Call Admission Control Dimensioning for VoIP Traffic over Wireless Access Networks: From Network to Application-specific Perspective

Kiril Kassev, Yakim Mihov, Adriana Kalaydzhieva, Boris Tsankov
Department of Telecommunication Networks
Technical University of Sofia
Sofia, Bulgaria
e-mail: kmk@tu-sofia.bg; yakim_mihov@abv.bg; akalaidjieva@abv.bg; bpt@tu-sofia.bg

Abstract—**Admission control is a key issue for quality of service (QoS) provisioning in both wired and wireless communication networks. Providing QoS for voice traffic transmission over packet-based network is a crucial task, which requires development of accurate resource estimation models. In order to perform bandwidth saving and improve the naturalness of the voice service contemporary voice coding schemes are equipped with the functionality of voice activity detection and background noise transmission during inactive speech periods. The adoption of traditional ON-OFF traffic model may cause significant errors in estimating the bandwidth required to meet the performance bounds of aggregated traffic flow. The aim of this article is to present a "new paradigm" of call admission control (CAC) dimensioning applicable to wireless access transmission media. The admission decision policy is codec-dependent and relies on the concept of user-perceived voice quality, since it could provide a tight connection with the QoS metrics that shall be guaranteed by the network. The well-known bufferless fluid-flow method is applied and new simple exact formulas for CAC performance evaluation are derived. The proposed methods are illustrated with numerical examples and some comparisons are made.**

Keywords-CAC; comfort noise generation; MOS; packet loss; perceived voice quality; VoIP over wireless access networks

I. INTRODUCTION

As next generation networks architecture is moving towards packet, also known as "all-IP" architecture, emerging wireless access technologies rely on fully shared radio resource allocation schemes, which allow scarce resource usage in an efficient way. Interactive services as well as real-time constraining services, such as voice or streaming applications, are supported over shared radio resource. Thus, it becomes critical that the emerging wireless access systems should be capable of employing efficient radio resource management schemes for packet data, in order to ensure that real-time services can be supported according to their stringent QoS requirements.

Providing QoS for real-time traffic flows in modern telecommunication networks is still a challenging task, which can be split up as follows: QoS guarantee in access network and QoS provisioning in the backbone. The former needs more difficult and more costly solutions, compared to the latter. It is due to the cheap available resource (e.g., fiber optic) and easy way of providing additional bandwidth in the core network, in comparison with the scarce and expensive resources in the wireless access networks (WAcN) as well as the specific features of the wireless medium. As a consequence, the QoS requirements may not be satisfied, even though a large amount of resources (i.e. bandwidth) is allocated to a certain connection.

The possibilities of contemporay WAcN, such as Wi-Fi (a set of standards IEEE 802.11), WiMAX (a set of standards IEEE 802.16), and LTE (3GPP standard of Long Term Evolution), to serve voice traffic are subject of particular attention. In spite of the increasing popularity of pure data services, the voice service demands still remain the biggest revenue contributor of telecommunication network operators [2]. VoIP traffic flows over WAcN encounter different problems and particularities like: a) stringent norms of admissible delay; b) the traffic flow is formed by relative short packets with high arrival rate; c) the scarce radio resource is wasted due to the relative long packet header, which can considerably exceed the packet payload, carrying voice frames; d) time-varying channel conditions.

Transmission of voice over packet-based networks is possible by applying different coding techniques. In order to prevent wasting available resources in WAcN voice codecs employing silence suppression techniques are preferred to be used. This is a reasonable solution in contrast to the implementation of constant bit rate (CBR) codecs, which are well accepted in systems where the network resources are not problematic. An additional option of the modern voice codecs is their ability of improving speech quality of parties participating in communication. This is done by generating a special frame, called Silence Insert Descriptor (SID), which describes the talker's background noise. As a result, the traffic profile of aggregated voice traffic should be addressed, by developing accurate models in which the generation of SID frames is included.

Since VoIP services continue to be commercially attractive for network operators, their adoption could be influenced by the users' satisfaction. The major challenge of the packet-based networks (either wireless access or core domain) serving streaming traffic is to provide QoS guarantee, such that consumer satisfaction is the same or similar to that of conventional fixed or cellular telephone services. This should imply methodologies and models for perceived voice quality prediction to be incorporated in algorithms for network resource management, with respect to QoS provisioning.

On the other hand, an important aspect of radio resource management and QoS provisioning in wireless access networks is towards the CAC mechanism implementation. The necessity of CAC arises with the wide deployment of connection-oriented packet switching technologies. Based on a source's traffic characteristics and required performance metrics, it encompass a set of tools, which has to take a decision whether or not a new connection can be accepted by the system, in addition of those connections in progress. If a new connection is admitted, it must not deteriorate the bandwidth usage and the performance of the connections already established. Since CAC is a fundamental mechanism for congestion control and QoS provisioning, it has been extensively studied in both wired [3] and wireless [4] network domains.

The CAC design and performance analysis became an inseparable part of ATM- and IP-based networks planning process [5], including different wireless networks as well. It can be classified based on various objectives and design options, such as QoS parameters (call-level and packet-level congestion probabilities, packet delay, bandwidth guarantee); throughput optimization, power allocation, fairness; controlling handover dropping probability, etc. Our research work is focused on both call blocking and packet dropping probabilities. Call admission decision is based on a simple threshold rule – an offered call is accepted if all the calls in progress are less than a pre-calculated threshold value.

Taking into consideration the decision time, CAC schemes can be classified as proactive (parameter-based) and reactive (measurement-based). In the former scheme the decision rule is based on a predictive analytical evaluation of the QoS constrains, while in the latter scheme, the CAC decision is based on certain QoS measurement. In order to get more efficient congestion control, a combination of both approaches can be realized.

This article is an extended version of the research work carried out in [1]. We investigate admission control schemes for streaming traffic and propose a new paradigm of a proactive CAC mechanism and VoIP dimensioning framework, based on perceived voice quality evaluation models. We aim at developing a solution that is applicable for a broad class of contemporary VoIP coding schemes.

II. RELATED WORK

The packet form of voice transmission differs considerably from other data transmissions. The VoIP traffic is streaming with stringent delay requirements. The activity periods (talk-spurts) are relatively long and the transmission rate of active voice frames (ACT) is not very high. The packets are relatively short with constant length and more often the packet header size is larger than the payload size. Due to the great interest in using IEEE 802.11 and 802.16 standards as well as emerging 3GPP LTE with E-UTRAN as access networks for packetized voice transmission, there exist numerous publications on modeling the VoIP call performance. References [6]-[8] (Wi-Fi), [9][10] (WiMAX) and [11]-[13] (LTE) are just a few among the latest ones.

In voice communications, speech is usually represented as a sequence of talk-spurts, interleaved with silence periods. This is a consequence of the widespread employment of voice codecs schemes with silence suppression or Voice Activity Detection (VAD) feature. In teletraffic point of view, this led to modeling the VoIP traffic pattern as an ON-OFF source. The analytical modeling of the ON-OFF voice traffic started with voice over ATM and continues with voice over IP applications [5]. Among analytical methods for performance evaluation, Markov-Modulated Poisson Process (MMPP) [14] and Fluid-flow model [15] are well-distinguished. Both models are described and compared in more details in [41][42]. The MMPP approach is more accurate, but at the cost of more computational efforts, when compared with the fluid-flow model, which is usually preferred as a less complex solution [16].

Although quite low bandwidth requirements, the packetized voice features lead to poor traffic performance in the high-speed WAcN [7][8][17]. In [6], the theoretical capacity for VoIP traffic of IEEE 802.11b WAcN is computed to be 15 calls. A comparative study shows the benefit of implementing silence suppression technique, which enhances the theoretical capacity to 38 calls. This is still a poor performance compared with the IEEE 802.11b radio channel PHY rate of 11 Mbps. Thus, it is beneficial to improve the performance and properly dimension the WAcN for VoIP traffic.

Along with the advantages of VAD feature for bandwidth reduction, its application may lead to sudden drop of the signal level during voice inactivity periods (OFF state), which is perceived unpleasant by the other dialogue party. Hence, to fill up this inactive period of time, a description of the background noise characteristics shall be sent from the inactive voice encoder. This is done by SID frames generated by the codec's algorithm. The corresponding output signal is referred to as comfort noise [18]. Hence, we should note that the presence of SID packets in VoIP traffic pattern can affect the traditional ON-OFF traffic evaluation [19]. Since the ON-OFF model is not valid for such a case, in [14] authors even went so far as to suggest the notation ON-SID instead of ON-OFF.

CAC techniques in modern wireless networks have been a subject of intensive study [4]. The design of CAC schemes for voice traffic has often been based on QoS performance metrics, which are treated separately (e.g., a design objective could be to achieve minimum packet loss rate for a specified delay constrains). On the other hand, in contrast to the TDM systems in which G.711 CBR codec can be only used, voice communications over packet networks can be realized by a number of low-bit-rate codecs for the purposes of bandwidth saving. Due to the complicated signal processing algorithms, user satisfaction strongly depends on the robustness of the particular type of codec to the instantaneous network performance (e.g., packet loss rate). In order to provide tight connection between both voice quality perceived by users and QoS parameters that shall be guaranteed by the network, either subjective or objective methods, developed through the years, should be incorporated. The Mean Opinion Score (MOS) is the most widely used subjective measure of voice quality expression and the ITU-T E-model is a computational model for predicting voice quality from network parameters [20]-[22].

MOS-based rate adaptation for VoIP sources has been incorporated in [23]. Architecture for adaptive control of a VoIP source coding rate, based on the state of the network, is proposed. The goal is to maximize the voice quality perceived by the receiver. Similar approach is presented in [24], where dynamic joint source and channel coding adaptation algorithm for the AMR speech codec is proposed. It is aimed at finding the optimum solution between packet loss recovery and end-to-end delay in either wired or wireless networks, in order to maximize the perceived voice quality. The ITU-T E-model is successfully incorporated in [25], which proposes an optimization algorithm that selects coding scheme, packet loss bound and maximum link utilization for a VoIP connection.

There are a large number of researches concerning the voice quality prediction models and their main application in playout buffer optimization algorithms. This is due to the fact that in the past, the choice of a buffer algorithm was entirely based on buffer delay and loss performance, treated separately (e.g., minimum end-to-end delay for a given packet loss).

In [26] a method for enhancing perceived quality of streaming applications and adaptive playout buffer control is proposed. The method is based on the assumption that the relationship between MOS and packet loss for codecs is linear, which is not correct, especially for newer codecs.

In [27] the assessment of how buffer algorithms affect perceived voice quality and how to choose the best algorithm and its parameters to obtain the optimal perceived quality has been carried out. The results are based on Internet trace data measurement and a new methodology for predicting speech quality, which combines the ITU-T Perceptual Evaluation of Speech Quality (PESQ) and ITU-T E-model.

Taking into consideration the widespread use of the E-model for voice quality prediction, Sun [28] points out that it is suitable for a limited number of codecs and network conditions. This is expressed by the necessity of performing time-consuming subjective test in order to derive model parameters. As a result, new accurate models for objective, nonintrusive voice quality prediction in packet networks are proposed. Based on a new methodology, authors of [28] propose efficient regression models, which can be applied for a number of modern codecs.

The packet loss probability evaluation and accurate VoIP bandwidth estimation for aggregate traffic, with respect to the perceived voice quality metrics is a crucial task in a CAC dimensioning. Comparative study shows that a little of research activities concerning this topic area have been carried out so far. A new approach to solving this problem is a subject of the present article, taking into account VAD and Comfort Noise Generation (CNG) features of the contemporary voice coding schemes (codecs).

III. System Model

The current section deals with the development of analytical models for call- and talk-spurt level performance evaluation of the proposed CAC mechanisms for VoIP traffic transmission.

Almost all known publications on packetized voice traffic performance over WAcN have a common feature – the investigations are carried out under overloaded (or nearly overloaded) system conditions. It is explained merely by the fact that the investigations are fulfilled by means of simulations. Good references are [29] and [30], where extensive research work and interesting characteristics of VoIP traffic service through a IEEE 802.16 system are obtained by means of simulations. Telecommunication service providers are mainly interested in system behavior under real traffic load, where the QoS measures such as call blocking, packet loss, etc. are rare events. The reason is that the QoS norms applied for commercial public networks restrict the traffic volume before an overload can occur. Direct simulation of rare events is time and resource consuming process. Thus, simulation acceleration methods are necessary to reduce the computing time. These methods may face certain problems. The proper setting of a particular method's parameters is critical for its performance and accurate results evaluation. An effective alternative is to employ analytical methods for system performance evaluation

Another reason to have a preference on analytical methods is the necessity of cross-layer design application due to the radio channel characteristics. Multi-layer simulations are often hard to be performed because they involve widely different time scales [31] (e.g., call request arrival, packet arrivals, packet processing at MAC level, etc.). Performance analysis by means of multi-layer simulation will take large amount of time and may be unpractical for a broad class of problems analysis.

The subject of interest is the traffic flow generated by multiple homogeneous sources and the bursty traffic in particular. Our considerations are restricted to streaming (real-time) traffic, generated by VoIP sources. We aim at developing an approach for Generalized VoIP (GVoIP) traffic source characterization and parameters estimation of a CAC for an access network.

A. Modeling of VoIP traffic with VAD and SID over wireless systems

A WLAN, WMAN or E-UTRAN, being an access network, is just one hop of a multi-hop end-to-end connection. The total one-way delay for good voice quality is fixed by ITU-T recommendations [32] to be less than 150 ms. Both observations and simulations show that the MAC functional delay is about 15 – 20 ms [9]. The same delay sustains at the other end. An additional delay of 50–60 ms can occur due to the necessity of jitter buffer [19]. Therefore, after extraction of possible queues delays of the backbone routers the remaining delay budget for a wireless access network is quite limited.

The stringent requirements on delays limit the ways in which the voice traffic losses (packet-scale and talk-spurt-scale) can be handled. The packet-scale losses are easy to be prevented by means of a short buffer. However, it is not practical to prevent talk-spurt-scale losses by means of a buffer. The buffer in this case would have to be large enough, and this would introduce an unacceptable delay. The talk-spurt-scale losses can be reduced to an admissible amount by selecting an appropriate medium transmission rate C during the wireless access network traffic planning process.

In this article, the bufferless fluid-flow approach is used for CAC parameters determination. The assumption that there is not a buffer at talk-spurt level leads to conservative estimates for packet losses and therefore to the safety side of CAC parameter determination.

The bufferless fluid-flow model is used quite a while ago [33]–[35] and [5, Chapter 12] due to its effectiveness and simplicity.

In the majority of investigations on VoIP traffic performance analysis, it is widely adopted the speech generation process to be modeled as a consequence of talk spurts and inactive periods, whose duration is usually assumed to be exponentially distributed. The so defined ON-OFF model does not take into consideration the existence of SID frames in the voice traffic pattern, generated by the voice encoder. When the source is in the ON state, it produces voice frames with a constant bit rate, which are encapsulated in voice packets. During the OFF state, the source sends no packets.

In order to determine the maximum number of calls (sessions) N admitted to the system by the CAC, meeting certain objectives, the bufferless fluid-flow or burst-scale loss approach is well-accepted.

An ON-OFF traffic source is usually characterized by means of the following parameters: the bit rate during a talk-spurt R; the mean bit rate r; the talk-spurt duration T_{ON} and the inactive period duration T_{OFF}. The mean bit-rate during ACT frames (packets) generation process is:

$$r = \frac{T_{ON}}{T_{ON} + T_{OFF}} R = \alpha R. \tag{1}$$

According to (1), an ON-OFF source could be characterized by means of any three out of four parameters. It should be noted that $1/\alpha$ measures the peak to mean ratio of the rate produced by a call and α denotes the voice activity factor.

By employing voice encoders that randomly generate SID frames (packets) during voice inactive periods, the OFF state bit-rate is $R_{OFF} > 0$. The VoIP traffic source with VAD and background noise transmission is called a *Generalized VoIP* (*GVoIP*) source [36]. A GVoIP traffic source is characterized by means of the above mentioned parameters plus the bit rate during the inactive period R_{OFF}. Hence, the overall mean bit rate r^G of a GVoIP source is derived as:

$$r^G = \alpha R + (1-\alpha) R_{OFF.} \tag{2}$$

Since SID frames generation has an influence on the overall packet flow, it could be quantitatively expressed as the ratio R_{OFF} / R, denoted by γ [36]. Thus, Equation (2) can be rewritten as follows:

$$r^G = \alpha R + (1-\alpha) R_{OFF} = \alpha R\left[1 + \frac{1-\alpha}{\alpha}\gamma\right]. \tag{3}$$

Comparison of (1) and (3) gives an increase of the voice mean bit rate due to the SID frames generation, i.e.

$$r^G = r\left[1 + \frac{1-\alpha}{\alpha}\gamma\right]. \tag{4}$$

The actual value of γ can vary, but usually the typical values are limited to 0.1 [37]. On the other hand, the typical values of α range from 0.35 to 0.45 [6]. Therefore, it could be expected an approximate increase of a GVoIP source mean bit rate up to 20 %.

For an aggregated model of multiple homogeneous VoIP sources, which generate SID frames, the arriving flow rate (AFR) of the multiplexed voice calls is:

$$AFR = jR + (i-j) R_{OFF,} \tag{5}$$

where, in the general case, i denotes the number of calls admitted to the system and j – the number of calls in talk-spurt.

B. System model and packet loss paradigm

Let us formulate the problem of aggregating a number of voice traffic sources over the radio interface of an access network. We suppose C represents the total bit rate link capacity allocated by a base station, of a particular wireless access technology, for the purposes of streaming traffic transmission. It is convenient to use the notation R (transmission rate during talk-spurt period) as a unit transmission resource. Therefore,

$$C = nR, \tag{6}$$

where

$$n = C / R \tag{7}$$

denotes the number of network resources (referred to as transmission resource units as well). In case of a classical ON-OFF model (without SID) n represents the maximum number of active calls that can be simultaneously served without any packet losses.

However, if n^G denotes the maximum number of active GVoIP traffic sources, including CNG feature that can be served simultaneously without packet losses, then:

$$C = n^G R + (i - n^G) R_{OFF}. \tag{8}$$

As a consequence:

$$n^{G} = \frac{\frac{C}{R} - i\gamma}{1-\gamma} = \frac{n - i\gamma}{1-\gamma}. \quad (9)$$

Hence, the maximum number of talk-spurts (active calls) that can be simultaneously served without packet losses is not constant any more, but depends on the number of the calls i admitted to the system. Recall that in the classical ON-OFF model, this value is constant and depends on C only (7).

The main parameter of a proactive CAC scheme is the maximum number of calls (sessions) admitted to the system by the CAC, denoted by N. In order to gain from the statistical multiplexing, it is obvious that $N > n^{G}$. On the other hand, the following expression $N < C / r^{G}$ shall be fulfilled in order to prevent a buffer from permanently overflowing. Hence, the following condition can be derived:

$$n^{G} < N < \frac{C}{r^{G}}. \quad (10)$$

The research efforts will be directed towards determining both the CAC threshold value N of maximum admissible calls (to keep the call blocking below the norm) and the radio-link resources n (necessary to keep the packet loss value below the norm).

The analytical evaluation is based on the following parameters: the offered traffic A, in terms of Erlangs; the GVoIP source parameters (such as, T_{ON}, T_{OFF}, R, and R_{OFF}); the QoS parameters – call blocking probability B and packet loss probability P_{PL}.

Our attention is paid to the packet loss evaluation, as a more intricate task. The analytical model derived can be used for evaluation of the necessary medium transmission rate C, with respect to the packet loss probability requirements.

The ratio of the packets loss rate to the arriving packets rate gives the probability of packet losses [5]

$$P_{PL} = \frac{E\,(AFR - C)^{+}}{E\,(AFR)}, \quad (11)$$

where E stands for the expectation operator, the numerator denotes the aggregated flow excess rate, and $(z)^{+}=\max(z,0)$.

Assuming the worst-case with N admitted calls (i.e. $i=N$) and using R as a transmission resource unit, the packet loss probability is derived by the following expression:

$$P_{PL} = \frac{\sum_{j=\left\lceil \frac{n-N\gamma}{1-\gamma} \right\rceil}^{N} P(j\,|\,N)\cdot\left[j+(N-j)\gamma - n\right]}{\sum_{j=0}^{N} P(j\,|\,N)\cdot\left[j+(N-j)\gamma\right]}, \quad (12)$$

where j represents the number of talk-spurts (active sources). In the case of multiple sources, we need to calculate the probability that j sources are active, given that i traffic sources are admitted. This conditional probability is given by the binomial distribution:

$$P(j\,|\,i) = \binom{i}{j}\cdot \alpha^{j}\cdot(1-\alpha)^{i-j}. \quad (13)$$

For the worst-case with N admitted calls, $i = N$.

Both the numerator and the denominator in (12) represent flows that are mixture of packets that carry voice (ACT frames) and packets that carry background comfort noise (SID frames). As Estepa [16][38] have correctly pointed out, since the loss of ACT packets forms the main influence on the perceived voice quality, it makes sense to account the loss of ACT packets only. The proportion between ACT and SID packets in the excess rate flow varies and depends on the number j of talk-spurts. In the aggregated excess rate flow exactly $\frac{j}{\left[j+(N-j)\gamma\right]}$ parts belong to ACT packets and the overall offered ACT flow is merely $\alpha \cdot N$. Therefore, ACT packets loss probability is expressed as:

$$P_{PL}^{ACT} = \frac{\sum_{j=\left\lceil \frac{n-N\gamma}{1-\gamma} \right\rceil}^{N} P(j\,|\,N)\cdot\left[j - \frac{n\cdot j}{j+(N-j)\gamma}\right]}{N\alpha}. \quad (14)$$

It should be noted that in [16][38] the following approximation formula for P_{PL}^{ACT} is used (rewritten here for the bufferless fluid-flow model and based on notations we have adopted):

$$P_{PL}^{ACT} \approx P_{PL}\cdot\left[1+\frac{(1-\alpha)\gamma}{\alpha}\right], \quad (15)$$

where P_{PL} is evaluated by (12).

A common feature, when N is determined, is to consider the case $i=N$ only and evaluate P_{PL} [5, pp. 141]. This corresponds to a system where all traffic sources suffer highest losses. For carrier-grade voice service the network operator has to guarantee QoS similar to that of circuit switched networks. For this reason, for a properly planned network the common assumption of $i=N$ tends to be an extremely rare event. As a consequence, it is significantly important to take into consideration all possible system states for which losses are encountered, when performing an analytical evaluation of P_{PL}.

A VoIP source is either in an idle or busy state. In case of VAD function implementation the VoIP source is alternating OFF and ON states after the call is setup (Fig. 1a). An assumption is made that any call originates or terminates when it is in silence period only. In case a call can originate or terminate during an active period (during a talk-spurt) the call state-transition diagram is modified as shown on Fig. 1b. Probabilities of transition from state *busy ON* to state *idle*

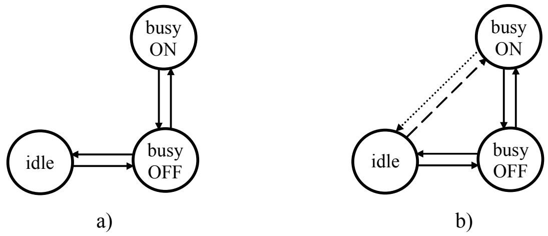

Figure 1. A VoIP source state-transition diagram

(dotted line on Fig. 1b) and particularly from state *idle* to state *busy ON* (dashed line on Fig. 1b) are negligible (normally there is a packet generation offset at the call start and seldom the call end will interrupt talk-spurt) and they will be omitted in our considerations.

At any time the radio link system can be in a state (i, j) where i $(i = 0,1,...,N)$ represents the number of accepted calls (N denotes the upper bound of the admitted calls) and j $(j = 0,1,...,i)$ is the number of active calls (number of talk-spurts in progress). The call flow forms a Poisson process with call rate Λ_c and call service time $1/\mu_c$, whereas the burst flow forms a Binomial process with single ON source burst arrival rate $\lambda_b = 1/T_{OFF}$ and single OFF source burst service rate $\mu_b = 1/T_{ON}$. The set of states (i, j) forms a two-dimensional Markov chain with the corresponding state-transition diagram shown on Fig. 2.

Following the assumption the call flow forms a Poisson process, the probability of exactly i sources being busy is given by

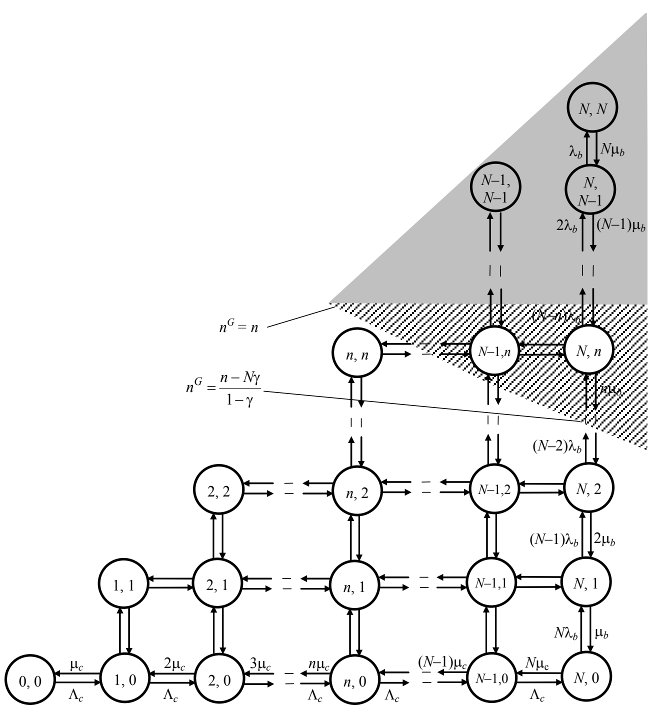

Figure 2. System state-transition diagram

$$P(i) = \frac{\frac{A_c^i}{i!}}{\sum_{x=0}^{N} \frac{A_c^x}{x!}}, \tag{16}$$

where

$$A_c = \frac{\Lambda_c}{\mu_c}.$$

The Markov process, which is represented by the two-dimensional state transition diagram on Fig. 2, is reversible. Thus, the performance measures of interest can be derived by the states probabilities, which are given on product form. A state joint probability $P(i, j)$, under the assumption of statistical equilibrium, is expressed as

$$P(i, j) = P(i) \cdot P(j \mid i), \tag{17}$$

where the conditional (burst flow) probability $P(j \mid i)$ is obtained by (13).

Let us first consider the case when the aggregated traffic flow is generated by multiple VoIP sources, where each one is represented as an ON-OFF traffic model.

Since we are interested in packet loss probability evaluation P_{PL}, in any state for which $j > n$ (the gray-filled area on Fig. 2) a packet is lost with certain probability. The offered rate in a state (i, j) is jR and the excess rate is $(j - n)R$. Therefore, the excess rate mean value is given by

$$\sum_{i=\lceil n \rceil}^{N} \sum_{j=\lceil n \rceil}^{i} R(j-n)P(i, j). \tag{18}$$

Based on (11) and after some rearrangements, the packet loss probability is given by the following relation:

$$P_{PL} = \frac{\sum_{i=\lceil n \rceil}^{N} \sum_{j=\lceil n \rceil}^{i} (j-n)P(i, j)}{\sum_{i=\lceil n \rceil}^{N} P(i) \sum_{j=0}^{i} jP(j/i)} = \frac{\sum_{i=\lceil n \rceil}^{N} \sum_{j=\lceil n \rceil}^{i} (j-n)P(i, j)}{\sum_{i=\lceil n \rceil}^{N} P(i) i\alpha}. \tag{19}$$

In order to perform a comparative analysis considering the worst-case scenario, i.e. $i = N$ [5, pp. 141], P_{PL} could be derived from (19)

$$P_{PL} = \frac{\sum_{j=\lceil n \rceil}^{N} (j-n)P(j/N)}{N\alpha}. \tag{20}$$

As a further step of the research, it should be noted that the traffic generation pattern of the contemporary VoIP coding schemes comprises of SID packets generated during the OFF periods. Thus, the presence of such packets affects the traditional ON-OFF traffic model and hence, the packet loss evaluation methodology when GVoIP sources are employed.

Based on the state-transition diagram (Fig. 2) as well as the main properties of GVoIP sources, in any state for which $j > n^G$ a packet loss with certain probability occurs. On the other hand, it could be seen that the parameter n^G is tightly coupled with the number of admitted calls, which is expressed by (9).

Following (11), it is more realistic to consider the ratio of lost packets to all arrived packets just for VoIP calls with packet losses only, i.e. $i > n$ (Fig. 2).

Hence, the packet loss evaluation of aggregated GVoIP traffic flow could be obtained by the following expression:

$$P_{PL} = \frac{\sum_{i=\lceil n \rceil}^{N} \sum_{j=\left\lceil \frac{n-i\gamma}{1-\gamma} \right\rceil}^{i} P(i,j) \cdot [j + (i-j)\gamma - n]}{\sum_{i=\lceil n \rceil}^{N} \sum_{j=0}^{i} P(i,j) \cdot [j + (i-j)\gamma]}. \quad (21)$$

Both the numerator and denominator in (21) represent a packet flow, which is a mixture of packets carrying voice (ACT packets) and those carrying background comfort noise pattern (SID packets). As it was pointed out earlier, we are interested in the ACT packets loss evaluation, due to their main influence on the perceived voice quality. The proportion between ACT and SID packets in the excess rate flow varies and depends on both the number of calls (i) and the number of talk-spurts (j). It should be noted that the total offered ACT flow at states with i calls is $\alpha \cdot i$. Therefore, we propose the following exact packet loss formula:

$$P_{PL}^{ACT} = \frac{\sum_{i=\lceil n \rceil}^{N} \sum_{j=\left\lceil \frac{n-i\gamma}{1-\gamma} \right\rceil}^{i} P(i,j) \cdot [j - \frac{nj}{j + (i-j)\gamma}]}{\sum_{i=\lceil n \rceil}^{N} P(i)\alpha i}. \quad (22)$$

The main contribution of the models we have developed is to propose a new paradigm of CAC dimensioning for VoIP traffic pattern generated from homogeneous sources with VAD and CNG features. As further part of the research, we address this problem from the users' perspective as well, by incorporating the broad term of perceived voice quality.

C. CAC Dimensioning with respect to the perceived voice quality

The concept of user-perceived QoS and its evaluation for real-time services, such as VoIP, is a key issue, since it could provide a tight connection with the QoS metrics that shall be guaranteed by the network. In VoIP applications the delay and packet losses are the main impairments that have direct impact on the perceived voice quality, and the MOS is the most widely used measure for this purpose.

Due to variety of voice coding schemes (codecs) and their robustness to the network conditions, all the information about the type of codec and packet losses is suitably represented by an appropriate equipment impairment model I_e, while the delay impairment model I_d encompasses a number of impairments due to one way delay of voice packets.

Since the models' parameters are usually calculated by a set of complex equations [22], efficient regression models for objective, nonintrusive voice quality prediction in packet networks are proposed in [28], which combine the ITU-T speech quality measurement algorithms (either PESQ or PESQ-LQ) and ITU-T E-model.

A non-linear equipment impairment I_e regression model for a number of contemporary low-bit-rate codecs is derived in [29] and has the following form:

$$I_e = a \ln(1 + b\rho) + c, \quad (23)$$

where ρ denotes the packet loss probability (in percentage), and the parameters (a, b, and c) are derived under the PESQ or PESQ-LQ algorithm and depend on the particular coding scheme used. Employing complicated signal processing algorithms, voice codecs can also have an impact on the perceived quality under zero packet loss and delay conditions. This is expressed by the parameter c in (23).

Unlike I_e which is codec dependent, the I_d factor is common to all codecs and it could be derived by a polynomial fitting of 6-th order [28]. Among all the impairments included in I_d, we take into consideration the following ones (assuming the rest to be in perfect condition):

(a) *packetization delay* d_{pack} – time taken to fill the packet payload, i.e. for the purposes of the packetization process it is necessary the number of codec frames N_{fpp} that will be encapsulated in each packet to be set. This option has a great importance, because for a particular type of codec, it can influence the bit rate the ACT and SID packets are generated. The packet mean bit rate is significantly increased, especially for low values of N_{fpp}, and vice versa. On the other hand, there is a trade-off between the packets mean bit rate and the packetization delay;

(b) *network delay* – contemporary packet-based networks are too complex and large-scale systems. Thus, standard techniques for delay analysis are not well-suited and a common practice for such studies lies in the scope of statistical delay distribution modeling, based on a network configuration and acquired trace data. This topic is covered in more depth in [39] and several statistical distribution functions are investigated and applied in [40] for a network delay description. As a consequence of applying statistical analysis, the network delay can be expressed by its mean value, denoted by $\overline{d_n}$. The overall mean packet delay is simply represented by $\overline{d} = d_{pack} + \overline{d_n}$.

Taking as a starting point the VoIP packet loss model in which all possible system states are taken into consideration (Fig. 2), we extend it to include the models for perceived voice quality prediction addressed previously. We aim at deriving the overall mean impairment factor $\overline{I_\Sigma}$, and hence the perceived voice quality, in terms of MOS, with respect to the parameters describing the underlying Markov process.

Carrying out the analysis, we can split up the state space in two non-overlapping areas, which are distinguished to each other by the packet loss occurrence. Thus, for $i \in (0, n]$, the packets are not lost and the mean impairment factor $\overline{I_\Sigma^{(0,n]}}$ includes the impairments of the codecs itself (under zero packet loss) and packets delay.

$$\overline{I_\Sigma^{(0,n]}} = (c + I_d)\sum_{i=0}^{\lfloor n \rfloor} P(i). \tag{24}$$

For $i \in (n, N]$ the influence of packet losses on perceived voice quality is quantitatively evaluated by applying the I_e model. The mean impairment factor $\overline{I_\Sigma^{(n,N]}}$ in such a case is obtained as follows

$$\overline{I_\Sigma^{(n,N]}} = \sum_{i=\lfloor n+1 \rfloor}^{N} P(i) \cdot [a \ln(1 + b\rho) + c + I_d], \tag{25}$$

where ρ denotes the mean packet loss rate (in percentage) for traffic flow of both ACT and SID packets and subject to $\forall\, i \in (n, N]$:

$$\rho = \frac{\sum_{j=\left\lceil \frac{n - i\gamma}{1-\gamma} \right\rceil}^{i} P(j \mid i)[j + (i - j)\gamma - n]}{\sum_{j=0}^{i} P(j \mid i)[j + (i - j)\gamma]} \cdot 100.$$

Therefore, for the overall mean impairment factor $\overline{I_\Sigma}$ we have

$$\overline{I_\Sigma} = (c + I_d)\sum_{i=0}^{\lfloor n \rfloor} \hat{P}(i) + \sum_{i=\lfloor n+1 \rfloor}^{N} \hat{P}(i)[a \ln(1 + b\rho) + c + I_d]. \tag{26}$$

If p denotes a threshold value of the set of states (at call-flow level) we are interested in during the perceived voice quality analysis, the following conditions shall be fulfilled: $0 \le p \le N$ and $p \le i \le N$. Hence, the common representation of the weighted factors in (26) is based on normalization of the set of probabilities $P(i)$, subject to the stated conditions, i.e.

$$\hat{P}(i) = \frac{P(i)}{\sum_{i=p}^{N} P(i)}, \qquad 0 < \hat{P}(i) \le 1. \tag{27}$$

$$\sum_{i=p}^{N} \hat{P}(i) = 1.$$

Having obtained the average impairment factor $\overline{I_\Sigma}$, and ignoring the other impairments (e.g., echo), the average R factor can be calculated as [22] and it commonly ranges from 50 (poor quality) to 100 (the best quality), i.e.

$$\overline{R} = R_0 - \overline{I_\Sigma} + A, \tag{28}$$

where R_0 incorporates the effect of noise, expressed by the basic SNR, and it does not depend on the network performance. For voice traffic it is assumed $R_0 = 93.2$. The advantage factor A takes into account the fact that some users could accept the voice quality reduction in return to the service convenience, for instance wireless access connection. Typical values of the advantage factor are $A = 0$ (wireline access) and $A = 10$ (GSM network access). Thus, the average R factor, expressed by (28), considers all the impairments as a result of particular network condition and the type of coding scheme employed.

The perceived voice quality is quantitatively expressed by the relationship between the R factor and MOS, as defined in ITU-T G.107 recommendation [22].

IV. Numerical Results

The current section is concerned with a quantitative analysis and comparative study of the methods for CAC dimensioning we have proposed.

In order to decrease the bandwidth usage the encoding scheme of each traffic source employs an activity detection function, which is quantitatively represented by the activity factor α. The offered traffic flow A_c is generated by multiple homogeneous (G)VoIP sources. The maximum number of calls (sessions) admitted to the system depends on the target call-level blocking probability B, which can be obtained by (16).

The commonly accepted ON-OFF model ignores the SID packets in the VoIP traffic pattern (γ = 0) and thus, its application could cause significant errors in estimating the bandwidth required to meet the performance bounds of aggregated traffic flow. Focusing on a more realistic case, where voice source traffic pattern includes both a silence suppression feature and comfort noise generation, we analyze a CAC mechanism assuming the worst-case scenario with N admitted calls, i.e. $i = N$. We set the ACT packets loss probability threshold (14) to $P_{PL}^{ACT} = 0.5\%$ [21]. The aim of the study is to determine the maximum number of admitted users N in case a fixed amount of network resource units n is allocated for the VoIP service. The quantitative relationship among the variables of interest is shown on Figure 3. In spite of improving the naturalness of conversations by introducing the CNG feature, SID packets generation during inactive periods (γ = 0.1) leads to additional consumption of allocated resources. This results in decreasing the number of users admitted to the system.

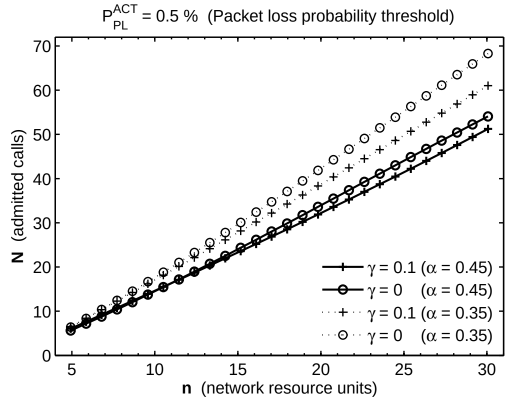

Figure 3. Number of admitted calls vs. allocated network resource units (a comparative analysis of multiplexing homogeneous VoIP and GVoIP sources, subject to $i = N$)

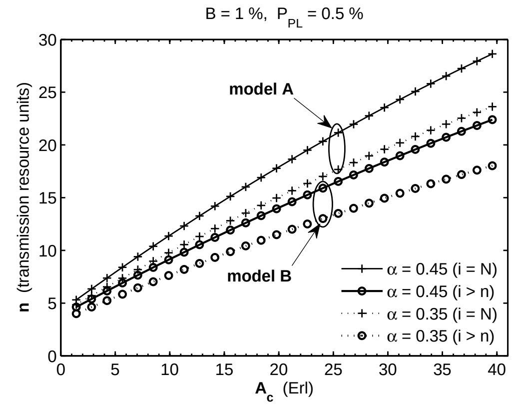

Figure 4. Access network dimensioning – CAC models comparison (homogeneous voice sources without CNG, subject to $i = N$ and $i > n$)

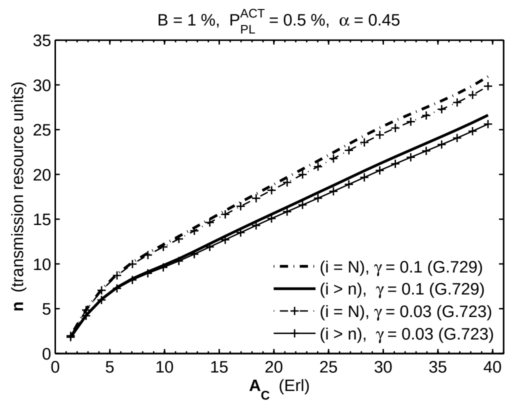

Figure 5. Comparative study of the dimensioning algorithm applicable for G.729 and G.723 coding schemes (homogeneous GVoIP sources, subject to $i = N$ and $i > n$)

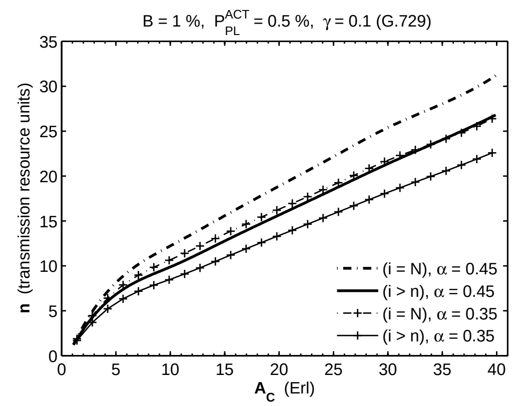

Figure 6. Comparative study of the dimensioning algorithm applicable for G.729 coding scheme

Based on the required performance thresholds, such as B and P_{PL}, as well as source traffic characteristics the task of CAC is to determine whether a connection can be accepted and, if accepted, the amount of network resources to be allocated. Fig. 4 depicts results of an access network dimensioning with typical values of P_{PL} and α, under assumption of offered traffic flow generated by homogeneous ON-OFF sources (without CNG feature). A comparative study has been carried out by applying both the model presented in [5], expressed by (20) (we will refer to it as "model A"), and proposed analytical model (19) (we will refer to it as "model B"). The "model A" corresponds to a system that encompasses the set of states for which $i = N$ is fulfilled. This is related to the most right column on the state-transition diagram (Fig. 2). Since network service providers are interested in more realistic system performance evaluation, it is necessary all possible system states to be taken into consideration (we can denote this condition as $i > n$). Numerical results show that this led to more efficient network resource (bandwidth) usage, because the system is not overdimensioned, as it is done by using (20).

For carrier-grade voice service it is of crucial importance a VoIP dimensioning framework for accurate estimation of the network resource, required to guarantee the performance bounds of aggregated GVoIP traffic sources, to be applied. This issue is addressed by the proposed analytical model (21) and (22), which is valid for any VoIP coding scheme, by setting both parameters α ($\alpha = 1$ corresponds to a CBR-type codec) and γ ($\gamma = 0$ – the model is valid for a VAD codec without CNG feature). We compare our dimensioning algorithm to the common approach of considering $i = N$. Comparative study includes the derived exact formula (14) for packet loss evaluation. We consider G.729 and G.723 coding schemes, both employing VAD and CNG features. Typical values of γ for both codecs are assumed to be 0.1 and 0.03 respectively [16]. Results depicted on Fig. 5 reveal the bandwidth (in term of transmission resource units) required in order to satisfy QoS constrains of aggregated traffic load A_c. It is demonstrated the bandwidth allocation margin that results from applying the proposed methodology for packet loss evaluation (22) ($i > n$).

On the other hand, silence suppression technology can considerably decrease the bandwidth usage needed. This is quantitatively represented on Fig. 6. Study results demonstrate that the same amount of network resource could be allocated to meet the call flow demands with higher value of activity factor ($\alpha = 0.45$) when the proposed approach is applied, compared to the case for which $i = N$.

Since we are interested in the voice (ACT) packets loss probability evaluation, in our analysis, we ignore the SID packet losses which are a part of the overall packet flow. This is represented by (22). On the other hand, SID packets are characterized with a small number and size (typically γ<0.1), which means they would not have great impact on the precision of voice packets loss probability evaluation, expressed by (21). The results presented on both Fig. 7 and Fig. 8 let us answer this question.

Fig. 7 depicts the comparison results of P_{PL} and P_{PL}^{ACT} versus transmission resource units n, for different values of offered traffic volume A_c and codec used. In order to get more accurate results, Fig. 8 shows the absolute difference $\Delta p = \left| P_{PL}^{ACT} - P_{PL} \right|$ of (21) and (22) for the same case. It could be concluded that there is no sense of using (22) instead of (21), when the offered traffic volume is high (above a certain threshold) as well as it is expected the packet loss probability is small enough (e.g., less than 1.10^{-3}).

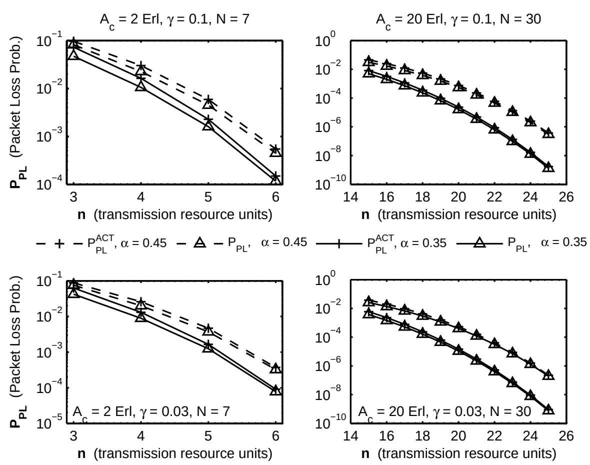

Figure 7. Comparative study of packet loss probability evaluation

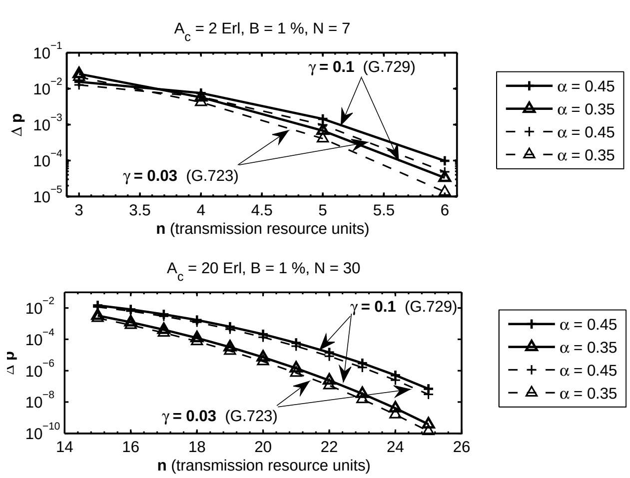

Figure 8. Comparative study of packet loss probability evaluation in terms of absolute difference (Δp)

The current trend of wireless networks dimensioning and performance evaluation is going towards application-specific quality measures, which takes into consideration the end user's satisfaction, rather than network parameters for QoS. This allows us to build a CAC mechanism that is capable of allocating the scarce network resource more precisely, based on the users' perceived QoS. The admission decision policy is codec-dependent and relies on the MOS value, which is expressed by the average impairment factor (26). We consider the following modern low-bit-rate coding schemes, supporting VAD and CNG: G.729, G.723.1, and the AMR codec (the highest mode – H, and the lowest – L), which has been adopted in the 3GPP networks. It is assumed the AMR codec maintains the either mode for an active call duration. The codec parameters for different values of N_{fpp} are calculated according to [16] and presented in Table I. The IP packets flow rate during a voice source active and inactive state is denoted by R and R_{off} respectively. A sequence of N_{fpp} consecutive codec frames (either ACT or SID) are sent in a single IP packet payload every $N_{fpp} \cdot T$ seconds, where the value T is codec-dependent and denotes the frame inter-arrival time.

The feature under investigation is towards a CAC mechanism dimensioning of an access point. We assume negligible packet delay ($I_d \approx 0$), which does not have a direct impact on the perceived voice quality evaluation. In order to achieve minimum packetization delay the number of codec frames that will be encapsulated in each IP packet is set N_{fpp} = 1 by the voice application.

Fig. 9 reveals the bandwidth (in terms of transmission resource units) required for offered aggregated GVoIP traffic flows to have a certain value of MOS, when a voice coding scheme of particular type is involved. The CAC decision policy is based on the target MOS value as well as the offered traffic load by setting call-level blocking probability B = 1 %. More generally, Fig. 10 depicts the link capacity

TABLE I. CODEC PARAMETERS

Codec	**AMR(H)**	**AMR(L)**	**G.723.1**	**G.729**
α	0.469	0.469	0.471	0.456
$\mathbf{N_{fpp}}$	1			
$\mathbf{N_{fpp} \cdot T}$ (ms)	20	29	30	10
R (bps)	19 600	20 800	17 067	40 000
$\mathbf{R_{off}}$ (bps)	1 662	2 400	900	4 583
γ	0.0852	0.1154	0.0527	0.1145
$\mathbf{N_{fpp}}$	2			
$\mathbf{N_{fpp} \cdot T}$ (ms)	40	58	60	20
R (bps)	14 069	12 800	11 733	24 000
$\mathbf{R_{off}}$ (bps)	1 662	2 410	789	4 583
γ	0.1181	0.1883	0.0672	0.191
$\mathbf{N_{fpp}}$	3			
$\mathbf{N_{fpp} \cdot T}$ (ms)	60	87	90	30
R (bps)	12 230	10 133	9 956	18 667
$\mathbf{R_{off}}$ (bps)	1 662	2 410	752	4 583
γ	0.1359	0.2378	0.0755	0.2456

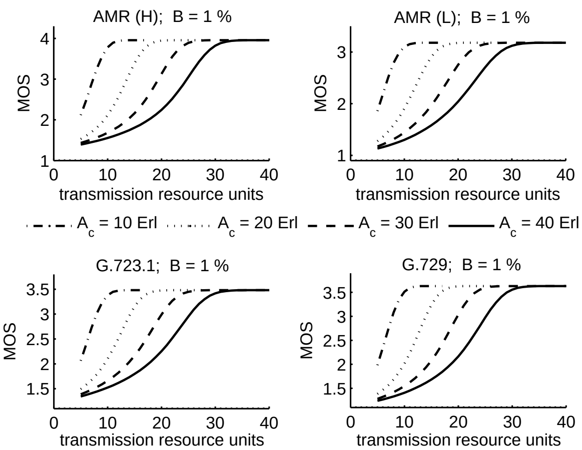

Figure 9. CAC dimensioning example – a common case

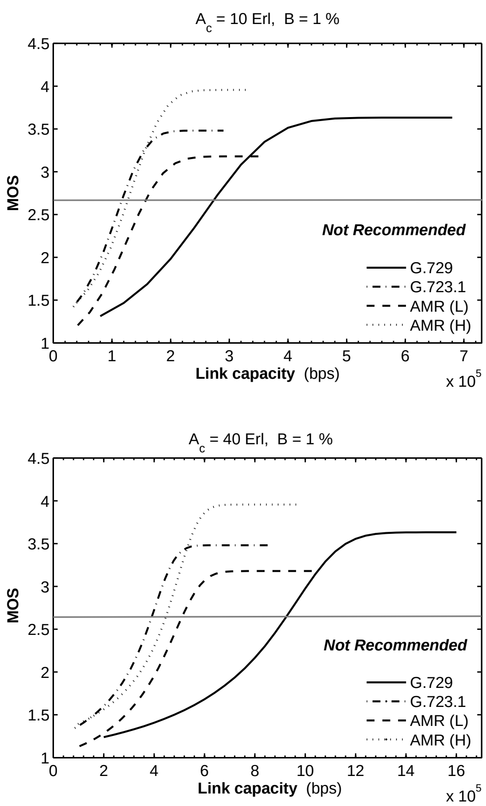

Figure 10. CAC dimensioning example for different coding schemes and GVoIP aggregated traffic load

that shall be guaranteed by an access point in order to reach certain perceived quality level. MOS rating is based on [22]. The method proposed could be incorporated into a codec negotiation procedure based on the allocated network resource and fulfilling the CAC decision criteria.

The VoIP traffic service over WAcN may encounter certain difficulties as a result of the wireless domain characteristics. The scarce radio resource may be wasted due to the traffic flow formed by packets with relative long headers, which could exceed the voice payload field several times. In most cases, this drawback can be minimized by adjusting the sequence of N_{fpp} consecutive codec frames building an IP packet payload field. From QoS perspective, this option leads to increase of packetization delay and it is not very appropriate, especially in case the voice packets could encounter considerable network delay. Applying the "new paradigm" of CAC dimensioning, Fig. 11 depicts the perceived voice quality, which could be achieved, under specific network parameters for a number of popular low-bit-rate voice codecs. It is assumed the aggregated GVoIP traffic load A_c = 40 Erl. It can be seen that for the same amount of consecutive voice frames per packet N_{fpp}, the access point shall allocate more resource when the average network delay d_n increases. This arises from the necessity of keeping the perceived voice quality at the same level when network conditions are getting worse. At the same time, the scarce resource is not wasted when the maximum possible value of MOS is requested to be obtained, since the capacity allocation to links could be accurately estimated by the model we propose.

V. Conclusion and Future Work

In this article, we have focused on a more realistic case where VoIP source traffic pattern includes not only a silence suppression feature, but comfort noise generation as well. The adoption of the traditional ON-OFF model may cause significant errors in estimating the bandwidth required in order to meet the performance bounds of aggregated VoIP traffic flow. Both call- and talk-spurt level considerations for a teletraffic design of access networks for voice services are proposed and comparative analysis has been carried out.

We propose a new paradigm of CAC dimensioning of wireless access networks serving packetized voice traffic. The new methodology has a broad area of application and is especially suitable for accurate link capacity estimation, taking into consideration the end user's service satisfaction, rather than network parameters for QoS treated separately. The approach is valid for a number of modern low-bit-rate voice codecs, employing VAD and CNG functionality, and is insensitive to the wireless technology in use.

Due to the great interest in emerging wireless access technologies, the research community is being interested in the system design, optimization and QoS requirements satisfaction of next-generation wireless access networks. The wireless environment features as well as user mobility draw the direction of the future research work. It will encompass the models we have developed and solve wireless communications resource management problems in order to maintain channel capacity and provide the performance guarantee. In order to achieve this goal, cross-layer adaptation and optimization mechanisms are going to be developed.

Acknowledgment

This article is a part of research work in the context of research project "Methods and models for resource management in convergent networks", funded by the Research and Development Sector of the Technical University of Sofia (RDS-TUS) under grant 102ni107-7.

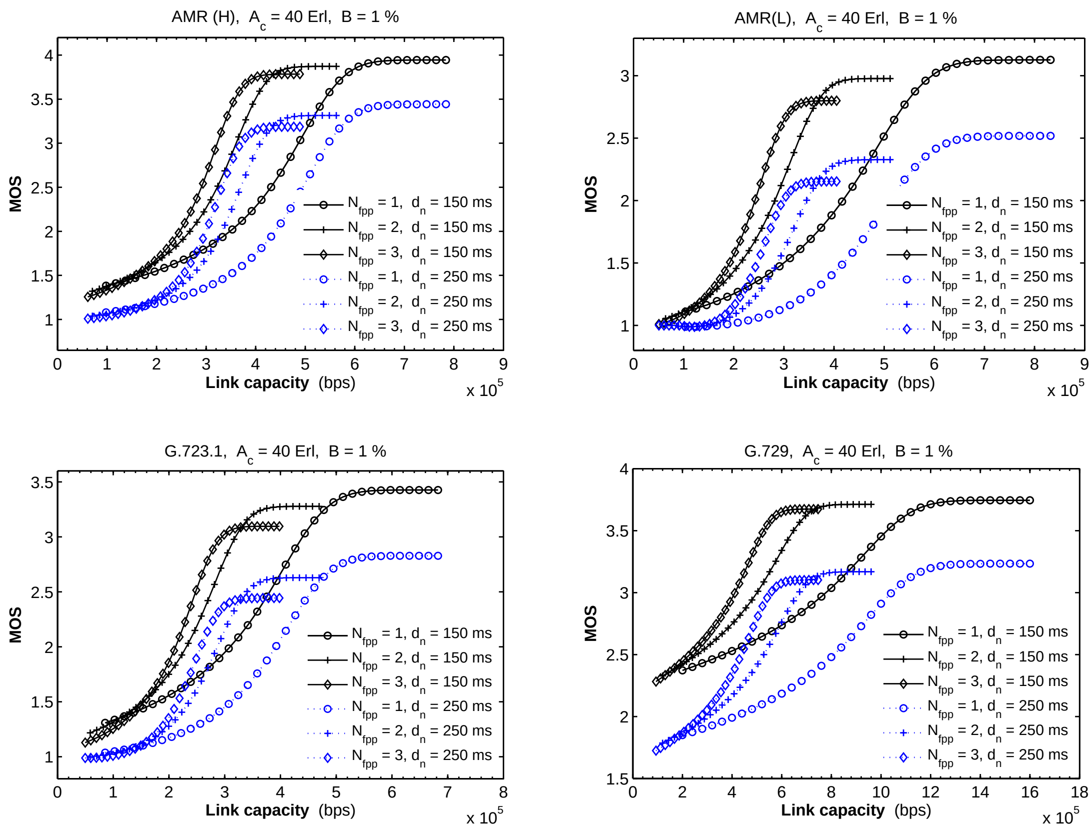

Figure 11. CAC dimensioning example for a number of low-bit-rate voice codecs, expressed in N_{fpp} and d_n

REFERENCES

[1] K. Kassev, Y. Mihov, A. Kalaydzhieva, and B. Tsankov, "A new paradigm of CAC dimensioning for VoIP traffic over wireless access networks," Proc. The Fourth International Conference on Digital Society (ICDS 2010), Feb. 2010, pp. 54-59.

[2] M. Chatterjee and S. Sengupta, "VoIP over WiMAX," in WiMAX Applications, S. Anshon and M. Ilyas, Eds. CRC Press, 2008, pp. 55-76.

[3] H. G. Perros and K. M. Elsayed. "Call admission control schemes: a review," IEEE Communications Magazine, vol. 34, no. 11, Nov. 1996, pp. 82-91.

[4] M. H. Ahmed, "Call admission control in wireless networks: A comprehensive survey," *IEEE Communications Surveys*, vol. 7, no. 1, 2005, pp. 50 – 69.

[5] J. M. Pitts and J. A. Schormans, Introduction to IP and ATM Design and Performance. Chichester, England: John Wiley & Sons, 2000.

[6] S. Shin and H. Schulzrinne, "Measurement and analysis of the VoIP capacity in IEEE 802.11 WLAN," IEEE Transactions on Mobile Computing, vol. 8, no. 9, pp. 1265-1279, Sept. 2009.

[7] A. Chan and S. C. Liew, "Performance of VoIP over multiple co-located IEEE 802.11 wireless LAN," IEEE Transactions on Mobile Computing, vol. 8, no. 8, pp. 1063-1076, Aug., 2009.

[8] S. Ramprashad and C. Pepin, "A study of silence suppression and real speech patterns amd their impact on VoIP capacity in 802.11 networks," Proc. IEEE Int. Conf. on Multimedia and Expo, 2007, pp. 939-942.

[9] E. Hossain, Y. Xiao, Li-Chun Wang and K. K. Leung, "Radio resource Management and protocol engineering for IEEE 802.16," (special issue), IEEE Wireless Communications, vol. 14, pp. 2-51, Feb. 2007.

[10] S. Shrivastava and R. Vannithamby, "Group scheduling for improving VoIP capacity in IEEE 802.16e networks," IEEE 69st Vehicular Technology Conf., VTC Spring, 2009, pp. 1-5.

[11] J. Puttonen et al, "Voice-over-IP performance in UTRA Long Term Evolution downlink," IEEE Vehicular Technology Conf 2008, VTC Spring 2008, pp. 2502-2506.

[12] J. Zhu, X. She and L. Chen, "Complementary recourse allocation for variable-size VoIP packet in E-UTRA," IEEE Wireless Communications and Networking Conf. (WCNC 2009), 2009, pp.1-5.

[13] H. Wang, J. Han and S. Xu, "Performance of TTI bundling for VoIP in EUTRAN TDD mode," IEEE 69st Vehicular Technology Conf., VTC Spring, 2009, pp. 1-5.

[14] A. Baiocchi, N. Mellazzi and A. Roveri, "Queuing performance and control in ATM," 13th International Teletraffic Congress, Copenhagen, 1991, pp. 13-18.

[15] R. C. F. Tucker, "Accurate method for analysis of a packet speech multiplexer," Electronics Letters, vol. 19, no. 14, pp. 536-537.

[16] A. Estepa and R. Estepa, "Accurate resource estimation for homogeneous VoIP aggregated traffic", Computer Networks, vol. 52, no. 13, Sep. 2008, pp. 2505-2517.

[17] W. Wang, S. C. Liew and V. O. K. Li, "Solutions to performance problems in VoIP over a 802.11 wireless LAN," IEEE Trans. Vehicular Techn., vol. 54, pp. 366-384, Jan. 2005.

[18] A. Benyassine et al, "ITU-T Recommendation G.729 Annex B: A silence compression scheme for V.70 digital simultaneous voice and data applications," IEEE Communications Magazine, vol. 35, no. 9, Sep. 1997, pp. 64-73.

[19] B. Goode, "Voice over Internet protockol (VoIP)," Proceedings of IEEE, vol. 90, no. 9, pp. 1495-1517, Sep. 2002.

[20] R. G. Cole and J. Rosenbluth, "Voice over IP performance monitoring," ACM Comput. Commun., vol. 31, no. 2, pp. 9-24, April 2001.

[21] A. P. Markopoulou, F. A. Tobagi, and M. Karam, "Assessing the quality of voice communications over Internet backbone," IEEE/ACM Transactions on Networking, vol. 11, no.5, pp. 747-760,. Oct. 2003.

[22] The E-Model, A Computational Model for Use in Ttransmission Planning, ITU-T Rec. G.107, Int. Telecommun. Union, July 2000.

[23] N. T. Moura *et al.*, "MOS-Based Rate Adaptation for VoIP Sources," Proc. IEEE international Conference on Communications (ICC2007), June.2007, pp. 628-633.

[24] J. Matta, C. Pepin, K. Lashkari, and R. Jain, "A Source and Channel Rate Adaptation Algorithm for AMR in VoIP Using the Emodel," Proc. 13th International Workshop on Network and Operating Systems Support for Digital Audio and Video, 2003, pp. 92-99.

[25] M. T. Gardner, V. S. Frost, and D. W. Petr, "Using Optimization to Achieve Efficient Quality of Service in Voice over IP Networks," Proc. IEEE Performance, Computing, and Communications Conference, Apr.2003, pp. 475-480.

[26] K. Fujimoto, S. Ata and M. Murata, "Adaptive playout buffer algorithm for enhancing perceived quality of streaming applications," in Proc. IEEE Globecom 2002, Nov. 2002.

[27] L. Sun and E. Ifeachor, "Prediction of perceived conversational speech quality and effects of playout buffer algorithms," in Proc. IEEE Int. Conf. Communications ICC'03, Anchorage, AK, May 2003, pp. 1-6.

[28] L. Sun and E. Ifeachor, "Voice quality prediction models and their application in VoIP networks," IEEE Transactions on Multimedia, vol. 8, no. 4, pp. 809-820, Aug. 2006.

[29] C. Cicconetti, A. Erta, L. Lenzini and E. Mingozzi, "Performance evaluation of the IEEE 802.16 MAC for QoS support," IEEE Transactions on Mobile Computing, vol. 6, pp. 26-38, Jan. 2007.

[30] D. Zhao and X. Shen, "Performance of packet voice transmission using IEEE 802.16 protocol," IEEE Wireless Communications, vol. 13, pp. 44-51, Feb. 2007.

[31] S. E. Elayoubi and B. Fourestie, "Performance evaluation of admission control and adaptive modulation in OFDMA WiMAX systems," IEEE Transaction on Networking, vol. 16, no.5, pp. 1200-1211, Oct. 2008.

[32] ITU-T "One Way Transmission Time," Recommendation G.114, May 2000.

[33] R. J. Gibbens, F. P. Kelly and P. B. Key, "A decision-theoretic approach to call admission control in ATM networks," ," IEEE J. on Selected Areas in Communications, vol. 13, no. 6, pp. 1101-1114,. Aug. 1995.

[34] M. Reisslein, K. W. Ross and Rajagopal, "Guaranteeng statistical QoS to regulated traffic: the single node case," Proc. 18th, IEEE INFOCOM, 1999, pp. 1061-1072.

[35] G. Mao and D. Habibi, "Loss performance analysis for heterogeneous ON-OFF sources with application to connection admission control," IEEE/ACM Transactions on Networking, vol. 10, no. 1, pp. 125-138,. Feb. 2002.

[36] A. Estepa and R. Estepa, "Accurete VoIP dimensioning for WAN links," Electronics Letters, vol. 43, no. 23. Nov. 2007.

[37] A. Estepa, R. Estepa and J. Vozmediano, "A new approach for VoIP traffic characterization," IEEE Communications Letters, vol. 8, no. 10, pp. 644-646, Oct. 2004.

[38] A. Estepa, R. Estepa, I. Campos, and A. Delgado, "Dimensioning aggregated voice traffic in MPLS nodes," Proc. of ONDM 2008, IFIP 2008, pp. 211-215.

[39] A. Mukherjee, "On the dynamics and significance of low frequency components of Internet load," Technical Report – University of Pennsylvania, 1992. (available at http://repository.upenn.edu/cis_reports/300). Last access: 01.2011.

[40] L. Sun, "Speech quality prediction for voice over Internet protocol networks," Ph.D. dissertation, University of Plymouth, U.K., Jan. 2004. (available at http://www.tech.plym.ac.uk/spmc/people/lfsun/publications/LSunPhDthesis.pdf). Last access: 01.2011.

[41] V. Frost, B. Melamed, "Traffic modeling for telecommunications networks," IEEE Communications Magazine, vol. 32, no. 3, Mar. 1994, pp. 70-81.

[42] A. Adas, "Traffic models in broadband networks," IEEE Communications Magazine, vol. 35, no. 7, Jul. 1997, pp. 82-89.