Mathematical and Computer Modelling 57 (2013) 2832–2845

Contents lists available at SciVerse ScienceDirect

Mathematical and Computer Modelling

journal homepage: www.elsevier.com/locate/mcm

Accurate modeling of VoIP traffic QoS parameters in current and future networks with multifractal and Markov models[☆]

Homero Toral-Cruz[a,*], Al-Sakib Khan Pathan[b], Julio C. Ramírez Pacheco[c]

[a] Department of Sciences and Engineering, University of Quintana Roo - UQROO, Chetumal, Quintana Roo, Mexico
[b] Department of Computer Science, International Islamic University Malaysia - IIUM, Kuala Lumpur, Malaysia
[c] Department of Basic Sciences and Engineering, University of Caribe - UNICARIBE, Cancún, Quintana Roo, Mexico

ARTICLE INFO

Article history:
Received 19 August 2011
Received in revised form 9 November 2011
Accepted 5 December 2011

Keywords:
VoIP
Packet loss
Jitter
Hurst parameter
Markov chains
Multifractality

ABSTRACT

In this paper, we analyze the jitter and packet loss behavior of Voice over Internet Protocol (VoIP) traffic by means of network measurements and simulations results. As result of these analyses, we provide a detailed characterization and accurate modeling of these Quality of Service (QoS) parameters. Our studies have revealed that VoIP jitter can be modeled by self-similar processes with short range dependence (SRD) or long range dependence (LRD). The discovery of LRD (a kind of asymptotic fractal scaling) in VoIP jitter was followed by a further work that shows the evidence of multifractal behavior. The implication of these behaviors for VoIP and other interactive multimedia services is that receiver de-jitter buffer may not be large enough to mask the jitter with LRD and multifractal characteristics. On the other hand, we use a description of VoIP packet loss based on microscopic and macroscopic packet loss behaviors, where these behaviors can be modeled by 2-state and 4-state Markov chains, respectively. Based on the above mentioned points, we present a methodology for simulating packet loss. Besides, we found relationships between Hurst parameter (H) with the packet loss rate (PLR); these relationships are based on voice traffic measurements and can be modeled by means of a power-law function, characterized by three fitted parameters. The proposed models can be used to: (i) design a de-jitter buffer, (ii) implement a synthetic generator of VoIP jitter data traces, where the synthetic jitter data traces can be used as test vectors to carry out the performance evaluation of a de-jitter buffer of VoIP system, and (iii) design effective schemes for packet loss recovery.
© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

In this era, voice can be transmitted by circuit switched (public switched telephone network — PSTN) and packet switched networks (Internet or IP network). Compared to traditional resource-dedicated PSTN, IP network is resource-shared. Therefore, the conditions in the PSTN are totally different from those in the IP network. In the last few years, VoIP has became one of the most attractive and important applications running on the Internet, poised to replace the PSTN in the future. There are several advantages in the case of voice transmission using VoIP technology: the reduced communication cost, the use of joined IP infrastructure, the use in multimedia applications, etc. Another interesting fact is that sending wireless phone calls over IP networks is considerably less expensive than that of sending over cellular voice networks. However, such types of communications must ensure good performance and quality of voice transmissions.

☆ This is the extended and revised version of all of our previous works.

* Corresponding author. Tel.: +52 9838350300; fax: +52 9838329656.
E-mail addresses: htoral@uqroo.mx (H. Toral-Cruz), sakib@iium.edu.my (A.-S.K. Pathan), jramirez@unicaribe.edu.mx (J.C. Ramírez Pacheco).

0895-7177/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.mcm.2011.12.007

Faster wireless networking technologies and more powerful mobile telephones promise to help solve these problems [1,2]. With the currently available technologies, VoIP over mobile networks is not yet very popular. However, the advantage of such technology is that keeping the wireless network as it is, with a higher Quality-of-Service (QoS) facility, the existing infrastructure of IP networks could be used for serving the users. If a company owns the WiFi connection, such VoIP communication could be very cheap. If either the WiFi or WiMAX network are unavailable, cellular technologies could be used to pass the voice traffic to the Internet. This is known as Mobile Voice over IP (mVoIP). mVoIP via cellular services could achieve QoS by prioritizing voice packets over those used for data and other traffic types.

The reality today is that the current Internet provides best-effort services in most of the cases and cannot guarantee the Quality of Service of real-time multimedia applications (such as VoIP). To achieve a satisfactory level of voice quality, the VoIP networks must be designed by using correct traffic models. Traffic modeling comprises the follows steps [3]: (i) Selection of one or more models that may provide a good description of the traffic type. In order to select an adequate traffic model, it is necessary to study the traffic characteristics. The main characteristics of a traffic source are its average data rate, burstiness, and correlation. The average data rate gives an indication of the expected traffic volume for a given period of time. Burstiness describes the tendency of traffic to occur in clusters. Data burstiness is manifested by the correlation function which describes the relation between packet arrivals at different times, and is an important factor in packet losses due to buffer and bandwidth limitations. (ii) Estimation of parameters for the selected model. Parameter estimation is based on a set of statistics (e.g., mean, variance, density function or autocovariance function, multifractal characteristics) that are measured or calculated from observed data traces. The set of statistics used in the inference process depends on the impact they may have on the main QoS parameters of interest. (iii) Statistical testing for election of one of the considered models and analysis of its suitability to describe the traffic type under analysis. One of the models which has been widely applied in classical teletraffic modeling is the Poisson model. However, the IP networks traffic has different characteristics and the Poisson approximation will be acceptable only under particular conditions [4]. In recent years several types of IP traffic behavior, that can have significant impact on network performance, were discovered [3]: long-range dependence, self-similarity and, more recently, multifractality. Long range dependent traffic produces a wide range of traffic volume away from the average rate. This great variation in the traffic volume leads to buffer overflow and network congestion that result in packet loss and jitter, which directly impact on the quality of VoIP applications.

Amongst the different quality parameters, packet loss is the main impairment which makes the VoIP perceptually most different from the PSTN. A number of studies have shown that packet losses exhibit a finite temporal dependency due to the multiplexing policy on the shared resources such as bandwidth and buffer through the transmission paths in the network. This temporal dependency can be modeled by a finite memory binary model. An example of such a model is the binary Markov chain model [5,6]. The objective of packet loss modeling is to characterize its probabilistic behavior.

Motivated by such concerns, we analyze the jitter and packet loss behavior of VoIP traffic by means of network measurements and simulations results. Furthermore, we provide a detailed characterization and accurate modeling of these main QoS parameters. We basically modeled the QoS parameters of VoIP traffic which could be related with regular VoIP communications over regular wired networks as well as for Mobile Voice over IP (mVOIP) (which functions as an application that runs over any wireless network technology that provides data access to the Internet). Hence, our modeling work is basically relevant to both regular VoIP and mVoIP technologies. These characterization and models can be used by other researches to design and implement de-jitter buffers, synthetic generators of VoIP jitter data traces and effective schemes for packet loss recovery.

The main contributions of this paper are summarized as follows:

1. A rich set of VoIP data traces was collected by means of network measurements.
2. A VoIP jitter model was proposed by self-similar and multifractal models.
3. A description of VoIP packet loss based on microscopic and macroscopic packet loss behaviors was suggested, where these behaviors can be modeled by 2-state and 4-state Markov models, respectively.
4. A methodology for simulating packet loss on VoIP traffic by means of a 4-state Markov model was proposed.
5. A relationship between the Hurst parameter and the Packet Loss Rate was found, where this relationship can be modeled by means of a power-law function with three fitted parameters.

The paper is organized as follows: Section 2 presents related works. Section 3 provides some mathematical background on the self-similar processes. Section 4 presents the measurements description. In Section 5, we discuss the jitter and packet loss behaviors. In Section 6, we propose that VoIP jitter can be modeled by self-similar and multifractal models. In Section 7, we present a methodology for simulating packet loss on VoIP traffic and propose a new model that allows relating the *H* parameter with the *PLR*. Finally, Section 8 concludes the paper highlighting the achievements from this work with possible future use of our findings.

2. Related works

There is an extensive amount of work oriented to the studies of Internet traffic. Different studies take a different perspective depending on their specific interest. The Internet research literature contains many analyses of data traffic measurements. However for VoIP, there are very few studies in the literature that report on the analysis of VoIP data traces captured on operational links of the Internet backbone. Such studies are needed in order to build accurate VoIP traffic models, which are a basis for the design of VoIP applications, network engineering and the development of Internet technology. VoIP requires strict QoS levels. The quality of a VoIP call depends mainly on low jitter and packet loss. Based on [7], QoS requirements for VoIP are constrained as follows: (i) jitter should be less than 40 ms, and (ii) packet loss should be below 3%. We are interested in the measurements, characterization and modeling of these QoS parameters of voice communications over the Internet.

In [7–9] the QoS parameters of VoIP applications are studied, though, the relationships between them to assess the overall effects of these parameters are not considered. Therefore, these studies are limited, because the impact of these QoS parameters is analyzed separately from the others. In this work, we show that some QoS parameters are intricately related to each other.

It has been shown through empirical studies that data traffic exhibits self-similar nature and long range dependence (LRD) [10–12]. The presence of LRD, is remarkably universal, and has become an indispensable part of traffic modeling, in particular for TCP/IP traffic in the Internet [13]. Furthermore, the discovery of evidence for multifractal behavior, a richer form of scaling behavior associated with non-uniform local variability, raised hopes that another "traffic invariant" had been found which could lead to a complete, robust model of aggregate wide area network (WAN) traffic over all time scales [13]. The multifractal behavior of network traffic was first noticed by Riedi and Véhel [14]. Subsequently, various studies have addressed the characterization and modeling of multifractal traffic, essentially within the framework of random cascades [15–17].

On the other hand, in the Internet, packet losses that occur due to temporary overloaded situations, are bursty in nature and they exhibit temporal dependency [5]. So, if packet n is lost, then normally there is a higher probability that packet $(n+1)$ will also be lost. Consequently, there is a strong correlation between consecutive packet losses, resulting in a bursty packet loss behavior. This temporal dependency can be effectively modeled by a finite Markov chain [5,6].

Previous works [18,19] presented a methodology for simulating packet loss. This methodology is restricted, because it incorporates only one microscopic period of packet loss by using a 2-state Markov chain. In order to generalize this methodology, in this work we proposed to incorporate "n" microscopic periods of packet loss by means of 4-state Markov chain.

In contrast to the above works, in this paper, we analyze VoIP traffic by means of network measurements and simulations results. As a result of these analyses, we provide a detailed characterization and accurate modeling of the main QoS parameters at packet level that could help both VoIP and mVoIP technologies.

3. Mathematical backgrounds and preliminaries

3.1. Self-similar processes

An attractive property of the self-similar processes for modeling a time series of IP traffic is the degree of self-similarity, which is expressed with a single parameter called the Hurst parameter. This parameter expresses the speed of decay of the autocorrelation function of the time series. In this section, a brief overview of self-similar processes is given from [10–12].

Continuous Self-similarity: A real-valued continuous time series $\{X(t), t \in \Re\}$ is self-similar with the exponent $0 < H < 1$ if, for any $a > 0$, the finite-dimensional distributions of $\{X(at), t \in \Re\}$ are identical to the finite-dimensional distributions $\left\{a^H X(t), t \in \Re\right\}$; i.e., $\{X(at), t \in \Re\} \stackrel{d}{=} \left\{a^H X(t), t \in \Re\right\}$, where $\stackrel{d}{=}$ denotes equality in distribution.

Discrete Self-similarity: Let $X_t = (X_t; t \in N)$ denote a discrete time series with mean μ_X, variance σ_X^2, autocorrelation function $r(k)$ and autocovariance function (ACV) $\gamma(k)$, $k \geq 0$; where X_t can be interpreted as the jitter, at time instance t.

When considering discrete time series, the definition of self-similarity is given in terms of the aggregated processes, as follows:

$$X_k^{(m)} = \left(X_k^{(m)}; k \in N\right) \tag{1}$$

where m represents the aggregation level and $X_k^{(m)}$ is obtained by averaging the original series X_t over non-overlapping blocks of size m, and each term $X_k^{(m)}$ is given by:

$$X_k^{(m)} = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i; \quad k = 1, 2, 3, \ldots. \tag{2}$$

Then it is said that X_t is self-similar $(H - ss)$ with self-similarity parameter $(0 < H < 1)$ if:

$$X_k^{(m)} \stackrel{d}{=} m^{H-1} X_t. \tag{3}$$

Second-order Discrete Self-similarity: X_t is called exactly second order self-similar with Hurst parameter H if the variance and covariance of the aggregated time series are defined by Eqs. (4) and (5), respectively:

$$\operatorname{var}\left(X_k^{(m)}\right) = \sigma_X^2 \cdot m^{2H-2} \tag{4}$$

$$\gamma_X^m(k) = \frac{\sigma_X^2}{2}\left((k+1)^{2H} - 2k^{2H} + (k-1)^{2H}\right) \quad k \geq 1. \tag{5}$$

The time series X_t is called asymptotically second-order self-similar if:

$$\lim_{m\to\infty} \gamma^m(k) = \frac{\sigma_X^2}{2}\left((k+1)^{2H} - 2k^{2H} + (k-1)^{2H}\right). \tag{6}$$

Second-order self-similarity (in the exact or asymptotic sense) has been a dominant framework for modeling IP traffic.

So far role of second-order self-similarity has been discussed but not much has been mentioned about the role of H and limiting values. The definition of long range dependence and its interconnection with the correlation factor $r(k)$ will now be discussed.

Let $r(k) = \gamma(k)/\sigma_X^2$ be the autocorrelation function of X_t with self-similarity parameter $0 < H < 1, H \neq 1/2$, then the asymptotic behavior of $r(k)$, is given by Eq. (7):

$$r(k) \sim H\,(2H-1)\,k^{2H-2} \quad k \to \infty. \tag{7}$$

In particular, if $\frac{1}{2} < H < 1$, $r(k)$ asymptotically behaves as $ck^{-\eta}$ for $0 < \eta < 1$, where $c > 0$ is a constant, $\eta = 2 - 2H$ and this also means that the correlations are nonsummable: $\sum_{k=-\infty}^{\infty} r(k) = \infty$. That is, the autocorrelation function decays slowly. When $r(k)$ obeys a power-law, the corresponding stationary process X_t is called long range dependent. On the other hand, X_t is short range dependent if the sum $\sum_{k=-\infty}^{\infty} r(k) < \infty$ does not diverge.

Generally speaking, time series with long-range dependence has a Hurst parameter $0.5 < H < 1$, on the other hand, time series with short-range dependence has a Hurst parameter $0 < H < 0.5$.

3.2. Haar wavelet-based decomposition and Hurst index estimation

The time series X_t can be decomposed into a set of time series [20], each defined by:

$$C_{X,t}^{2,i} = X_t^{(2^{i-1}E)} - X_t^{(2^iE)} \quad i \in N \tag{8}$$

where $X_t^{(2^iE)}$ is the time series X_t after two operations, which are:

1. Aggregation at level 2^i, as defined by Eqs. (1) and (2). i.e., $m = 2^i$.
2. Expansion of level 2^i, which consists of 'repeating' each element of a time series 2^i times.

i.e., $X_j^{(2^iE)} = X_k^{(2^i)}$ for $k = 1 + \left\lfloor \frac{j-1}{2^i} \right\rfloor$ and $j \in N$.

These zero mean components $C_{X,t}^{2,i}$ have three important properties:

1. They synthesize the original time series without loss, i.e.,
$$X_t = \sum_i C_{X,t}^{2,i}. \tag{9}$$
2. They are pairwise orthogonal:
$$\left\langle C_{X,t}^{2,i_1}, C_{X,t}^{2,i_2} \right\rangle = 0; \quad \text{for } i_1 \neq i_2. \tag{10}$$
3. If X_t is exactly self-similar, then the variance of the components comply:
$$\mathrm{var}(C_{X,t}^{2,i}) = 2^{2H-2} \cdot \mathrm{var}(C_{X,t}^{2,i-1}). \tag{11}$$

Properties 1, 2 and 3 imply Eqs. (12) and (13):

$$\sigma_X^2 = \sum_i \mathrm{var}(C_{X,t}^{2,i}) \tag{12}$$

$$\sigma_X^2 = \frac{1}{1-2^{2H-2}} \cdot \mathrm{var}(C_{X,t}^{2,1}). \tag{13}$$

Then, the variance of the ith component is related to the variance of X_t as following:

$$\mathrm{var}(C_{X,t}^{2,i}) = (1-r) \cdot r^{i-1} \cdot \sigma_X^2 \tag{14}$$

where

$$r = 2^{2H-2}. \tag{15}$$

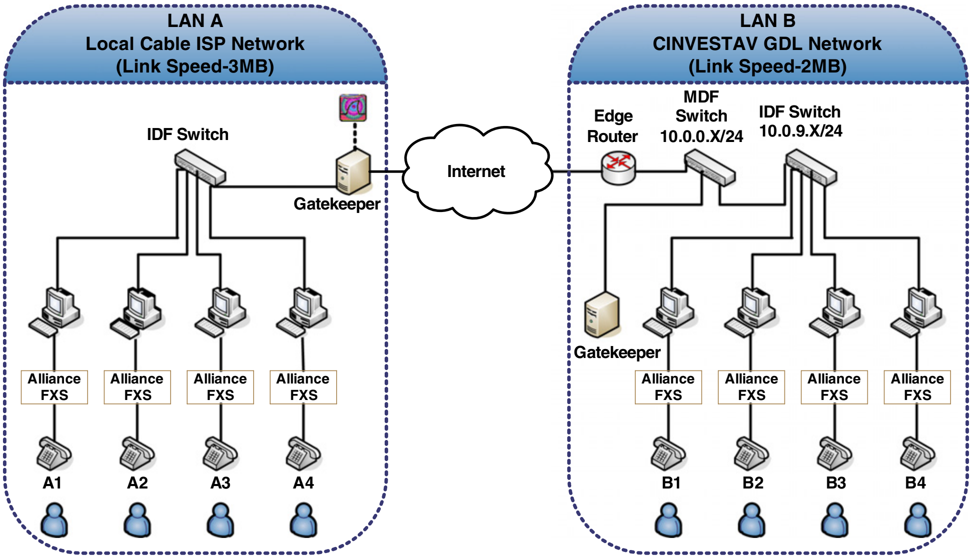


Fig. 1. Measurement scenario.

Table 1
Parameter configuration employed in the test calls.

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

The plot $\log_2[\text{var}(C_{X,t}^{2,i})]$ vs. i is equivalent to the wavelet-based diagram proposed in [17], the Logscale Diagram (LD-Diagram); i.e., $\text{var}(C_{X,t}^{2,i}) = \frac{E|d_X(j,\cdot)|^2}{2^j}$ when using the Haar family of wavelet basis functions $\psi_{j,k}(t) = 2^{-j/2}\psi_0(2^{-j}t - k)$ (see [21]) where:

$$\psi_0(t) = \begin{cases} +1 & 0 \le t < 1/2 \\ -1 & 1/2 \le t < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

For a finite-length time series with "L" octaves, the number of octaves (j) of the LD-Diagram is related to index i of Eq. (8), according to Eq. (17):

$$j = i. \tag{17}$$

The LD-Diagram of an exactly self-similar time series is a straight line. Then, a linear regression can be applied in order to estimate the Hurst index.

4. Network measurements

In order to accomplish our analysis, extensive jitter and packet loss measurements were collected, as follows:

- Test calls were established by a VoIP application called "Alliance Foreign eXchange Station" (FXS) [22].
- The jitter and packet loss were measured by Wireshark [23] to obtain a set of data traces.
- The measurement scenario was based on a typical H.323 architecture; see Fig. 1.
- The parameter configuration employed in the test calls is shown in Table 1:
  - Four simultaneous test calls were established between A1/B1, A2/B2, A3/B3 and A4/B4 endpoints, see Fig. 1 and Table 1.
  - The configurations used in the test calls are based on two parameters: CODEC type (G.711 and G.729), and voice data length (10 ms, 20 ms, 40 ms and 60 ms), see Table 1.
  - The measurement periods were one hour for each test call (call duration time).

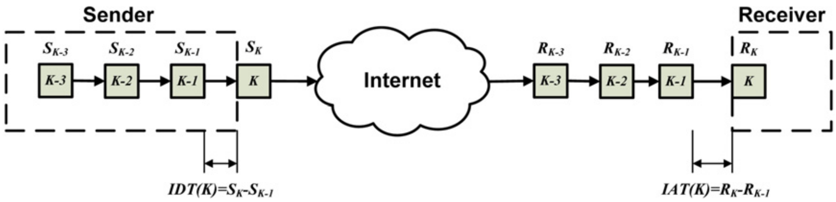


Fig. 2. Jitter experienced across Internet paths.

– For each measurement period (one hour), four jitter and packet loss data traces were obtained.
– The four configuration sets contain more than one hundred-thirteen million voice packets corresponding to 710 jitter and 710 packet loss data traces, measured during typical working hours, between 2004 and 2010 [24].

5. Jitter and packet loss behavior

The voice quality of VoIP applications depends on many parameters; such as bandwidth, one way delay (OWD), jitter, packet loss rate, codec, voice data length, and de-jitter buffer size. In particular, jitter and packet loss have an important impact on voice quality. In this section, we describe the jitter and packet loss behavior.

5.1. Jitter

When voice packets are transmitted from source to destination over IP networks, packets may experience variable delay, called jitter. The packet inter-arrival time (IAT) on the receiver side is not constant even if the packet inter-departure time (IDT) on the sender side is constant. As a result, packets arrive at the destination with varying delays (between packets) referred to as jitter. We measure and calculate the difference between arrival times of successive voice packets that arrive on the receiver side, according to RFC 3550 [25], this is illustrated in Fig. 2. This figure shows the jitter measurement between the sending packets and the receiving packets.

Let S_K be the RTP timestamp and R_K be the arrival time in RTP timestamp units for packet K. Then, for two packets K and $K-1$, the OWD difference between two successive packets, K and $K-1$ is given by the Eq. (2):

$$J(K) = (R_K - S_K) - (R_{K-1} - S_{K-1}) = (R_K - R_{K-1}) - (S_K - S_{K-1}) = \mathrm{IAT}(K) - \mathrm{IDT}(K) \tag{18}$$

$$\mathrm{IAT}(K) = J(K) + \mathrm{IDT}(K) \tag{19}$$

where, $\mathrm{IDT}(K) = (S_K - S_{K-1})$ is the inter-departure time (in our experiments, IDT $= \{10$ ms, 20 ms, 40 ms, and 60 ms$\}$) and $\mathrm{IAT}(K) = (R_K - R_{K-1})$ is the inter-arrival time or arrival jitter for the packets K and $K-1$. In the current context, IAT(K) is referred to as jitter.

On the other hand, the voice data lengths of 10 ms, 20 ms, 40 ms and 60 ms are used and the successive voice packets are transmitted at a constant rate, i.e., 1 packet/10 ms, 1 packet/20 ms, 1 packet/40 ms and 1 packet/60 ms, respectively. However, when voice packets are transported over IP networks, they may experience delay variations and packet loss. From Eq. (19), a relationship between jitter and packet loss can be established using the following equations:

If packet $K-1$ is lost,

$$\mathrm{IAT}(K) = J(K) + (2) \cdot \mathrm{IDT}(K). \tag{20}$$

Therefore, if n consecutive packets are lost,

$$\mathrm{IAT}(K) = J(K) + (n+1) \cdot \mathrm{IDT}(K). \tag{21}$$

Therefore, Eq. (21) describes the packet loss effects in the VoIP jitter. The relationship expressed by this equation can be used for simulating packet loss (see Section 7.2).

5.2. Packet loss

There are two main transport protocols used on IP networks, user datagram protocol (UDP) and transmission control protocol (TCP). While UDP protocol does not allow any recovery of transmission errors, TCP includes an error recovery process. However, the voice transmission over TCP connections is not very realistic. This is due to the requirement for real-time (or near real-time) operations in most voice related applications. As a result, the choice is limited to the use of UDP which involves packet loss problems. Packet loss can occur in the network or at the receiver side, for example, due to excessive network delay in case of network congestion.

Owing to the dynamic, time varying behavior of packet networks, packet loss can show a variety of distributions. The loss distribution most often studied in speech quality tests is random or Bernoulli-like packet loss. Random loss here means independent loss, implying that the loss of a particular packet is independent of whether or not previous packets were lost.

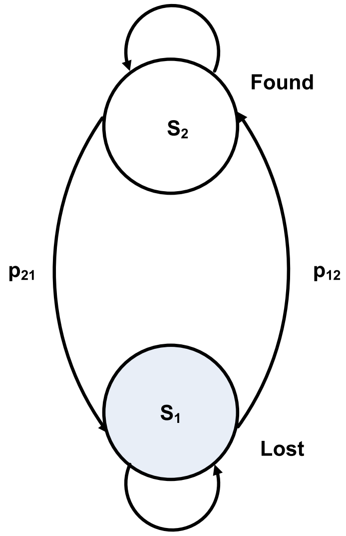


Fig. 3. 2-state Markov chain.

However, random loss does not represent the loss distributions typically encountered in real networks. For example, losses are often related to periods of network congestion. Hence, losses may extend over several packets, showing a dependency between individual loss events. In this work, dependent packet loss is often referred to as bursty. The packet loss is bursty in nature and exhibits temporal dependency [5]. As noted earlier in the Introduction, if packet n is lost then normally there is a higher probability that packet $n+1$ will also be lost. Consequently, there is a strong correlation between consecutive packet losses, resulting in a bursty packet loss behavior. Hence, this temporal dependency can be effectively modeled by a finite Markov chain [5,6]. In this work, the 2-state and 4-state Markov chains are used.

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 state S_i to the state S_j. The probabilities of transitions between states can be represented by the transition matrix P:

$$P = \begin{bmatrix} S_1 & S_2 & \cdots & S_m \\ S_1 & S_2 & \cdots & S_m \\ \vdots & \vdots & \ddots & \vdots \\ S_1 & S_2 & \cdots & S_m \end{bmatrix} \qquad (22)$$

such that $S_1 + S_2 + \cdots + S_m = 1$.

In the 2-state Markov chain (see Fig. 3), one of the states (S_1) represents a packet loss and the other state (S_2) represents the case where packets are correctly transmitted or received. The transition probabilities in this model are represented by p_{21} and p_{12}. In other words, p_{21} is the probability of going from S_2 to S_1, and p_{12} is the probability of going from S_1 to S_2. Different values of p_{21} and p_{12} define different packet loss conditions that can occur on the Internet.

The steady-state probability of the chain to be in the state S_1, namely the *PLR*, is given by Eq. (23) [26]:

$$S_1 = \frac{p_{21}}{p_{21} + p_{12}} \qquad (23)$$

and clearly $S_2 = 1 - S_1$.

The collected data traces in real IP networks can be modeled accurately with a higher number of states, i.e., n-state Markov chains. However, for network planning, a trade off is desirable between very accurate modeling of data traces and a low number of model input parameters, in order to yield a model still usable for network planners with reasonable effort. Therefore, we used a simplification of an n-state chain, i.e., the 4-state Markov chain. Fig. 4 shows the state diagram of this 4-state Markov chain. In this model, a '*good*' and a '*bad*' state are distinguished, which represent periods of lower and higher packet loss, respectively. Both for the '*bad*' and the '*good*' state, an individual 2-state Markov chain represents the dependency between consecutively lost or found packets.

The two 2-state chains can be described by four independent transition probabilities (two for each one). Two further probabilities characterize the transitions between the two 2-state chains, leading to a total of six independent parameters for this particular 4-state Markov chain.

In the 4-state Markov chain, states S_1 and S_3 represent packets lost, S_2 and S_4 packets found and 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. The four steady-state probabilities

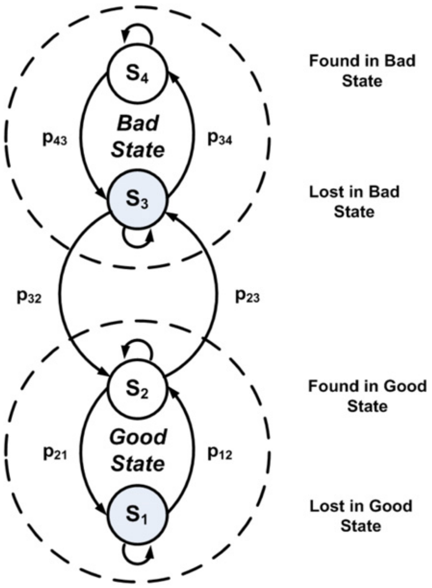


Fig. 4. 4-state Markov chain.

of this chain are [26]:

$$S_1 = \frac{1}{1 + \frac{p_{12}}{p_{21}} + \frac{p_{12}p_{23}}{p_{21}p_{32}} + \frac{p_{12}p_{23}p_{34}}{p_{21}p_{32}p_{43}}} \tag{24}$$

$$S_2 = \frac{1}{1 + \frac{p_{21}}{p_{12}} + \frac{p_{23}}{p_{32}} + \frac{p_{23}p_{34}}{p_{32}p_{43}}} \tag{25}$$

$$S_3 = \frac{1}{1 + \frac{p_{34}}{p_{43}} + \frac{p_{32}}{p_{23}} + \frac{p_{21}p_{32}}{p_{12}p_{23}}} \tag{26}$$

$$S_4 = \frac{1}{1 + \frac{p_{43}}{p_{34}} + \frac{p_{32}p_{43}}{p_{23}p_{34}} + \frac{p_{21}p_{32}p_{43}}{p_{12}p_{23}p_{34}}}. \tag{27}$$

The probability of the chain to be either in S_1 or in S_3, which corresponds to *PLR*, is then $r = S_1 + S_3$ [26].

6. Jitter modeling

6.1. Self-similarity, SRD and LRD

Time-dependent statistics (e.g., correlation) are important for performance evaluation of IP networks. These statistics can be used to measure the impact of specific impairments. Several studies have found that self-similarity and long-range dependence can have a negative impact in data traffic, because they give rise to great losses and/or delays. For this reason, it is important to analyze the correlation structures (SRD and LRD) of the VoIP traffic. The Hurst parameter is used to measure the degree of self-similarity and LRD. Generally speaking, time series with long-range dependence have a Hurst parameter of $0.5 < H < 1$, on the other hand, time series with short-range dependence have a Hurst parameter of $0 < H < 0.5$.

Unlike other statistics, the Hurst parameter, although mathematically well defined, cannot be estimated unambiguously from real-world samples. Therefore, several methods have been developed in order to estimate the Hurst parameter. Examples of classical estimators are those based on the R/S statistic [27] (and its unbiased version [28]), detrended fluctuation analysis (DFA) [28,29], maximum likelihood (ML) [30], aggregated variance (VAR) [27], wavelet analysis [17], etc. In [31], Clegg developed an empirical comparison of estimators for data in raw form and corrupted in various ways. An important observation is that the estimation of the Hurst parameter may differ from one estimator to another, which makes the selection of the most adequate estimator a difficult task. It seems to depend on how well the data sample meets the assumptions the estimator is based on. However, through analytical and empirical studies it has been discovered that the estimators that have the best performance in bias and standard deviation, and, consequently, in mean squared error (MSE), are the Whittle ML and the wavelet-based estimator proposed by Veitch and Abry in [17]. From these two estimators, the wavelet-based is computationally simpler and faster [27,17].

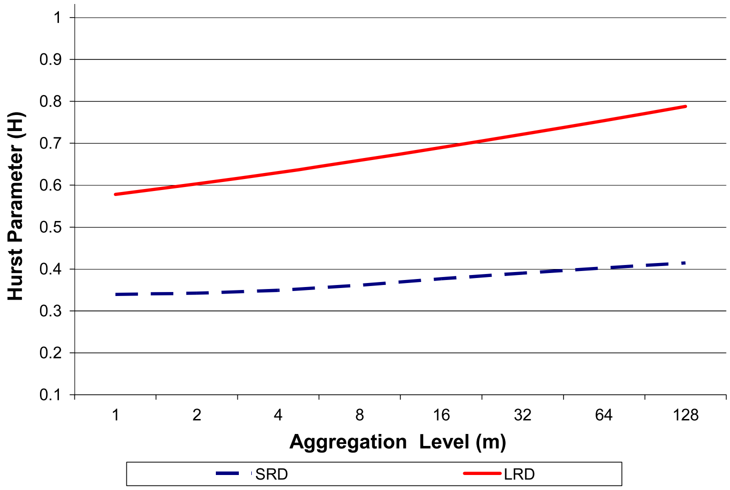


Fig. 5. Hurst parameter for VoIP jitter data traces.

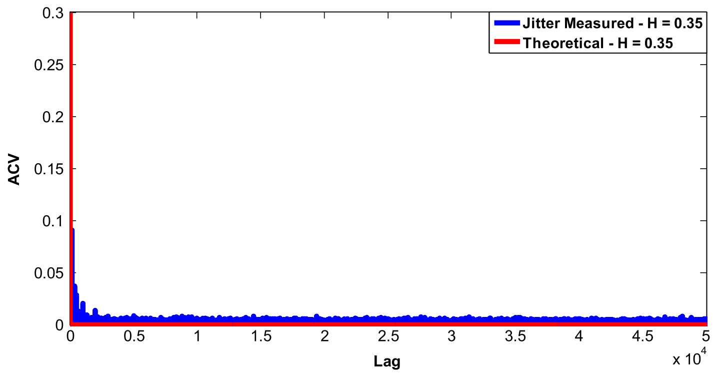


Fig. 6. Autocovariance function for VoIP jitter data traces with SRD.

Motivated by the above points, and following the methodology proposed in [32] to find correlations and long-range dependence, the Hurst parameter is estimated by the wavelet-based estimator [17] of jitter data traces as a function of the aggregation level m ($m = \{1, 2, 4, 8, 16, 32, 64, 128\}$). Fig. 5 shows the Hurst parameter of representative jitter data traces to different aggregation levels m. Generally, Hurst parameters larger than 0.5 for all aggregation levels are a strong indication of long-range dependence. It can be observed from Fig. 5 that a set of jitter data traces has Hurst parameters larger than 0.5 for all aggregation levels. This indicates a high degree of long-range dependence (LRD). In contrast, other sets of jitter data traces have Hurst parameters lower than 0.5. These results are thus not a strong indication of long-range dependence. This indicates that the autocovariance functions decay quickly to zero, indicating no memory property (SRD).

In order to complement the above analysis, a study was made of the behavior of autocovariance functions of the corresponding jitter data traces in Fig. 5.

Fig. 6 shows the comparison between the autocovariance function of a measured data trace with $H = 0.35$ and the theoretical autocovariance function. It can be observed that the autocovariance function of the measured data trace behaves similarly to the ideal model and decay quickly to zero. This result confirms the presence of short-range dependence or memoryless property in VoIP jitter data traces.

In contrast, a comparison was made between the autocovariance function of a measured data trace with $H = 0.58$ and the theoretical autocovariance function, as shown in Fig. 7. In this figure, a similar behavior can be observed between these autocovariance functions and a very slow decaying from them. This behavior indicates the presence of a high degree of long-range dependence or long memory property.

These results show that VoIP jitter exhibits self-similar characteristics with short- or long-range dependence, therefore, a self-similar process can be used to model the jitter behavior. However, we can see that the data traces with SRD have a higher degree of self-similarity in contrast to the data traces with LRD. In the next section we give an explanation of this behavior.

6.2. Multifractal behavior

The network analysis was transformed by the discovery of scale invariance properties in the data traffic (self-similarity). So also, the presence of long range dependence has become an important part of traffic modeling, in particular for data traffic.

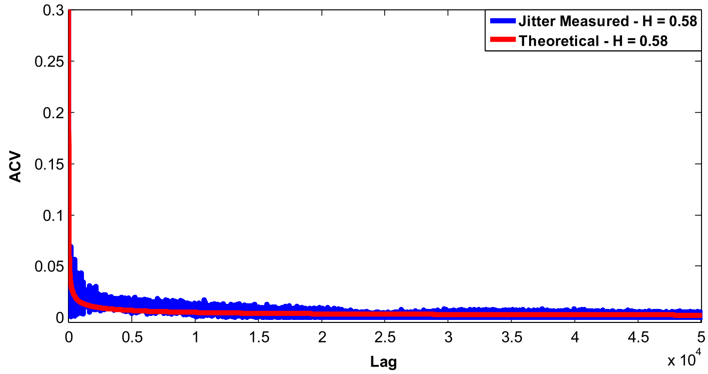


Fig. 7. Autocovariance function for VoIP jitter data traces with LRD.

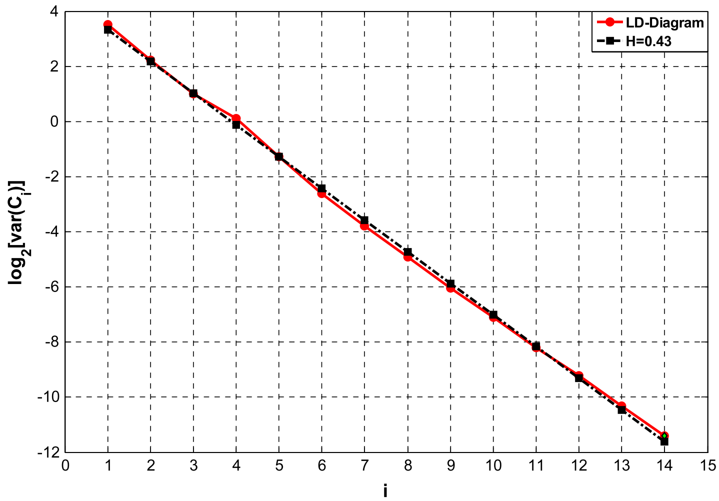


Fig. 8. Components behavior of VoIP jitter data traces: monofractal behavior.

The discovery of LRD and weak self-similarity in the VoIP jitter data traces was followed by a further work that shows the evidence for multifractal behavior. The discovery of evidence for multifractal behavior is a richer form of scaling behavior associated with non-uniform local variability, which could lead to a complete and robust model of IP network traffic over all time scales.

In this section, we review the evidence for multifractal behavior of VoIP jitter. In order to accomplish this analysis, we decomposed the time series of VoIP jitter into a set of time series or components $C_{X,t}^{2,i}$ as it is defined in Eq. (8). The behavior of these components is used to determine the kind of asymptotic fractal scaling. If the variance of the components of a time series is modeled by a straight line, the time series exhibits monofractal behavior. Then, a linear regression can be applied in order to estimate the Hurst parameter. On the other hand, if the variance of the components cannot be adequately modeled with a linear model, then the scaling behavior should be described with more than one scaling parameter, i.e., the time series exhibits multifractal behavior. [33]. In Figs. 8 and 9, we show the components behavior of the collected VoIP jitter data traces.

Fig. 8 shows the components behaviors of a VoIP jitter data trace that belong to the data sets with SRD. It is observed that the variance of the components of this time series is modeled by a straight line; therefore, the time series exhibits monofractal behavior.

Fig. 9 shows the components behaviors of a VoIP jitter data trace that belong to the data sets with LRD. It is observed that the variance of the components of this time series cannot be adequately modeled with a linear model, and the scaling behavior should be described with multiple scaling parameters (biscaling), therefore, this time series exhibits multifractal behavior.

These results show that VoIP jitter with SRD or LRD, exhibit monofractal or multifractal behavior, respectively. This phenomenon explains the behavior of the data traces with SRD and high degree of self-similarity (scale invariance), because the self-similarity is defined for a single scale parameter. On the other hand, the data traces with LRD, exhibit weak self-similarity because have associated non-uniform local variability (multifractal behavior).

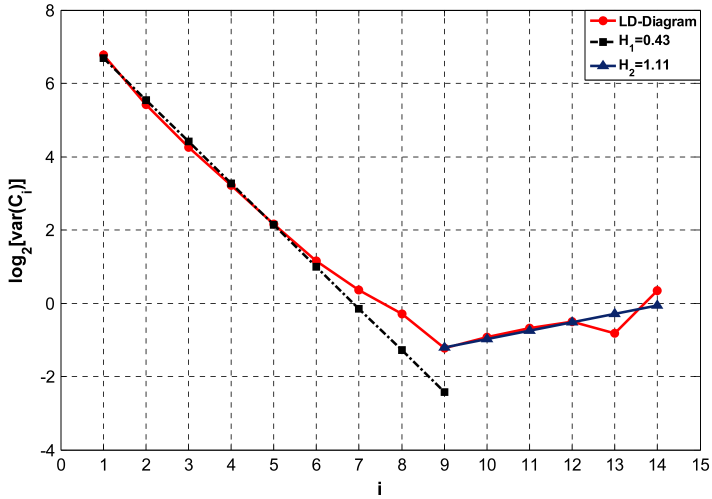


Fig. 9. Components behavior of VoIP jitter data traces: multifractal behavior.

7. Packet loss modeling: A power law model

7.1. Packet loss model framework

In this paper, a description of VoIP packet loss based on narrow and wide time windows was used. The packet loss behavior over a narrow time window is called here microscopic, and the packet loss behavior over wide time windows is called macroscopic [34]. Microscopic behavior refers to a packet loss period observed on a "time window" W_1 of the packet loss data trace; where, this packet loss period has a specific PLR_1.

On the other hand, macroscopic behavior refers to a set of microscopic periods ($W_1, W_2, W_3, \ldots, W_n$) that are observed on all packet loss data traces; where, each microscopic period has a particular $PLR(PLR_1, PLR_2, PLR_3, \ldots, PLR_n)$, as shown in Fig. 10(a). Fig. 10(a) illustrates different levels of PLR for each microscopic period. Therefore, the packet losses do not occur homogeneously. Instead, they are concentrated in some time intervals (i.e. the packet loss is bursty).

Microscopic behavior can be effectively modeled by a Markov chain with a low number of states [34]. On the other hand, macroscopic behavior can be effectively modeled by a Markov chain with a higher number of states [34]. Here, sub-states represent phases of a given microscopic behavior. Ideally, an n-state Markov chain is required to capture the macroscopic behavior. Especially for network planning, a trade off is desirable between very accurate modeling of data traces and a low number of model input parameters, in order to yield a model still usable for network planners with reasonable effort. Therefore, the 2-state Markov chain proposed in [34] for modeling the microscopic and macroscopic periods was used.

Fig. 10(b) and (c) show some packet loss patterns extracted from VoIP test calls. In Fig. 10(b), we can see that packet loss behavior is homogeneous, i.e., the packet loss pattern is represented by a microscopic period. On the other hand, in Fig. 10(c) the packet loss is non homogeneous, i.e., the packet loss pattern is represented by a concatenation of two microscopic periods.

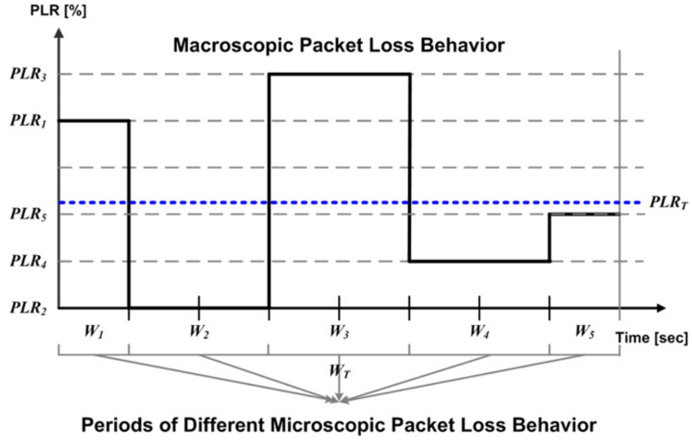
In order to simplify this packet loss description, the microscopic periods can be classified in two sets, one for low and one for high packet loss rates. The threshold used to delimit between a low or high level of packet loss, is a function of the perceived quality, good or poor respectively, according to the computed MOS values. In Fig. 10(b) and (c), the microscopic period with lower packet loss rate is delimited by the solid square, while the microscopic period with higher packet loss rate is delimited by the dashed square.

7.2. Methodology for simulating packet loss

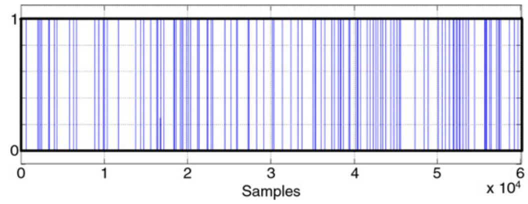
The current methodologies for simulating packet loss consist only of generating packet loss patterns by Markov chains of different orders [34]. Therefore, the studies based on this methodology are limited, because the impact of this parameter is analyzed separately from the others. A new methodology to simulate packet loss in two stages is proposed: first, by generating packet loss pattern by a 4-state Markov chain and secondly, by applying this packet loss pattern to a VoIP jitter data trace, i.e., the simulation of the effect of this packet loss pattern in the VoIP jitter by the relationship between packet loss and jitter, as shown in Eq. (21).

In order to simplify this methodology and achieve the trade-off between very accurate modeling of data traces and a low number of model input parameters, we consider the following:

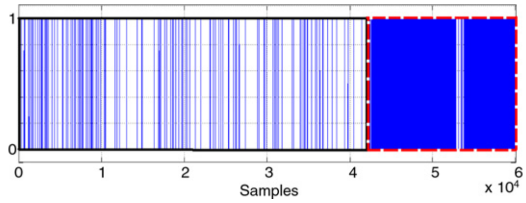
- The relationship between packet loss rate and jitter, summarized by Eq. (21).
- The description of VoIP packet loss based on microscopic and macroscopic behaviors.



(a) Microscopic and macroscopic behavior.



(b) Homogeneous PLR.



(c) Non homogeneous PLR.

Fig. 10. Packet loss descriptions from VoIP test calls.

Table 2
Algorithm for applying the packet loss patterns.

```
FOR n = 2 to N
    IF (P[n] = 1)
X[n] = X[n] + X[n − 1]
       END IF
     END FOR
       i = 1
FOR n = 2 to N
   IF (P[n] ≠ 1)
     X̂[i] = X[n − 1]
       i = i + 1
    END IF
   END FOR
```

- A set of real jitter and packet loss data traces, where the jitter data traces exhibit self-similar characteristics and the packet loss data traces are homogeneous, as shown in Fig. 10(a). Then, these packet loss patterns can be represented by only one microscopic period with a low level of packet loss rate PLR_0 (near zero).
- In order to incorporate "n" microscopic packet loss periods with high level of PLR, as it is shown in Fig. 10(b), we generated packet loss patterns by means of a 4-state Markov chain.

Let $X = \{X_t : t = 1, \ldots, N\}$ be a VoIP jitter data trace with a length of N, self-similar (H parameter $0 < H_0 < 0.5$), and with a low packet loss rate PLR_0. The packet loss patterns are generated by means of a 4-state Markov chain, and are represented as the binary sequences $P = \left\{P_t^\tau : t = 1, \ldots, N;\ \tau = 0, 1, 2, \ldots, T-1\right\}$, where $P_t^\tau = 1$ means a packet loss, $P_t^\tau = 0$ means a success or received packet, N is the length of the packet loss pattern and T is the number of packet loss patterns used. The relationship between jitter and packet loss from Eq. (21) is used to apply the packet loss patterns to X_t by the algorithm shown in Table 2.

As a result of using the above algorithms, the new time series $\hat{X}_t^\tau$ were obtained, for $t = 1, \ldots, N$ and $\tau = 0, 1, 2, \ldots, T-1$. For each $\hat{X}_t^\tau$ the PLR and H parameter were calculated, and the functions $f\ (PLR_\tau, H_\tau)$ were generated.

7.3. Simulation results

In this section, applying the methodology proposed above, simulation results are presented. The simulations are accomplished over the measured VoIP jitter data traces. Fig. 11 illustrates the relationships between the packet loss rate

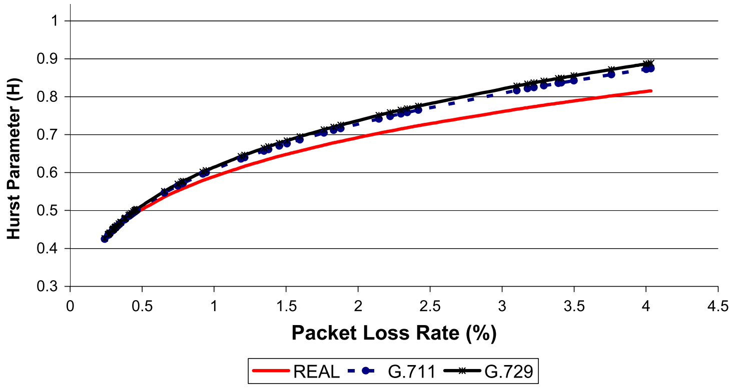


Fig. 11. Relationship between PLR and H parameter: $f(PLR_\tau, H_\tau)$ vs. $f_{REAL}(PLR_\varepsilon, H_\varepsilon)$.

Table 3
Fitted parameters of Fig. 11.

$f(PLR_\tau, H_\tau)$	$\hat{H}_0$	$\hat{a}$	$\hat{b}$	MSE
G.711	0.0428	0.5659	0.2760	0.001474
G.729	0.0430	0.5716	0.2805	0.002305
$f_{REAL}(PLR_\varepsilon, H_\varepsilon)$	0.0429	0.5471	0.2475	

and the Hurst parameter. The figure shows the empirical functions $f(PLR_\tau, H_\tau)$ that were obtained from simulation results and the function $f_{REAL}(PLR_\varepsilon, H_\varepsilon)$.

The functions $f(PLR_\tau, H_\tau)$, resulted from applying "T" packet loss patterns to representative VoIP jitter data traces X_t. In these functions, each point represents the PLR_τ and H_τ of a particular new time series $\hat{X}_t^\tau$. The function $f_{REAL}(PLR_\varepsilon, H_\varepsilon)$ is generated by "E" jitter data traces. In this function each point represents the PLR_ε and H_ε of a particular jitter data trace X_t^ε, where $t = 1, \ldots, N$, $\varepsilon = 1, 2, \ldots, E$, and "$E$" is the number of representative jitter data traces used.

The respective differences between the functions corresponding to simulation results $f(PLR_\tau, H_\tau)$ and the function $f_{REAL}(PLR_\varepsilon, H_\varepsilon)$, were quantified in terms of MSE.

Table 3 shows the fitted parameters and MSE between $f(PLR_\tau, H_\tau)$ and $f_{REAL}(PLR_\varepsilon, H_\varepsilon)$. The results are summarized in Fig. 11 and Table 3. As a result of this analysis, we conclude that the presented methodology for simulating packet loss, achieved the trade-off between a very accurate modeling and a low number of model input parameters. Based on these results, it is proposed that the relationships found between packet loss rate and Hurst parameter can be modeled by a power-law function, as shown in section the 7.4.

7.4. Proposed model based on the obtained results

From the simulations that have been carried out, it was found that the relationship between the H parameter and the PLR can be modeled by a power-law function, characterized by three fitted parameters, as follows:

$$H_M = \hat{H}_0 + \hat{a}\,(PLR)^{\hat{b}} \tag{28}$$

where H_M is the H parameter of the model found; $\hat{H}_0$, $\hat{a}$, and $\hat{b}$ are the fitted parameters; $\hat{H}_0$ is the H parameter when $PLR = 0$. The fitted parameters are estimated by linear regression. The strategy to find, $\hat{H}_0$, $\hat{a}$, and $\hat{b}$ is such that it minimizes the mean squared error, i.e. MSE $= \int_r (\hat{H}_0^+ \hat{a} r^{\hat{b}} - H_\tau)^2 dr$, and the validity of the proposed model corresponds to those ranges of $r = PLR$ (e.g. 0%–4%).

8. Conclusions and future research direction

In this paper, the jitter and packet loss behavior of VoIP traffic have been analyzed which could be useful both for today's networks and future networks supporting VoIP/mVoIP technologies. As a result of our findings and thorough analyses, a detailed characterization and accurate modeling of the main QoS parameters was provided. Firstly, we proposed that VoIP jitter can be properly modeled by means of self-similar and multifractal models. Secondly, a methodology for simulating packet loss on VoIP Jitter was presented. This methodology is based on:

- The relationship between PLR and jitter, summarized by Eq. (21).
- The packet loss description of VoIP based on microscopic packet loss behavior (modeled by a 2-state Markov chain) and macroscopic packet loss behavior (modeled by a 4-state Markov chain).
- The self-similar behaviors of VoIP jitter.

Finally, a relationship between the Hurst parameter and the Packet Loss Rate was found, where this relationship can be modeled by means of a power-law function with three fitted parameters, summarized by Eq. (28). Simulation results show the effectiveness of our models in terms of MSE.

The proposed models can be used in future to: (i) design a de-jitter buffer, (ii) to implement a synthetic generator of VoIP jitter data traces, where the synthetic jitter data traces can be used as test vectors to carry out the performance evaluation of a de-jitter buffer of VoIP system, and (iii) design effective schemes for packet loss recovery. As a whole, we hope that side by side benefitting the VoIP communications modeling of present and future networks, our work would serve as a rich reference paper in this area.

References

[1] D. Geer, The future of mobile VoIP in the enterprise, IEEE Computer 42 (6) (2009) 15–18.
[2] S.K. Chui, O.-C. Yue, W.C. Lau, Impact of handoff control messages on VoIP over wireless LAN system capacity, in: Proceedings of the 14th European Wireless Conference, EW, 22–25 June, 2008, pp. 1–5.
[3] A. Nogueira, P. Salvador, R. Valadass, A. Pacheco, Modeling network traffic with multifractal behavior, Telecommunication Systems 24 (2–4) (2003) 339–362.
[4] O.I. Sheluhin, S.M. Smolskiy, A.V. Osin, Self-Similar Processes in Telecommunications, John Wiley & Sons, Ltd., 2007 (Chapters 1 and 3).
[5] O. Hohlfeld, Stochastic packet loss model to evaluate QoE impairments, PIK Journal 1 (2009) 53–56.
[6] G. Haßlinger, O. Hohlfeld, The Gilbert–Elliott model for packet loss in real time services on the Internet, in: Proceedings of the 14th GI/ITG Conference on Measuring, Modelling and Evaluation of Computer and Communication Systems, MMB, 31 March–2 April 2008, pp. 269–286.
[7] T. Hussaina, S. Habib, Assessing and redesigning enterprise networks through NS-2 to support VoIP, Procedia Computer Science 5 (2011) 742–748.
[8] S. Karapantazis, F.-N. Pavlidou, VoIP: a comprehensive survey on a promising technology, Computer Networks 53 (12) (2009) 2050–2090.
[9] J. Klink, B. Miazga, P. Sieradzki, M. Marcisz, On VoIP quality assessment, in: Proceedings of the Poznañ Telecommunication Workshop, PWT, 11 December, 2008, pp. 1–4.
[10] K.M. Rezaul, V. Grout, A survey of performance evaluation and control for self-similar network traffic, in: Proceedings of the Second International Conference on Internet Technologies and Applications, ITA, 4–7 September, 2007, pp. 514–524.
[11] G. Samorodnitsky, Long range dependence, Foundations and Trends in Stochastic Systems 1 (3) (2007) 163–257.
[12] G. Zhang, G. Xie, J. Yang, D. Zhang, Self-similar characteristic of traffic in current metro area network, in: Proceedings of the 15th IEEE Workshop on Local and Metropolitan Area Networks, 10–13 June, 2007, pp. 176–181.
[13] D. Veitch, N. Hohn, P. Abry, Multifractality in TCP/IP traffic: the case against, Computer Networks 48 (3) (2005) 293–313.
[14] R. Riedi, J. Véhel, Multifractal properties of TCP traffic: a numerical study, Technical Report No. 3129, INRIA Rocquencourt, France, 1997. www.dsp.rice.edu/~riedi.
[15] J. Gao, Y. Cao, W.-W. Tung, J. Hu, Multiscale Analysis of Complex Time Series: Integration of Chaos and Random Fractal Theory, and Beyond, John Wiley & Sons, Inc., 2007 (Chapters 8–10).
[16] O.I. Sheluhin, A.A. Atayero, A.V. Garmashev, Detection of teletraffic anomalies using multifractal analysis, International Journal of Advancements in Computing Technology 3 (4) (2011) 174–182.
[17] D. Veitch, P. Abry, A wavelet based joint estimator for the parameters of LRD, IEEE Transactions on Information Theory 45 (3) (1999) 878–897.
[18] K. Tarnay, G. Adamis, T. Dulai, Advanced Communication Protocol Technologies: Solutions, Methods, and Applications, IGI Global, 2011 (Chapter 17).
[19] H. Toral, D. Torres, L. Estrada, Simulation and modeling of packet loss on VoIP traffic: a power-law model, WSEAS Transactions on Communications 8 (10) (2009) 1053–1063.
[20] L. Estrada, D. Torres, H. Toral, Variance error for finite-length self-similar time series, in: Proceedings of the 7th International Conference on Computing, Communications and Control Technologies, CCCT, in the Context of the 2nd International Multi-Conference on Engineering and Technological Innovation, IMETI, 2009, pp. 193–198.
[21] M.V. Wickerhauser, Adapted Wavelet Analysis from Theory to Software, IEEE Press, 1994, pp. 213–235.
[22] Advanced information CTS (Centro de Tecnología de Semiconductores) property, Alliance FXO/FXS/E1 VoIP System. www.cts-design.com.
[23] Wireshark, A network protocol analyzer. http://www.wireshark.org/.
[24] H. Toral, QoS parameters modeling of self-similar VoIP traffic and an improvement to the E model, Ph.D. Thesis, Electrical Engineering, Telecommunication Section, CINVESTAV, Guadalajara, Jalisco, Mexico, 2010.
[25] RFC 3550, RTP: a transport protocol for real-time applications, Internet Engineering Task Force, 2003.
[26] Al-Sakib Khan Pathan, Mukaddim Pathan, Hae Young Lee, Advancements in Distributed Computing and Internet Technologies: Trends and Issues, IGI Global, 2011 (Chapter 1).
[27] Hae-Duck J. Jeong, Jong-Suk R. Lee, K. Pawlikowski, Comparison of various estimators in simulated FGN, Simulation Modelling Practice and Theory 15 (9) (2007) 1173–1191.
[28] J. Mielniczuk, P. Wojdyllo, Estimation of Hurst exponent revisited, Computational Statistics & Data Analysis 51 (9) (2007) 4510–4525.
[29] P. Shang, Y. Lu, S. Kamae, Detecting long-range correlations of traffic time series with multifractal detrended fluctuation analysis, Chaos, Solitons and Fractals 36 (2008) 82–90.
[30] G. Horn, A. Kvalbein, J. Blomskold, E. Nilsen, An empirical comparison of generators for self similar simulated traffic, Performance Evaluation 64 (2007) 162–190.
[31] R.G. Clegg, A practical guide to measuring the Hurst parameter, in: Proceedings of the 21st UK Performance Engineering Workshop, School of Computing Science, Technical Report Series, 2006, pp. 43–55.
[32] F.H.P. Fitzek, M. Reisslein, MPEG-4 and H. 263 video traces for network performance evaluation, IEEE Network 15 (6) (2001) 40–54.
[33] S. Stoev, M.S. Taqqu, C. Park, J.S. Marron, On the wavelet spectrum diagnostic for Hurst parameter estimation in the analysis of Internet traffic, Computer Networks 48 (3) (2005) 423–445.
[34] A. Raake, Short- and long-term packet loss behavior: towards speech quality prediction for arbitrary loss distributions, IEEE Transactions on Audio Speech and Language Processing 14 (6) (2006) 1957–1968.