Analytical Description of a Parameter-based Optimization of the Quality of Service for VoIP Communications

LEOPOLDO ESTRADA[1], DENI TORRES[1], HOMERO TORAL[1,2]
[1]Department of Electrical Engineering
CINVESTAV IPN
GUADALAJARA, JALISCO, MEXICO
[2]Department of Postgraduate
ITSCH
LAS CHOAPAS, VERACRUZ, MEXICO
lestrada@gdl.cinvestav.mx, www.gdl.cinvestav.mx

Abstract: - In this work, the perceived quality of VoIP communications is studied. The distributions of the number of consecutive received and lost packets, respectively named gap and burst, of a VoIP communication are modeled with discrete two-state and four-state Markov chains. Algorithms for estimating the transition probabilities between states and from these, the packet loss rate and the respective gap and burst length distributions, are described. Through a study of monitored VoIP calls, it is shown that these models can adequately represent the geometric-type decay of these distributions and that although two-state model performs well for homogeneous losses, for non-homogeneous losses the four-state model fits better. An analysis of the performance of a packet-level FEC scheme, based on N-packet redundancy, is presented. The perceived packet loss rate that results of applying this correction scheme is quantified. For the studied measurements, 1-packet redundancy is sufficient to decrease the perceived loss rate below 1%. Also, the impairments of the perceived quality of voice after the FEC technique and a de-jitter buffer is quantified. The resulting equations can be used to optimize the adjust parameters of the VoIP call, e.g., level of redundancy, type of codec used and de-jitter buffer size. The proposed methodology can be extended if other types of improvements are included.

Key-Words: - VoIP, Packet loss distribution, Packet reception distribution, N-packet FEC, QoS.

1 Introduction

Internet became the point of convergence of information and media transmission. Data, voice, video, etc., are transmitted through the same communication channel. The service provided by the Internet is named "best effort", which means that the devices between links generally do not differentiate between the types of traffic and there is neither resource reservation nor prioritization. The exceptions are those networks where special services are provided [1] [2]. Congestion due to the high demand of network resources is a cause of the impairment of its quality of service, which consists of delay problems, i.e., the delay and its variation (delay *jitter*) are higher, and packet loss. For time critical applications, like VoIP, end-to-end delay can have high impact on quality of service [3]. The a*utomatic repeat request* (ARQ) technique, the correction scheme of the *transmission control protocol* (TCP), is used to eliminate (or reduce) packet losses, but it is not suitable for many real-time and near real-time applications, which have tighter delay tolerance. Then, other types of error correction techniques, adequate for these applications, are needed, e.g. *multiple packet transmission* (MPT) or *forward error correction* (FEC), to assure certain quality of service.

In this work, modeling of packet loss of a VoIP communication through a wide area network (WAN) is developed. Discrete finite-state Markov chains are used to represent how these losses occur in the communication channel, which consists of a sequence of routers connected by links through which the packets traverse.

Consecutive packet receptions and losses are named *gaps* and *bursts*, respectively. Due to the time-correlated occupancy of the network, packet losses commonly occur in bursts such that their lengths follow a geometric-type distribution, as well as gaps [4] [5].

At small time scales, i.e. a few seconds or minutes, a two-state Markov chain can reproduce this phenomenon, but a non-homogeneous behavior becomes noticeable at larger scales and, in this case, the two-state Markov chain is insufficient, thus a more general model is necessary. The four-state Markov chain seems to capture or simulate better this widely known non-homogeneous behavior of the characteristics of network traffic. The four-state model approach allows us to represent and simulate those periods with low and high *packet loss rate* (PLR) that alternate in sequence according to certain probability.

MPT consists of sending copies of packets when high losses occur. In order to maximize the probability of reception, these copies must be equally spaced in the time [6]. Although this technique has the advantage that it is very easy to implement, unless a low bit rate coded is used, it has the disadvantage of high bandwidth requirement consumption.

The N-packet FEC technique consists of sending information about packet n along with later packets, i.e., with packets $n+1$, $n+2$, …, $n+N$, in order to reconstruct packet n in the case it is lost. With this correction scheme, the last N packets of a burst can be recovered and then, the perceived PLR of the end user is lower than the real PLR due to the network. Generally, the amount of redundancy is defined as a function of the PLR [7], e.g., it is not efficient to send redundant information if there are no missing packets. This correction technique, which is performed at packet level, is the scope of the analysis presented in section 5.1. The FEC technique also reduces the burstiness of packet loss, which affects the quality of a VoIP communication [8].

Many codification schemes of the redundant information on later packets have been proposed in the literature [9]. The authors of [10] developed an algorithm to estimate the optimal value of w for wireless communications. Also, they remark that a signaling protocol is needed for a VoIP call in order to ascertain media encoding and packetization parameters end-to-end.

In this work, an estimation of the quality of the VoIP communication is presented. It consists of an estimation of the E-model's R factor, which considers a codification scheme that codes voice packets using G.711 or G.729, a packet level FEC technique that codes redundant information with G.729 and a resizable de-jitter buffer. An advantage of this model is that the equations for the codec-loss impairments can be easily obtained, as combining the separately defined impairments of codecs G.711 and G.729, according to the E-model's R factor. The packet loss burstiness is considered by modeling network losses with finite-state Markov chains, which allow us to predict the perceived PLR when applying the N- packet FEC technique. The proposed model can be used for the estimation of the optimal adjust parameters, e.g., de-jitter buffer size, the level of redundancy and the voice data length that maximize the quality of the communication.

2 Contributions

The contributions of this work are summarized as follows:

1. A statistical description of the two-state and four-state Markov chains, assuming that it is time-homogeneous (i.e., the probabilities of transition between states are constant) is presented.
2. An analytical description of the performance of n-packet FEC scheme is given, i.e., the perceived PLR as a function of the network loss rate, the burst length distribution and the level of redundancy.
3. An estimation of the impairments due to low bit rate codec and packet loss is proposed for combined codification, i.e., when normal packets are redundant information with a different type codec.
4. A methodology to estimate the adjust parameters (codec type, level of redundancy and de-jitter buffer size), based on the estimation of the E-model's R factor, is presented.
5. A set of measurements, which consists of monitored VoIP calls from which loss sequences are obtained, is studied in order to verify the proposed models.

3 Finite-state Markov Chains

3.1 Matrix Representation of the Steady-state

Let $S = S_1, S_2, \ldots, S_m$ be the m states of an m-state Markov chain and let p_{ij} be the probability of the chain to pass from the state S_i to the state S_j, i.e., $p_{ij} = P(X_i = x_i | X_{i-1} = x_{i-1})$. Having the Markov property means that, given the present state, future states are independent of the past states, i.e., $P(X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, \ldots) = P(X_{n+1} = x_{n+1} | X_n = x_n)$. The Markov chains used in this work also are time-homogeneous, which means that the probabilities of transition between states are constant over time, i.e., $P(X_{n+1} = x_{n+1} | X_n = x_n) = P(X_n = x_n | X_{n-1} = x_{n-1})$.

All states communicate (are reachable from) each other, which makes the chain irreducible. Also, the chain is aperiodic, i.e., state S_i can be reached from itself in any number of steps ($n = 1,2,3, \ldots$).

The probabilities of transitions between states can be represented by a *transition matrix*. The elements of the one-step $m \times m$ transition matrix $\underline{T}$ are $T_{ij} = p_{ij}$. To obtain the n-step transition matrix it is necessary to multiply the matrix itself n times [11], i.e.,

$$\underline{T_n} = \underline{T}^n \quad (1)$$

As the number of steps (n) increases, the probability of the matrix to be in the state S_i from an initial state depends less on this one. i.e., as n tends to ∞, the matrix $\underline{T_n}$ converges to a matrix with the next form:

$$\underline{T_\infty} = \lim_{n\to\infty} \underline{T_n} = \begin{bmatrix} s_1 & s_2 & \cdots & s_m \\ s_1 & s_2 & \cdots & s_m \\ \vdots & \vdots & \ddots & \vdots \\ s_1 & s_2 & \cdots & s_m \end{bmatrix} \quad (2)$$

such that

$$s_1 + s_2 + \cdots + s_m = 1 \quad (3)$$

In (2) and (3), s_i represents the named *steady* probability of state S_i. The steady-state transition matrix $\underline{T_\infty}$ can be obtained then by solving (3) and (4) [12]:

$$\overline{S}\underline{T} = \overline{S} \quad (4)$$

where $\overline{S} = [s_1 \quad s_2 \quad \dots \quad s_m]$.

Assuming that the chain is irreducible and aperiodic, the matrix $\underline{T_\infty}$ is well defined and unique.

3.1.1 Numerical Approximation

Obtaining analytical expressions for the elements of $\underline{T_\infty}$ (i.e., s_1, s_2…) can be difficult when the number of states is large. In this case, a numerical approximation is more suitable, which is described as follows:

Let $\underline{T}$ be a $m \times m$ transition matrix, which has a unique steady-state solution, and let $\{(\lambda_i, \bar{v}_i); i = 1, \dots, m\}$ be its pairs of eigenvalues and eigenvectors (i.e., $\underline{T}\bar{v}_i = \lambda_i \bar{v}_i$), such that $\lambda_i > \lambda_j$ for $i < j$. This matrix $\underline{T}$ can be decomposed into the special form

$$\underline{T} = \underline{P}\underline{D}\underline{P}^{-1} \quad (5)$$

where $\underline{P}$ is a matrix composed of the eigenvectors of $\underline{T}$, $\underline{D}$ is the diagonal matrix constructed from the corresponding eigenvalues and $\underline{P}^{-1}$ is the inverse of $\underline{P}$. Then $\underline{T_n}$ can be calculated easily as

$$\underline{T_n} = \underline{P}\underline{D}^n\underline{P}^{-1} \quad (6)$$

As all elements of the diagonal of the matrix $\underline{D}$ are lower than 1 except $D_{1,1}$, then

$$\underline{T_\infty} = \underline{P}\underline{D}^\infty\underline{P}^{-1} = \underline{P}\underline{D}'\underline{P}^{-1} \quad (7)$$

where the only non-zero element of $\underline{D}'$ is $D_{1,1} = 1$.

This method is also useful when obtaining short-term approximations, i.e., $\underline{T_n}$ for small n.

3.2 Two-state Markov Chain

The two-state Markov chain is shown in Fig. 1. State S_1 represents packet loss and S_2, packet reception. Two substitutions ($p_{11} = 1 - p_{12}$ and $p_{22} = 1 - p_{21}$) are made in order to represent the chain with the lowest number of parameters. The steady-state probability of the chain to be in the state S_1, namely the PLR, is given by (8) [7]:

$$s_1 = \frac{p_{21}}{p_{12} + p_{21}} \quad (8)$$

and clearly $s_2 = 1 - s_1$.

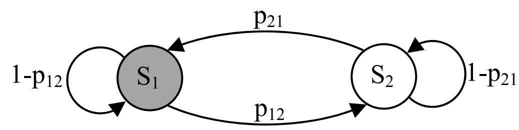

Fig. 1: Two-state Markov chain. White and shady circles represent correct and erroneous states, respectively.

The burst and gap length distributions ($f_b(k)$ and $f_g(k)$, respectively) can be expressed in terms of p_{12} and p_{21}, as expressed by (9) and (10):

$$f_b(k) = p_{12}(1 - p_{12})^{k-1} \quad (9)$$

$$f_g(k) = p_{21}(1 - p_{21})^{k-1} \quad (10)$$

which have also respective means $E\{f_b(k)\} = 1/p_{12}$ and $E\{f_g(k)\} = 1/p_{21}$. It is easy to proof (9), as $\sum_{k=1}^{\infty} f_b(k) = 1$ and $f_b(k+1) = f_b(k) \cdot (1 - p_{21})$; and similarly for (10).

3.3 Four-state Markov Chain

The four-state Markov chain is shown in Fig. 2. Missing arrows indicate zero probability. States S_1 and S_3 (shady circles) represent packet losses (erroneous); S_2 and S_4 (white circles), packet reception (correct).

Six parameters ($p_{21}, p_{12}, p_{43}, p_{34}, p_{23}, p_{32} \in (0,1)$) are necessary to define all the transition probabilities. Without loss of generality, probabilities of transitions between correct states, as well as transitions between erroneous ones, have been set to zero.

The four steady-state probabilities of this chain are:

$$s_1 = \frac{1}{1 + \frac{p_{12}}{p_{21}} + \frac{p_{12}p_{23}}{p_{21}p_{32}} + \frac{p_{12}p_{23}p_{34}}{p_{21}p_{32}p_{43}}} \quad (11)$$

$$s_2 = \frac{1}{1 + \frac{p_{21}}{p_{12}} + \frac{p_{23}}{p_{32}} + \frac{p_{23}p_{34}}{p_{32}p_{43}}} \quad (12)$$

$$s_3 = \frac{1}{1 + \frac{p_{34}}{p_{43}} + \frac{p_{32}}{p_{23}} + \frac{p_{21}p_{32}}{p_{12}p_{23}}} \quad (13)$$

$$s_4 = \frac{1}{1 + \frac{p_{43}}{p_{34}} + \frac{p_{32}p_{43}}{p_{23}p_{34}} + \frac{p_{21}p_{32}p_{43}}{p_{12}p_{23}p_{34}}} \quad (14)$$

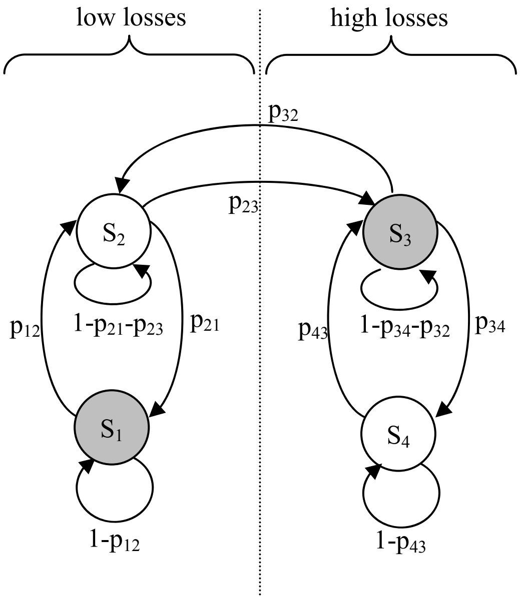

Fig. 2: Four-state Markov chain. Only two types of transitions between different states are allowed: from correct to erroneous and from erroneous to correct.

The probability of the chain to be either in S_1 or in S_3, that corresponds to PLR, is then:

$$r = s_1 + s_3 \quad (15)$$

The average burst length ($\bar{b}$) is calculated as the quotient of the probability of loss and the probability of transition from a lossless state to a loss state (16), that is:

$$\bar{b} = \frac{s_1 + s_3}{s_2(p_{21} + p_{23}) + s_4(p_{43})} \quad (16)$$

Similarly, the average gap length is:

$$\bar{g} = \frac{s_2 + s_4}{s_1(p_{12}) + s_3(p_{34} + p_{32})} \quad (17)$$

Note that the transitions from error state to correct state and vice versa have equal probability, i.e. $s_2(p_{21} + p_{23}) + s_4(p_{43}) = s_1(p_{12}) + s_3(p_{34} + p_{32})$.

The distribution of the burst length can be derived the following the next procedure:

Let $f_b(k)$ denote the probability that the burst length is k; $C_1(k)$, the probability that the burst length is k or greater and the k^{th} transmission is from state S_1 and $C_3(b)$, the probability that the burst length is k or greater and k^{th} transmission is from state S_1 and $C_b(k)$, the probability that the burst length is k or greater such that $C_b(k) = C_1(k) + C_3(k)$ and $f_b(k) = C_b(k) - C_b(k+1)$. Clearly $C_b(k) = \sum_{i=k}^{\infty} f_b(i)$. Also, as transitions between states S_1 and S_3 have zero probability, $C_1(k+1) = C_1(k)(1-p_{12}) = C_1(1)(1-p_{12})^k$ and $C_3(k+1) = C_3(k)(1-p_{34}-p_{32}) = C_3(1)(1-p_{34}-p_{32})^k$. Then to calculate $f_b(k)$ it is necessary to obtain $C_1(1)$ and $C_3(1)$, whose respective values are $C_1(1) = s_2p_{21}/[s_2(p_{21}+p_{23}) + s_4p_{43}]$ and $C_3(1) = (s_2p_{23} + s_4p_{43})/[s_2(p_{21}+p_{23}) + s_4p_{43}]$.

As the minimum burst length is 1, $C_b(1) = C_1(1) + C_3(1) = 1$. Then, the distribution of the burst length is:

$$f_b(k) = C_1(1)Q_1(k) + C_3(1)Q_3(k) \quad (18)$$

where $Q_1(k) = (1-p_{12})^{k-1} - (1-p_{12})^k = p_{12}(1-p_{12})^{k-1}$ and $Q_3(k) = (1-p_{34}-p_{32})^{k-1} - (1-p_{34}-p_{32})^k = (p_{34}+p_{32})(1-p_{34}-p_{32})^{k-1}$. As expressed by (18), $f_b(k)$ is the sum of two geometric series with respective rates $1-p_{12}$ and $1-p_{34}-p_{32}$; this implies that $f_b(k)$ is a decreasing function of k, i.e., bursts of greater length have lower probabilities than shorter ones.

A similar procedure can be followed to obtain the gap length distribution ($f_g(k)$), which is:

$$f_g(k) = C_2(1)Q_2(k) + C_4(1)Q_4(k) \quad (19)$$

where $C_2(1) = (s_1p_{12} + s_3p_{32})/[s_1p_{12} + s_3(p_{32}+p_{34})]$, $C_4(1) = (s_3p_{34})/[s_1p_{12} + s_3(p_{32}+p_{34})]$, $Q_2(k) = (1-p_{21}-p_{23})^{k-1} - (1-p_{21}-p_{23})^k = (p_{21}+p_{23})(1-p_{21}-p_{23})^{k-1}$ and $Q_4(k) = (1-p_{43})^{k-1} - (1-p_{43})^k = p_{43}(1-p_{43})^{k-1}$. Also note that $C_2(1) + C_4(1) = 1$.

Note that, although the resulting equations correspond to the four-state model of Fig. 2, this procedure can be applied for any finite-state Markov chain, which consists of finding firstly the cumulative density functions (CDF), i.e., $C_b(k)$ and $C_g(k)$.

4 Modeling from a Loss Sequence

Let us define the loss sequence as follows:

$$Y_k = \begin{cases} 0; & \text{if packet } k \text{ is received} \\ 1; & \text{if packet } k \text{ is lost} \end{cases} \quad (20)$$

From the loss sequence, the probabilities of transitions were also estimated using the algorithms explained in sections 4.1 and 4.2, for the two-state and the four-state models, respectively.

4.1 Two-state Parameters Estimation

The estimations of p_{12} and p_{21} are: $p_{12} = t_{c\to e}/n_1$ and $p_{21} = t_{e\to c}/n_0$, where $t_{c\to e}$ and $t_{e\to c}$ are the respective number of transitions from correct states to error states (i.e., when $Y_k = 0$ and $Y_{k+1} = 1$) and from error states to correct states (i.e., when $Y_k = 1$ and $Y_{k+1} = 0$), and n_0 and n_1 are the respective number of received and lost packets (i.e., the respective numbers of zeros and ones of Y_k).

4.2 Four-state Parameters Estimation

In this case the values of the sequence Y_t are divided into regions of two types: the first with lower loss rate (whose first and last values are zeros) and the second with higher loss rate (whose first and last values are ones) than certain threshold, e.g. 1%. Then, from the first region, p_{12} and p_{21} are estimated as explained in section 4.1. Similarly, p_{43} and p_{34} are estimated from the second region. Finally, let $t_{1st\to 2nd}$ be the number of transitions from the first region to the second; $t_{2nd\to 1st}$, the number of transitions from the second to the first; n_{1st}, the number of received packets in the first region (zeros) and n_{2nd}, the number of lost packets in the second region (ones), then $p_{23} = t_{1st\to 2nd}/n_{1st}$ and $p_{32} = t_{2nd\to 1st}/n_{2nd}$.

5 Performance Metrics

The International Telecommunication Union (ITU) defined the E-model in the ITU-T Recommendation G.107 [13], as a tool for quality measurement for planning purposes. The E-model provides a prediction of the expected quality as perceived by the end user. This model is based on impairment factors, as expressed by (E-model's R factor):

$$R = R_0 - I_s - I_d - I_e + A \quad (21)$$

where R_0 is the signal-to-noise ratio, I_s represents all impairments which occur more or less simultaneously with the voice signal, I_d sums all delay impairments due to delay and echo effects, I_e represents the impairments that are caused by low-bit rate codecs and A represents and advantage factor which certain systems provide in comparison to conventional systems.

A simplified version of (21), that represents the impairment as a function of the packet delay and PLR, is:

$$R = 93.2 - I_d - I_e \quad (22)$$

where I_d, the delay impairment, is defined as [14]:

$$I_d = 0.024d + 0.11(d - 177.3)H(d) \quad (23)$$

where d is the mouth-to-ear delay in ms and

$$H(d) = \begin{cases} 0; & d < 177.3 \\ 1; & d \geq 177.3 \end{cases} \quad (24)$$

The quantity I_e is the impairment caused by low bit rate codecs [13], and its general expression is:

$$I_e(r) = \gamma_1 + \gamma_2 \ln(1 + \gamma_3 r) \quad (25)$$

where r is the packet loss probability and the values of γ_1, γ_2 and γ_3 are constants that depend on the type of codec used [15], e.g.,

$$I_{e,G.711}(r) = 0 + 30\ln(1 + 15r) \quad (26)$$

and

$$I_{e,G.729}(r) = 11 + 40\ln(1 + 10r) \quad (27)$$

The mean opinion score (MOS), a numerical indication of the perceived quality of the received media after compression and/or transmission, can be estimated from the R factor as: $MOS = 1$ for $R < 0$, $MOS = 1 + 0.035R + R(R - 60)(100 - R)7 \cdot 10^{-6}$ for $0 \leq R < 100$ and $MOS = 4.5$ for $R > 100$ [13]. In order to achieve the best perceived quality as possible, the R factor and, as a consequence, the MOS must be maximized.

In order to do this, the impairments factors I_d and I_e must be adjusted. I_d depends on many parameters, many of them cannot be controlled by the end user directly (e.g., constant and variable network delays) but others can, e.g. the size (in ms) of the de-jitter buffer. In its turn, I_e depends on the codec used and the packet loss rate. It can be adjusted by selecting the type of codec used and by applying an error correction technique, e.g. N-packet FEC described in section 5.1.

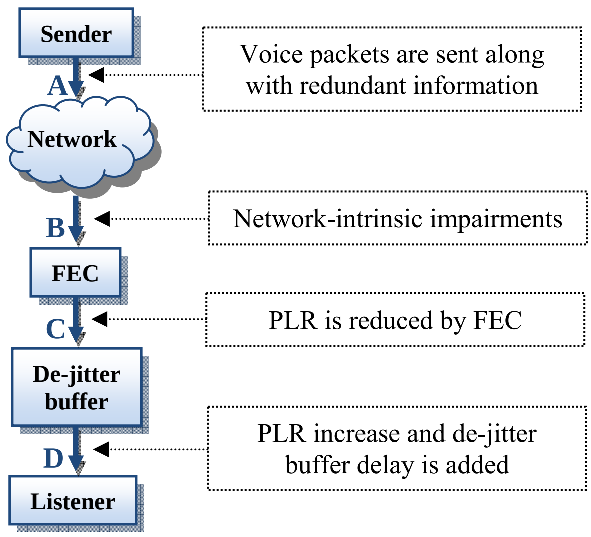

Fig. 3: VoIP communication. The main sources of impairments are indicated.

Fig. 3 represents the path followed by the packets from sender to receiver. Coders and decoders are implicit in sender and listener, respectively. There are two main sources of packet losses: network congestion (from A to B) and de-jitter buffer (from C to D). The PLR due to network congestion is reduced by the N-packet FEC. The network also adds variable delays, i.e., delay jitter, that are eliminated by the de-jitter buffer, but at the expense of an additional delay for all (or most) packets and an increase in packet loss: since long delayed packets, although they successfully arrive to the receiver, are discarded and, consequently, lost from the point of view of the listener. Furthermore, bit-level errors that may be present in received but corrupted packets are an important source of errors, especially for wireless communications.

In order to quantify the voice quality by means of the E-model's R factor, the performance of the N-packet FEC block and the de-jitter buffer must be described analytically.

5.1 N-packet FEC Performance

N-packet FEC consists of that packet $n+1$ contains information about packet n, so that if packet n is lost, it can be approximately reconstructed from the associated information. Packet n cannot be reconstructed if there is no redundant information, i.e. when packet $n+1$ is also lost. The 1-packet FEC technique performance can be described as: it reduces the size of a burst of length k to $k-1$. The perceived PLR ($r_1{}'$) is proportional to the perceived average burst length, which in this case decreases by 1 (packet), then it is equal to:

$$r_1{}' = \frac{(\bar{b}-1)r}{\bar{b}} \tag{28}$$

where $\bar{b}$, the average burst length, is $\bar{b} = \sum_{k=1}^{\infty} k f_b(k)$ and $f_b(k)$ is the burst length distribution.

If the redundancy level extends to N packets, i.e. packet n has information about $n+1$, $n+2$, …, $n+N$ packets, the length of all bursts decreases from k to $\max(0, k-N)$ packets, then the new burst length distribution $f_b^{'}(k)$ is:

$$f_b^{'}(k) = \begin{cases} \sum_{i=1}^{N} f_b(i); & k = 0 \\ f_b(k+N); & k > 0 \end{cases} \tag{29}$$

Note that (29) considers bursts of zero length. The interpretation of this is as follows: bursts do really occur in the network but, as they are corrected by a N-packet FEC technique, they are diminished (when $k > N$) or eliminated (when $k \leq N$) in the receiver. Then, $f_b'(k)$ is the new burst length distribution and its mean can be calculated as:

$$\bar{b}' = \sum_{k=0}^{\infty} [k f_b{}'(k)] \tag{30}$$

$$\bar{b}' = \bar{b} - N + \sum_{k=1}^{N} (N-k) f_b(k) \tag{31}$$

Consequently, the perceived PLR is:

$$r_N' = \frac{\left[\bar{b} - N + \sum_{k=1}^{N-1}(N-k) f_b(k)\right] r}{\bar{b}} \tag{32}$$

which is a generalized form of (28).

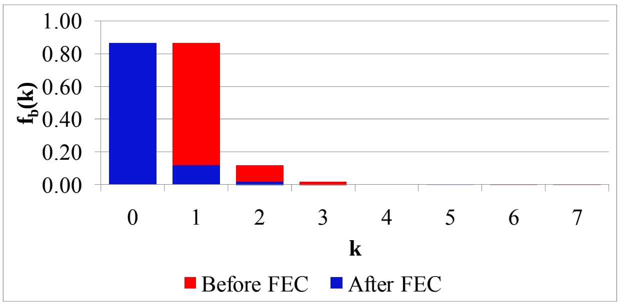

Fig. 4: Example of the burst length distribution before and after 1-packet FEC. PLR is reduced from 4% to 0.55%, approximately.

Fig. 4 shows an example of the burst length distribution, as expressed by (18) with $p_{21} = 0.001350$, $p_{12} = 1.000000$, $p_{43} = 0.054507$, $p_{34} = 0.845146$, $p_{23} = 0.001968$ and $p_{32} = 0.016989$, and how it is modified, from the point of view of the receiver, after applying 1-packet FEC. In this case, all burst are reduced in 1 packet, as is defined by (29) and, as a consequence, the perceived PLR is reduced from 4.022% to 0.548%. Similarly, Fig. 5 shows the comparison of the burst length distribution before and after 2-packet FEC. In this case the PLR is reduced to 0.075%.

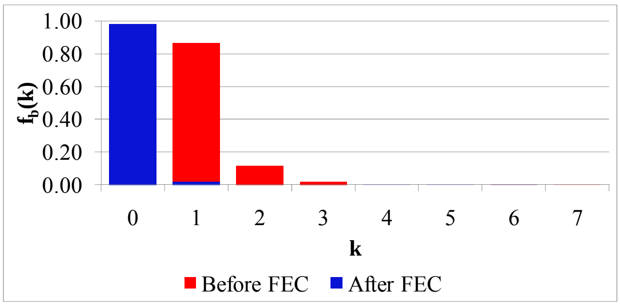

Fig. 5: Example of the burst length distribution before and after 2-packet FEC. PLR is reduced from 4% to 0.075%, approximately.

Note that (32) expresses the perceived PLR of the receiver without considering other sources of losses, e.g., additional perceived losses occur if packets are delayed more than certain threshold (i.e., de-jitter buffer size).

5.2 Estimation of the Codec Impairments after FEC reconstruction

As explained in section 5.1, the N-packet FEC technique sends additional information of the immediate previous packets within the current sending packet. The amount of redundant information (i.e., in bytes) of one packet must be less (or at most equal) than N times the packet size without redundancy. Let us estimate the codec impairment (I_e) for these two cases:

1. Normal packets coded using G.711 and redundant information coded using G.729.
2. Normal packet and redundant information both coded using G.729.

For the first case, if the PLR due to the network is r and the perceived PLR after the FEC block, then the impairment after this block (i.e., at point C in Fig. 1) depends on the original loss rate (r) and the percent of reconstructed packets ($r - r'$). Note that this function is neither equal to (26) nor to (27) and must satisfy the following conditions:

i. If $r' = r$ (i.e., no packets reconstructed):

$$I_{e,G.711,G.729}(r, r) = I_{e,G.711}(r) \quad (33)$$

ii. If $0 < r' < r$ (i.e., some packets are reconstructed):

$$I_{e,G.711,G.729}(r, r') > I_{e,G.711}(r') \quad (34)$$

iii. If $r' = 0$ (i.e., all packets are reconstructed):

$$I_{e,G.711,G.729}(r, 0) > I_{e,G.711}(0) \quad (35)$$

iv. For a fixed $r > 0$, it must be a non-decreasing function of r'.

According to (34), for a fixed value of r', the impairment increases as the percent of reconstructed packets is greater. Based on these properties, we define the estimation of the codec impairment as:

$$I_{e,G.711,G.729}(r, r') = I_{e,G.711}(r') + (r - r')\Delta_e \quad (36)$$

where Δ_e is:

$$\Delta_e = I_{e,G.729}(r') - I_{e,G.711}(r') \quad (37)$$

This estimation (36) satisfies the properties described previously.

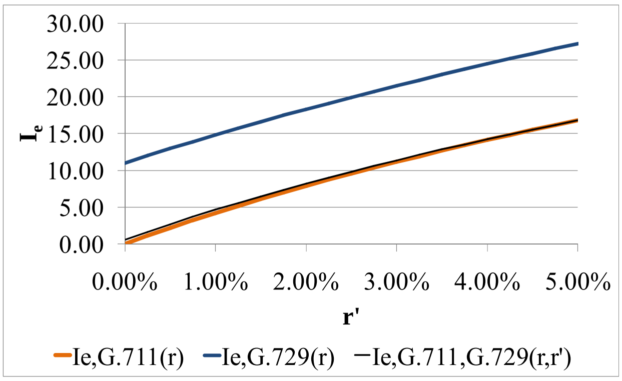

Fig. 6: Estimation of the codec impairments (I_e) for G.711, G.729 and combined G.711-G.729.

Fig. 6 shows the proposed model (36), compared with the estimations for codecs G.711 and G.729. Note that the percent of reconstructed packets ($r - r'$) depends on the level of redundancy (N) and the burst distribution ($f_b(k)$).

For the second codification scheme (i.e., using G.729 for normal packets and redundant information), the estimation of I_e is the same as (27), substituting r with r':

$$I_{e,G.729,G.729}(r, r') = I_{e,G.729}(r') \quad (38)$$

Note that this estimation is not including those additional losses caused by de-jitter buffer, which discard long delayed packets. To estimate the impairment after the de-jitter buffer block, the values of r and r' must be updated in the case of that any packet is discarded.

5.3 Impact of the De-jitter Buffer Size on the Perceived Packet Loss Rate

Let $F_D(t)$ be the packet delay (OWD) distribution and w, the maximum waiting time for each packet (i.e., the maximum delay for a packet that is not discarded by the receiver). Then, the probability for a single packet to be discarded (p_d) is:

$$p_d = 1 - F_D(w) \quad (39)$$

Note that equation (39) considers the packet delay process as stationary, otherwise it would be time dependent.

Let r'' be the perceiver PLR after FEC and the receiver buffer, in this order. It is the sum of two probabilities: the probability of a packet to be lost due to network congestion and the probability of the packet to arrive at the receiver and be discarded by the receiver buffer, then an approach to r'' is:

$$r'' = r' + (1 - r')[1 - F_D(w)] \quad (40)$$

where r' is the same than r'_N and defined by (32).

As the theoretical waiting time that minimizes the perceived PLR is ∞ (which implies that the packets would never be sent to the listener), a more adequate estimation of the optimal w must consider the expected voice quality of the communication, e.g., by means of the E-Model's R-Factor.

5.4 Parameter-optimizable Quality of the VoIP Communication

The maximum value for the MOS, which indicates the maximum quality of the communication, is achieved when the R factor is also maximized. The strategy is then to set the adjust parameters to their respective optimal values, e.g., the redundancy level (N), the de-jitter buffer size (w), the type of codec (in this case, G.711 or G.729) and the voice data length (inter-departure time or IDT). Many of these adjust parameters are easier to optimize than the others, as they are independent (or almost) of the others. E.g., the level of redundancy depends only on the PLR, it is increased as the PLR become greater than certain thresholds to decrease it to acceptable levels.

The E-model's R factor can be estimated then in terms of these adjust parameters as follows:

As it is defined by (22), it is the sum of the impairments due to coded-PLR (I_e) and those due to mouth-to-ear delay (I_d). The first one depends on the types of codec used, as described in section 5.2, and the type and level of redundancy, as the perceived PLR decreases by packet reconstruction. The second one depends of the mouth-to-ear delay, which can be expressed as the sum of two delays, i.e.,

$$d = w + \delta \tag{41}$$

where w is the delay caused by the de-jitter buffer (which is equal to its size) and δ is the sum of all other delays, e.g., packet transmission, queuing, coding/decoding, etc. The delay impairment is then estimated as expressed by (23) for $d = w + \delta$. Obviously the value of d is random, but its distribution can be known, as it depends on the distribution of δ and the value of w.

Finally, the E-model's R factor is estimated as a function of the following parameters:

a) Communication-intrinsic parameters: the network PLR (r), the burst length distribution ($f_b(k)$) and the delay due to network and the devices except the de-jitter buffer .
b) Adjust parameters: the codec used, the level of redundancy (N), and the de-jitter buffer size w.

A methodology to obtain the maximum quality is to estimate the communication-intrinsic paramenters and then to find the optimum adjust parameters that maximize the estimation of the E-model's R factor, and consequently, the MOS. This analysis can be extended if other improvement techniques are implemented, e.g., bit-level error correction [16].

6 Study Case: Characterization of Packet Loss of VoIP Calls

In order to verify the proposed models, a set of VoIP calls was studied. These calls were established with the Alliance FXS PCI Voice Cards developed at CTS CINVESTAV with the following characteristics:

- H.323 architecture
- Four ports
- Codec G.711-A law[17]/G.729[18]

The voice data length used for the VoIP calls is shown in Table 1.

Table 1: Used codec types and voice data length.

Voice data length		
ms	$Bytes$	
	G.711	G.729
10	80	10
20	160	20
40	320	40
60	480	60

6.1 Measurement Scenario

The measurement scenario consists of two LANs [19]:

- LAN A: CINVESTAV IPN
- LAN B: Local Cable-ISP Network

Both LANs are in Guadalajara, Mexico, they have different *Internet service providers* (ISP) and are interconnected by the Internet backbone.

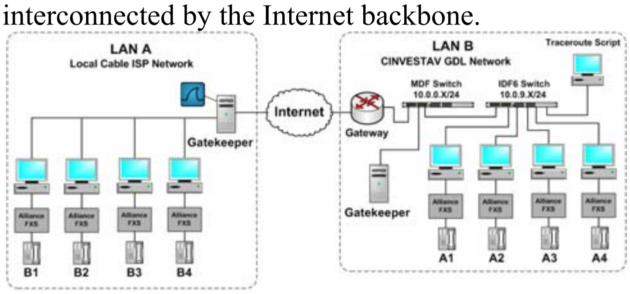

Fig. 7: Measurement scenario.

Table 2: Measurement protocol.

Set	A1/B1	A2/B2	A3/B3	A4/B4
1	G.711-10ms	G.711-20ms	G.711-40ms	G.711-60ms
2	G.729-10ms	G.729-20ms	G.729-40ms	G.729-60ms
3	G.711-10ms	G.711-20ms	G.729-10ms	G.729-20ms
4	G.711-40ms	G.711-60ms	G.729-40ms	G.729-60ms

As shown in Fig. 7, the H.323 zone is composed by the endpoints $A1$, $A2$, $A3$ and $A4$ located in LAN A, the gatekeeper and the endpoints $B1$, $B2$, $B3$ and $B4$, both located in LAN B, each endpoint has an Alliance FXS PCI Voice Card and a conventional cord phone. The measurements protocol is shown in Table 2. The measurements were monitored at LAN A using the Network Protocol Analyzer Wireshark [20].

Additionally, a Traceroute-based script [21] was implemented in LAN B, in a parallel fashion to the VoIP measurements, in order to sample the path followed by the VoIP packets.

6.2 Collected Data Sets

The measurement protocol is described in Table 3. The number of packets sent and the total payload (in bytes) for each set is approximated.

Table 3: Description of the VoIP calls.

Set	Period of measurement	Number of Data traces	Total number of sent packets*	Total payload (Bytes)*
1	Sep/07/2007 10:00→16:00	24	4140000	910620000
2	Sep/10/2007 10:00→16:00	24	4140000	305820000
3	Sep/11/2007 10:00→16:00	24	6480000	840240000
4	Sep/12/2007 10:00→16:00	24	1800000	484200000
*Values are approximated				

6.3 Post-processing and Filtering of the Measurements

The captured RTP streams were processed with Wireshark and filtered with a script to obtain the respective series of sequence numbers (for received packets) of each call (each one representing an 1-hour VoIP call) and the series of inter-arrival time (namely arrival jitter) of consecutive (in sequence number) packets. From the series of sequence numbers of each call, the loss sequence (Y_k), the PLR and the respective gap and burst length distributions ($f_g(k)$ and $f_b(k)$) were obtained.

As explained in section 4, the probabilities of transitions were estimated from the sequence Y_k. From this probabilities, the gap and burst length distributions of both two-state and four-state models, defined by (9) and (10) for two-state model and by (18) and (19) for the four-state model were obtained. The square root of the mean squared error between the respective distributions obtained from measured and theoretical models is also calculated.

6.4 Results

The results presented in this work correspond to the 48 VoIP data traces of sets 3 and 4 which were the ones that presented higher PLR [19].

The burst and gap length distributions of one of the captured traces obtained from a VoIP call with codec G.711 and packet inter-departure time of $20ms$ are shown in Fig. 8 and Fig. 9, respectively.

In Fig. 8 it is shown that the burst length decays rapidly, e.g., to zero probability for burst of length lower than 5 packets. It is also observed that both two-state and four-state models can characterize this decay.

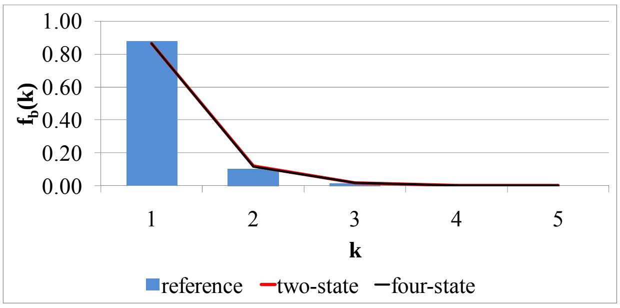

Fig. 8: Burst length distribution of one of the loss sequences.

The gap length distribution decays slower than the burst length distribution. There exist gaps of tens and hundreds of packets with non-negligible probability and, in this case, the less flexible one-parameter formula of the two-state model cannot fit the measured distribution, in contrast with the four-state model, which fits it adequately.

The SMSE for burst length distribution of both two-state and four-state model is quite similar (less than 0.01) for most traces, as seen in Fig. 10. But there is a remarkable difference between both models in the gap distribution. In Fig. 11 it can be observed that the SMSE four-state model fits remarkably better the gap distribution for most traces (its maximum SMSE is 0.002).

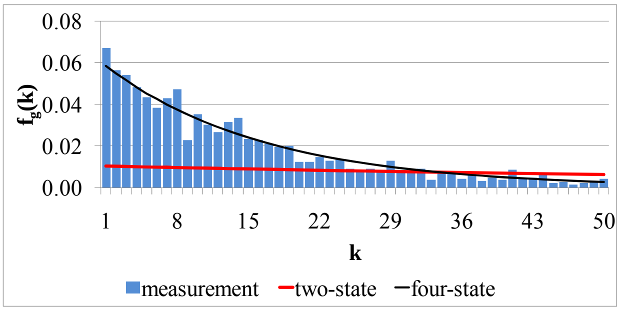

Fig. 9: Gap length distribution of one of the loss sequences.

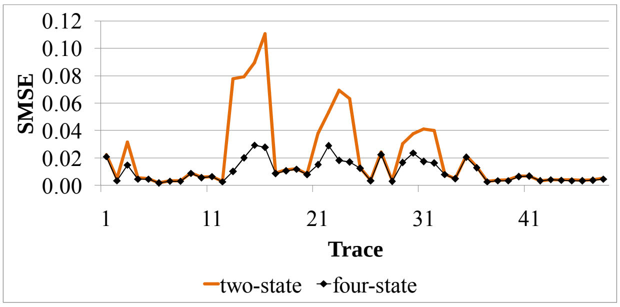

Fig. 10: SMSE of two-state and four-state burst length distribution.

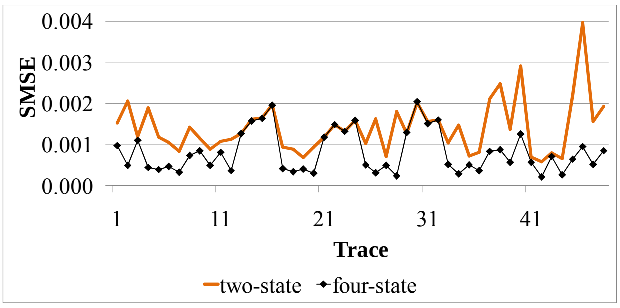

Fig. 11: SMSE of two-state and four-state gap length distribution.

Fig. 12 shows the PLR of the 48 studied data traces, which is calculated as the quotient of the number of lost packets and the number of sent packets. Also, by applying (32), the perceived PLR after a N-packet FEC is estimated for $N = 1$, 2 and 3.

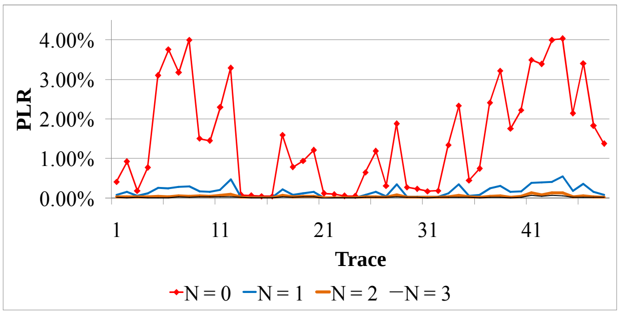

Fig. 12: Perceived PLR for redundancy of $N = 0,1,2,3$ packets.

To determine how the performance is improved when increasing the level of redundancy (N), the relative gain is calculated, which defined as:

$$\Delta_{r'}(N) = \frac{-(r'_N - r'_{N-1})}{r}; N > 0 \quad (42)$$

Fig. 13 shows the relative gain for the studied traces for the redundancy levels $N = 1$, 2 and 3. The major relative gain (approximately 80%) is obtained by adding redundancy of one packet, i.e., for $N = 1$. In this case the perceived PLR decreases below 0.55% for all studied traces, which is acceptable for VoIP calls. Although PLR constraints can be lower than 0.1% for Internet backbone routers or public telephony systems, a less strict limitation applies for VoIP providers and user local's ISP networks, where losses up to 1% are considered undetectable [22].

Although other communication scenario may need a different level of redundancy, e.g., $N = 0$ (no redundancy) or $N = 2$, these results are still significant. It is shown that, as the losses occur in burst of short length (e.g., one or two packets), the major gain is obtained with the first level of redundancy.

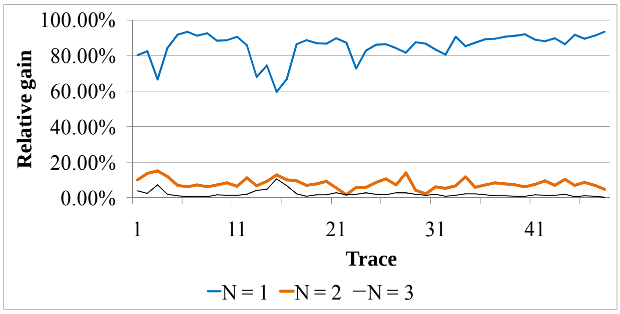

Fig. 13: Relative gain of the perceived PLR.

7 Conclusion

In this work, modeling and characterization of packet loss for a VoIP communication is presented.

The performance of the communication is measured by means of the MOS and the E-model's R factor (as the first can be expressed as a function of the second). Packet reception and loss is modeled by finite-state Markov chains. An innovative contribution of this study corresponds to the models based on two-state and four-state Markov chains: the equations for theoretical gap and burst length distributions, as a function of the probabilities of transitions for both models are proposed. The strategy used to obtain the gap and burst length distributions for the four-state model presented in Section 3.3 exemplifies the generalized methodology for a m-state Markov chain model, which consists of finding firstly their respective CDF, i.e., $C_b(k)$ and $C_g(k)$.

Algorithms for reconstructions, i.e., estimation of the probabilities of transitions between states for two-state and four-state models, are also described.

It is shown through an evaluation based on SMSE that both two-state (at least for most cases) and four-state models can capture the geometric-type decay of the distribution of the burst length, but the two-state model fails to model the gap length distribution when non-homogeneous losses are present. I.e., the gap length distribution is the sum of two geometrical series, as defined by (19), not only one, as defined by (10).

An analysis of the N-packet FEC scheme is also presented. The expected perceived PLR obtained with this correction scheme is quantified, as expresses (32). This resulting general formula applies for the N-packet FEC scheme, regardless of the shape of the burst length distribution.

Through the study of the measurements and the computation of the perceived PLR and relative gain, it is shown that 1-packet FEC is generally sufficient to improve the quality of the communication to an acceptable level, e.g., where the PLR is lower than 1%.

The equations for the estimation of the impairments for combined codification, i.e., when normal packets and redundant information may be coded differently, are derived. These estimations satisfy certain properties that result of the combination of two different codification schemes, as described in section 5.2.

References:

[1] R. Dobrescu, D. Hossu, S. Mocanu, M. Nicolae, New algorithms for QoS performance improvement in high speed networks, *WSEAS Transactions on Communications*, Vol. 7, No. 12, 2008, pp. 1192-1201.

[2] T. Palade, E. Puschita, Requirements for a New Resource Reservation Model in Hybrid Access Wireless Network, *WSEAS Transactions on Communications*, Vol. 7, No. 3, 2008, pp. 377-382.

[3] R. Libnik, A. Svigelj, G. Kandus, R. Libnik, Performance evaluation of SIP based handover in heterogeneous access networks, *WSEAS Transactions on Communications*, Vol. 7, No. 5, 2008, pp. 448-458.

[4] H. Schulzrinne, J. F. Kurose, D. F. Towsley. Loss Correlation for Queues with Bursty Input Streams, *Proc. IEEE ICC '92, Chicago II*, 1992, pp.219-224.

[5] J. R. Yee, E. J. Weldon, Evaluation of the Performance of Error Correcting Codes on a Gilbert Channel, *IEEE transactions on Communications*, Vol. 43 No. 8, 1995.

[6] J. Bolot, S. Fosse-Parisis, D. Towsley. Adaptive FEC-based Error Control for Interactive Audio in the Internet, *Proceedings of IEEE INFOCOM*, 1998, pp. 1453-1460.

[7] H. Lee, H. Lee. A Packet Loss Recovery Scheme Based on the Gap Statistics, *C. Kim (Ed.): ICOIN 2005, LNCS 3391, Springer-Verlag Berlin Heidelberg*, 2005, 2005, pp. 627-634.

[8] H. Zhang, L. Xie, J. Byun, P. Flynn, C. Shim, Packet Loss Burstiness and Enhancement to the E-Model, *Proceedings of the Sixth International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing and First ACIS International Workshop on Self-Assembling Wireless Networks (SNPD/SAWN'05)*, 2005.

[9] N. Erdöl, C. Castellucia, Ali Zilouchian. Recovery of Missing Speech Packets Using the Short-time Energy and Zero-crossing Measurement, *IEEE Transactions on Speech and Audio Processing*, Vol. 1 No. 3, July 1993.

[10] S. Madhani, S. Shah, A. Gutierrez, Optimized Adaptive Jitter Buffer Design for Wireless Internet Telephony, *Global Telecommunications Conference, 2007. GLOBECOM '07*. IEEE, 2007, pp. 5248-5253.

[11] W. Ching, M. K. Ng., *Markov Chains: Models, Algorithms and Applications*, Springer, 2006.

[12] G. G. Yin, Q. Zhang, *Discrete-Time Markov Models, Two-Time-Scale Methods and Applications*, Springer, 2005.

[13] ITU-T Recommendation G.107, The E-model, a computational model for use in transmission planning, 2003.

[14] G. L. Choudhurya, R G. Cole, Design and analysis of optimal adaptive de-jitter buffers, *Computer Communications 27*, 2004, pp. 529-537.

[15] R.G. Cole, J.H. Rosenbluth, Voice over IP performance monitoring, *ACM J. Computer Communications Review 31*, 2001.

[16] H. D. Wacker, J. Boercsoek, H. Hillmer, Redundant Data Transmission and Nonlinear Codes, *WSEAS Transactions on Communications*, Vol. 7, No. 6, 2008.

[17] ITU-T Recommendation G.711, Pulse Code Modulation (PCM) of Voice Frequencies, 1993.

[18] ITU-T Recommendation G.729, "Coding of speech at 8 kbit/s using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP)," 1998.

[19] H. Toral, D. Torres, C. Hernandez, L. Estrada. Self-similarity, Packet Loss, Jitter, and Packet Size: Empirical Relationships for VoIP, *18th International Conference on Electronics, Communications and Computers, CONIELECOMP*, March 2008.

[20] "Wireshark: A Network Protocol Analyzer", http://www.wireshark.org/.

[21] V. Jacobson, "Traceroute", ftp://ftp.ee.lbl.gov/traceroute.tar.gz, February 1989.

[22] K. Maheswari, M. Punithavalli, A Survey of Packet Loss in VoIP, *International Journal of Computational Intelligence Research*, Vol. 5, No. 1, 2009, pp. 57–66.