Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2008, Article ID 328089, 14 pages
doi:10.1155/2008/328089

*Research Article*

# Dimensioning Method for Conversational Video Applications in Wireless Convergent Networks

**Alfonso Fernandez-Duran,[1] Raquel Perez Leal,[1] and José I. Alonso[2]**

[1] *Alcatel-Lucent Spain, Ramirez de Prado 5, 28045 Madrid, Spain*
[2] *Escuela Tecnica Superior de Ingenieros de Telecomunicación, Universidad Politécnica de Madrid, Ciudad Universitaria, 28040 Madrid, Spain*

Correspondence should be addressed to Alfonso Fernandez-Duran, afd@telefonica.net

Received 1 March 2007; Revised 19 June 2007; Accepted 22 October 2007

Recommended by Kameswara Rao Namuduri

New convergent services are becoming possible, thanks to the expansion of IP networks based on the availability of innovative advanced coding formats such as H.264, which reduce network bandwidth requirements providing good video quality, and the rapid growth in the supply of dual-mode WiFi cellular terminals. This paper provides, first, a comprehensive subject overview as several technologies are involved, such as medium access protocol in IEEE802.11, H.264 advanced video coding standards, and conversational application characterization and recommendations. Second, the paper presents a new and simple dimensioning model of conversational video over wireless LAN. WLAN is addressed under the optimal network throughput and the perspective of video quality. The maximum number of simultaneous users resulting from throughput is limited by the collisions taking place in the shared medium with the statistical contention protocol. The video quality is conditioned by the packet loss in the contention protocol. Both approaches are analyzed within the scope of the advanced video codecs used in conversational video over IP, to conclude that conversational video dimensioning based on network throughput is not enough to ensure a satisfactory user experience, and video quality has to be taken also into account. Finally, the proposed model has been applied to a real-office scenario.



## 1. INTRODUCTION

A large number of technological changes are today impacting on communication networks which are encouraging the introduction of new end-user services. New convergent services are becoming possible, thanks to the expansion of IP-based networks, the availability of innovative advanced coding formats such as H.264, which reduce network bandwidth requirements providing good video quality, and the rapid growth in the supply of dual-mode WiFi cellular terminals. These services are ranging from the pure voice, based on unlicensed mobile access (UMA) or voice call continuity (VCC) standards, to multimedia including mobile TV and conversational video communications. The new services are being deployed in both corporate and residential environments. In the corporate environments, conferencing and collaboration systems could take advantage of the bandwidth available in the private wireless networks to share presentation material or convey video conferences efficiently at relatively low communication costs. In the residential segment, mobile TV and IP conversational video communications are envisaged as key services in both mobile and IP multimedia subsystem (IMS) contexts. The success of these scenarios will depend on the quality achievable with the service, once a user makes a handoff from one network to the other (vertical handoff) and stays in the wireless domain. Video communications are usually a relatively high wireless resource-demanding service, because of the amounts of information and the real-time requirements. Services of a broadcast nature usually go downlink in the wireless network, and therefore contention and collision have a reduced effect on the network performance and capacity, while conversational video goes in both directions, suffering from the statistical behavior of the wireless contention protocol. The wireless network performance will depend on the particular video and audio settings used for the communications, and therefore the network will need to be designed and dimensioned accordingly to ensure a satisfactory user experience.

Video transmission over WLAN has been analyzed under different perspectives. An analysis of different load conditions using IEEE802.11 is presented in [1]; the study makes an assessment of the video capacity by measuring capacity in a reference testbed, but the main focus is on video streaming not on bidirectional conversational video. Although no dimensioning rules are proposed, it is interesting to mention that the measurements shown mix both contention protocol and radio channel conditions. The implications of video transmission over wireless and mobile networks are described in [2]. Although dimensioning is not targeted, the paper discusses effects of frame slicing on different numbers of packets per frame commonly used in wireless networks. The study shows that a rather low slicing in the order of 6 to 10 packets per frame is a good approach for the packet error concealment. However, it is not directly applicable to conversational video over wireless networks, since the resulting packet size could be so small that the radio protocol efficiency is severely affected. Performance and quality in terms of peak signal-to-noise ratio (PSNR) under radio propagation conditions are shown in [3], and a technique to improve the performance under limited coverage conditions is proposed, but capacity-limited conditions necessary for dimensioning are not analyzed. A discussion on the packet sizes and the implications in the PSNR is presented in [4]. The results shown are based only on simulations, and no model is proposed to predict the system performance.

The performance of conversational video over wireless networks, to be used for network dimensioning purposes, has to be analyzed under the radio access protocol perspective to evaluate the implications of the wireless network on the conversational video. The present study is based on the analysis of the effects of the medium access protocol used in IEEE802.11 on the video performance. In a first step, performance is analyzed by considering the protocol throughput as a consequence of contention and collisions. In a second step, a video quality indicator based on effective frame rate is used to assess the actual video performance beyond the protocol indicator, so as to arrive at more realistic dimensioning figures. In the present study, the availability of standardized (IEEE802.11e) techniques is assumed for traffic prioritization. The standard reference framework for IP network impairment evaluation is G.1050, and H.264 is assumed for service profiles, both from ITU-T [5, 6].

The following sections introduce the framework for conversational video applications, a new and simple model of conversational video over wireless LAN dimensioning, and show that different results are achieved using throughput and video quality approaches. Both discrepant results could be conciliated for proper network dimensioning, as it is also shown in a real-office scenario.

## 2. CONVERSATIONAL VIDEO APPLICATIONS

Today's communication networks are greatly affected by a number of technological changes, resulting in the development of new and innovative end-user services.

One of the key elements for these new applications is video services that impact on the appearance of new multimedia services. Voice services are complemented with video and text (instant messaging and videoconference, etc.) services; services can be combined and end-users can change from one type of service to another. Likewise, multiparty communication is becoming more and more popular. Services are being offered across a multitude of network types. Examples are multimedia conferences and collaborative applications that are now enhanced to support nomadic (traveling employees with handheld terminals) and IP access (workers with an SIP client on their PC and WLAN access). On the other hand, new devices are being introduced to enable end-users to use a single device to access multiple networks. Examples include the dual-mode phones, that can access mobile networks or fixed networks, or handheld devices which support fixed-mobile convergence and conversational video applications [7, 8]. As a consequence of the evolution of the technologies and applications stated in the previous paragraphs, a new analysis of conversational video applications in wireless convergent networks is required. To do that, ITU-T Rec. H.264 | ISO/IEC 14496-10 and H.264 advanced coding techniques have been considered as a video coding format. Moreover, ITU-T G.1050 recommendations have been taken into account as a reference framework for the evaluation of an IP wireless network section in terms of delay and random packet loss impact.

### 2.1. ITU-T G.1050 model considerations

ITU-T G.1050 recommendation [5] specifies an IP network model and scenarios for evaluating and comparing communication equipment connected over a converged wide-area network. This recommendation describes services' test profiles and applications, and it is scenario-based. In order to apply it to conversational video applications conveyed over wireless networks, the following services' profiles and end-to-end impairment ranges should be taken into account as a reference framework.

The contribution to delay (one-way latency) and random packet loss of the wireless LAN section, analyzed in this paper, should be compatible with the corresponding end-to-end impairment detailed in Table 1. This should be a boundary condition, in a first step, towards the analytical results.

On the other hand, taking into account the kind of applications proposed, that is, multivideo conference, fixed-mobile convergent video applications over a single terminal, and so forth, the typical scenario location combination will be business-to-business. However, business-to-home and home-to-business scenarios should be also considered in the case of teleworking. Even more new scenarios, not included in the recommendation, such as business-to-public areas and vice versa (i.e., airport and hotels) in the case of nomadic use, could take place. End-user terminals will be PCs and/or handheld terminals with video capabilities.

### 2.2. H.264 profiles and levels to be used

H.264 "represents an evolution of the existing video coding standards (H.261, H.262, and H.263) and it was developed in response to the growing need for higher compression of

Table 1: Service test profiles and impairment ranges.

| Service test profile | Application (examples) | One-way latency range (min–max) (ms) | Random packet loss range (min–max) (%) |
|---|---|---|---|
| Profile A: well-managed IP network | High-quality video and VoIP, conversational video (real-time applications, loss-sensitive, jitter-sensitive, high interaction) | 20–100 (regional) 90–300 (intercontinental) | 0–0.05 |
| Profile B: partially managed IP network | VoIP, conversational video (real-time applications, jitter-sensitive, interactive) | 50–100 (regional) 90–400 (intercontinental) | 0–2 |

moving pictures for various applications such as videoconferencing, digital storage media, television broadcasting, Internet streaming, and communication" [6].

The H.264 defines a limited subset of syntax called "profiles" and "levels" in order to facilitate video data interchange between different applications. A "profile" specifies a set of coding tools or algorithms that can be used in generating a conforming bit stream, whereas a "level" places constraints on certain key parameters of the bit stream. The last recommendation version defines seven profiles (baseline, extended, main, and four high-profile types) and fifteen "levels" per "profile." The same set of "levels" is defined for all "profiles."

Just as an example, the H.264 standard covers a broad range of applications for video content including real-time conversational (RTC) services such as videoconferencing, videophone, and so forth, multimedia services over packet networks (MSPNs), remote video surveillance (RVS), and multimedia mailing (MMM), all of them are very suitable to be deployed over convergent networks.

In this paper, video applications have focused on baseline and extended profiles and low rates (64, 128, and 384 Kbps) corresponding to levels 1, 1.b, and 1.1 of the H.264 standard. The new capabilities and increased compression efficiency of H.264/AVC allow for the improvement of the existing applications and enable new ones. Wiegand et al. remark the low-latency requirements of conversational services in [9]. On the other hand, they state that these services follow the baseline profile as defined in [10]. However, they pointed out the possibility of evolution to the extended profile for conversational video applications.

## 3. THROUGHPUT-BASED CAPACITY IN WLAN

This section describes a simple method to estimate the video capacity in IEEE802.11 networks by estimating the effect of collisions on the air interface. This method is based on the principles described in [11], further developed in [12], and adapted to voice communications in [13].

### 3.1. Principles for throughput estimation

In general, a station that is going to transmit a packet will need to wait for at least a minimum contention window, following a distributed interframe space (DIFS) period in which the medium is free. If the medium is detected as busy, the packet transmission is delayed by a random exponential backoff time measured in slot times (timing unit). Looking at the IEEE802.11 family, there are differences in duration for the same parameter. This set of values is very relevant since it defines the performance of the network for each of the PHY standards [13–16].

The first step in the analysis of the protocol, CSMA-CA, is to determine the time interval in which the packet transmission is vulnerable to collisions. Looking at the distributed coordination function (DCF) timing scheme based on CSMA-CA with request-to-send-clear-to-send (RTS-CTS), it appears that during the time interval of DIFS and an RTS packet, a collision could take place. This assumption is true in the case of a hidden node. This hidden node effect is likely to happen with certain frequency. For example, in an access network using directional antennas, most of the nodes cannot see each other, that is, most of the nodes are hidden. If we denote the period in which the protocol is vulnerable to collisions as $\tau$, this could be expressed as $\tau = \eta(t_{\text{DIFS}} + t_{\text{RTS}} + t_{\text{SIFS}}) + t_p$, where $t_{\text{DIFS}}$ is the duration of DIFS interval, $t_{\text{RTS}}$ is the duration of the signaling packet, $t_{\text{SIFS}}$ is the duration of short intraframe space (SIFS) interval, $t_p$ is the propagation time, and $\eta$ is the proportion of hidden nodes. The packet transmission has several parts: the packet transmission itself, made up of the packet duration $T$, part of which is the vulnerability period $\tau$, and the waiting intervals part in which no transmission takes place. In the case of very few hidden nodes, it is possible not to use the RTS-CTS protocol, but CTS-to-self. The new vulnerability period could be estimated as $\tau = \eta(t_{\text{DIFS}} + t_{\text{CTS}} + t_{\text{SIFS}} + T) + t_p$. The relationship between the vulnerability period and the duration of the packet transmission is $\alpha = \tau/T$. This value is key in estimating the network efficiency. Following the notation introduced in [17], it is possible to obtain the basic expressions that could be developed to obtain the parameters that influence the video network throughput and the estimated quality.

### 3.2. Contention window

The contention window as defined in IEEE802.11 is a mechanism with big influence on the network behavior since it has a significant impact on the collision probability.

Let the probability of collision or contention in a first transmission attempt be

$$P_c = (1 - P_{\text{ex}})P_{\text{CW0}} = (1 - e^{-g\tau})P_{\text{CW0}} = (1 - e^{-\alpha G})P_{\text{CW0}}, \tag{1}$$

where $P_{\text{CW0}}$ is the probability that another station selects the same random value for the contention window, $P_{\text{ex}}$ is the probability of a successful packet transmission, $g$ is the packet arriving rate in packets/sec, and $G$ is the offered traffic.

Extrapolating to $n$ transmission attempts

$$P_c = \sum_{i=1}^{n} \left(1 - e^{-\alpha G}\right)^i P_{\mathrm{CW}(i-1)}, \tag{2}$$

where $P_{\mathrm{CW}(i)}$ will be given by

$$P_{\mathrm{CW}i} = \frac{N-1}{\mathrm{CW}i + 1}, \tag{3}$$

with CW$i$ being the maximum duration of the contention window in the current status according to [17] and $N$ the number of simultaneous users.

Combining previous equations, the expected value of the contention window will be given by

$$\mathrm{CW} = \mathrm{CW}_0 + \sum_{i=1}^{n} \frac{\mathrm{CW}i}{\mathrm{CW}i + 1}(N-1)\left(1 - e^{-\alpha G}\right)^{2i+1}. \tag{4}$$

### 3.3. Throughput estimation

Taking the approach introduced by [11] and further developed in [12, 13], following the sequence of activity and inactivity periods, and because the packet streaming process is memoryless, we can consider the process of busy transmission duration times $\widetilde{B}$ and $B$ as its average value. Similarly, we could call $\widetilde{U}$ the process of durations in which transmissions are successful (with average value $U$), and $\widetilde{I}$ the duration of waiting times with average $I$; therefore the process for the transmission cycles will be $\widetilde{B} + \widetilde{I}$, and the throughput will be obtained from

$$S = \frac{U}{B + I}. \tag{5}$$

Let us consider first the inactivity period. This duration is the same as the duration of the interval between the end of a packet transmission and the beginning of the next one. Since the packet sequencing is a memoryless process, we could express

$$F_I(x) = \mathrm{prob}[\widetilde{I} \le x] = 1 - \mathrm{prob}[\widetilde{I} > x] \tag{6}$$

$$= 1 - \mathrm{P}[\text{No packet sent during } x] = 1 - e^{-gx}. \tag{7}$$

This means that "$I$" has an exponential distribution with average

$$I = \frac{1}{g}. \tag{8}$$

Following [11–13], and introducing the effect of the contention window described above, we obtain

$$U = T\left(1 - P_{\mathrm{CW}} + P_{\mathrm{CW}} e^{-g\tau}\right),$$
$$B = T + \tau + P_{\mathrm{CW}}\left(\tau - \frac{1 - e^{-g\tau}}{g}\right). \tag{9}$$

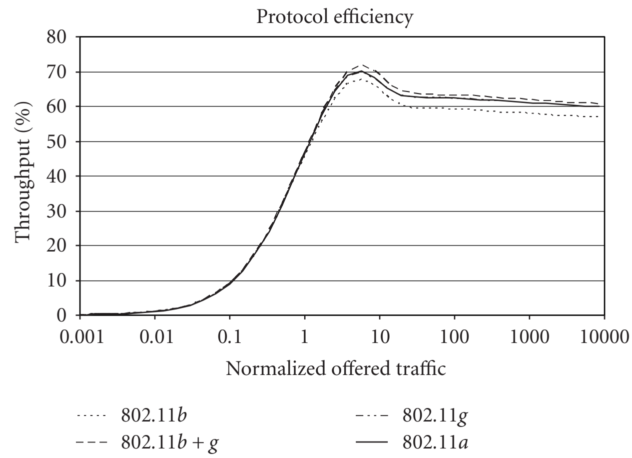


Figure 1: Throughput efficiency in IEEE802.11.

Table 2: Example of vulnerability factors for packet lengths of 1024 bytes in the different variants of the IEEE802.11 standard.

| | 802.11$b$ | 802.11$b+g$ | 802.11$g$ | 802.11$a$ |
|---|---|---|---|---|
| Vulnerability factor ($\alpha$) | 0.129 | 0.120 | 0.123 | 0.122 |

Combining the above results, we finally obtain

$$S = \frac{U}{B + I} = \frac{T\left[1 - P_{\mathrm{CW}}\left(1 - e^{-g\tau}\right)\right]}{T + \tau + P_{\mathrm{CW}}\left(\tau - \left(1 - e^{-g\tau}\right)/g\right) + 1/g}. \tag{10}$$

To turn $S$ into a more manageable format, we can normalize (10) with respect to the duration of the packet transmission period. If $G$ represents the average packet transmission time measured in packets per packet transmission period, that is,

$$G = gT, \tag{11}$$

using the vulnerability factor defined above, we obtain

$$\alpha = \frac{\tau}{T}. \tag{12}$$

The throughput expression results in the following:

$$S = \frac{G\left[1 - P_{\mathrm{CW}}\left(1 - e^{-\alpha G}\right)\right]}{1 + P_{\mathrm{CW}} + G\left(1 + \alpha\left(1 + P_{\mathrm{CW}}\right)\right) + P_{\mathrm{CW}} e^{-\alpha G}}. \tag{13}$$

Taking the timing values defined in the standard, the throughput versus the normalized offered load $G$ for the different values of vulnerability factor in the different variants of IEEE802.11 is shown in Figure 1.

As can be seen, the relative efficiency in the four functional variants of the standard is practically the same. This is due to the fact that the resulting vulnerability factor for all of them is very similar as shown in Table 2.

If we compare the efficiency of this protocol with the cable protocols such as Ethernet, we see that the latter are more efficient. This is because the vulnerability factor resulting in the radio protocol is larger. While the IEEE802.11 protocolreaches maximum throughput at around 70%, as shown in

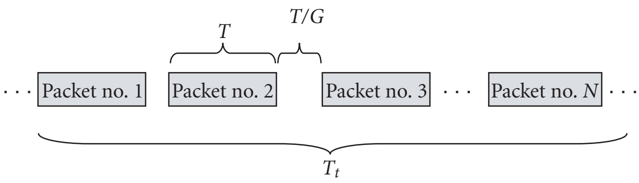


FIGURE 2: Time representation of offered traffic ($G$), number of simultaneous users ($N$), and expected packet ($T$) and frame ($T_t$) durations.

Figure 1, partially because of the contention window, Ethernet networks reach 90% in the same packet size (1024 bytes). It has to be noted that the throughput decreases with the decrease in the packet size since the relative transmission time decreases, and then the vulnerability factor increases. It is therefore evident that the protocol is more efficient with larger packets than with smaller ones.

### *3.4. Throughput-based service dimensioning in WLAN*

The analysis described so far shows the system's behavior in normalized terms, with no relationship to the transport of a specific service.

The offered traffic $G$ has to be associated with the number of users to whom a given capacity is offered, in such a way that the throughput is represented as a function of the number of users requiring a given type of service.

For a type of service characterized by a bit rate $r_b$ and an IP packet size $n_b$, we will get an average data frame duration for the service:

$$T_t = \frac{n_b}{r_b}. \tag{14}$$

If we also assume that the packet duration is $T$, the relationship between the traffic offered ($G$) and the number of sources ($N$) will be given by the following expression:

$$G = \frac{1}{T_t/NT - 1}. \tag{15}$$

Or in other terms,

$$N = \frac{GT_t}{T(1+G)}. \tag{16}$$

Figure 2 shows a representation of the timing involved to obtain the relationship between $G$ and $N$.

As can be seen, the saturation point of the system is reached for values of $N$ that are very sensitive to the packet size required by the service, regardless of the total bandwidth required. This fact is very relevant since it anticipates that the system performance will be very dependent on the service information structure apart from its bandwidth requirements.

Combining expressions (13) and (15), it is possible to represent the throughput as a function of the number of simultaneous users. Figure 3 shows an example of the behavior of the IEEE802.11 family of physical layers for given service bandwidth and packet size.

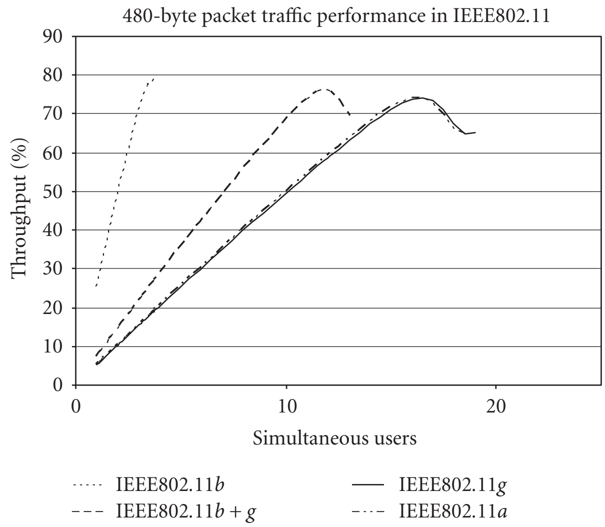


FIGURE 3: Throughput versus number of users for 480-byte packets in IEEE802.11 variants, for a service of 384 Kbps.

TABLE 3: Maximum number of users for the variants of IEEE802.11.

| | 802.11$b$ | 802.11$b+g$ | 802.11$g$ | 802.11$a$ |
|---|---|---|---|---|
| | | 480-byte packets | | |
| $N_{max}$ | 3 | 11 | 16 | 16 |

As shown in Figure 3, the throughput reaches a maximum for a specific value of $N$ depending on the service characteristics. Specifically, the maximum values reached in the above figure are shown in Table 3.

It appears to be clear that both the throughput and the maximum number of users are very sensitive to the packet size used. As a reference value, 480-byte packets have been used to determine the maximum system throughput. The maximum system throughput turns out to be below 80% of the capacity offered by the physical interface.

As apparent, it is necessary to estimate the maximum number of simultaneous users that yields the maximum throughput for a given service configuration. To obtain this value ($N_{max}$), it is necessary to calculate the maximum of the expression $S(N)$. Unfortunately, the maximum of $S$ versus $N$ leads to an expression without an exact analytical solution. To arrive at an approximate solution, it is necessary to carry out a polynomial development of one of the terms, which eventually yields the following expression:

$$G_{max} \approx \frac{\sqrt{5+4/\alpha} - 1}{\alpha + 1}. \tag{17}$$

Since $N$ is an integer value, we can assume that the approximated expression matches the exact solution with a reasonable number of terms in the development. The combination of (16) and (17) produces the following result:

$$N_{max} = \text{Int}\left\{\frac{T_t(\sqrt{5+4/\alpha} - 1)}{T(\sqrt{5+4/\alpha} + \alpha)}\right\}. \tag{18}$$

With this expression, we could apply the figures of the typical multimedia services, that is, data, voice, and video.

Table 4: Conversational video service parameters for H.264 at 384 Kbps.

| | 802.11$b$ | 802.11$b+g$ | 802.11$g$ | 802.11$a$ |
|---|---|---|---|---|
| Phy. capacity (Mbps) | 11 | 54 | 54 | 54 |
| IP packet size (bytes) | 240 | 240 | 240 | 240 |
| IP+UDP+RTP headers (bytes) | 40 | 40 | 40 | 40 |
| $T$(s) | $2,31\cdot10^{-3}$ | $6,73\cdot10^{-4}$ | $4,53\cdot10^{-4}$ | $4,59.10^{-4}$ |
| Concatenation | 1 | 1 | 1 | 1 |
| $T_t$(s) | 0,005 | 0,005 | 0,005 | 0,005 |
| $\alpha$ | $1,74\cdot10^{-1}$ | $1,19\cdot10^{-1}$ | $1,28\cdot10^{-1}$ | $1,26.10^{-1}$ |
| $G_{\max}$ | 3,66 | 4,66 | 4,45 | 4,49 |
| $N_{\max}$ | 2 | 6 | 9 | 9 |

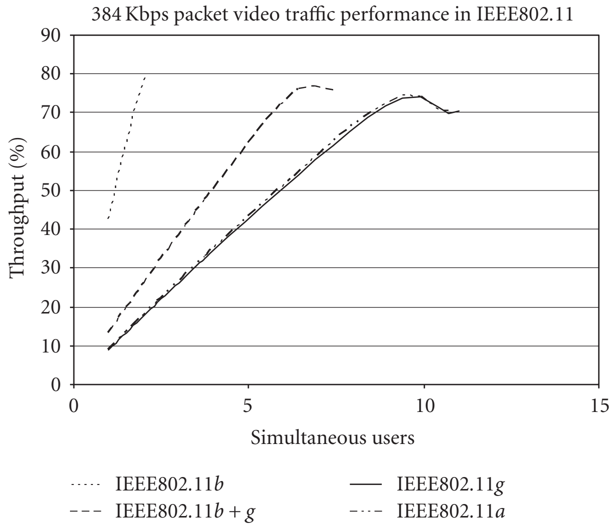


Figure 4: 384 Kbps conversational video over IP performance with IEEE802.11e protocol using 240-byte packets.

### 3.5. *Throughput of conversational video over IP service*

Conversational video over IP introduces additional restrictions to the system, mainly resulting from the average bandwidth required. Since voice and telephony traffic characteristics are well known, their Poisson process characteristics match the contention model described in the previous sections very well. This is still valid even with the introduction of prioritization mechanisms from IEEE802.11e.

Considering the use of 384 Kbps video codec, the service will be defined by the parameters shown in Table 4.

With the service configuration defined as shown in Table 3, it results in a behavior as shown in Figure 4.

As it becomes apparent, the differences between the IEEE802.11 physical layer variants are significant. In addition, it has to be noted that, depending on the particular operating conditions, that is, on the maximum capacity offered by the physical layer, the video codec used, video frame size, and so forth, the results could differ significantly. For example, Table 5 shows the maximum number of simultaneous users using 240-byte packets which is a typical expected packet size in conversational video applications.

Table 5: Throughput-based video capacity.

| | Throughput-based video conversations | | | |
|---|---|---|---|---|
| | 802.11$b$ | 802.11$b+g$ | 802.11$g$ | 802.11$a$ |
| 64 Kbps | 11 | 36 | 52 | 51 |
| 128 Kbps | 5 | 19 | 27 | 27 |
| 384 Kbps | 2 | 6 | 9 | 9 |

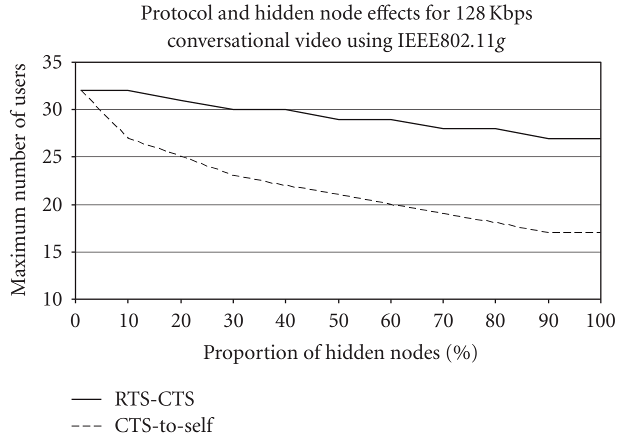


Figure 5: Effect of hidden node on the system capacity.

### 3.6. *RTS-CTS versus CTS-to-self approach for conversational video capacity*

In the normal use of WLANs, it is possible to select the use of the RTS-CTS protocol or leave it in the default CTS-to-self mode. In conditions in which many users share the air interface, and in conditions in which a hidden node effect appears, it is reasonable to use the RTS-CTS protocol to ensure system performance in terms of delay and capacity. On the other hand, when few users share the medium, it could happen that the use of the CTS-to-self mode provides some improvement. This section compares both protocols belonging to the IEEE802.11, to determine the optimal conditions of use as a function of the hidden nodes in the network.

According to the definition of the protocol in Section 3, the protocol performance in terms of throughput is conditioned by the vulnerability time period. The difference between RTS-CST and CTS-to-self is twofold. On the one hand RTS-CTS uses extra resources to manage the protocol, but it has a reduced vulnerability time period, and on the other hand, CTS-to-self uses fewer network resources, but it has a longer vulnerability period. Comparing the two approaches, there is a tradeoff between use of resources and vulnerability to collisions.

To illustrate the effect of hidden nodes on the system capacity, the two protocol variants RTS-CTS and CTS-to-self have been compared, and the results are shown in Figure 5.

As the proportion of hidden nodes increases, the probability of collision also increases regardless of the protocol scheme selected. In the case of RTS-CTS scheme, mechanisms to maintain the vulnerability time period under certain limits are available. This is why the reduction in capacity is in the order of 16%. In the case of CTS-to-self, the vulnerability time period is extended to the complete packet, and that is why it experiences a capacity reduction in the order of 47%. Depending on the particular network conditions and services, the figures could differ, but in general terms, the use of RST-CTS protocol appears to be advantageous for services with moderate number of users.

A simple and approximated approach to estimate the performance in hidden node conditions consists of substituting $\alpha$ in (18) with $\alpha'\eta$, where $\eta$ is the estimated proportion of hidden nodes.

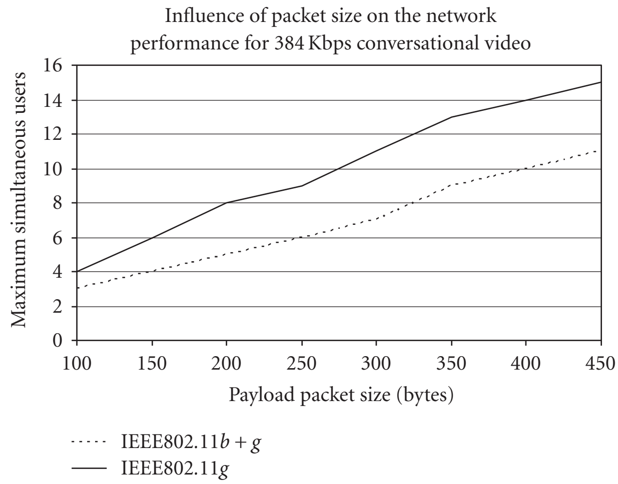


Figure 6: Influence of packet size on the system performance.

### 3.7. *Influence of conversational video packaging on the performance*

According to the system performance estimation shown in Section 3.3, the total system capacity and throughput could depend significantly on the expected packet size used by the service. Because of the nature of video payload, it cannot be guaranteed that IP packets are of a given size; nevertheless the video IP packet sizes correspond to a multimodal distribution in which only certain values are possible. Therefore, the analysis has to be based on the expected packet size values. Depending on the profile and the group of video objects (GoV) scheme selected, the expected packet size delivered could differ, unless measures are taken to ensure an average packet size. The larger the packet size used for a given services, the higher the throughput and capacity of the system will be.

Let us take the expected packet size as

$$E(s) = S = \sum_k s_k P_k, \qquad (19)$$

where $s$ is the distribution of packet sizes, $s_k$ are the discrete values associated to $s$, and $P_k$ is the probability of occurrence of each type of packet sizes.

To illustrate the effect of the packet size on the performance, Figure 6 shows how the maximum number of simultaneous users could increase the expected payload packet size.

This behavior follows two principles: the larger the expected payload packet sizes are, the fewer number of packets will be needed to maintain the average bit rate, and therefore less collision events may occur, and the larger the packet is, the lower the impact of the necessary headers will be. As a counterpart, if radio channel conditions suffer from degradation (increase in the packet error rate), the total video PSNR could be reduced as described in [4]. In general, the frame slicing of conversational video will be rather small.

### 3.8. *Audio and video performance interaction*

Audio could also play an important role in conversational video communications. There are two possibilities to convey the audio: including the audio as part of the audio and video payload, taking advantage of packet grouping and synchronization, or taking audio and video streams separately through the network, making it possible to have a greater diversity of end-user profiles and devices. Both schemes are equally used for conversational video calls, but from the wireless protocol point of view, the interleaving of audio and video has advantage of performing close to the case of video only.

The case of separate audio and video streams is very common for multiconference environments where user terminals could have audio only or audio and video capabilities, and all could take part in the same conference. Many cases of separate audio and video flows come from the fact that part of the communication is conveyed through a network (e.g., audio through the cellular mobile) and the other part is conveyed through an IP wireless network (e.g., video part through IEEE802.11 interface). These conditions are easily experienced using dual-mode cellular wireless terminals. In the case of using strictly separate audio and video streams, some extra room has to be allowed to allocate the audio streams in the network. Fortunately, both audio and video behaviors are sufficiently linear below the maximum throughput, and this allows the combination of the two services using simple proportion rules. For example, if under certain conditions an IEEE802.11$g$ network can afford 21 G.729 calls or 9 H.264 384 Kbps video calls. However, if we combine separate audio and video streams, the total audio plus video calls will be in the order of 6. More information of the voice capacity estimation can be found in [17, 18, 20, 21].

## 4. VIDEO QUALITY ESTIMATION PRINCIPLES

As described in previous sections, the maximum number of simultaneous users running conversational video applications can be estimated from the maximum throughput of the IEEE802.11 protocol. Moving to a more user-centric approach, it would be convenient to estimate the conversational video capacity also based on the video quality. By following this approach, it appears that the maximum number of simultaneous users could be different.

The first step is to select a reasonable video quality indicator that relates quality to the wireless network conditions. The two main potential indicators of the network conditions are the delay for packet delivery and the packet loss probability.

A usual approach to estimate video quality is the peak signal-to-noise ratio (PSNR), or more recently video quality rating (VRQ); both are usually estimated from the mean square error (MSE) of the video frames after the impairments (e.g., packet loss) with respect to the original video frames [22, 23]. From these values, there is some correlation to video mean opinion score (MOS). Unfortunately, the relationship between packet loss and MSE is not straightforward since not all packets conveyed through the wireless network have the same significance. Alternatively, a relatively simpler quality indicator introduced in [24] is proposed. This indicator is the effective frame rate that is introduced and discussed in later sections of this paper.

As packet errors occur in the wireless network, video frames are affected, making some of them unusable, and therefore the total frame rate is reduced. Video quality will be acceptable if the expected frame rate of the video conversations is kept above certain value.

### 4.1. *Delay in conversational video over IP in wireless networks*

A very important characteristic in conversational video communications is the end-to-end delay, since it could have a direct impact on the perceived communication quality, by producing buffering or synchronization problems between audio and video in case of separate streams.

To estimate the delay contribution introduced by the wireless network, let us proceed as in Sections 3.2 and 3.3. According to (1), the probability of success for a packet transmission is given by the following expression:

$$P_{ex} = e^{-g\tau}. \tag{20}$$

The probability of a packet transmission being unsuccessful will be

$$P_c = 1 - P_{ex} = 1 - e^{-g\tau}. \tag{21}$$

Because of the IEEE802.11 operation, we know that in the case of unsuccessful packet delivery at the first attempt, the backoff time is increased to the next integer power of two. This in turn will be the window to generate a random waiting time before the transmission or retransmission takes place. Although many retransmissions could take place before a packet is successfully delivered, there is a dominant influence on the first retransmission to the total delay. Since the rest of the packet transmissions are not necessarily in the same contention window backoff, the nominal delay will be given by

$$C = (b + (N-1)T)P_c = (b + (N-1)T)(1 - e^{-g\tau}), \tag{22}$$

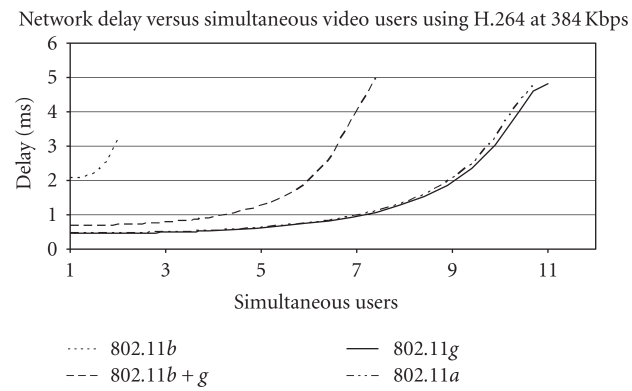


Figure 7: Video communication delay for the different variants of IEEE802.11 as a function of the number of simultaneous users.

Table 6: Video transmission delay values in the maximum throughput conditions for H.264 at 384 Kbps.

| | 802.11$b$ | 802.11$b+g$ | 802.11$g$ | 802.11$a$ |
|---|---|---|---|---|
| $N_{max}$ | 2 | 6 | 9 | 9 |
| Expected delay (ms) | 1.9 | 1.3 | 1.2 | 1.1 |

where $N$ is the number of simultaneous voice communications and $b$ is the backoff time.

The expected value for the duration of a transmission will be given by the time associated to the successful transmissions and the time associated to the retransmissions. Therefore the expected delay will be

$$D = C + Te^{-g\tau} = (b + (N-1)T)(1 - e^{-g\tau}) + Te^{-g\tau} \tag{23}$$

or

$$D = (b + (N-2)T)(1 - e^{-g\tau}) + T. \tag{24}$$

Combining (11), (12) and (15), (24) can be also expressed using the service variables as

$$D = (b + (N-2)T)\left(1 - e^{-\alpha(NT/(T_t - NT))}\right) + T. \tag{25}$$

As can be seen in Figure 7, the delay grows monotonically with the number of simultaneous video communications, until the point at which network saturation is reached. Under these conditions, the delays that are achieved are those corresponding to the maximum throughput conditions. An example of these results is shown in Table 6.

The delay values shown so far take into account only the delay introduced by the wireless network in the uplink. To consider the total delay, it is necessary to introduce the delays introduced by the voice codecs, the concatenation delay, and any other delay resulting from the video processing.

The total delay contribution from the wireless network could be comparatively smaller than the delay resulting from the rest of the actions taking place in the conversational video transmission. As an example, it is common that video codecs introduce delays for frame buffering and video processing. In the case of 16 frames per second, the delay of a single frame

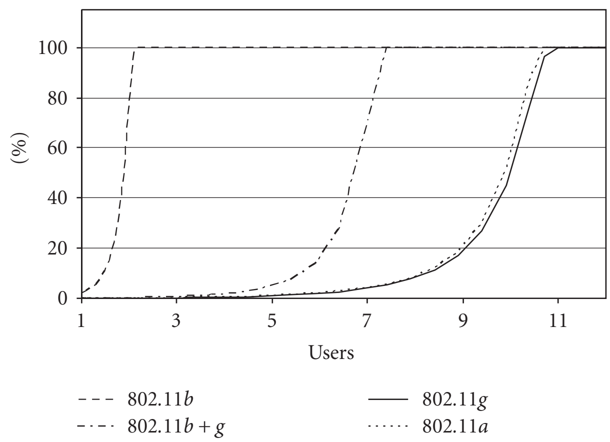


Figure 8: Packet loss probability as a function of the number of simultaneous users for H.264 at 384 Kbps.

buffering will be 62.5 milliseconds, which is about one order of magnitude longer than the delay introduced by the wireless protocol. As it is detailed in [25], additional processing delay has to be added to the buffering delay. Therefore, by comparison, there is no impact on the display deadline violations caused by the protocol contention or collision. On the other hand, the contribution of the protocol to the resulting delay is compatible with the one-way latency expected in Profile B, partially managed IP networks defined in G.1050 [5], and even in most of the scenarios, it will also be compatible with Profile A; so the impact on quality should be low.

### 4.2. Packet loss of conversational video over IP in wireless networks

The network throughput behavior is not monotonic as shown in previous sections. This effect is a result of the increase in the number of retransmissions that produces an avalanche effect. In the case of conversational video over IP, the maximum number of retransmissions to deliver the same voice packet should remain limited. According to [4], the increase in the number of transmission attempts from two to three has a maximum improvement of 2 dB in PSNR, and further increase has practically no effect on the PSNR. This in addition avoids an unnecessary increase in delay and jitter. Following this rationale, the maximum number of packet reattempts could be limited to two, and after that, the packet is dropped.

If the probability of successful packet transmission is given by (7) or equivalently by $P = e^{-\alpha G}$, then the probability of two consecutive packets being unsuccessful will be given by

$$P_{\text{pl}} = \left(1 - e^{-\alpha G}\right)^2 = \left(1 - e^{-\alpha(NT/(T_t - NT))}\right)^2. \quad (26)$$

Taking the values shown in previous sections for $G$ and $\alpha$, the packet loss probability becomes as shown in Figure 8.

Although packet loss probabilities shown could reach relatively high values, the maximum acceptable limit is around 20%. These values will be used later to estimate their influence on the resulting video conversation quality.

### 4.3. Packet loss and radio propagation channel

Although the discussion in the previous sections is mainly focused on the radio protocol performance, radio propagation conditions have a decisive impact on the performance of the conversational video. To illustrate this fact, it is necessary to understand the behavior of the signal strength. In complex propagation scenarios, such as indoor ones, small changes in the spatial separation between wireless access points and observation points bring about dramatic changes in the signal amplitude and phase. In typical wireless communication systems, the signal strength analysis is based on the topologies of combined scenarios that experience fading, produced by several causes. Several propagation studies assume that the fading can be modeled with a random variable following a lognormal distribution as described in [26–28] in the form of

$$f_i(\hat{s}) = \frac{1}{\sigma_i\sqrt{2\pi}} e^{-(\hat{s}-\mu_i')^2/2\sigma_i^2}, \quad (27)$$

where $\hat{s}$ is the received path attenuation represented in dB, $\mu_i'$ is the average signal losses received at the mobile node from the wireless access point $i$ and could be expressed as

$$\mu_i' = k_1 + k_2 \log(d_i). \quad (28)$$

$\mu_i'$ represents the propagation losses at the observation point from the access point (AP), AP $i$, and $d_i$ represents the distances from the observation point to the wireless access point $i$. Constants $k_1$and $k_2$ represent frequency-dependent and fixed attenuation factors and the propagation constant, respectively. Finally $\sigma_i$ represents the fading amplitude.

The received signal strength could be similarly expressed as

$$\mu_i = P_{tx} - (k_1 + k_2 \log(d_i)), \quad (29)$$

where $P_{tx}$ is the transmitted power.

Following this principle for dimensioning purposes, video packets will be lost in conditions in which the signal strength falls below a sensitivity threshold $s_T$. Therefore the probability of having a signal strength outage would be

$$P_T = P(s > s_T) = 1 - P(s < s_T) = 1 - F(s_T), \quad (30)$$

where $F$ is the cumulative distribution function of $f$, commonly represented as

$$F_i(\hat{s}) = \frac{1}{\sigma_i\sqrt{2\pi}} \int_0^{\hat{s}} e^{-(r-\mu_i')^2/2\sigma_i^2} dr. \quad (31)$$

Equation (31) does not provide information on the duration and occurrence rate of the fading; nevertheless extensive measurement campaigns have shown that fading tends to occur in lengthy periods and at a low frequency as described in [28], rather than short isolated and frequent events. On the

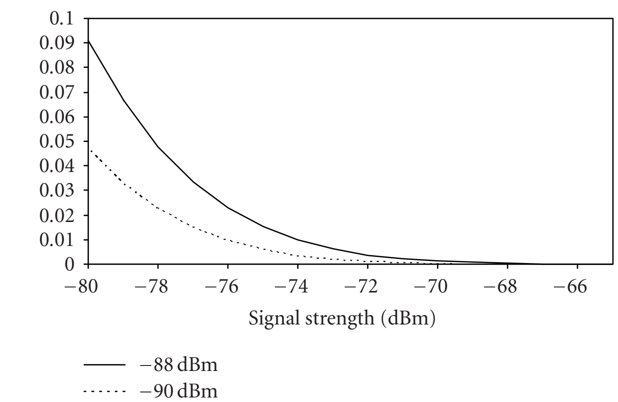


Figure 9: Packet error probability as a function of the signal strength for 6 dB lognormal fading.

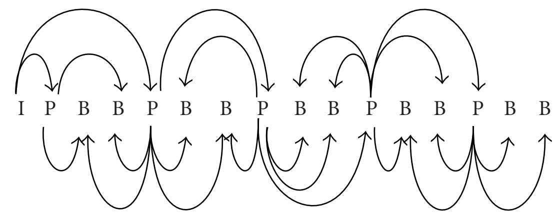


Figure 10: Compressed video frame-type interrelations.

other hand, since conversational video packets are of relatively short duration (e.g., 5 milliseconds), the probability of outage as provided in (30) could be taken as an estimate of the packet loss probability for a set of channel conditions.

In selecting a set of typical working conditions, it is possible to estimate the probability of packet errors due to the fading. For instance, selecting $\sigma = 6$ dB lognormal fading and sensitivity thresholds of −88 dBm and −90 dBm, the results are shown in Figure 9. The signal strength shown on the horizontal axis is the average power estimated using a propagation model like the one described in ITU-R recommendation P.1238-3, once the average power is obtained. It is then possible to estimate the link performance in terms of packet losses. In the case of network deployment with good coverage (−76 dBm to −73 dBm average), the packet loss is kept below 1%.

## 5. CONVERSATIONAL VIDEO CAPACITY OVER WIRELESS LAN BASED ON QUALITY

As mentioned in Section 4, a good approach to estimate the video quality is based on evaluating the frame rate drop due to the impact of packet losses on the video frame integrity.

The sources of packet losses are on one hand the contention protocol, and on the other hand the losses caused by the radio channel conditions. For specific scenarios, both should be taken into account. Since both processes are statistically independent, the total packet loss could be obtained as the addition of both. Nevertheless, to analyze the effect of the contention protocol, it is assumed that the radio propagation conditions are sufficiently good to consider the contention protocol as the dominant effect.

The consequence of a packet loss in a generic video sequence depends on the particular location of the erroneous packet in the compressed video sequence. The reason for this is related to how compressed video is transmitted through the IP protocol. The plain video source frames are compressed to form a new sequence of compressed video frames. The new sequence, depending on the H.264 service profile applied, could be made up of three types of frames: I (Intra) frames that transport the content of a complete frame with lower compression ratio, P (Predictive) frames that transport basic information on prediction of the next frame based on movement estimators, and B (Bidirectional) frames that transport the difference between the preceding and the next frames. These new frames are grouped in the so-called group of pictures (GoP) or groups of video objects (GoV) depending on the standard. The GoV could adopt many forms and structures, but for our analysis, we assume a typical configuration of the form IPBBPBBPBBPBBPBB. This means that every 16 frames there is an Intra followed by Predictive and Bidirectional frames. IP video packets are built from pieces of the aforementioned frame types and delivered to the network. If a packet error has been produced in a packet belonging to an Intra frame, the result is different from the same error produced in a packet belonging to a Predictive [29] or Bidirectional frame. A model is proposed in [24] to characterize the impact of packet losses on the effective frame rate of the video sequence.

There are some characteristics that are applicable to the case of conversational video, and in particular to portable conversational video, and that are not necessarily applicable to other video services like IPTV or video streaming. The first important characteristic is the low-speed and low-resolution formats (CIF or QCIF) that in turn produce a very low number of packets per frame, especially if protocol efficiency is taken into account increasing the average packet size (see Figure 6). In these conditions, a single packet could convey a substantial part of a video frame. The second important characteristic comes from the portability and low consumption requirement at the receiving end that in turn requires a lighter processing load to save battery life. The combination of the two aforementioned characteristics makes those packet losses impact greatly on the frame integrity, and concealment becomes very restrictive. In conversational video, it could be better, for instance, to maintain a clear fixed image of the other speaker on the screen than to try error compensation at the risk of severe image distortions and artifacts. Following these characteristics, every time a packet is lost in a frame, the complete frame becomes unusable, and some action could be taken at the decoder end, to mitigate the effect such as freezing or copying frames, but the effective frame rate has been reduced, and it has to afford some form of video quality degradation.

Following the notation in [24], the observed video frame rate will be $f = f_0(1 - \phi)$, where $\phi$ is the frame drop rate and $f_0$ is the original frame rate.

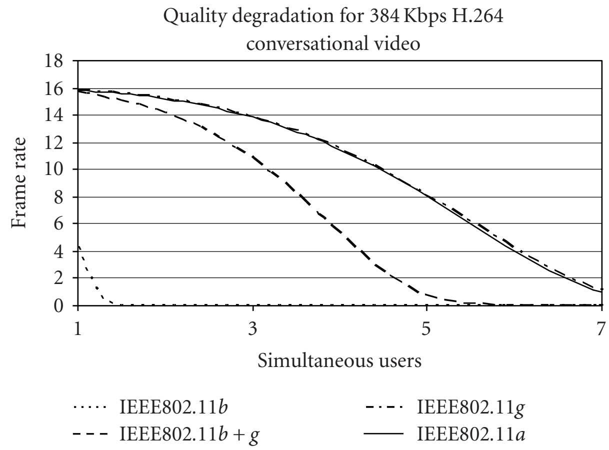


Figure 11: Frame rate as a function of simultaneous users for 384 Kbps H.264 CIF.

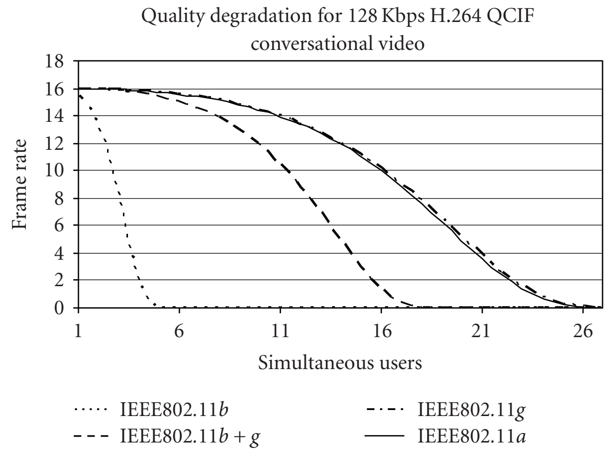


Figure 12: Frame rate as a function of simultaneous users for 128 Kbps H.264 QCIF.

In turn, the frame drop rate comes from

$$\phi = \sum_i P(f_i)P(\overline{F} \mid f_i), \tag{32}$$

where "$i$" denotes the different types of I, P, and B frames, $P(f_i)$ is the fraction of bit streams from each type of frames, and $P(\overline{F} \mid f_i)$ is the conditional probability of having a packet loss on a frame of type $i$. The latter can be estimated by taking into account the interrelation between frames as shown in Figure 10.

As a consequence of this, [24] arrives at closed form expressions for the conditional probabilities as follows:

$$P(\overline{F}|@I) = 1 - (1-p)^{S_I}, \tag{33}$$

$$P(\overline{F} \mid P) = 1 - \frac{(1-p)^{S_I}}{N_P\left(1-(1-p)^{S_P}\right)}\left(1-(1-p)^{S_P N_P}\right), \tag{34}$$

$$P(\overline{F} \mid B) \le 1 - \frac{(1-p)^{S_I+S_P+S_B}}{N_P\left(1-(1-p)^{S_P}\right)}\left(1-(1-p)^{S_P N_P}\right), \tag{35}$$

where $S_I$, $S_P$, and $S_B$ are the number of packets in the respective frame types and $N_P$ is the number of $P$ frames. For simplicity, in this estimation, it is considered that a packet loss within a frame produces an unusable frame. Although the assumption is very restrictive, it yields a conservative indicator of quality.

Combining (26) to (35), it is possible to obtain a simple video quality estimator based on the effective frame rate resulting from the packet losses. In Figures 11–13, the effective frame rate reduction as the number of simultaneous conversational video users increases is shown, for the cases of H.264 at 384 Kbps CIF format, and 128 Kbps and 64 Kbps QCIF formats.

From the results above, it is possible to obtain the maximum number of simultaneous users that ensures a minimum frame rate, and therefore dimensioning could be related to a quality estimator. As a quality limit indicator, a frame rate of 6 frames per second has been selected, because even in conversational video lower frame rate results are too annoying for the user's perception. Table 7 summarizes the different cases. If we compare Tables 3 and 7, it appears that selecting the peak throughput as an indicator for conversational video capacity could result in low video quality, since even at that point some packet losses take place, and although the losses are small, the potential impact could be significant. The frame rate quality indicator places the system at a point in which user experience could be more satisfactory.

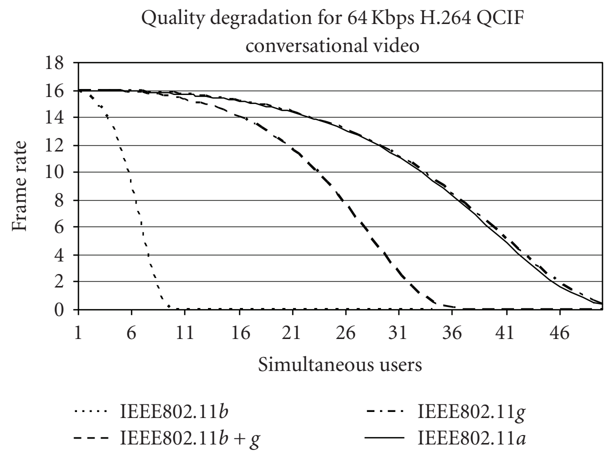


Figure 13: Frame rate as a function of simultaneous users for 64 Kbps H.264 QCIF.

Table 7: Quality-based conversational video capacity.

| | Quality-based video conversations | | | |
|---|---|---|---|---|
| | 802.11b | 802.11b+g | 802.11g | 802.11a |
| 64 Kbps | 7 | 28 | 40 | 39 |
| 128 Kbps | 3 | 13 | 19 | 19 |
| 384 Kbps | 1 | 3 | 5 | 5 |

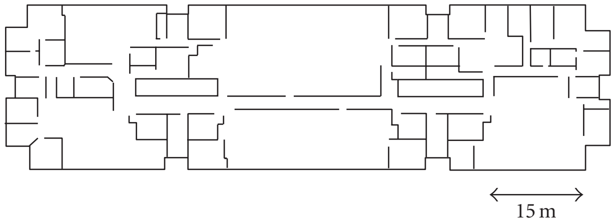


Figure 14: Simulation scenario layout.

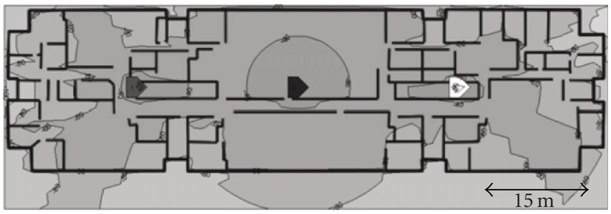


Figure 15: Simulation scenario power levels.

## 6. SCENARIO SIMULATION

To illustrate the principles shown in the previous sections, an example has been selected as shown in Figure 14. The scenario corresponds to a small office environment of 103 × 28 meters covered by three access points represented with the lighter and darker shapes within the layout. The access points are supposed to operate in IEEE802.11$b + g$, that is, in $g$ mode with backward compatibility to $b$. To simulate the conditions as realistically as possible, the walls have been modeled to introduce attenuation in the signal propagation and edge diffraction.

The power levels that result in the scenario are also shown as isolines in Figure 15 where darker shading corresponds to higher signal strength. As can be seen, most of the interest area in the scenario is within the −80 dBm line, which is sufficient to obtain maximum capacity with IEEEE802.11$b + g$.

To highlight the critical areas in the scenario, the peak capacity has been estimated and shown in Figure 16. The lighter areas represent the weaker signal conditions, that in turn result in lower peak capacity in the scenario.

For the conversational video service, the approach followed in previous sections could be compared with simulation results of video capacity. The simulation consists of running iterations that place a number of users in random positions, ranging from a minimum number of users, to a maximum one, and evaluating the number of call attempts rejected by the call admission control for each value, because of the congestion conditions and the field strength. The ratio between the call attempts and the rejections is the outage probability.

The overall process followed in the simulations is represented in Figure 17.

For each of the video configurations, the process starts with a minimum number of simultaneous users. Uniform random coordinates of users are generated to meet the number of users within the perimeter of the layout in Figure 14. The next step in the process is to associate the clients to the

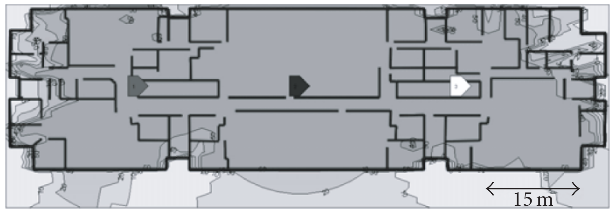


Figure 16: Simulation scenario peak available capacity.

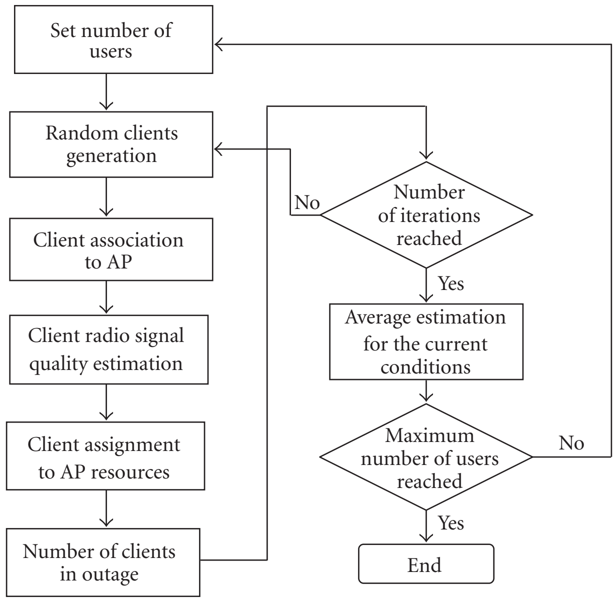


Figure 17: Simulations flow diagram.

access points. The association is based on signal strength; that is, each of the clients will be linked to the AP with the better radio link. Since the scenario is indoor, the association is not necessarily correlated with the distance to the AP. For each of the user's coordinates $p_i$, it is possible to find the power received from the access points using

$$\mu'_j(x_i, y_i) = \sum_k \lambda_k + 20\log\{\max[1, d_j(x, y)]\}, \qquad (36)$$

where $\mu'_j(x_i, y_i)$ represents the propagation losses at the user's coordinates $p_i$ to the $j$ access points, $\lambda_k$ the attenuation of the $k$ crossed walls in the path from the user's coordinates to the base station, and $d_j$ the distances from the user's coordinates $p_i$ to the $j$ access points. The average signal strength will then be expressed as

$$S_j = P_{\mathrm{TX}} + G_{\mathrm{TA}} - l_{\mathrm{T}} - \mu'_j + G_{\mathrm{RA}} - l_{\mathrm{R}}, \qquad (37)$$

where $P_{\mathrm{TX}}$ is the transmission power in dBm, $G_{\mathrm{TA}}$ is the gain of the transmission antenna in dBi, $l_T$ are the additional losses in the transmission path before the antenna, $G_{\mathrm{RA}}$ is the gain of the receiving antenna, and $l_R$ are the additional losses between the receiving antenna and the receiver.

The association of a user located at point $p_i$ will be to the access point from which it receives the higher power level

Table 8: Simulation results for several outage probabilities.

| Outage probability | H.264 @ 64 Kbps | H.264 @ 128 Kbps | H.264 @ 384 Kbps | 128 Kbps quality | 384 Kbps quality |
|---|---|---|---|---|---|
| 1% | 94 | 48 | 12 | 29 | 6 |
| 5% | 102 | 51 | 13 | 32 | 7 |
| 10% | 105 | 54 | 15 | 34 | 8 |
| | | | Average simultaneous users per access point | | |
| Reference | 36 | 19 | 6 | 13 | 3 |
| 1% | 31 | 16 | 4 | 10 | 2 |
| 5% | 34 | 17 | 4 | 11 | 2 |
| 10% | 35 | 18 | 5 | 11 | 3 |

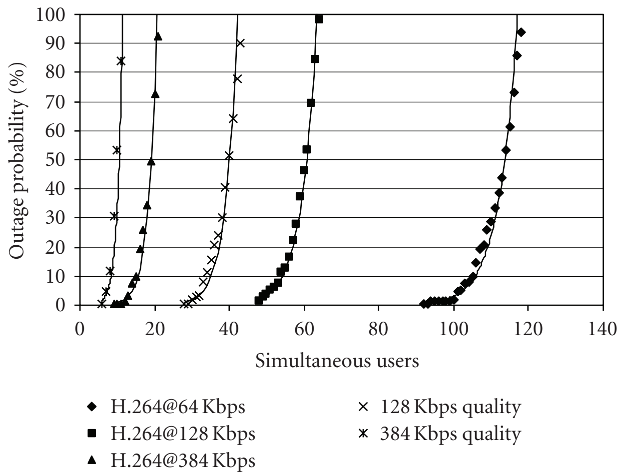


Figure 18: Scenario simulation results.

between the $j$ access points. This simple process can be modeled as the $j$ value for which power is

$$S = \max\{S_j\}. \tag{38}$$

The association process in a multicell environment will also fix the frequency channel used by the user client, and therefore it will condition the interaction with the rest of the cells.

The signal quality at each of the user positions is estimated by using the resulting signal strength from (37) and (38) with respect to the associated access point. The signal quality in turn provides the working point for the particular client and then the resources needed for the video transmission. In the case where the radio quality does not reach a minimum threshold, the client is considered to be in an outage. The remaining clients are considered at their respective working conditions to evaluate whether the total available resources are enough to allocate all the video conversations. If the resources are not enough, the number of clients is reduced till the video conversations fit in the system. The difference from the initial number of clients and the allocated video conversations is the outage fraction of users. This process is repeated several times for each number of clients ranging in an interval that depends on the video mode used.

Following this approach, the results obtained are shown in Figure 18, for the different types of video codecs.

Every point in Figure 18 is the average result of running similar condition over a large number of realizations, and therefore the distributions of points follow exponential trends as could be expected. The solid lines represent the best-fit exponential trend for the sets of points.

From the outage probability curves, the maximum number of simultaneous video conversations could be obtained for a given value of outage probability (typically between 1% and 10%). Figure 18 shows the network behavior for the main video codecs, with the use of quality-related metrics in the admission control (reduction of video frame rate).

Table 8 shows a summary of results from the scenario simulation, where it can be seen that the total number of simultaneous video calls estimated using the capacity approach (marked as reference) a bit higher than the one is obtained in the simulation results. This is caused by the random placement of video users in weaker propagation areas of the scenario. It has to be noted that these results could be very dependent on the particular scenario, and this is only taken as an example.

## 7. CONCLUSIONS AND FUTURE WORK

Conversational video over wireless LAN dimensioning has been addressed under the optimal network throughput and the perspective of video quality. The paper proposes a simple and new dimensioning model that incorporates both key aspects of throughput and quality.

The maximum number of simultaneous users due to throughput is limited by the collisions taking place in the shared medium with a statistical contention protocol. The video quality is conditioned by the packet loss in the contention protocol. Both approaches have been analyzed under the scope of common conversational video profiles used in conversational video applications over wireless LANs. The approach presented is compatible and could be complemented with radio propagation effects.

To illustrate the applicability of the dimensioning model proposed, an office scenario has been simulated.

Subsequent research steps will be taken, to incorporate the effects of handover in the conversational video performance, as well as video quality indicators applicable to video streaming and IPTV over WLAN.

## ACKNOWLEDGMENT

The authors are thankful to the support of the Spanish Ministry of Education and Science within the framework of Project no. TEC2005-07010-C02-01/TCM.

## REFERENCES

[1] Y. Koucheryavy, D. Moltchanov, and J. Harju, "Performance evaluation of live video streaming service in 802.11b WLAN environment under different load conditions," in *Proceedings of the 1st International Workshop on Multimedia Interactive Protocols and Systems (MIPS '03)*, vol. 2899 of *Lecture Notes in Computer Science*, pp. 30–41, Napoli, Italy, November 2003.

[2] T. Stockhammer, M. M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 657–673, 2003.

[3] A. Miu, J. G. Apostolopoulos, W. Tan, and M. Trott, "Low-latency wireless video over 802.11 networks using path diversity," in *Proceedings of the International Conference on Multimedia and Expo (ICME '03)*, vol. 2, pp. 441–444, Baltimore, Md, USA, July 2003.

[4] E. Masala, C. F. Chiasserini, M. Meo, and J. C. De Martin, "Real-time transmission of H.264 video over 802.11-based wireless ad hoc networks," in *Proceedings of Workshop on DSP in-Vehicular and Mobile Systems*, pp. 193–207, Nagoya, Japan, April 2003.

[5] ITU-T Recommendation G.1050, "Quality of service and performance-generic and user related aspects network model for evaluating multimedia transmission performance over Internet Protocol," November 2005.

[6] ITU-T Recommendation H.264, "Infrastructure of audiovisual services-coding of moving video. Advanced video coding for generic audiovisual services," March 2005.

[7] Alcatel, "Open IMS solutions for innovative applications," White Paper, Ed. 7, November 2005.

[8] R. Pérez Leal and P. Cid Fernández, "Aplicaciones innovadoras en el entorno IMS/TISPAN," in *Proceedings of Telecom I+D*, p. 9, Madrid, Spain, November-December 2006.

[9] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, 2003.

[10] ITU-T Recommendation H.241, "Infrastructure of audiovisual services-communication procedures extended video procedures and control signals for H.300-series terminals," May 2006.

[11] F. A. Tobagi, "Multiaccess protocols in packet communication systems," *IEEE Transactions on Communications Systems*, vol. 28, no. 4, pp. 468–488, 1980.

[12] R. Rom and M. Sidi, *Multiple Access Protocols: Performance and Analysis*, Springer, New York, NY, USA, 1990.

[13] A. F. Fernandez and J. I. Alonso, "Method to estimate voice over IP and data traffic capacity in IEEE WLANs," in *International Workshop on Networks and Applications Towards a Ubiquitously Connected World (EUNICE '05)*, pp. 204–208, Madrid, Spain, July 2005.

[14] S. W. Kim, B.-S. Kim, and Y. Fang, "Downlink and uplink resource allocation in IEEE 802.11 wireless LANs," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 1, pp. 320–327, 2005.

[15] M. S. Gast, *802.11 Wireless Networks: The Definitive Guide*, O'Reilly Media, Sebastopol, Calif, USA, 2002.

[16] ISO/IEC 8802-11ANSI/IEEE Std 802.11 First Edition 1999-00-00, "Information technology–telecommunications and information exchange between systems–local and metropolitan area networks–specific requirements–part 11: wireless LAN medium access control (MAC) and physical Layer (PHY) specifications," 1999.

[17] K. Medepalli, P. Gopalakrishnan, D. Famolari, and T. Kodama, "Voice capacity of IEEE 802.11b, 802.11a and 802.11g wireless LANs," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '04)*, vol. 3, pp. 1549–1553, 2004.

[18] O. Awoniyi and F. A. Tobagi, "Effect of fading on the performance of VOIP in IEEE 802.11 a WLANs," in *Proceedings of IEEE International Conference on Communications*, vol. 6, pp. 3712–3717, Paris, France, June 2004.

[19] D. P. Hole and F. A. Tobagi, "Capacity of an IEEE 802.11b wireless LAN supporting VoIP," in *Proceedings of IEEE International Conference on Communications*, vol. 1, pp. 196–201, Paris, France, June 2004.

[20] K. Medepalli, P. Gopalakrishnan, D. Famolari, and T. Kodama, "Voice capacity of IEEE 802.11b and 802.11a wireless LANs in the presence of channel errors and different user data rates," in *Proceedings of the 60th IEEE Vehicular Technology Conference (VTC '04)*, vol. 60, pp. 4543–4547, Los Angeles, Calif, USA, September 2004.

[21] N. Hegde, A. Proutière, and J. Roberts, "Evaluating the voice capacity of 802.11 WLAN under distributed control," in *Proceedings of the 14th IEEE Workshop on Local and Metropolitan Area Networks (LANMAN '05)*, pp. 6, September 2005.

[22] J. Hu, S. Choudhury, and J. D. Gibson, "PSNRr,f: assessment of delivered AVC/H.264 video quality over 802.11a WLANs with multipath fading," in *Proceedings of the 1st Multimedia Communications Workshop (MULTICOMM '06)*, pp. 1–6, Istambul, Turkey, June 2006.

[23] S. Winkler, *Digital Video Quality: Vision Models and Metrics*, John Wiley & Sons, New York, NY, USA, 2005, Genista Corporation, Montreux (Switzerland).

[24] N. Feamster and H. Balakrishnan, "Packet loss recovery for streaming video," in *Proceedings of the 12th International Packet Video Workshop*, pp. 11, Pittsburgh, Pa, USA, April 2002.

[25] M. Horowitz, A. Joch, F. Kossentini, and A. Hallapuro, "H.264/AVC baseline profile decoder complexity analysis," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 704–716, 2003.

[26] T. K. Sarkar, Z. Ji, K. Kim, A. Medouri, and M. Salazar-Palma, "A survey of various propagation models for mobile communication," *IEEE Antennas and Propagation Magazine*, vol. 45, no. 3, pp. 51–82, 2003.

[27] H. Hashemi, M. McGuire, T. Vlasschaert, and D. Tholl, "Measurements and modeling of temporal variations of the indoor radio propagation channel," *IEEE Transactions on Vehicular Technology*, vol. 43, no. 3, pt. 1-2, pp. 733–737, 1994.

[28] F. Babich and G. Lombardi, "Statistical analysis and characterization of the indoor propagation channel," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 455–464, 2000.

[29] O. A. Lotfallah, M. Reisslein, and S. Panchanathan, "A framework for advanced video traces: evaluating visual quality for video transmission over lossy networks," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 42083, 21 pages, 2006.