Abstract: This paper proposes a new real-time communication scheme for IEEE 802.11e HCCA networks. The proposed communication scheme is called Group Sequential Communication (GSC) and it aims the reduction of the polling overhead associated to the real-time message transmission when using the HCCA function. The GSC scheme partially eliminates the polling inefficiency, by means of a virtual token passing procedure among polled stations. Thus, it reduces the number of exchanged messages between the HCCA controller and the polled stations. The GSC uses a real-time group concept, where the real-time members of the group are granted a sequential access to the communication medium.

Keywords: Real-Time Communication, IEEE 802.11e HCCA function

1. INTRODUCTION

Currently, the industrial communication infrastructure is partially moving from wired to wireless environments (Willig et al., 2005). It is expected that in the near future wireless communication will play a major role in the overall industrial environments.

According to Moyne and Tilbury (Moyne and Tilbury, 2007), Ethernet and WiFi networks will continue to grab larger and larger shares of the industrial networks installed base, driven largely by lower cost through volume, the spreading use of Internet protocols, the higher availability of solutions and tools for these network types, etc. However, it is well known that timing constraints imposed by office or domestic environments are significantly different from those found in traditional industrial environments. Additionally, there is a recent trend towards the use of the same industrial communication infrastructure to transfer real-time control traffic, multimedia traffic and background traffic (Sauter and Vasques, 2006). This is specially true in wireless environments. That is, the same communication infrastructure must be able to handle traffic with contradictory requirements, in what concerns reliability and timing requirements.

Consequently, multiple approaches and techniques are being developed to enable the support of real-time communication services upon wireless communication environments. Within this context, the IEEE 802.11 family of protocols is one of the main wireless contenders for supporting industrial communications. One of the main reasons for is that this family of protocols is easily able to replace Ethernet in a transparent way, implementing the two lowest layers of the ISO/OSI model, the IEEE 802.11 protocol provides all the required functionalities to enable the support of
the Internet Protocol (IP), that is virtually the basis for applications over Ethernet networks.

The IEEE 802.11 protocol was standardized in 1999 by the IEEE as the IEEE 802.11 standard, which was later reaffirmed in 2003 (IEEE Std 802.11, 2003). The basic service set (BSS) is the building block of an IEEE 802.11 WLAN, which actually provides two types of configurations: independent BSS (IBSS) and infrastructure. The IBSS is the most basic type for an IEEE 802.11 WLAN, which may be composed of, at least, two stations. This operation mode is often referred to as ad hoc. The infrastructure mode includes one or more access points (AP) that convey the communication among wireless stations.

The IEEE 802.11e (IEEE Std 802.11e, 2005) standard has been published as an amendment to the original standard. This amendment is intended to provide differentiated levels of Quality of Service (QoS) to the supported applications, including voice and video over WLANs. It incorporates an additional coordination function called hybrid coordination function (HCF), that is only used in QoS network configurations.

The HCF functions provide two mechanisms to support applications with QoS requirements: the Enhanced Distributed Channel Access (EDCA), which delivers traffic based on differentiating user priorities (UPs) and; the Hybrid Coordination Function (HCF) Controlled Channel Access (HCCA), which allows the reservation of transmission opportunities (TXOPs) with the hybrid coordinator (HC). The TXOP is the time interval during which the station keeps the medium access control. Consequently, a station can transmit multiple frames within an acquired TXOP.

The HCCA mechanism operates in the infrastructure mode was proposed to guarantee upper bounded medium access delays and, it was however, some preliminary research studies (Casetti et al., 2005; Garg et al., 2003) have already reported that the original HCCA mechanism may not be adequate to support real-time communication, mainly due to the high polling overhead caused by stations that do not have messages to be transmitted. Nevertheless, we believe that an improved HCCA mechanism will have the potential to provide an adequate real-time communication service. The major challenge concerning the improvement of the original HCCA mechanism is the development of adequate scheduling algorithms, enabling the provision of efficient RT communication services.

In this paper, a new communication scheme called “Group Sequential Communication (GSC)” is proposed. The GSC scheme reduces the HCCA polling overhead when dealing with typical industrial communication scenarios. It compels the sequential transfer of messages from the participants of a real-time group (RT-group), without the traditionally required polling overhead. Whenever a RT-group transfer is triggered, the RT-group participants serialize the transfer of their messages without the need of any extra message exchange. This serialization is achieved by means of a Virtual Token Passing (VTP) among the RT-group members. The VTP approach was firstly proposed by Carreiro et al. (Carreiro et al., 2005) as a field level communication scheme for an Ethernet-based industrial communication architecture. Afterwards, the basic idea of the VTP approach was extended to support real-time communication services in IEEE 802.11/802.11e networks operating upon both DCF and EDCA mechanisms (Moraes et al., 2007).

The remainder of this paper is organized as follows. Section 2 briefly introduces the IEEE 802.11e medium access control protocol. Section 3 describes the state-of-the-art for real-time communication in IEEE 802.11e networks. Section 4 shows how the proposed GSC mechanism works. Section 5 adapts the GSC mechanism to deal with typical Industrial Communication Scenarios. Afterwards, in section 6 both the admission control and the worst-case timing analysis are presented. Finally, in section 7 some conclusions are drawn.

2. IEEE 802.11E STANDARD

The IEEE 802.11 Medium Access Control (MAC) protocol implements two complementary MAC mechanisms. The Distributed Coordination Function (DCF) is the basic access control mechanism of IEEE 802.11. It is based on Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA). The Point Coordination Function (PCF) is a polling protocol that allows access to a free contention environment. The PCF mechanism was proposed to support the real-time traffic transfer. However, most part of the WLAN network cards never actually implemented the PCF mechanism, due to complexity reasons.

The quest for real-time communication in IEEE 802.11 networks lead to the establishment of the IEEE 802.11e task group in July 1999. In December 2005, the Task Group E published the IEEE 802.11e amendment (IEEE Std 802.11e, 2005). This amendment targets real-time communication in wireless networks. It incorporates an additional coordination function called hybrid coordination function (HCF), that is only used in QoS network configurations.

As mentioned before, the HCF provides two mechanisms for supporting applications with QoS requirements: the EDCA that delivers traffic based
Fig. 1. Example of CFP repetition interval.

The HC is able to capture the control of the wireless medium, as it waits a shorter time between transmissions than stations using the EDCA or DCF mechanisms. The HC may include a Contention Free (CF) parameter element in the Beacon frame, which informs all stations to set their Network Allocation Vectors (NAVs) to the end of the controlled phase. During the CFP, the HC controls the access to the channel by polling all the stations in the polling list. For each polled station it is granted a transmission opportunity (TXOP) to transfer its own traffic. On the other hand, the HC is also allowed to start a TXOP in the CP immediately after the channel to be idle for one PIFS (PCF Interframe Space) period, the called Controlled Access Phase (CAP). A CAP ends when the HC does not reclaim the channel after a duration of PIFS, after the end of a TXOP. However, the QoS stations (QSTAs) must send QoS reservation requests using the special QoS management frame, called Traffic Specification (TSPEC). The TSPEC frame contains the set of parameters that define the QoS characteristics of the network traffic that may be needed, or may be available, for any particular instance of parameterized QoS traffic. Such fields are set to zero if there are no specified parameter values. Figure 1 illustrates an example of the CFP repetition interval.

Similarly to the PCF scheme, the HC also polls all the stations in the polling list, even though some stations may have no messages to transmit. When the HC polls a station that has no packets to transfer, the station will transmit a null frame, after the QoS CF-poll. Thus, the polling overhead is roughly equal to the time interval from sending the polling frame till the end of the ACK frame (Son et al., 2005). As a consequence, there may be a low channel utilization and a high implementation overhead (Lim et al., 2004).

3. REVIEW OF PREVIOUS RELEVANT WORKS

In this section, the most relevant proposals to improve real-time traffic support in IEEE 802.11e networks are presented. Although this standard supports both ad hoc and infrastructure modes, only infrastructure-based approaches are presented, as the focus of this work is on infrastructure-based networks.

The state-of-the-art communication approaches can be classified in two main groups: the IEEE 802.11e compliant solutions and the IEEE 802.11e non-compliant solutions. The latter group presents some interesting ideas but are not economically viable today. Moreover, it is not realistic to assume the existence of IEEE 802.11 free environments, which impair the use of non-compliant solutions. Therefore, we will only analyze IEEE 802.11e compliant solutions. Such IEEE 802.11e compliant solutions can still be divided into two sub-groups: solutions that use the original HCCA polling scheme and solutions that modify or enhance the HCCA polling scheme.

Solutions that use the original HCCA polling mechanism suffer the consequence of the polling overheads. Therefore, multiple state-of-the-art works propose new scheduling algorithms to improve the HCCA performance for constant bit rate (CBR) and variable bit rate (VBR) multimedia application (Voice and Video).

Yang, in (Yang, 2005), proposes an enhanced HCCA scheme (E-HCCA), based on prediction for real-time multimedia communication. E-HCCA balances the delay and the queue length of TS (Traffic Stream) for CBR and VBR traffic flows in WLAN. The scheme evaluates the mean applications data rate during a TXOP and predicts the data rate in the next round TXOP.
grates the benefits of both EDCA mode and HCF polling mode. The network capacity is effectively increased to better support real-time multimedia communication.

Grilo, Macedo and Nunes in (Grilo et al., 2003), propose a scheduling algorithm for the IEEE 802.11e Hybrid Coordination Function. The proposed algorithm is compatible with the link adaptation mechanisms implemented in commercial WLANs, as it bounds the amount of time during which each station may control the wireless medium access. The performance of the algorithm is evaluated through computer simulations and compared with the original IEEE 802.11e scheduler.

Rashid, Hossain and Bhargava in (Rashid et al., 2006), introduce a novel queueing analytic framework that is useful to analyze the performance of the HCCA functions when supporting VBR traffic applications. The proposed frameworks is useful to improve the HCCA scheduler and admission controller designs.

Noh, Suzuki and Tasaka in (Noh et al., 2006), propose a new scheduling scheme for audio-video transmission with HCCA. In the proposed scheduling scheme, the HC firstly calculates the TXOP duration in a service interval for each wireless station on the basis of its mean data rate; it then adds additional TXOPs for each wireless station which had queued audio or video packets at the end of the previous TXOP. Simulation results show that the proposed scheduling scheme enhances the video quality and inter-stream media synchronization quality under low traffic conditions.

Son, Lee, Yoo and Park (Son et al., 2005), propose an effective and simple polling scheme, that reduces the number of polling cycles for stations that have no packets to transmit. The proposed scheme uses two parameter values: the number of valid data transmission and the schedule periodicity. When the polled station transmits a data holding packet, the station becomes more frequently polled by reducing the polling period. If the polled station transmits a null data packet, the polling period of the station is increased to the maximum polling period according to the data characteristics. The simulation results show that the throughput increase 35.8% when fifteen out of thirty stations have packets to transmit. If all the polled stations have messages to transfer, its behavior is similar to the original HCCA scheme. That is, it keeps the same overhead problems, when dealing with typical industrial applications (small packet sizes).

Lo, Lee and Chen (Lo et al., 2003) designed a multipolling mechanism called Contention Period Multipoll (CP-Multipoll), which incorporates the DCF access scheme into the polling scheme. This approach uses different backoff values for the multiple message streams in the polling group. The contending order of polled stations is the same as the ascending order of the assigned backoff values. The station in the polling list initializes its transmission immediately after receiving the CP-Multipoll frame. This action avoids the interference from other stations performing the backoff procedures in the DCF mode. The main advantage of CP-multipoll proposal is the high channel utilization and low implementation overhead.

4. THE GROUP SEQUENTIAL COMMUNICATION (GSC) PROPOSAL

In this section, we propose the Group Sequential Communication (GSC) scheme. The GSC scheme intends to partially eliminate the HCCA polling inefficiency when dealing with typical industrial applications (small packet sizes). It considers the use of a virtual token passing procedure among polled stations. Thus, it reduces the number of exchanged messages between the HCCA controller and the polled stations.

4.1 Rationale

The GSC scheduler scheme operates during the CFP period. It does not use a polling scheme to grant TXOPs to each of the real-time (RT) stations previously registered in a RT-group. Instead, it sequentially grants the medium access right to each of the RT-group stations through a Virtual Token Passing (VTP) procedure. Once the schedule of a RT-group is triggered, a VTP procedure is started among the RT-group members (hereafter also referred as GSC stations). This procedure serializes the message transfers from the RT-group members, without the need to exchange any extra polling message between the HC and the RT-group stations. The VTP procedure uses a set of local counters that are incremented at specific slot boundaries. Such local counters implement a distributed variable that defines which station has granted the medium access. Whenever a GSC station does not have a frame to be transmitted, it immediately transfers the medium access right to the next RT-group station without wasting bandwidth with extra message exchanges. It is worth noting that the original HCCA mechanism transmits a null-frame whenever a polled station does not have any message ready to be transmitted.
4.2 Specification of the GSC procedure

The GSC procedure considers a sequence group \( SG \) (RT-group) with \( np \) members. The membership is represented as \( L = \{GI_1, GI_2, ..., GI_{np}\} \), where \( GI_i \) is used as station identification (ID) in this paper. The GSC procedure circulates a virtual token in \( L \). Specifically, all members of group \( SG \) maintain a local Sequence Counter (SC) that are images of the distributed variable \( SC \). The GSC station \( GI_i \) captures the virtual token when \( SC \) equals \( GI_i \).

The GSC procedure (Figure 2) works as follows. In the beginning of the CFP repetition interval, which is also called Service Interval (SI), the HC sends a Beacon frame. Such a Beacon frame sets all NAVs to the end of the controlled phase and also sets the value of the distributed variable \( SC \) to 1. This Beacon frame uses information from the TSPEC management frames. Whenever a GSC station has a packet ready to be transferred and its image of \( SC \) \((SC_i)\) is equals its ID \((GI_i)\), then it is authorized to transfer real-time messages immediately after the medium being idle during at least SIFS time \((aSIFS\_time)\). Conversely, if the GSC station holding the token does not have any RT message to transmit, the SC value will be incremented after \( aSlot\_time \) in all the GSC stations, allowing the next GSC station to transfer its messages. This operation is repeated successively until the end of all GSC stations. More precisely, whenever the GSC station holding the token transmits its messages during its reserved TXOP interval, each GSC station will increment its local \( SC \) value in the first slot time (ST) following SIFS. Conversely, whenever the GSC stations do not have messages to transfer, each GSC station will also increment its local \( SC \) value in the first available ST following an initial SIFS interval. Afterwards, at the beginning of a new SI interval, another Beacon frame will be sent to set all the parameter values of the GSC group.

Figure 3 illustrates the behavior of the GSC scheduler during a CFP repetition interval. In this example, it is considered a \( SG \) group with 4 stations \((np = 4)\).

According to the GSC scheme, whenever the HC transmits a Beacon frame, station \( GI_1 \) must start a transmission just after the medium being idle during SIFS time interval in order to grant the medium access. However, if the station \( GI_1 \) does not start a transmission at this time, the HC will send another Beacon frame in the next slot time, i.e., at instant \( aSIFS\_time + aSlot\_time \). This operation will be repeated up to the maximum number of retransmissions attempts. If there is no answer from station \( GI_1 \), the RT-group will not be able to transfer its own messages. Most likely station \( GI_1 \) will be the HC. Considering that all stations received the Beacon frame and station \( GI_1 \) started its transmission, it will be allowed to access the medium during up to its TXOP. At the end of the transmission, all GSC stations increment its local \( SC \) value passing the virtual token to the next station \((GI_2)\) \(^1\). Then, station \( GI_2 \) will transfer its messages, also during the allowed TXOP. Afterwards, station \( GI_3 \) is able to start its transmission. However, considering

\(^1\) The end of a TXOP period can be detected by all stations whenever the medium is detected idle during \( aSIFS\_time + aSlot\_time \).
that station $GI_3$ does not have any message to be transferred the $SC$ counter is incremented in all the GSC stations after one $a$SlotTime time interval. Thus, it allows station $GI_4$ to start its transmission. At the end of the TXOP of station $GI_4$ ($SC > np$), the HC send a CF-end frame resetting all NAVs, thus allowing the initialization of the contention period (CP), where standard EDCA stations can start contending for the medium access.

5. NO-ACK GSC: ADAPTATION OF THE GSC MECHANISM TO DEAL WITH TYPICAL INDUSTRIAL COMMUNICATION

Nowadays, there are typically two types of communication paradigms associated to industrial applications: master-slave and publisher-subscriber. The first paradigm uses unicast mechanisms, where the master always starts the communication to a specific slave and this slave just answers that master questions. The second communication paradigm uses multicast mechanisms, where the publisher component always starts the communication publishing its message upon the network. In such a case, a message is not addressed to any specific host. Instead, each message has an identifier, thus one or more subscriber components can consume it. For example, two of the most popular fieldbus networks, Profibus and Foundation Fieldbus use master-slave and publisher-subscriber approaches, respectively.

The master-slave paradigm is the simplest communication mechanism, where the sender just needs to know the address of the receiver to send its messages. However, this approach is completely centralized and it does not offers multicast/broadcast data exchanges. Due to this aspect, master-slave mechanisms have a severe performance inefficiency with respect to scalability.

![Fig. 4. no-ACK GSC repetition interval.](image)

Conversely, publisher-subscriber approach is more decentralized and flexible than master-slave. Beyond that, its communication mechanisms are intrinsically multicast. One of the main advantages of the publisher-subscriber mechanism is that multicast data transmission mechanisms save bandwidth network. Thus, the publisher-subscriber paradigm is adequate solution to distributed real-time industrial applications, where small packets must be periodically transferred between sensors, controllers and actuators according to strict transfer deadlines.

As a consequence of the publisher-subscriber approach, the use of confirmation messages (ACK) is no longer applicable. Therefore, in order to cope with typical industrial scenarios, the GSC proposal must be used without ACK. The no-ACK GSC scheduling is similar to the previously described GSC scheme. The main difference is that the local values of the $SC$ distributed variable are always incremented after the medium being idle during SIFS time interval. Therefore, the TXOPs must allow the transmission of just one message per GSC station in each SI interval. This scheme is shown in Figure 4. It is easy to conclude that the no-ACK GSC scheduling is simpler than the ACK version and it presents smaller overhead.

6. TIMING ANALYSIS

In this section, both the admission control and the worst-case timing analysis of the GSC scheme are presented. Specifically, schedulability analysis equations for the publisher-subscriber communication paradigm (without ACK) are deduced, where each GSC station transmits its own messages with different data payloads.

The proposed GSC scheme uses the TSPEC as defined in the IEEE 802.11e HCCA standard. The TSPEC is an element sent through a management frame that contains information about the characteristics and QoS expectation of a traffic stream. As mentioned before, similarly to HCCA, it defines two contention periods (Figure 4). The first part is used as a contention-free period by GSC stations that have reserved TXOPs and the second part is used as a contention period for low priority stations (DCF or EDCA stations). The maximum Service Interval (SI) specifies the maximum time interval between the start of two consecutive service periods ($SI = CFP + CP$).

The scheduling and the admission control of a new traffic stream is made by the HC, that evaluates the scheduled SI and the TXOP duration. For the sake of simplicity the TXOP considering that each GSC station transmits only one RT message. Considering that the CFP is given by a percentage $(\alpha)$ of the SI interval and, during a CFP some GSC stations may have no messages to transfer, the CFP can be defined as follows:

$$CFP \leq \alpha \ast SI$$  \hspace{1cm} (1)
The maximum CFP is equal to the sum of all TXOPs assigned to the \( np \) members of the group \( SG \) separated by SIFS intervals. Thus,

\[
CFP_{\text{max}} = \sum_{i=1}^{np} (TXOP(i) + SIFS) + \beta \tag{2}
\]

where \( \beta \) is the maximum time to transfer both the Beacon and the CF-end frames.

Therefore, considering \( np \) admitted traffic streams, a new stream \((np+1)\) can be admitted if it satisfies the following inequality:

\[
np+1 \sum_{i=1}^{np} (TXOP(i) + SIFS) + \beta \leq SI - CP \tag{3}
\]

Equation 3 only checks if there is enough available space in the CFP repetition interval to schedule all the requested traffic streams.

From equations 1 and 2, it is also possible to obtain the following relation:

\[
\alpha \geq \frac{\sum_{i=1}^{np} (TXOP(i) + SIFS) + \beta}{SI} \tag{4}
\]

Similarly to the HCCA mechanism, the GSC proposal also allows that standard DCF or EDCA stations to contend for the medium access during the CP interval. Therefore, there will be an induced jitter, as the beacon frame that triggers the transfer of the RT-group must wait that the transmission medium becomes free.

An important assumption that must be considered to determine the maximum jitter of the CFP repetition interval \( SI_{\text{max}} \) is that the wireless communication medium is essentially an open communication environment. That is, any new participant can try to access the communication medium at any instant (according to the MAC rules) and establish its own communication channels. Therefore, the worst-case jitter will occur whenever a default station acquires the transmission medium and uses all the allowed TXOP time immediately before the end of the CP interval. Such occurrence will cause the maximum delay in the beginning of the next SI (Figure 5).

Therefore, the longest CFP repetition interval \( SI_{\text{max}} \) is given by:

\[
SI_{\text{max}} = SI + TXOP_{\text{max}} + PIFS \tag{5}
\]

where \( TXOP_{\text{max}} \) is the maximum TXOP interval defined for the set of standard EDCA stations.

Analyzing equations (2) and (3), it is possible to verify that the maximum number of real-time stations \((np)\) admitted by a GSC scheduler increases linearly with \( \alpha \) and \( SI \). On other a hand, the increase of the \( \alpha \) parameter has a significative impact over the no real-time traffic, as the \( \alpha \) parameter is a percentage of the transmission bandwidth.

Considering that \( \alpha \) values up to 0.4 can be easily achievable, then for a typical industrial scenario it is possible to support up to 383 real-time stations \((np)\), when \( SI \) is 50ms and each real-time station sends only one packet with 64 bytes per cycle, for an error free channel with a transmission rate of 36 Mbps (PHY), in accordance with (802.11a, 1999).

If the supported applications allow the use of 100ms periodicity for SI, then it is possible to support the same \( np \) with 0.2 for the \( \alpha \) parameter. For applications whose dynamics are faster than those \((SI= 20ms, for example)\), the GSC can support up 75 real-time stations with \( \alpha \) parameter set to 0.2.

7. CONCLUSION

This paper proposed a new scheduling scheme for HCCA in the IEEE 802.11e. The proposed scheme is called Group Sequential Communication Algorithm (GSC). This mechanism enhances the inefficient HCCA polling scheme, by means of a group sequence communication based on virtual token passing procedure among RT stations. This technique guarantees a smaller polling overhead, when compared to the IEEE 802.11e original scheme.

This technique is particularly adequate to typical industrial communication scenarios, where RT-stations transfer multiple small-sized packets at periodic intervals.
8. ACKNOWLEDGMENTS

The authors would like to give a special thanks to CAPES (Brazil) and GRICES (Portugal) agencies for their support.

REFERENCES


