multiplexing scheme sets the sending interval duration equal to $TXOP_{AC}$, (the transmission time computed during the admission control phase and based on mean values of transmission parameters and QoS requirements), it does not reflect data rate fluctuations, inducing degraded network performance. This fixed capacity is tailored for constant bit rate TSs, whereas it results insufficient to deliver the enqueued traffic when data rate increases with respect to its mean value. Starting from these considerations the basic idea of IDTH and IDTH+ is that the missing resources could be recovered, without any recomputation of accorded transmission time, taking advantage from the same preallocation mechanism, but in the opposite situation, i.e., when data rate decreases. Indeed, in this case the actual transmission duration can be shorter than $TXOP_{AC}$ and, in general, its residual portion would be wasted. This interval of time should be used to tackles the lack of capacity when data rate increases.

Adopting the imprecise computation model, the scheduler suitable to provide the negotiated QoS can be built distinguishing the required activity of the main centralised scheduler, that deals with the needed service, and that of an additional refined centralised scheme, suitable to introduce a dynamic capacity allocation as provided, for instance, by the statistical multiplexing. This further scheduler can be activated optionally or partly to limit the system overload. Thus, considering the performance degradation in the case of VBR TSs, the additional module can deal with the management of the unused resources to integrate the missing capacity. In particular, in our case, as shown in Figure 1, the centralised scheduler is the *mandatory* part J_i with execution time m_i, suitable to provide an acceptable QoS as negotiated during the admission control, whereas the additional module, i.e., IDTH or IDTH+, corresponds to the *optional* part O_i (whose execution time is o_i), suitable to improve the offered QoS, in particular in stressing scenarios.

Figure 1 Imprecise computation model

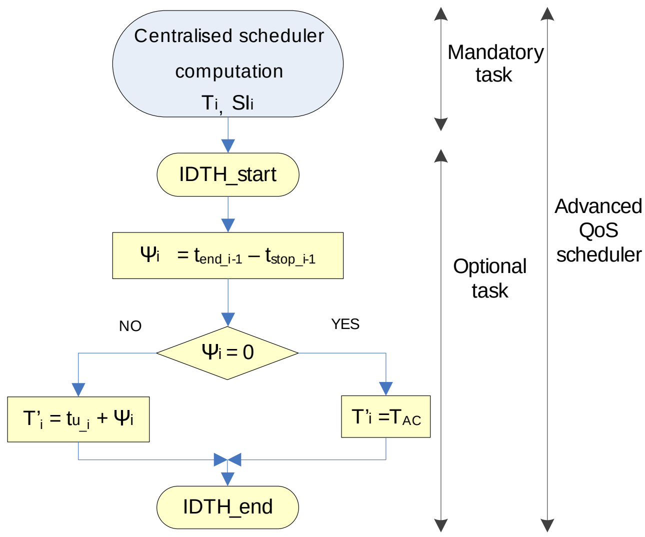


Furthermore IDTH and IDTH+ tune the allocation of recovered capacity through the monitoring of effectively used transmission time to provide an amount of resources tailored to the station needs. As illustrated in the following, since they are able to make variable the transmission time, they are suitable to change the resource grant from a fixed preallocation to an instantaneous dynamic resource assignment that takes the advantages of statistical multiplexing.

3.1 IDTH

IDTH refines the resource assignment performed by centralised schedulers taking into account instantaneous stations requirements through a one-step monitoring of used capacity. In particular, it performs two actions: the reclaiming of residue capacity and the evaluation of station requirements along with the transmission time assignment. Without interfering with the main scheduling engine, as it will be shown in Sections 4 and 5, it integrates a dynamic mechanism to instantly increase the current assigned $TXOP$ (from now T), when needed. With respect to a 'pure' statistical multiplexing approach, where the transmission interval should be recomputed at each polling time impacting on the admission control, IDTH simply recovers the resources unused by previously polled stations, already considered in the admission control computation and that would be lost, and assigns that to the next polled QSTA. This assignment is requirements-aware in the sense that it is performed when needed, i.e., if during its previous polling the station required a transmission time greater than its $TXOP_{AC}$. In particular, as will be explained below, the control of instantaneously used resources, suitable to make the capacity assignment more tailored to station requirements, is performed by keeping memory of the transmission time used during the previous polling of the scheduled station. In Figure 2 a flow graph summarises the tasks of the global scheduler, obtained through the integration of IDTH with a centralised algorithm, highlighting the relationship between these two components. The centralised scheduler, that deals with the mandatory activity, computes the main protocol parameters, SI and T. Then, before each polling, the optional part, carried out by IDTH, performs the additional task suitable to refine the transmission time computation with the aim to improve the service where the centralised scheduler is deficient. IDTH executes its bandwidth reclaiming supported by the transmission duration monitoring, then the new transmission time T' is passed to the centralised scheduler, that polls the next station in the list and communicates the value of the transmission time that will be used during the current session.

Figure 2 IDTH scheduling (see online version for colours)



Going deep inside, at the polling time of QSTA$_i$ IDTH adds the following instantaneous computation of the current transmission time T_i':

$$T_i' = \begin{cases} T_{i_{AC}} & \text{if } \psi_i = 0 \\ t_{u_i} + \psi_i & \text{if } \psi_i > 0 \end{cases} \tag{2}$$

where $T_{i_{AC}}$ is the fixed transmission duration assigned during the admission control. ψ_i is the recovered residue transmission time of the preceding polled station QSTA$_{(i-1)}$ expressed as:

$$\psi_i = t_{\text{end}_{i-1}} - t_{\text{stop}_{i-1}}, \tag{3}$$

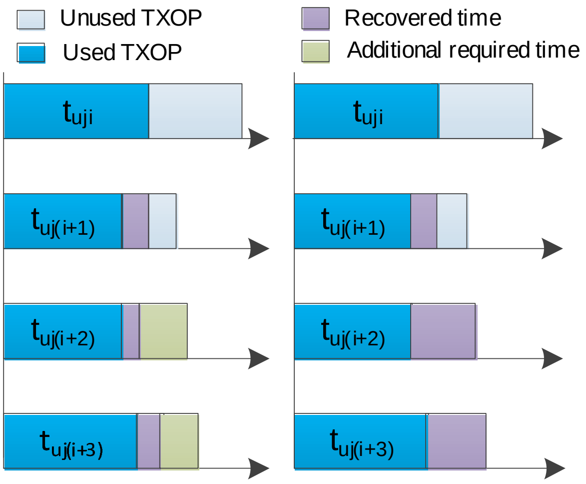
where $t_{\text{end}_{i-1}} = t_{p_{i-1}} + T_{i-1}$ is the ending time of the transmission of QSTA$_{(i-1)}$ when T_{i-1} is completely exhausted, $t_{p_{i-1}}$ is its polling time, and $t_{\text{stop}_{i-1}}$ is its actual ending transmission time. Finally, t_{u_i} is the transmission time used by QSTA$_i$ during its previous polling and is the parameter considered for the monitoring of the instantaneous used capacity: IDTH adds $\psi_i > 0$ to t_{u_i} instead of $T_{i_{AC}}$ to take into account the actual traffic profile.

This approach is based on the elementary consideration that, if during the previous polling a station has shown a data rate lower than its mean value, and if the traffic maintains a constant slope, supposedly in the subsequent polling this station will does not need an additional capacity. Obviously, this is true except in the points where the data rate slope changes its sign, but in this case ψ_i allows to provide a further capacity and the monitoring scheme itself, assigning t_{u_i} at the next polling, enables the system to increase gradually the instantaneous transmission time. This simple one-step monitoring of the effective duration of the previous transmission makes the resource assignment more tailored to the instantaneous station requirements and prevents the accumulation of unnecessary resources. Moreover, it enables IDTH take advantages from the statistical multiplexing approach, whose dynamic resource allocation is particularly usefulness in the case of unpredictable traffic. Otherwise, when $\psi_i = 0$ IDTH sets $T_i' = T_{i_{AC}}$, since assigning t_{u_i} can generate an irreversible decreasing of subsequent transmissions duration, as shown in Cecchetti et al. (2012b).

3.2 *IDTH+*

As illustrated above, IDTH introduces a mechanism for the computation of an adjustable transmission interval to make dynamic a deterministic time division multiplexing allocation; this is useful when data rate varies. However, it shows some drawbacks in the case of changes in the sign of data rate variations. Figure 3(a) illustrates the scheduling of a QSTA$_i$ performed by IDTH when, without loss of generality, data rate decreases from ith to $(i+1)$th polling, and suddenly increases from $(i+2)$th to $(i+3)$th polling.

Figure 3 Comparison of IDTH and IDTH+ in the case of data rate variations: (a) IDTH and (b) IDTH+ – recovered time + $t_u < T_{AC}$ (see online version for colours)



(a) (b)

From ith to $(i+1)$th polling IDTH assigns the recovered time (in violet) and the polled station, owing to the data rate decreasing, simply lives the exceeding transmission time (in azure). When date rate changes the sign of its variation and begins increasing from $(i+2)$th to $(i+3)$th polling, IDTH is able to provide only ψ, that is related to the traffic of the previous polled station, and whose amount is not predictable. Thus, two situations can occur: the current polled station could immediately receive the required transmission interval dispatching all the arriving packets or it could wait for a greater amount of capacity (in figure, the missing time is in green), thus it experiences a slow increase of sending time with respect to its requirements. This randomness, owing to traffic profile and owing to the position of the station in the polling list, makes the attempt to follow traffic variations less reactive.

Taking into account the previous considerations, IDTH+ is proposed to limit this hysteresis in the answer of IDTH when data rate decreases with respect to its mean value and to guarantee a minimum resource threshold to cope with data rate changes, especially when there is a sudden increase after a continuous decreasing, that is the illustrated scenario where IDTH shows its lack. To do that, maintaining the same approach of IDTH, IDTH+ introduces a further condition to verify during the instantaneous computation of the current T_i':

$$T_i' = \begin{cases} T_{i_{AC}} & \text{if } \psi_i = 0 \vee t_{u_i} + \psi_i < T_{i_{AC}} \\ t_{u_i} + \psi_i & \text{if } \psi_i > 0 \wedge t_{u_i} + \psi_i > T_{i_{AC}} \end{cases} \tag{4}$$

The following pseudo-code listed in Algorithm 1 illustrates the IDTH+ scheduling activity.

Algorithm 1 IDTH+ scheduling.

IDTH+ `start`
$\psi_i \leftarrow t_{end_{i-1}} - t_{stop_{i-1}}$
if $(\psi_i = 0)$ **or** $(t_{u_i} + \Psi_i \leq T_{AC})$ **then**
  $T_i' \leftarrow T_{AC}$
else
  $t_{u_i} + \psi_i$
end if
IDTH+ `end`

The difference with IDTH is in the assignment of $T_{i_{AC}}$, that now in IDTH+ is subject to two different conditions. The first one is when there is no residue time from previous transmissions ($\psi_i = 0$), whereas the second one is when $\psi_i \neq 0$ and the summation of this remaining time and of the monitored used transmission time is less than $T_{i_{AC}}$. This simple modification allows to provide a greater transmission time without falling below a threshold that can delay the increase of T' when needed. Proposition 1 in the following Section 4 will highlight the theoretical motivations besides the introduction of IDTH+.

The more meaningful advantage of IDTH+ is that it is suitable to improve the performance of IDTH simply modifying the conditions to verify when recovered resources are allocated, without any additional computational load, as will be shown in Section 4. In Section 5, the improvements introduced by IDTH+ with respect to IDTH will be analysed through simulation.

Finally, IDTH and IDTH+ can cooperate with the centralised scheduler, whose activity is untouched, as will be shown in Section 4, and are activated when needed and when there are unused resources. Furthermore, as it will be demonstrated, their integration is not affected by some drawbacks of the imprecise computation method.

4 IDTH and IDTH+ analysis

In this section IDTH and IDTH+ are studied from a theoretical point of view, considering their properties and that of the global scheduler obtained through their cooperation with a preexistent centralised algorithm. The focus is on the evaluation of the effective advantage of adding these algorithms taking into account two different points of view, that of the centralised scheduler and that of the resulting global one, and considering for both of them positive and negative aspects.

Some fundamental real-time properties of IDTH have been investigated in Cecchetti et al. (2012b), showing that it does not jeopardise the admission control feasibility condition and the respect of deadlines, along with the demonstration that the assignment of $T_{i_{AC}}$ avoids the scheduling misbehaviour owing to the hysteresis in the dynamic transmission allocation. These considerations can be extended to IDTH+ owing to the similarities of the reclaiming mechanisms. These results are meaningful since they show as the reclaiming methods of these algorithms are able to provide the stations resources in addition to the capacity accorded by the admission control, and without violating its feasibility test. Indeed, the number of admitted stations does not change but only the recovered resources, already considered in the initial admission control computation, are moved between stations, with only the effect of polling in advance some stations. Considering the respect of deadlines, it is highlighted as the redistribution of resources done by IDTH and IDTH+ does not get worse the real-time behaviour of the global scheduler compared with the centralised one.

In the following, first of all, it is investigated if the modification introduced by the scheduling rules of IDTH+ is suitable to improve the reactivity of IDTH. Then, the imprecise computation model is taken as a guideline for the analysis of the global scheduler: this approach is suitable to distinguish the main scheduling component from the additional IDTH and IDTH+ and, then, to highlight their contribution and impact compared to the performance of the centralised and of the global schedulers. Indeed, in the imprecise computation model a global scheduler is considered as the product of an *optional* task, that is isolated to distinguish the scheduling part that can experience service degradation, and the *mandatory* one that deals with the basic QoS. Both of them contribute to an advanced and refined QoS. Moreover the distinction between the mandatory part, that is the strictly necessary component, and the optional one, whose execution is optional, is useful to lightening the system load. The impact of these features will be illustrated below. It is discussed if IDTH and IDTH+ experience the drawbacks of the imprecise computation from the scheduling activity point of view. Indeed, considering this theoretical approach from a different point of view, focusing on the system load, it is evaluated if the obtained global scheduler is affected by an increased complexity compared to the centralised one. Furthermore the impact of IDTH and IDTH+ on the scheduling error and on the time loss that characterise the scheduling behaviour is analysed.

Finally, the QoS offered by the global scheduler is investigated, to evaluate if adding IDTH and IDTH+ effectively improves the guaranteed service of the main scheduler, i.e., if the global scheduler is able to provide an improved QoS compared to the centralised one. In particular, the impact of IDTH and IDTH+ on the service rate is studied as QoS parameter; this analysis is meaningful since the service rate directly affects the performance experienced by the stations. Consequently, the effect on queues length is theoretically examined, that is dependent by the service rate and is strictly related to other performance parameters such as delay and packets loss, as will be shown by simulations results in Section 5.

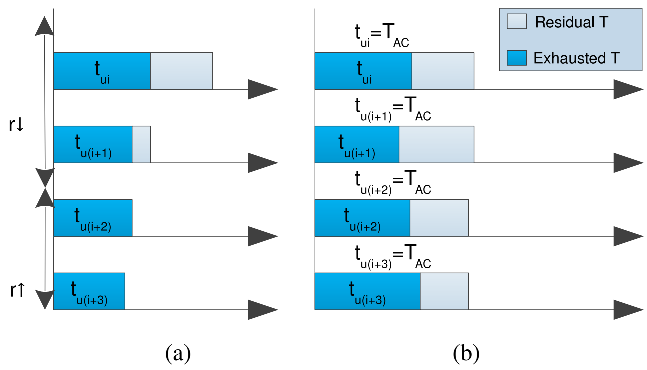
The following proposition clarifies the theoretical motivations besides the introduction of IDTH+.

Proposition 1: *IDTH+ improves the responsiveness of IDTH reducing its inertia in following rapid data rate changes.*

Proof: Without loss of generality Figure 4 shows the effect of assigning $T'_i = t_{u_i}$ in the case of VBR TSs whose data rate changes the sign of its variations. Figure 4(a) displays that, assigning $T'_i = t_{u_i}$ when, for instance, the data rate firstly decreases and then increases, IDTH tries to follow this change but it could be unable to provide the needed transmission time. Indeed, in the figure it is illustrated that in this case the accorded transmission time is exhausted and no more resources are available to deliver enqueued traffic.

Now, focusing on Figure 4(b), by means of the simple modification of the computation of T' when $t_{u_i} + \psi_i < T_{i_{AC}}$, i.e., when the total resources that can be allocated are less than the ones accorded by the admission control, IDTH+ immediately offers a greater transmission interval equal to the maximum that can be assigned by IDTH (T_{AC}). This simple choice avoids an excessive reduction of given resources, that can make following the data rate increase slow and not responsive. Thus, IDTH+ is able to fill the need of capacity when data rate rapidly grows up. In the case of slow increase or when data rate decreases again, the unspent time is simply returned, without any modification of the IDTH mechanism. Hence, although IDTH provides a dynamic T', IDTH+ is suitable to overcome its inertia in following rapid data rate changes, improving the recovery mechanism. □

Figure 4 Scheduling example when $\psi_i = 0$: (a) $T'_i = t_{u_i}$ and (b) $T'_i = T_{AC}$ (see online version for colours)



4.1 IDTH and IDTH+ vs. imprecise computation model

In the following the analysis of IDTH and IDTH+ is deepened taking into account the features of the imprecise computation. First of all, since one of the aims of this model is lightening the system load, distinguishing the mandatory part of the scheduling from the optional part, it is discussed if the integration of IDTH and IDTH+ has a negative impact on the computational load imposed to the obtained global scheduler, i.e., if this global scheduler is affected by an increased complexity with respect to the centralised one.

Proposition 2: *The computational complexity of the global scheduler $O_g \equiv O_c$, where O_c is the computational complexity of the centralised scheduler.*

Proof: IDTH and IDTH+ are independent and do not interfere with the scheduling activity of the centralised scheduler that they are simply added to. Consequently the computational complexity of the global scheduler $O_g = \max\{O_c, O_{IDTH,IDTH+}\}$, where O_c is the computational complexity of the centralised scheduler. Thus, since the computation of IDTH and IDTH+ is limited to $t_{u_i} + \psi_i$, their computational complexity is equal to *O(1)* and, consequently, $O_g \equiv O_c$.

From the imprecise computation model point of view, the global scheduler takes the advantages of the integration of IDTH and IDTH+, providing an instantaneous dynamic resource allocation and improving the scheduling activity of the centralised one, without the drawback owing to the integration itself, that appears to be 'transparent'.

Moreover, the reduced computational complexity of IDTH and IDTH+ is the main advantage of the use of these algorithms compared to other solutions, such as the introduction of a new improved QoS algorithm, since, without an additional computational cost with respect to the centralised scheduler, they are able to improve the efficiency of the resource management adding more resources when needed, without overprovisioning and avoiding waste of resources. Furthermore, they succeed in adding dynamicity in the computation of T without a recalculation of transmission time that may weight the scheduling activity.

For the same reasons than above, the randomness of number of reclaiming does not affect the computational complexity of the global scheduler. □

As previously mentioned, in the imprecise computation model the *optional* task O_i is suitable to isolate the scheduling part that can experience performance degradation. Indeed, since the *optional* task has the same absolute deadline d_i of the *mandatory* one J_i, as shown in Figure 5(a), it is interesting to investigate from a different point of view if the obtained global scheduler is able to respect this deadline or the real-time performance are degraded and a scheduling error is introduced.

Proposition 3: *The global scheduler obtained by the integration of IDTH and IDTH+ has scheduling error $\epsilon_i = 0$ and time loss $\lambda_i = 0$ if a local upper bound Ψ_{ik} of the current recovered time is set as $\Psi_{ik} = d_i - t_{\text{end}_{ik}}$.*

Proof: In general, at polling kth of a station QSTA$_i$, if its previously polled station did not use its assigned T_{i-1}, this will be completely transferred by IDTH and IDTH+ to QSTA$_i$ minus the time interval $\delta = t_{\text{SIFS}} + t_{\text{NULL}} + t_{\text{SIFS}}$, that takes into account, respectively, the waiting time before a station starts its transmission (t_{SIFS}), and the sending time of `CF-Null` frame, that alerts the QAP that the transmission queue is empty. Thus, $\psi_{ik} \leq T_{i-1} - \delta$. As demonstrated in Cecchetti et al. (2012b), the redistributions of resources performed by IDTH and IDTH+ simply implies that the polling time of the subsequent station is anticipated. If all the stations preceding QSTA$_i$ in the polling queue did not have traffic to send, then

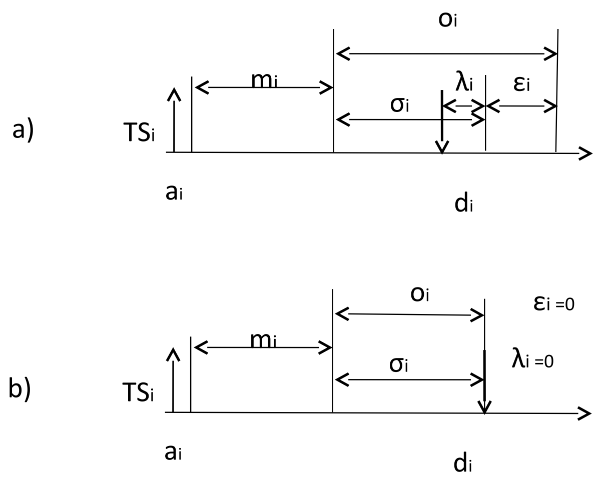
$$0 \leq \psi_{ik} \leq \sum_{j=1}^{i-1}(T_j - \delta) \leq \sum_{j=1}^{i-1}(T_j) - (i-1)\delta.$$

In this case an excessive accumulation of resources could affect the respect of the temporal requirements but IDTH and IDTH+ assign ψ_i to the next station in the polling list, without distinguishing the QSTAs requirements. Moreover, since the optional tasks IDTH and IDTH+ have the same deadline d_i of the main centralised scheduler, due to the model formulation, they can experience a performance degradation and, for instance, miss this deadline if it is assigned a computation time $\sigma_i < o_i$, where o_i is its required computation interval. In this case a scheduling error $\epsilon_i > 0$ can occur and the *time loss*, i.e., the computation interval exceeding the deadline, $\lambda_i > 0$. Thus, to prevent this deadline miss, in particular focusing on stations with strict real-time expectations, at each polling k it is useful to set the following upper bound Ψ_{ik}:

$$0 \leq \psi_{ik} < d_i - t_{\text{end}_{ik}} = \Psi_{ik}.$$

In this case, as shown in Figure 5(b), since the *optional* task O_i (IDTH and IDTH+) has the same deadline d_i of the *mandatory* one J_i (the centralised scheduler), the error $\epsilon_i = 0$ and the time loss $\lambda_i = 0$.

Figure 5 Imprecise computation: (a) general model and (b) model with IDTH and IDTH+



This result shows that the cooperation with IDTH and IDTH+ is not affected by the general faults owing to the integration of additional modules and it confirms that this further scheduling contribution does not increase the number of deadlines miss compared to the centralised scheduler. □

Summarising, the properties studied above highlight how the global scheduler, obtained integrating IDTH and IDTH+, does not differ from the original centralised algorithm in terms of computational complexity, respect of the threshold of available resources and deadline miss. Thus, compared to the imprecise computation approach and maintaining the same goal of QoS improvement, the global scheduler does not have the faults of this model and IDTH and IDTH+ result to be suitable to integrate the scheduling activity in a seamless way.

4.2 IDTH and IDTH+ impact on QoS: service rate and transmission queues length analysis

Finally, to deepening the analysis of the effect of IDTH and IDTH+, this section investigates if their integration allows to obtain a global scheduler with an improved QoS. In the previous section it has been shown as their cooperation does not affect the scheduling properties of the centralised scheduler and the obtained global one does not undergo a deterioration. In the following it is investigated if they are able to really improve the offered QoS to corroborate their usefulness.

Since IDTH and IDTH+ act on the transmission time duration, their impact on service rate is studied. Then a deeper investigation on their effect on the stations queues length is provided, since the number of enqueued packets is strictly related to the dispatching rate. This is an interesting analysis owing to the fact that any variation on the transmission buffers length can influence also experienced delay and packets loss.

In the following the effect of the recovered time ψ on VBR TSs is analysed in the case of bursty traffic, where, in general, statistical multiplexing shows good performance owing to its flexibility.

Theorem 4: *Assuming to model the transmission queue of a node in the presence of bursty traffic through an M/G/1 queue, the use of ψ allows to serve bursts of traffic with an increased data rate $\lambda' = \lambda \cdot \frac{E\{s\}+E\{\psi\}}{E\{s\}}$ without modify the admission control.*

Proof: With the purpose of generality, in the following the effect of ψ is analysed independently from the centralised scheduler. It is assumed to model the transmission queue in a node in the presence of bursty traffic and with a general HCCA scheduling policy with an M/G/1 queue, where the incoming traffic is modelled by a Poisson process and the data interarrival time follows an exponential distribution. In this case the mean queue length L is expressed by the Pollaczek-Khinchine formula as function of the incoming traffic:

$$L = \rho + \frac{\rho^2 + \lambda^2 Var\{s\}}{2(1-\rho)}, \tag{5}$$

where $\rho = \frac{\lambda}{\mu}$ is the utilisation factor, λ is the arrival rate of the Poisson process, $1/\mu$ is the mean of the service time s, (referring to a general scheduler, s is represented by a stochastic variable), whose variance is *Var{s}*; initial conditions imply no enqueued packets. In the following the queue length is analysed in two different situations:

a s instantly increased by ψ owing to the integration of IDTH or IDTH+ and data rate λ unchanged, and

b centralised scheduler without IDTH or IDTH+ and increased data rate.

Case a. When ψ is provided to a centralised scheduler, the current assigned transmission time is $T' = s + \psi$. Assuming arrival data rate λ unchanged, equation 5 becomes:

$$\begin{aligned} L' &= \lambda E\{s+\psi\} + \frac{\lambda^2 E^2\{s+\psi\} + \lambda^2 Var\{S+\psi\}}{2(1-\lambda E\{s+\psi\})} \\ &= \frac{\lambda E\{s+\psi\}[2(1-\lambda E\{s+\psi\})]}{2(1-\lambda E\{s+\psi\})} \\ &\quad + \frac{\lambda^2 E^2\{s+\psi\} + \lambda^2 Var\{S+\psi\}}{2(1-\lambda E\{s+\psi\})} \\ &= \frac{2\lambda E\{s+\psi\} - \lambda^2 E^2\{s+\psi\}}{2(1-\lambda E\{s+\psi\})} \\ &\quad + \frac{\lambda^2 Var\{s+\psi\}}{2(1-\lambda E\{s+\psi\})}. \end{aligned} \tag{6}$$

Case b. When only the centralised scheduler is used and the arrival data rate grows up to λ', equation 5 becomes as follows:

$$\begin{aligned} L'' &= \lambda' E\{s\} + \frac{\lambda'^2 E^2\{s\} + \lambda'^2 Var\{S\}}{2(1-\lambda' E\{s\})} \\ &= \frac{\lambda' E\{s\}[2(1-\lambda' E\{s\})] + \lambda'^2 E^2\{s\} + \lambda'^2 Var\{s\}}{2(1-\lambda' E\{s\})} \\ &= \frac{2\lambda' E\{s\} - \lambda'^2 E^2\{s\} 2(1-\lambda' E\{s\}) + \lambda'^2 Var\{s\}}{2(1-\lambda' E\{s\})} \end{aligned} \tag{7}$$

Equating term by term equations (6) and (7), the queue length is the same in these two different scenarios, if:

$$\lambda E\{s+\psi\} \equiv \lambda' E\{s\}, \tag{8}$$

$$\lambda'^2 Var\{s\} = \lambda^2 Var\{s+\psi\}. \tag{9}$$

Replacing the expression of λ' derived by the second equation above, the following relationship is obtained:

$$\frac{\mathrm{Var}\{s+\psi\}}{E^2\{s+\psi\}} = \frac{\mathrm{Var}\{s\}}{E^2\{s\}},$$

that is true since $\mathrm{Var}\{s\} = m^{(2)}\{s\} - E^2\{s\}$. Thus, considering the service globally offered, adding IDTH or IDTH+ has the same effect than serving a data source with a greater rate and applying only the centralised scheduler. The global scheduler is able to serve a traffic source with a data rate λ' higher than that considered during the admission control and whose expression derived from equation (9) is equal to:

$$\lambda' = \lambda \frac{E\{S\} + E\{\psi\}}{E\{S\}}. \tag{10}$$

A confirmation of the previous result can be derived focusing on the utilisation factor ρ. In case *(a)*, independently from the scheduling rules, the current transmission time is $s' = s + \psi$, and $E\{s'\} = 1/\mu = E\{s + \psi\} = E\{s\} + E\{\psi\}$, owing to the linearity of the expected value, and the new utilisation is

$$\rho' = \lambda(E\{s\} + E\{\psi\}).$$

In case *(b)*, the utilisation is $\rho'' = \lambda' E\{s\}$. Comparing these two expressions,

$$\rho' \equiv \rho'' \iff \lambda' = \lambda \frac{E\{s\} + E\{\psi\}}{E\{s\}},$$

that confirms equation (10), as expected.

This result highlights how the adding of ψ, recovered from unexhausted transmission intervals, provides an instantaneous overprovisioning without impacting on the admission control. Moreover this makes the transmission time dynamic as required by VBR traffic, taking the advantage of statistical multiplexing in an easier way. The use of an M/G/1 queue ensures the generality of the theorem, true for each scheduling policy and confirms that the effects of IDTH and IDTH+ are the same for each centralised scheduler. □

Since the previous theorem shown that IDTH and IDTH+ impact on the service rate used to dispatch the enqueued traffic and the global scheduler is suitable to serve TSs with a greater data rate, as a consequence of these results in the following Theorem the length of transmission queues when IDTH and IDTH+ are used is analysed.

Theorem 5: *The global scheduler obtained integrating IDTH and IDTH+ is suitable to reduce the transmission queues length to* $Q'_{ik_{\min}} = \Pi_i - k \cdot r \cdot (T_{i_{AC}} + \Psi_{i_{\max}})$.

Proof: Without loss of generality, a station $QSTA_i$ with backlogged traffic is considered as worst case since it may need a transmission time greater than its assigned $T_{i_{AC}}$ to deliver the enqueued streams. Moreover the non integration (case *(a)* and the integration of IDTH and IDTH+ (case *(b)* are evaluated.

Case a. When IDTH and IDTH+ are not integrated, during the kth polling of $QSTA_i$, assuming that the accorded transmission interval is exhausted, its queue length expressed in terms of enqueued traffic is as follows:

$$Q_{i_k} = Q_{i_{k-1}} + \pi_{ik} - \tau_{ik} = Q_{i_{k-1}} + \pi_{ik} - T_{i_{AC}} \cdot r,$$

where $Q_{i_{k-1}}$ is the enqueued traffic deriving from the previous $(k-1)$th polling phase, π_{ik} is the current incoming traffic, τ_{ik} is the current delivered traffic and r is the mean transmission rate. Iterating

$$\begin{aligned} Q_{i_{k-1}} &= Q_{i_{k-2}} + \pi_{i(k-1)} - \tau_{i(k-1)} \\ &= Q_{i_{k-2}} + \pi_{i(k-1)} - T_{i_{AC}} \cdot r, \end{aligned}$$

where $Q_{i_{k-2}} = Q_{i_{k-3}} + \pi_{i(k-2)} - \tau_{i(k-2)}$. In particular, assuming that $QSTA_i$ exhausts its assigned transmission time, the maximum traffic delivered by the centralised scheduler without IDTH and IDTH+ is:

$$\tau_{i_{\max}} = T_{i_{AC}} \cdot r \tag{11}$$

Isolating the total amount of traffic Π_i arriving at station $QSTA_i$ until the kth polling, that is independent from the transmission intervals duration, and assuming initial empty queues, the minimum queue length at the kth polling is

$$Q_{ik_{\min}} = \Pi_i - krT_{i_{AC}}. \tag{12}$$

Case b. Instead, when IDTH and IDTH+ are integrated, during the kth polling the queue length of $QSTA_i$ is as follows:

$$Q_{i_k} = Q_{i_{k-1}} + \pi_{ik} - (t_{ui} + \psi_{ik}) \cdot r.$$

Owing to the current upper bound Ψ_{ik} of ψ_{ik}, the delivered traffic τ_{ik} can reach the following maximum amount

$$\tau_{ik_{\max}} = (t_{u_i} + \Psi_{ik}) \cdot r \geq (t_{u_i} + \psi_{ik}) \cdot r = \tau_{ik}$$

Moreover, if it is assumed that the station exhausted its maximum transmission time $T_{i_{AC}}$ during its previous polling, the maximum becomes

$$\tau'_{ik_{\max}} = (T_{i_{AC}} + \Psi_{ik}) \cdot r \geq \tau_{i_{\max}}.$$

Finally, if a global upper bound $\Psi_{i_{\max}}$ of ψ_{ik} is set $\forall k$ as follows

$$\Psi_{i_{\max}} = D_i - T_{i_{AC}} \geq \psi_{ik}, \forall k \tag{13}$$

where D_i is the relative deadline, thus the maximum delivered traffic is

$$\tau'_{i_{\max}} = r \cdot (T_{i_{AC}} + \Psi_{i_{\max}}) \geq \tau_{i_{\max}} \tag{14}$$

where $\tau_{i_{\max}}$ is the maximum traffic delivered without IDTH and IDTH+ expressed by equation (11). Consequently the minimum queue length is

$$Q'_{ik_{\min}} = \Pi_i - kr(T_{i_{AC}} + \Psi_{i_{\max}}) \geq Q_{ik_{\min}}. \tag{15}$$

Thus, it is possible to conclude that the global scheduler obtained integrating IDTH and IDTH+ outperforms the original centralised one delivering a greater amount of traffic without increasing the service rate. Hence it is suitable to further reduce the transmission queues length. □

The previous result will be corroborated in Section 5 by the simulations results on queues length. Moreover, since the previous Theorem shows that IDTH and IDTH+ are suitable to reduce the length of the queues where the packets wait to be transmitted, it suggests that the obtained global scheduler can outperform the performance of the centralised one in terms of packets loss and experienced delay, as will be illustrated by simulations in Section 5.

5 Performance analysis

This section investigates through simulations the effect of the integration of IDTH and IDTH+ with the reference scheduler and WCBS, (described in Section 2), chosen as example of a centralised scheduler. In particular since WCBS, being based on EDF, it is underperforming with respect to the reference one in the case of VBR traffic, it is suitable to highlight if IDTH and IDTH+ are able to improve network performance. Indeed, although the performance of the global schedulers, obtained integrating IDTH and IDTH+, are biased by that of the centralised one, the illustrated results allow to understand the effects of the proposed reclaiming mechanisms and how and where they improve the global scheduling behaviour. Moreover IDTH+ is evaluated with respect to IDTH to highlight the effects of its modifications in the scheduling rules.

In the simulation scenario it is assumed that the stations communicate directly without hidden node problem, and RTS/CTS mechanism, MAC level fragmentation and multirate support are disabled. IEEE 802.11g Physical layer, with the mandatory *Orthogonal Frequency-Division Multiplexing* modulation, is implemented, see Table 2 for MAC and PHI parameters.

Table 2 MAC/PHY simulation parameters

Parameter	*Value*	*Parameter*	*Value*
SIFS (μs)	10	PLCP header (b)	24
DIFS (μs)	28	Preamble (b)	72
PIFS (μs)	19	Data rate ($Mbit/s$)	54
Slot time (μs)	9	Basic rate ($Mbit/s$)	1

In NS-2 network simulator (1996) independent replications of 700 s with a warm-up time of 100 s are run until the 95% of the confidence interval for each measure. To analyse the effect of IDTH and IDTH+ on the management of different types of traffic (CBR and VBR) and to stress the schedulers with an increasing variable data rate, the network scenario is composed by one station with G.729A VoIP traffic (see Table 3), three stations with video streaming high quality MPEG4 trace files of 60 minutes extracted by the Video Trace Library (2005) video traces archive, and one with a video conference pre-encoded trace file *LectureHQ-Reisslein*, whose parameters are listed in Table 4. Lastly one backlogged data station sends Service Data Unit of 1500 bytes using IEEE 802.11 Distributed Coordination Function that performs a best effort service. In the following transmission queues length, mean access delay and packets drop rate are chosen as performance metrics since they are strictly related to the accorded transmission time duration. Indeed, if queues length is affected by the introduction of the presented algorithms, as studied in Section 4, then reasonably the packets drops rate, i.e., the rate of discarding packets from the queues owing to the expiration of their delay bound, is influenced as well as the access delay experienced by the TSs. The displayed results try to clarify this aspect.

Table 3 G.729A VoIP parameters

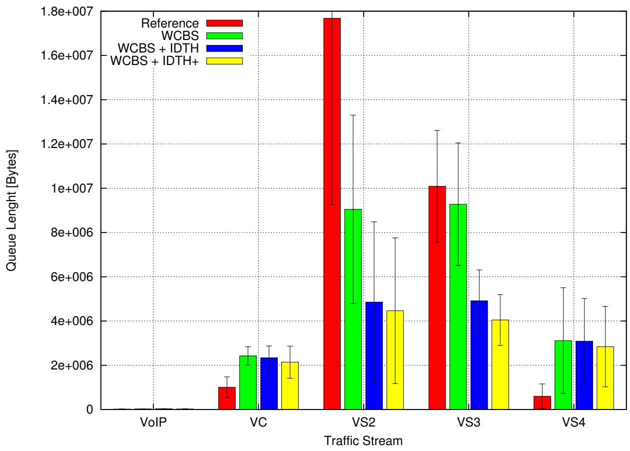
Parameter	*Value*	*Parameter*	*Value*
Frame size *(B)*	10	Payload size *(B)*	20
Frames per packet	2	IP/UDP/RTP	
Period *(s)*	0.02	Header size *(B)*	40
Data rate *(kb/s)*	24	SDU size *(B)*	60

Table 4 Video streaming and video conference parameters

Parameter	*VC*	*VS2*	*VS3*	*VS4*
Mean frame size (B)	600	2900	2900	3500
Max. frame size (B)	11386	22239	15251	16960
Period (s)	0.033	0.040	0.040	0.040
Mean data rate (Kb/s)	158	580	580	700
Max. data rate (Kb/s)	2733	4400	3100	3400

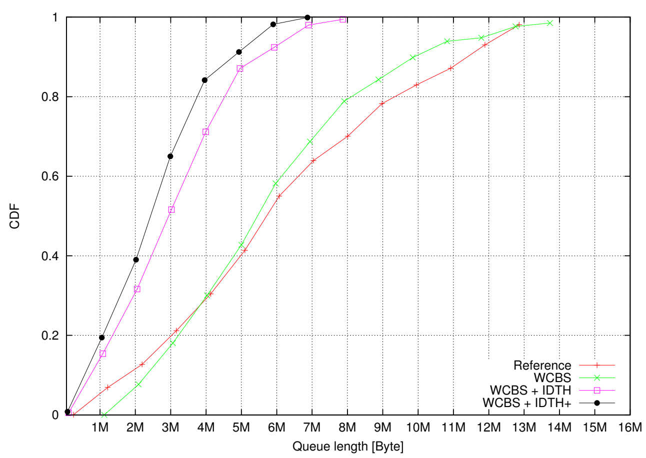
Figure 6 shows the 99th percentile of the transmission buffer length for reference, WCBS, WCBS combined with IDTH and IDTH+ schedulers. Compared to the reference scheduler WCBS is underperforming when serving TSs with less variable data rate (VC and VS4), as explained in Cecchetti and Ruscelli (2011). As highlighted at the beginning of this section the performance of the global schedulers, obtained by the integration of IDTH and IDTH+, are biased by that of the underlying main scheduler, whose behaviour they tune. The centralised algorithm maintains the management of polling period and transmission time allocation, according to its preexistent scheduling rules. This means that the ‘substrate’ where IDTH and IDTH+ work are the performance reached by the centralised algorithm, here WCBS, and starting from that they try to improve the scheduling behaviour. Furthermore this consideration implies that IDTH and IDTH+ can be evaluated and their impact perceived only in comparison with that of the considered centralised scheduler. In the case of queues length, the integration of IDTH and IDTH+ cannot fully compensate the misbehaviour of WCBS owing to the postponing deadlines mechanism that delays the polling and, consequently, increases the queues length. This is particularly evident in the case of low VBR TSs, such as VoIP and VS4 stations. Indeed in this case, since there are low variations in data rate, the reclaiming mechanism is activated rarely, only when data rate diverges from its mean value, used during the admission control, but, owing to the limited amplitude of variations, the recovered resources are poor and, then, the effect of IDTH and IDTH+ is limited. On the other hand, the considered figure highlights the regions where IDTH and IDTH+ work better, i.e., where they give their best: in the illustrated result with VS2 and VS3, that are characterised by high VBR. In this situation the collaboration with IDTH and IDTH+ allows the centralised scheduler to follow more closely the traffic variability, reducing the queues length, as proven in Theorem 5. Indeed, with high VBR TSs the variations of data rate with respect to its mean values are frequent and can be considerable, hence the amount of recovered resources can be greater and the frequency of recovery higher than in the case of low VBR traffic. Looking at the simulation results, IDTH allows a reduction of buffer length of VS2 about 72% in comparison with the reference scheduler, and about 50% in comparison with WCBS. Moreover, as far as IDTH+ is concerned, it is able to outperform IDTH thanks to its mechanism suitable to speed up the scheduling reactivity stressed by the rapid changes of data rate. In particular, it allows a further improvement of about 75% in comparison with the reference scheduler with an increase of 3% point, and of about 12% with respect to IDTH. In the case of VS3, the improvement of IDTH+ in comparison with the reference is about 75%, and with WCBS about 63%; finally, the difference with IDTH is greater, about 20%.

Figure 6 Ninety-ninth percentile of transmission buffer length (see online version for colours)



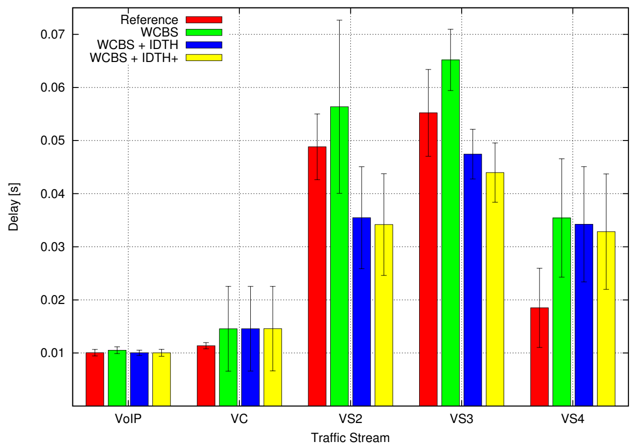
These deductions are confirmed by the analysis of the *cumulative distribution function* (CDF) of mean access delay in the case of VS3, illustrated in Figure 7: when integrated with WCBS, IDTH has 50% of probability to halve from 6 MB to 3 MB the queues length of reference, and from 5,5 MB in the case of WCBS. From the same reasons than above, IDTH+ is able to provide a further reduction until 2,5 MB, with an improvement of about 17% in comparison with IDTH.

Figure 7 CDF of buffer length in the case of VS3 TS (see online version for colours)



These results corroborate the advantage of the introduction of IDTH+ as improvement of IDTH, and this is particularly significant taking into account that the modification is only about the conditions applied to the recovered resource allocation, without any additional computational load. Figure 8 illustrates the impact on mean access delay, defined as the interval between the time when a packet reaches the MAC level and when the corresponding `ACK` is received. In general, the injection of resources of IDTH and IDTH+ reduces the mean access delay with respect to WCBS and reference schedulers, confirming the analytical deductions in Section 4. As in the case of queues length and for the same motivations, this effect is more relevant with highly VBR TSs, where, owing to the large variations of the recovered time, the current T' can be greater than T_{AC} used by WCBS and reference schedulers. For instance, in the case of VS2 WCBS is underperforming with respect to the reference scheduler of about 12%. Applying IDTH the situation is reversed with a reduction of the mean access delay of about 29% in comparison with the reference scheduler. Moreover, IDTH+ is able to refine this effect allowing a further improvement of 3% of performance obtained by IDTH. In the case of VS3 WCBS underperforms the reference scheduler of about 25%, whereas IDTH outperforms the reference of 15% and WCBS of 28%; IDTH+ reduces the mean access delay of reference of about 19% and of WCBS of about 31%.

Figure 8 Mean access delay (see online version for colours)



The CDF of the access delay experienced by VS3 TS, shown in Figure 9, confirms these considerations: IDTH is suitable to reduce the mean delay especially in the case of highly VBR TSs, with 50% of probability to halve from 65 ms of reference and 0.75 ms of WCBS to 42 ms. From the same reasons than above, IDTH+ is able to provide a further reduction until 36 ms. Finally, Figure 10 illustrates the number of packets per second discarded from the queues when, in ideal conditions, their waiting time overcomes their delay bound. Since the proposed algorithms allow to reduce the time a packet remains into the transmission queue decreasing the mean delay, the discarded rate is reduced, especially in the case of high VBR TSs, where simulation results have shown the greater improvement: IDTH is able to reduce over half the number of discarded packets of VS2 and VS3. With the integration of IDTH+ the reduction reaches about 59%. The illustrated results show that IDTH and IDTH+ are a useful and efficient mechanism to improve the behaviour of a centralised scheduler, filling into the gap of the provided QoS in presence of VBR streams. Furthermore they highlight how IDTH+, using a simple variation, is suitable to tune the behaviour of IDTH, overcoming its performance.

Figure 9 CDF of access delay of VS3 traffic stream (see online version for colours)

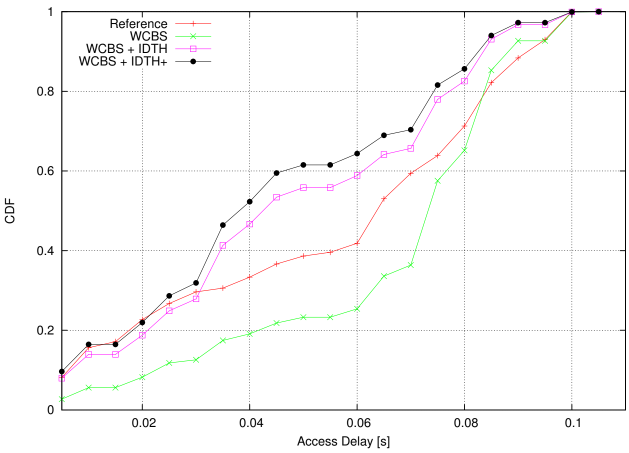
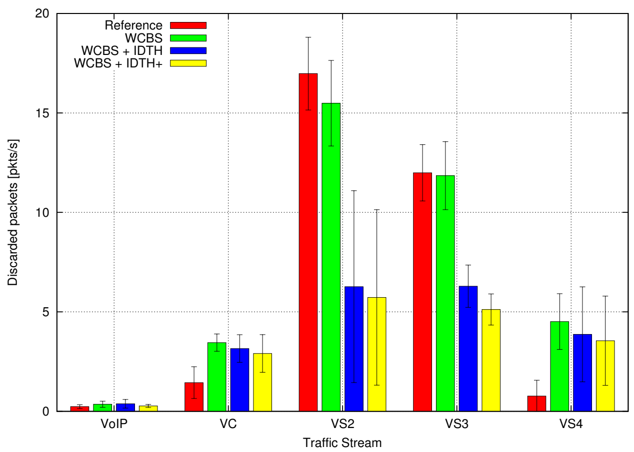


Figure 10 Packets discard rate (see online version for colours)



6 Conclusions

In this paper the problem of improving the performance of centralised QoS schedulers in the case of VBR TSs is investigated. *immediate dynamic TXOP HCCA* (IDTH) scheduler is deeply analysed and its refined version *immediate dynamic TXOP HCCA plus* (IDTH+) is introduced. They are studied both theoretically, using the imprecise computation model, than by simulation showing that their integration with a centralised scheduler is beneficial in general, and in particular in the case of schedulers underperforming with highly VBR traffic. IDTH and IDTH+ provide an instantaneous overprovisioning without a recomputation of transmission time and are suitable to take advantages from the statistical multiplexing approach in a 'light' way, facing off the waste of resources typical of algorithms based on a fixed preallocation of resources. Indeed the global scheduler, obtained integrating IDTH and IDTH+, is able to provide an increased service rate and experiences improved performance in terms of transmission queues length, mean access delay and packets drop rate, that are achieved with an *O(1)* computational complexity cost. In particular, as far as IDTH+ is concerned, it is suitable to increase the responsiveness of IDTH in the case of rapid changes of data rate, allowing a further performance improvement by means of a simple algorithm modification.

References

Al-Maqri, M., Othman, M., Ali, B. and Hanapi, Z. (2013) 'Adaptive TXOP assignment for QoS support of video traffic in IEEE 802.11e networks', *Proceedings of IEEE RFM', Penang, Malaysia*, pp.144–149.

Ansel, P., Ni, Q. and Turletti, T. (2006) 'FHCF: a simple and efficient scheduling scheme for IEEE 802.11e wireless LAN', *Mobile Networks and Applications*, Vol. 11, No. 3, pp.391–403.

Arora, A., Yoon, S-G., Choi, Y-J. and Bahk, S. (2010) 'Adaptive TXOP allocation based on channel conditions and traffic requirements in IEEE 802.11e networks', *IEEE Trans. Vehic. Tech.*, Vol. 59, No. 3, pp.1087–1099.

Baker, T.P. (1991) 'Stack-based scheduling for real-time processes', *Real-Time Syst.*, Vol. 3, No. 1, pp.67–99.

Boggia, G., Camarda, P., Grieco, L.G. and Mascolo, S. (2007) 'Feedback-based control for providing real-time services with the IEEE 802.11e MAC', *IEEE/ACM Trans. Net.*, Vol. 2, No. 15, pp.323–333.

Cecchetti, G. and Ruscelli, A. L. (2008) 'Performance evaluation of Real-Time schedulers for HCCA function in IEEE 802.11e wireless networks', *Proceedings of ACM Q2SWinet*, Vancouver, Canada.

Cecchetti, G. and Ruscelli, A.L. (2011) 'Real-time support for HCCA function in IEEE 802.11e networks: a performance evaluation', *Wiley Security and Communication Networks*, Vol. 4, No. 3, pp.299–315. Special Issue on Security for QoS Assured Wireless Networks.

Cecchetti, G., Ruscelli, A.L., Mastropaolo, A. and Lipari, G. (2012a) 'Dynamic TXOP HCCA reclaiming scheduler with transmission time estimation for IEEE 802.11e real-time networks', *Proceedings of ACM MSWiM*, Paphos, Cyprus.

Cecchetti, G., Ruscelli, A.L., Mastropaolo, A. and Lipari, G. (2012b) 'Providing variable TXOP for IEEE 802.11e HCCA real-time networks', *Proceedings of IEEE WCNC*, Paris, France.

Cicconetti, C., Lenzini, L., Mingozzi, E. and Stea, G. (2007a) 'Design and performance analysis of the Real-Time HCCA scheduler for IEEE 802.11e WLANs', *Computer Networks*, Vol. 51, No. 9, pp.2311–2325.

Cicconetti, C., Lenzini, L., Mingozzi, E. and Stea, G. (2007b) 'An efficient cross layer scheduler for multimedia traffic in wireless local area networks with IEEE 802.11e HCCA', *ACM Mob. Comput. and Commun. Reviews*, Vol. 11, No. 3, pp.31–46.

Cowling, J. and Selvakennedy, S. (2004) 'A detailed investigation of the IEEE 802.11e HCF reference scheduler for VBR traffic', *Proceedings of 13th ICC*, Paris, France, pp.453–459.

Fan, Y. and Huang, C. (2005) 'Real-time traffic scheduling algorithm in WLAN', *Proceedings of 4G Mobile Forum (4GMF)*, 7–9 July, San Diego, USA.

Fattah, H. and Leung, C. (2002) 'An overview of scheduling algorithms in wireless multimedia networks', *IEEE Wireless*, Vol. 9, No. 5, pp.76–83.

Ferrari, D. and Verma, D. (1990) 'A scheme for real-time channel establishment in wide-area networks', *IEEE JSAC*, Vol. 8, No. 3, pp.368–379.

Grilo, A. and Nunes, M. (2002) 'Performance evaluation of IEEE 802.11e, *Proceedings of 13th IEEE PIMRC*, Vol. 1, pp.511–517.

Grilo, A., Macedo, M. and Nunes, M. (2001) 'A service discipline for support of IP QoS in IEEE 802.11 networks', *Proceedings of PWC*, Laapenranta, Finland.

Grilo, A., Macedo, M. and Nunes, M. (2003) 'A scheduling algorithm for QoS support in IEEE 802.11e networks', *IEEE Wireless Comm.*, Vol. 10, No. 3, pp.36–43.

Grow, R. (1982) 'A Timed-Token protocol for local area networks', *Proceedings of Electro/82, Number Paper 17/3, 'Token Access Protocols*, I. Electronic Conventions, El Segundo, CA.

Hayajneh, T. and Al-Mashaqbeh, G. (2014) 'Multimedia traffic over wlans: Qos support and performance evaluation', *5th IEEE International Conference on Information and Communication Systems (ICICS)*, Irbid, Jorndan.

IEEE802.11 (2007) *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification*, http://standards.ieee.org/getieee802/download/802.11-2007.pdf

Inanc, I., Keceli, F. and Ayanoglu, E. (2006) 'An adaptive multimedia QoS scheduler for IEEE 802.11e wireless LANs', *Proceedings of IEEE ICC*, Istanbul, Turkey.

Jackson, J. R. (1955) *Scheduling a Production Line to Minimize Maximum Tardiness*, Research Report 43, University of California.

Ju, K. and Chung, K. (2013) 'Dynamic TXOP allocation for multimedia QoS providing over wireless networks', *Proceedings of IEEE International Conference on Information Networking (ICOIN)*, Bangkok, Thailand.

Kuan, C. and Dimyati, K. (2007) 'Utilization model for HCCA/EDCA mixed mode in IEEE 802.11e', *ETRI Journal*, Vol. 29, No. 6, pp.829–831.

Lai, W.K., Shien, C. and Jiang, C. (2009) 'Adaptation of HCCA/EDCA ratio in IEEE 802.11 for improved system performance', *Int. J. on Innovative Computing, Inf. and Contr.*, Vol. 5, No. 11(B), pp.4177–4188.

Larcheri, P. and LoCigno, R. (2006) 'Scheduling in 802.11e: Open-loop or closed-loop?', *Proceedings of IFIP WONS*, Les Ménuires, France.

Lee, J., Shin, I., Park, G., Song, W., Kim, J. and Hong, J. (2007) 'An efficient bandwidth reclaim scheme for the integrated transmission of real-time and non real-time messages on the WLAN', *Proceedings of ICCS*, Vol. 4490, Beijing, China, pp.925–932.

Lee, K., Seong, K. and Ryu, W. (2011) 'Efficient QoS scheduling algorithm for multimedia services in IEEE 802.11e WLAN', *Proceedings of ICC*, Kyoto, Japan.

Lin, K., Natarjan, S. and Liu, J. (1987) 'Concord: a system of imprecise computation', *Proceedings of IEEE COMPSAC*, Tokyo, Japan.

Liu, C.L. and Layland, J.W. (1973) 'Scheduling algorithms for multiprogramming in a hard-real-time environment', *J. ACM*, Vol. 20, pp.46–61.

Liu, J., Lin, K. and Natarjan, S. (1987) 'Scheduling real-time, periodic jobs using imprecise results', *Proceedings of IEEE RTSS*, San Jose, California.

Liu, J., Lin, K., Shih, A., Chung, C., Yao, J. and Zhao, W. (1991) 'Algorithms for scheduling imprecise computations', *IEEE Computer*, Vol. 24, No. 5, pp.58–68.

LoCigno, R., Palopoli, L. and Colombo, A. (2007) 'Analysis of different scheduling strategies in 802.11e networks with multi-class traffic', *Proceedings of IEEE LCN*, Dublin, Ireland.

Lu, S., Bharghavan, V. and Srikant, R. (1999) 'Fair scheduling in wireless packet networks', *IEEE/ACM Trans. Net.*, Vol. 7, No. 4, pp.473–489.

Mangold, S., Choi, S., May, P., Klein, O., Hiertz, G. and Stibor, L. (2002) 'IEEE 802.11e wireless LAN for quality of service', *Proceedings of European Wireless*, Vol. 1, Florence, Italy, pp.32–39.

NS-2 network simulator (1996) http://www.isi.edu/nsnam/ns/

Pastrav, A., Puschita, E. and Palade, T. (2012) 'HCCA support in IEEE 802.11 networks', *Proceedings of Int. Symposium on Electronic and Telecommunications*, Timisoara, Romania.

Ruscelli, A.L. and Cecchetti, G. (2014) 'A IEEE 802.11e HCCA scheduler with a reclaiming mechanism for multimedia applications', *Advances in Multimedia*, Vol. 2014, Article ID 372693, 22 pages.

Ruscelli, A.L., Cecchetti, G., Alifano, A. and Lipari, G. (2012) 'Enhancement of QoS support of HCCA schedulers using EDCA function in IEEE 802.11e networks', *Ad Hoc Networks*, Vol. 10, No. 2, pp.147–161. Special Issue on Recent Advances in Analysis and Deployment of IEEE 802.11e and IEEE 802.11p Protocol Families.

Ruscelli, A.L., Cecchetti, G., Mastropaolo, A. and Lipari, G. (2011) 'A greedy reclaiming scheduler for IEEE 802.11e HCCA networks', *Proceedings of ACM MSWiM*, Miami Beach, Florida.

Saheb, S., Bhattacharjee, A., Dharmasa, P. and Kar, R. (2012) 'Enhancement hybrid coordination function controlled channel access-based adaptive scheduler for delay sensitive traffic in IEEE 802.11e networks', *IET Networks*, Vol. 1, No. 4, pp.281–288.

Siddique, M., Wenning, B., Timm-Giel, A., Görg, C. and Mühleisen (2010) 'Generic spectrum sharing method applied to IEEE 802.11e WLANs', *Proceedings of Sixth Advanced International Conference on Telecommunications (AICT)*, Barcelona, Spain.

Siris, V. A. and Courcoubetis, C. (2006) 'Resource control for the EDCA and HCCA mechanisms in IEEE 802.11e networks', *Proceedings of WiOpt', Boston, Massachusetts*, pp.22–27.

Skyrianoglou, D., Passas, N. and Salkintzis, A. K. (2006) 'ARROW: an efficient traffic scheduling algorithm for IEEE 802.11e HCCA', *IEEE Transaction on Wireless Communications*, Vol. 5, No. 12, pp.3558–3567.

Tsao, S-L. (2000) 'Extending earliest due date scheduling algorithms for wireless networks with location-dependent errors', *Proceedings of 52nd IEEE VTC-Fall*, Vol. 1, Boston, Massachusetts, pp.223–228.

Video Trace Library (2005) http://trace.eas.asu.edu/