Source Models for Speech Traffic Revisited

Michael Menth, *Member, IEEE*, Andreas Binzenhöfer, and Stefan Mühleck

***Abstract*—In this paper, we analyze packet traces of widely used voice codecs and present analytical source models which describe their output by stochastic processes. Both the G.711 and the G.729.1 codec yield periodic packet streams with a fixed packet size, the G.723.1 as well as the iLBC codec use silence detection leading to an on/off process, and the GSM AMR and the iSAC codec produce periodic packet streams with variable packet sizes. We apply all codecs to a large set of typical speech samples and analyze the output of the codecs statistically. Based on these evaluations we provide quantitative models using standard and modified on/off processes as well as memory Markov chains. Our models are simple and easy to use. They are in good accordance with the original traces as they capture not only the complementary cumulative distribution function (CCDF) of the on/off phase durations and the packet sizes, but also the autocorrelation function (ACF) of consecutive packet sizes as well as the queueing properties of the original traces. In contrast, voice traffic models used in most of today's simulations or analytical studies fail to reproduce the ACF and the queueing properties of original traces. This possibly leads to underestimation of performance measures like the waiting time or loss probabilities. The models proposed in this paper do not suffer from this shortcoming and present an attractive alternative for use in future performance studies.**

***Index Terms*—Correlation, queueing behavior, traffic models, voice codecs.**

I. Introduction

SPEECH is usually sampled at a frequency of 8 kHz, where each probe is encoded by one byte resulting in a bit rate of 64 kbit/s. This information has traditionally been transmitted continuously over circuits in the public switched telephone network. In packet-switched networks several probes are collected from intervals of fixed length T, put into a packet equipped with header information, and transmitted. This saves transmission overhead for individual probes. However, the packetization delay contributes to the end-to-end delay seen by the application and, therefore, it cannot be chosen arbitrarily large. Typical values for T are 20 or 30 ms depending on the voice codec.

Due to the high redundancy in human speech, voice data can be well compressed whereas different voice codecs exploit this fact to a different degree. The G.711 and the G.729.1 codec simply encode speech into packets of fixed size. The G.723.1 and the iLBC codec detect silence phases during which they suppress the generation of data leading to an on/off process on the packet level. Although this creates streams of variable bit rate from a macroscopic viewpoint, we refer to them by on/off flows as they either periodically sent packets of fixed size or no packets at all. Finally, the GSM AMR and the iSAC codec take additional advantage of speech characteristics and compress the sampled data into packets of different size. We call their output variable bit rate (VBR) streams.

Source models for speech traffic are an old and well-studied topic as they are often needed by engineers. While some simulation studies in the telecommunication area use original data traces of voice traffic, artificial models for the generation of compressed digitized speech are preferred to avoid the repetition of limited traces in long simulation runs. Analytical studies even depend on source models as they require a mathematical description of the traffic. A survey of current literature reveals that a large number of simulative or analytical studies [1]–[9] use an on/off model with exponentially distributed on/off phases with mean durations of 352 ms and 650 ms. Most of them refer to [10] which cites "private work" [11]. We tried to track the latter, but without success. The work of Brady [12]–[14] seems to be the next popular source which reports mean durations for on/off phases of about 1.3 s and 1.7 s, respectively. Also wide-spread simulation tools like OPNET or NS-2 model the output of G.711 with the same values 352 ms and 650 ms, and 1.0 s and 1.35 s [15], respectively. However, they all ignore that these data were measured in the 1960s and 1970s based directly on human speech and not on compressed digitized voice. The old models fail to capture important characteristics of packet traces created by compressed digitized voice. In particular, they underestimate the autocorrelation function (ACF) of consecutive packet sizes by far. As a consequence, simulative or analytical performance studies using these old models run the risk of underestimating waiting times, loss probabilities, and other performance measures.

This paper presents simple stochastic models for different voice codecs which are currently used in a wide variety of applications. They produce on/off or VBR traffic. To the best of our knowledge, this is the first paper presenting a source model for VBR voice at all. We use a large set of representative speech samples, encode them by the codecs under study, analyze their output, and parameterize the different models. We show that they well capture both the complementary cumulative distribution function (CCDF) of the on/off phase durations of the original traces and the packet sizes. Furthermore, the models also reproduce the ACF of consecutive packet sizes as well as the queueing behavior of the original packet traces when several voice streams are fed to a single server queue. Since the classical voice traffic models fail in this point, we recommend to use the new models in future performance studies.

The paper is structured as follows. Section II briefly reviews related work. In Section III we present measurement results and derive quantitative stochastic models for typical vocoder output. We validate them by comparing the statistical properties and the queueing behavior of synthetic traces to those of the original traces. Section IV summarizes our work.

Manuscript received November 23, 2007; revised June 09, 2008; approved by IEEE/ACM Transactions on Networking Editor M. Hofmann.

The authors are with the Institute of Computer Science, University of Würzburg, Informatik 3, 97074 Wuerzburg, Germany (e-mail: menth@informatik.uni-wuerzburg.de).

Digital Object Identifier 10.1109/TNET.2008.2006222

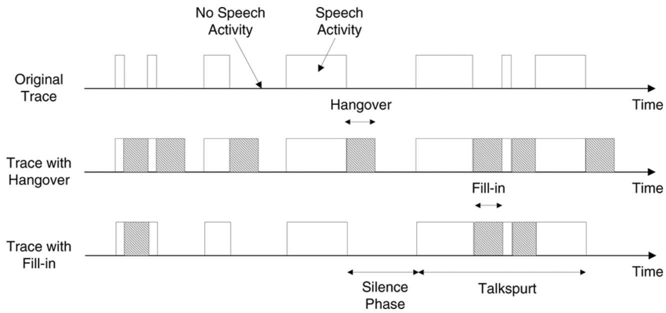


Fig. 1. Illustration of the hangover and fill-in techniques. With hangover all talkspurts are extended by t_h. With fill-in all silence periods that have a duration of less than t_f are filled in.

II. Related Work

A good introduction to traffic models in general can be found in [16]. The paper describes the basic ability of different models to reproduce characteristics of the original process like a non-exponentially decaying autocorrelation. However, model parameters are not given.

Deriving source models for CBR speech traffic is a rather trivial task as CBR sources send packets of fixed size in regular intervals. In contrast, codecs using silence detection generate typical on/off packet processes which have often been characterized in literature. Silence or voice activity detectors (SD, VAD) suffer from "front-" and "end-clipping", e.g., the coder might clip the end of a sentence when the speaker's voice goes down. Hence SDs and VADs may implement hangover or fill-in techniques to avoid "clipping" [17]. When using hangover, a coder switches from the on-state to the off-state with a certain delay t_h. The fill-in technique bridges a short gap between two intervals of voice activity if the gap is not longer than t_f. Fig. 1 compares the effect of both techniques applied to some sample input. Both techniques prolong the duration of the on-phases. For $t_h = t_f$, hangover increases the average on-phase length more than fill-in, but fill-in introduces an extra delay of at least t_f. Thus, the output of vocoders depends on their parameterization. Most papers characterize the duration of uninterrupted activity or silence. Older papers measure analog voice while newer papers measure the generation or suppression of speech packets. Most of them study the duration of the on/off phases depending on the VAD sensitivity, the hangover, and the fill-in. They use an exponential or geometric approximation of the duration of the on/off phases, but point out that this simple model is not a good fit.

Early work [18] introduces the notion of talkspurts which are defined as the duration of the speech of one party that may contain pauses. In later work, a talkspurt describes a contiguous interval of recognizable speech activity, i.e., several talkspurts of a single party may follow each other. The work of Brady [12], [13] has reported different mean values for the duration of on/off phases depending on the sensitivity of the VAD: 1.31 s and 1.70 s for −45 dBm, 1.3 s and 1.72 s for −40 dBm, and 0.9 s and 1.66 s for −35 dBm. Altogether 137.4 min of two-way conversations were investigated, i.e., 274.8 min of speech. Parameters for a discrete-time Markov chain with two states are given in the paper to model the resulting output, but Brady also states that this is not a good fit. [14] presents an exponential model for generating on/off speech patterns in two-way conversations and reports a duration for the on/off phases of 1.2 s and 1.8 s. These parameters are used, e.g., in [19]. A more complex model for two-way conversations is presented in [20].

Most simulative and analytic studies use the values 352 ms and 650 ms for the duration of the on/off phases. They are reported in a paper of Sriram and Whitt [10] who cite the "private work" of May and Zebo [11]. Interestingly, many papers [1]–[9] use these values and some of them wrongly refer to some of Brady's works instead of citing [10] or [11]. In this paper, we refer to this traditional model by *Geom-352/650*.

The ITU P.59 [21] recommendation specifies an artificial on/off model to generate human speech. The durations of the talkspurts and silence intervals are 227 ms and 596 ms without hangover and 1.004 s and 1.587 s with hangover. Jiang and Schulzrinne investigated the G.729 Annex B VAD and the NeVoT SD [17] which use dynamic sensitivity thresholds to detect talkspurts and silence intervals. In addition, they discuss the impact of hangovers. They reported mean spurts and gaps of 293 ms and 306 ms for G.729B and 267 ms and 272 ms for NeVoT SD. They compared the queueing behavior of the empirical data with the one of an exponential model and showed that this is a bad approximation. The dependence of the talkspurt duration on the hangover interval has also been studied in [22], [23].

In [24] Deng *et al.* observe that the assumption of exponentially distributed talkspurts and silence intervals is not a good approximation. They tested packet voice from early VoIP tools like vat, NeVoT, Maven and recognized silence phases only if they are larger than 3 packets. As a consequence, Deng reports mean on/off phase durations for conversational speech of 7.24 s and 5.69 s which is already by an order of magnitude larger than the most widely used traditional *Geom-352/650* model. In [25] the distribution of the on/off phases of the codec traces is approximated by a Weibull distribution. Only on/off phases larger than 100 ms were recognized. The work reports mean talkspurt and silence durations of 1.58 s and 0.87 s. In [26], the codecs G.723.1, G.729B, and GSM FR were investigated. Their call level analysis provides a mean holding time of 114 s. Their packet level analysis reports mean durations of 2.28 s and 1.48 s, 2.37 s and 1.56 s, and 2.50 s and 1.55 s for the duration of the on/off phases for the three codecs. They propose to model their duration by a generalized Pareto distribution and found a long range dependency in the rate of the superposition of several voice calls.

None of the above models considers the autocorrelations of the output of the codecs even though they are known to have a significant influence on different performance measures like the queueing behavior [27]. Li and Mark study the queue length distribution of multiplexed sources in [28]. Each source is modelled as a discrete-time on/off process with geometrically distributed on/off phases. The large impact of positive autocorrelations on the waiting time in queueing systems is mentioned but not expressed in terms of a quantitative measure.

In our work, we use a different interpretation of on/off phases which is similar to the one of [18]. On/off phases are recognized as such only if they are sufficiently long, otherwise we interpret them just as short breaks or noise within on- or off-phases. As a consequence, we report mean durations of the on/off phases in the order of 11 s which is an order of magnitude larger than those reported in the papers above. Thus, simple exponential or geometric models can further be applied, but analytical or simulative studies should use appropriate mean values for the duration of the on/off phases. We will further show that, in contrast to previous work, our more elaborated models not only provide a good fit for the distribution function of the phase durations, but are also able to capture the autocorrelation of consecutive packet sizes and at the same time yield a good approximation for the queueing properties of voice traffic.

Although many papers model VBR video traffic [29]–[31], we are not aware of any source models for VBR voice codecs in the literature. A preliminary short version of this paper has been presented as a poster [32].

III. Source Models for Speech Traffic

In this section, we consider two representatives of each of the three different vocoder types: constant bit rate (CBR) codecs, codecs with silence detection, and variable bit rate (VBR) codecs. We apply each codec to a large set of typical telephone conversations (3.5 h = 7 h speech) from [33], a publicly available database of English speech sources which were specifically designed to be used in research and speech technology. We then analyze the original packet traces and provide quantitative models describing the codec output. To validate the accordance of the stochastic models and the original traces, we compare the cumulative distribution function (CDF) of the packet sizes, the complementary CDF (CCDF) of the on/off phase durations, the autocorrelation function (ACF) of consecutive packet sizes, and the CCDF of the packet waiting time when several voice streams are fed to a single server queue.

A. Voice Codecs With Constant Bit Rate

CBR codecs send a bit stream of constant rate which is independent of the voice input. This leads to a very simple behavior on the packet level which can be seen in Fig. 2. The ITU G.711 [34] codec is mainly applied in digital telephony and uses pulse code modulation (PCM) sampled at a rate of 8 kHz and 8 bits per sample which results in a 64 kbit/s stream. The algorithmic complexity is very low and due to the relatively high bandwidth usage, the voice quality is very good and often used as a reference. The ITU G.729.1 standard [35] was also designed for voice communication and adds wideband functionality to the G.729 standard by offering different bit rates from 14 to 32 kbit/s in steps of 2 kbit/s. To analyze the behavior of the codecs in practice, we measured the output stream of the G.711 and the G.729.1 codec. Voice packets are usually transmitted using UDP over IPv4 entailing a header overhead of 8 and 20 bytes, respectively. However, it is also possible to use additional or alternative headers like RTP (at least 12 bytes) or IPv6 (40 bytes). To make our results independent of the network layer, we concentrate on the plain output of the codecs disregarding any headers.

We measured the G.711 codec using CounterPath's X-Lite [36], a freely available SIP based softphone which produces a main stream of 68.8 kbit/s. The implementation of the codec sends its control information separately as well as piggybacked

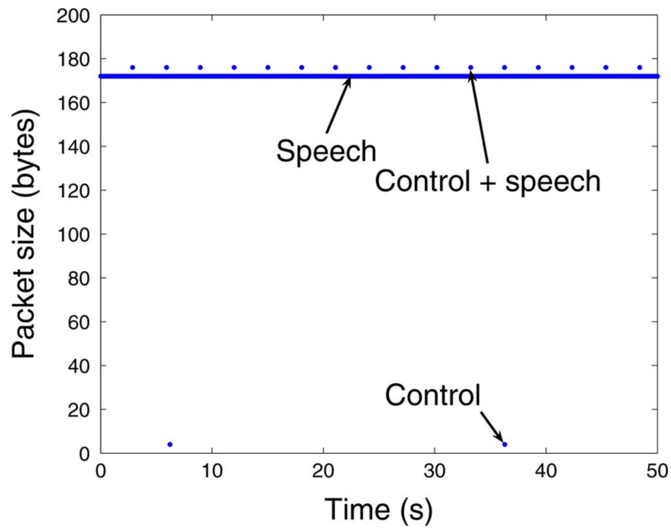


Fig. 2. Packet trace of G.711 voice coder. The coders output is highly regular and independent of the voice input.

TABLE I
Packet Types of the G.711 and the G.729.1 Codec.

Codec	Type	Packet size	Period
G.711	Control	4 bytes	30s
	Speech	172 bytes	20ms
	Speech + control	176 bytes	3s
G.729.1	Control	5 bytes	1s
	Speech	38 bytes	20ms

on regular data packets. The trace of the G.729.1 codec was obtained using SkypeOut [37] to call a regular landline user from Skype. The codec strictly differentiates between control information and actual speech data. Thus, both codecs send periodic control information in addition to their main audio stream. Table I gives a detailed description of the packet sizes and the periods at which they are sent. Due to the simplicity of the codecs in this category and the fact that their output rates are independent of the input, Table I suffices to easily generate synthetic streams for simulations or analytical studies.

B. Voice Codecs With Silence Detection

Voice codecs with silence detection are able to detect voice activity in terms of "speech on" or "speech off" and transmit packets of fixed size only while the user is talking. Thus, the output on the network layer consists of contiguous talkspurts and silence intervals, the so-called on- and off-phases. Two prominent examples for such codecs are the G.723.1 and the iLBC vocoder. The G.723.1 codec is an ITU-T standard since 1995 [38]. It was specially designed for voice encoding at low bandwidth and is mostly used in VoIP applications, e.g., in Netmeeting or Picophone. G.723.1 can operate in two different modes generating 6.4 kbit/s with 24 byte chunks or 5.3 kbit/s with 20 byte chunks every 30 ms. We generate packet traces using the Picophone software [39] which relies on the reference implementations of Microsoft. G.723.1 produces a main audio stream of packets with fixed size and sends additional control information of 1 byte every 3 s (cf. Fig. 3).

The Internet Low Bit Rate Codec (iLBC) [40] developed by Global IP Sound (GIPS) is suitable for robust voice communication over IP. It is designed for narrow band speech and

TABLE II
STATISTICS ABOUT ON/OFF PHASE DURATIONS BASED ON W0-MEASUREMENTS INCLUDING MLE-BASED PARAMETERS FOR THE CORRESPONDING *NBIN* AND THE *GEOM* APPROXIMATION IN PACKETS.

Codec	G.723.1		iLBC	
Phase	on	off	on	off
$E[D^{real}]$	1.304 s	1.480 s	3.113 s	3.279 s
$c_{var}[D^{real}]$	1.7938	2.9858	0.7697	1.9152
r^{NBin}	0.31302	0.11243	1.71571	0.27330
p^{NBin}	$7.14891 \cdot 10^{-3}$	$2.27372 \cdot 10^{-3}$	$1.62657 \cdot 10^{-2}$	$2.49416 \cdot 10^{-3}$
p^{Geom}	$2.24859 \cdot 10^{-2}$	$1.98671 \cdot 10^{-2}$	$9.54523 \cdot 10^{-3}$	$9.06599 \cdot 10^{-3}$

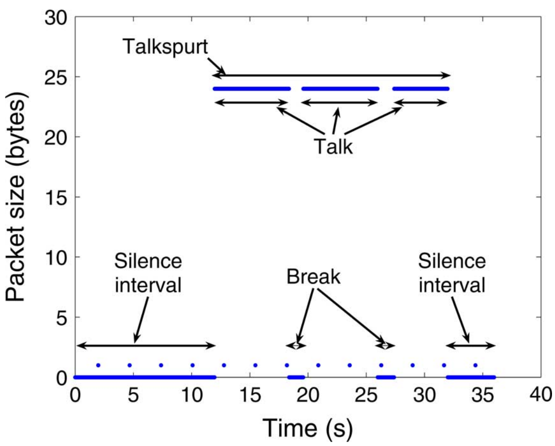
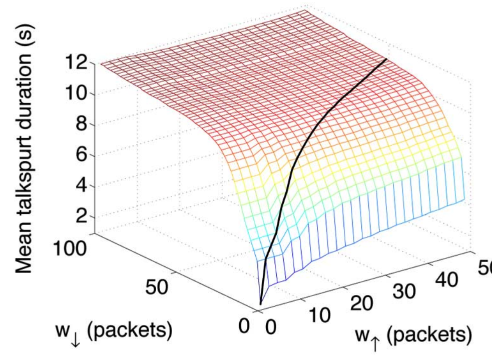


Fig. 3. G.723.1 outputs an audio stream and control information. The audio stream consists of silence intervals and main talkspurts that are interrupted by short breaks and noise.

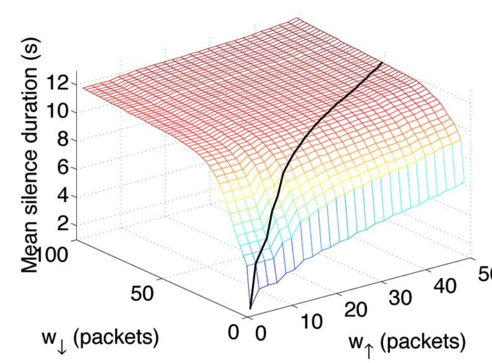
results in a payload bit rate of 13.33 kbit/s for 30 ms frames and 15.20 kbit/s for 20 ms frames. The codec enables graceful speech quality degradation in the case of lost frames, which occurs in connection with lost or delayed IP packets. We used the implementation of XLite sending 62 bytes every 30 ms resulting in a bit rate of 16.53 kbit/s which is slightly larger than the one indicated in the standard. Thus, some control information seems to be piggybacked.

1) Traffic Model for Codecs With Silence Detection: Fig. 3 shows a typical packet trace of the G.723.1 codec. Hardly any audio packets are transmitted during a silence interval. The talkspurts, however, are interrupted by short breaks which arise from short pauses a speaker makes while talking. Obviously, the codec detects these pauses and temporarily stops the transmission of voice packets. As a consequence actual talk phases are separated by breaks which are relatively short compared to the silence intervals separating consecutive talkspurts. This grouping of talk phases leads to an autocorrelation on packet level which must be preserved by a good traffic model since long range dependence (LRD) is decisive for the model's queueing properties [41]. Therefore we propose to measure the length of whole talkspurts instead of strictly contiguous on/off phases. Due to the noisy structure, the automatic detection of the beginning and end of major talkspurts is difficult. We discuss three different approaches for their recognition.

- **(W0)** We take contiguous on- and off-phases as observed in the original trace such that major talkspurts are cut in pieces. This method has been applied in previous work.
- **(W1)** We require that on- and off-phases start with at least w consecutively generated or suppressed packets, which can easily be controlled by a single moving window.



(a) Talkspurts.



(b) Silence intervals.

Fig. 4. Impact of the window parameters $w\downarrow$ and $w\uparrow$ on the measured mean duration of the on/off phases for the G.723.1 codec.

- **(W2)** We require that on-phases start with at least $w\uparrow$ consecutively generated packets and that off-phases start with at least $w\downarrow$ consecutively suppressed packets, which can be controlled by two different moving windows.

Figs. 4(a) and (b) show the measured mean duration of the on-phases obtained using the W2-approach depending on the values of $w\downarrow$ and $w\uparrow$. The mean talkspurt durations $E[D_{\text{on}}^{\text{real}}]$ and the mean silence duration $E[D_{\text{off}}^{\text{real}}]$ increase with both window sizes, but we recognize a stable plateau for $w\downarrow \geq 50$ and $w\uparrow \geq 15$. This means that neglecting short breaks within on-phases of less than 1.5 s and short noise within off-phases of less than 450 ms leads to relatively stable measured mean values for the durations of on/off phases. We consider the window parameters $w\downarrow = 50$ and $w\uparrow = 15$ useful and use them in the following as default setting for the W2-approach.

The measured mean values for $w\downarrow = 1$ and $w\uparrow = 1$ in both figures correspond to the W0-approach and the solid lines correspond to measured mean values for the W1-approach. The measured mean value for the W0-approach underestimates the

TABLE III
STATISTICS ABOUT ON/OFF PHASE DURATIONS *BASED ON W2-MEASUREMENTS* ($w\downarrow = 50$ AND $w\uparrow = 15$)
INCLUDING MLE-BASED PARAMETERS FOR THE CORRESPONDING *NBIN* AND THE *GEOM* APPROXIMATION IN PACKETS.

Codec	G.723.1		iLBC	
Phase	on	off	on	off
$E[D^{real}]$	11.54 s	11.98 s	11.23 s	11.31 s
$c_{var}[D^{real}]$	0.61003	0.60261	0.58344	0.61887
r^{NBin}	2.70609	2.77289	2.96094	2.62917
p^{NBin}	$6.98575 \cdot 10^{-3}$	$6.89591 \cdot 10^{-3}$	$7.84782 \cdot 10^{-3}$	$6.92564 \cdot 10^{-3}$
p^{Geom}	$2.59291 \cdot 10^{-3}$	$2.49792 \cdot 10^{-3}$	$2.66430 \cdot 10^{-3}$	$2.64550 \cdot 10^{-3}$

length of "almost contiguous" on-phases by an order of magnitude. For a small window w, the W1-approach misinterprets short breaks for silence intervals, for large windows w, it often fails to recognize the first talk phases of a talkspurt. Finally, the W2-approach offers the flexibility to ignore short breaks in a talkspurt using a relatively large value for $w\downarrow$ and to simultaneously detect the beginning of all talkspurts using a small value for $w\uparrow$.

A similar behavior is observed for the iLBC codec for the same values of $w\downarrow$ and $w\uparrow$. The statistical properties of the on/off phase durations are given for the G.723.1 and the iLBC codec in Tables II and III for the W0- and the W2-approach.

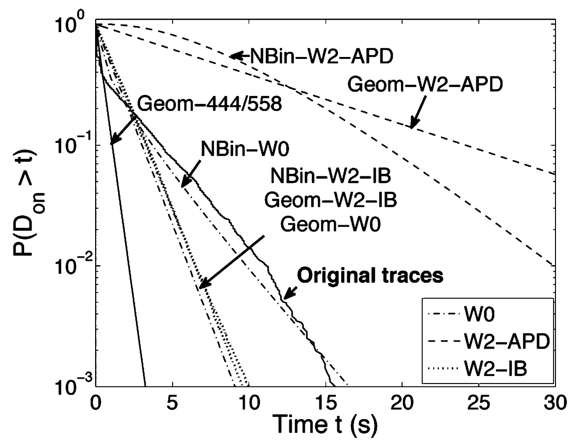
The voice activity factor (VAF) α is the fraction of the number of generated packets and the number of generated and suppressed packets. For the G.723.1 we get $\alpha = 0.443$ from our measurements and for the iLBC we get $\alpha = 0.488$. We approximate the distribution of the length of the talkspurts and the silence intervals in packets with the geometric distribution (*Geom*), i.e., $P(X^{Geom} = k) = p^{Geom} \cdot (1-p^{Geom})^k$, and the negative binomial distribution (*NBin*), i.e.,

$$P(X^{NBin} = k) = \frac{\Gamma(r^{NBin} + k)}{k! \cdot \Gamma(r^{NBin})} (p^{NBin})^{r^{NBin}} (1-p^{NBin})^k$$

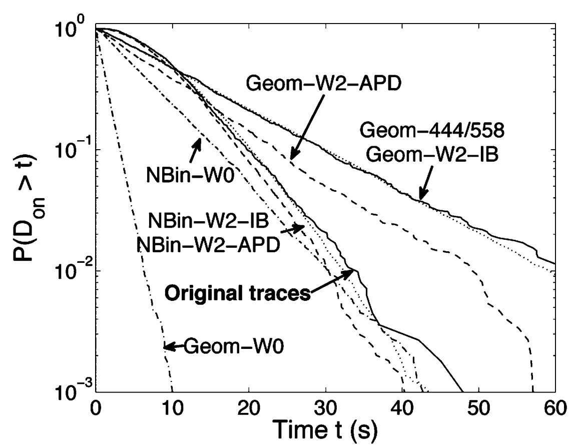
where Γ is the Gamma function. Modelling truly contiguous on- and off-phase durations based on the measured mean values $E[D_{\text{on}}^{\text{real}}]$ and $E[D_{\text{off}}^{\text{real}}]$ given in Table III neglects the many breaks within a talkspurt resulting in an overestimated VAF. We propose two different approaches to tackle this problem:

- **(APD)** Adapt phase durations: we adjust the mean duration of the on/off phases measured by the W2-approach in such a way that the original VAF is met, i.e., we use $E[D_{\text{on}}^{\text{APD}}] = \alpha \cdot (E[D_{\text{on}}^{\text{real}}] + E[D_{\text{off}}^{\text{real}}])$ and $E[D_{\text{off}}^{\text{APD}}] = (1-\alpha) \cdot (E[D_{\text{on}}^{\text{real}}] + E[D_{\text{off}}^{\text{real}}])$ to model the durations of contiguous on/off phases.
- **(IB)** Introduce breaks: we use $E[D_{\text{on}}^{\text{real}}]$ and $E[D_{\text{off}}^{\text{real}}]$ of Table III to model the length of the major talkspurts and silence intervals and generate talk and break phases within the talkspurts as observed in Fig. 3 by geometric distributions. To that end, we measure the average durations of the talk and break phases observed within talkspurts and obtain $E[D_{\text{talk}}^{\text{real}}] = 1.464$ s and $E[D_{\text{break}}^{\text{real}}] = 0.102$ s for G.723.1 and $E[D_{\text{talk}}^{\text{real}}] = 3.128$ s and $E[D_{\text{break}}^{\text{real}}] = 0.103$ s for iLBC.

We denote on/off phase durations generated by the geometric and negative-binomial distribution based on measurement results from the W0-approach by *{Geom, NBin}-W0*. If they are based on the measurement results from the W2-approach, we denote them by *{Geom, NBin}-W2-{APD, IB}* to indicate



(a) Measured by the W0-approach.



(b) Measured by the W2-approach.

Fig. 5. CCDFs of the on-phase durations for the original traces of the G.723.1 codec and different traffic models.

how the VAD is corrected. The parameters for the generation of the on- and off-phases (in packets) are derived by maximum likelihood estimators (MLE) and summarized in Table II for *{NBin, Geom}-W0* and in Table III for *{Geom, NBin}-W2-APD*. A fair comparison of the traditional model *Geom-352/650* and the G.723.1 output requires that both have the same VAD. Therefore, we adapt the average length of its on/off phases and get *Geom-444/558*. We then produce packet streams according to this model and measure them either with the W0-approach or W2-approach.

To compare the measurement results, Figs. 5(a) and (b) show the complementary cumulative distribution functions (CCDF) of the on-phase duration for the original traces and different models for the G.723.1 codec. The complementary CDF as well as a logarithmic y-axis are chosen so that the differences for large on-phases become better visible. Fig. 5(a) is obtained with W0-measurement while Fig. 5(b) is obtained with W2-measurement. Looking at Fig. 5(a) on the one hand, the traditional model *Geom-444/558* has significantly shorter on-phase durations and *{Geom, NBin}-W2-APD* have significantly longer on-phase durations compared to the original traces. The accordance of the curves for *{Geom, NBin}-W2-IB*, *Geom-W0* with the curve for the original traces is acceptable, but not good. Only the *NBin-W0* approach seems to be a good fit.

Looking at Fig. 5(b) on the other hand, the on-phase durations are clearly longer due to the W2-measurement method. The traditional *Geom-444/558* model surprisingly overestimates the on-phase durations ($E[D_{\text{on}}^{W2}] = 12.59$ s and $E[D_{\text{off}}^{W2}] = 2.98$ s) as off-phases are recognized only if they are longer than 450 ms. That is, using the W2-approach many silence phases are wrongly interpreted as talkspurt breaks. As a consequence, the VAF of the W2-measured trace is $\alpha = 0.808$ instead of $\alpha = 0.443$ for the W0-measured original traces. *Geom-W0* heavily underestimates the durations ($E[D_{\text{on}}^{W2}] = 1.39$ s and $E[D_{\text{off}}^{W2}] = 1.58$ s) and so does *NBin-W0* ($E[D_{\text{on}}^{W2}] = 7.05$ s and $E[D_{\text{off}}^{W2}] = 6.41$ s) although hardly visible in Fig. 5(b). The CCDFs of the *Geom-W2-{APD, IB}* do not well approximate the CCDF of the original traces, but *NBin-W2-{APD, IB}* lead to a fairly good match. Combining the results of the W0 and W2 measurement, *NBin-W2-IB* provides the best fit for the original traces on different time scales.

2) Autocorrelation of the Modified On/Off Models: The empirical autocorrelation function (ACF) for lag j can be calculated from m consecutive random variables (RV) X_i $(0 \leq i < m)$ by $r_m(j) = \frac{\hat{C}_m(j)}{S_m^2}$, where S_m^2 is the empirical variance and

$$\hat{C}_m(j) = \frac{1}{m-j} \cdot \sum_{0 \leq i < m-j} (X_i - \bar{X}) \cdot (X_{i+j} - \bar{X})$$

the empirical autocovariance of the m RVs. The values of $r_m(j)$ range between -1 and 1. If $r_m(j)$ is close to 1, RVs X_i and X_{i+j} have almost perfect correlation, if it is close to -1, they have almost perfect anti-correlation. If consecutive RVs are independent and identically distributed (iid), an ACF of $r_m(j) \approx 0$ can be expected for any lag $j > 0$.

We consider the ACFs of consecutive packet sizes which are either zero or the standard packet size in case of on/off streams. The mean on/off phase durations have a significant impact on the ACFs since the ACFs are a measure for the similarity of consecutive packet sizes with lag j. As a result, streams with long on/off phases have ACFs that are slowly decaying with lag j while streams with short on/off phases have ACFs that are rather quickly decaying with lag j. Therefore, we use the ACFs for further validation of the different models. Fig. 6 shows that the original traces reveal strong positive ACF values even for large lags. The ACF values for the traditional model *Geom-444/558* and *Geom-W0* are significantly lower than those of the original traces. The same holds for *NBin-W0* but to a minor degree. *{Geom, NBin}-W2-APD* clearly overestimate the ACF of the original traces, but *{Geom, NBin}-W2-IB* match them fairly well.

3) Queueing Behavior: To further compare the discussed models we compare their queueing properties. To that end, we

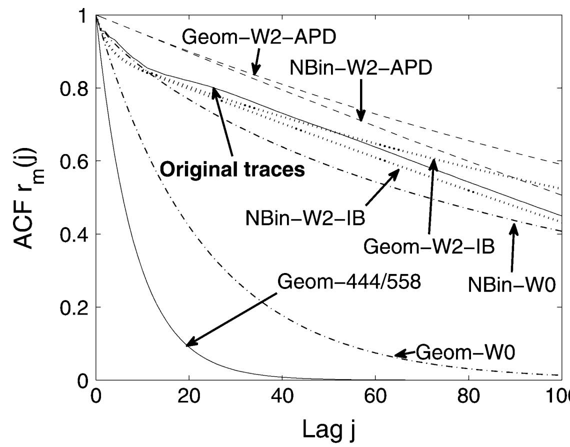


Fig. 6. ACF of consecutive packet sizes of different traffic models for the G.723.1 codec.

multiplex $n = 20$ periodic flows onto a single link, i.e., we consider an $n \cdot D/G/1{-}\infty$ queue. Different flows are independent of each other. Packet sizes within a flow are driven by one of the discussed models or by original packet traces. We choose the link capacity such that a link utilization of 60% is achieved and measure the resulting waiting time W of the packets. The simulation of periodic systems is challenging as periodic processes are not ergodic. The simulation outcome depends on the transmission instants (phases) of the flows within a period, the so-called phase pattern. We randomly initiate the phase pattern for each simulation run and perform sufficiently many runs in order to obtain reliable results. We compute the CCDF of the packet waiting time from 50000 runs. Each simulation run is 2050 periods long. Statistics are collected only after cutting off the initial warmup phase of 50 periods. The 95 percent confidence intervals for the independently derived CCDFs are negligibly small and are omitted in the figures for the sake of clarity.

The CCDFs of the packet waiting times are presented in Fig. 7 for the G.723.1 codec. We choose a relatively large utilization of 60% to make the differences of the queueing properties among the different source traffic models well visible. The CCDF values decrease rather quickly for increasing waiting times, but remain almost constant at a level of $2 \cdot 10^{-3}$. These long waiting times occur due to transient overload which leads to significant queueing and possibly even to packet loss. This happens only with a small probability when sufficiently many flows are in the on-phase such that the bandwidth does not suffice to carry the traffic. Reducing the utilization by increasing the bandwidth in the experiment decreases the probability for transient overload and very long waiting times.

To achieve the same activity factor with the traditional model as with our measurement, we apply the APD method to *Geom-352/650*. The adapted traditional model *Geom-444/558* heavily underestimates the waiting times of the original traces and so does *Geom-W0* albeit to a minor degree. *{Geom, NBin}-W2-APD* overestimate them slightly, thus, they provide a conservative approximation for the queueing properties of the original traces. The packet waiting times of *{Geom, NBin}-W2-IB* are in good accordance with those of the original traces.

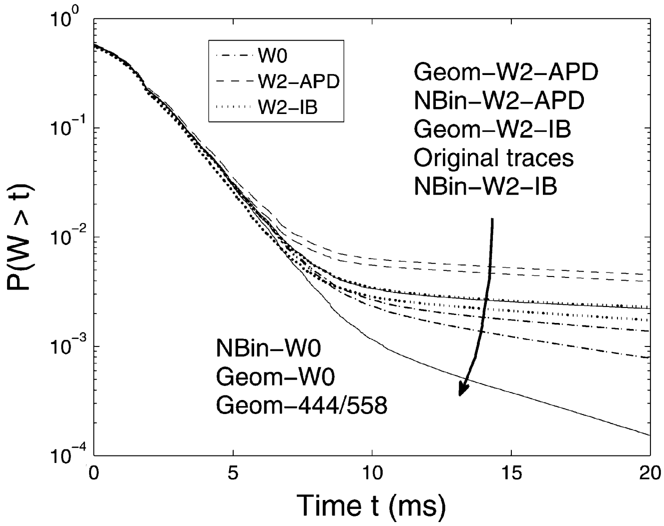


Fig. 7. CCDF of packet waiting times W of an $n \cdot D/G/1-\infty$ queue based on different source traffic models for the G.723.1 codec.

To summarize, the widely used source model for speech traffic *Geom-352/650* underestimates the voice activity. Adapting the voice activity while keeping the on/off phase durations similar leads to too optimistic queueing properties as the durations of the on/off phases are too short. In case a simple model is needed, *Geom-W2-APD* provides a better yet conservative approximation of the queueing properties of the original traces. However, *Geom-W2-APD* overestimates the ACF and the queueing properties of the original traces and can serve as a conservative approximation only. The more complex model *Geom-W2-IB* is more accurate. To additionally meet the CCDF of the on/off phase durations, *NBin-W2-IB* should be used as a source traffic model.

C. Voice Codecs With Variable Bit Rate

Finally, we consider more sophisticated audio codecs. Depending on the speech input they produce packets of different size which leads to VBR traffic. The GSM Adaptive Multi-Rate (AMR) codec [42] is the default speech codec for third generation wireless systems and operates at a rate between 4.75 kbit/s and 12.2 kbit/s. GSM is the most widely used standard for mobile phones and measurements were obtained using the 3GPP reference implementation of GSM AMR. The iSAC [43] is a proprietary codec by Global IP Sound (GIPS) which produces a bit rate between 10 kbit/s and 32 kbit/s. It is one of several codecs being used by the VoIP client Skype [37]. Both codecs adapt their transmission rates to the quality of the communication channel. While GSM AMR decreases the size of its speech packets in times of bad transmission quality to save bandwidth, the Skype implementation of the iSAC codec increases its bit rate, possibly to counteract packet loss by increasing information redundancy. In this paper, however, we concentrate on the behavior of the codecs under perfect network conditions.

1) Traffic Models for VBR Codecs: Figs. 8(a) and (b) show typical traces for the GSM AMR and the iSAC codec. GSM AMR has a VAD, but in contrast to previous codecs, it sends empty packets for synchronization purposes instead of omitting

TABLE IV
STATISTICAL INFORMATION ABOUT PACKET SIZES AND PERIODS FOR ORIGINAL TRACES AND THE MODELLING MMC FOR BOTH THE GSM AMR AND THE iSAC CODEC.

Codec	GSM AMR
No data	0 bytes
SID update	5 bytes
SID first	5 bytes
Speech	31 bytes
Speech period	20 ms
$E[B^{\text{real}}]$	14.097 bytes
$c_{var}[B^{\text{real}}]$	1.0727
$E[B^{\text{MMC}}]$	14.096 bytes
$c_{var}[B^{\text{MMC}}]$	1.0728
n_s	3
n_a	10
W_a	15

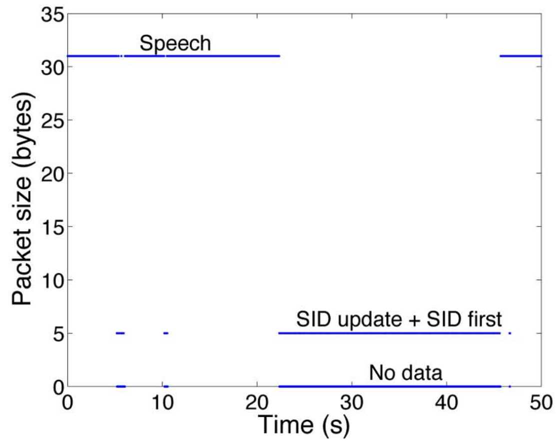
Codec	iSAC
Control size	3 bytes
Control period	20 s
min(packet size)	21 bytes
max(packet size)	166 bytes
Speech period	30 ms
$E[B^{\text{real}}]$	71.319 bytes
$c_{var}[B^{\text{real}}]$	0.625
$E[B^{\text{MMC}}]$	71.321 bytes
$c_{var}[B^{\text{MMC}}]$	0.621
n_s	7
n_a	20
W_a	20

them when there is no data to send. In addition, it produces silence descriptor (SID) packets which describe the recorded background noise to create adequate comfort noise at the receiver side in phases of silence. Therefore, the packet stream on the network layer does not result in an on/off process. The iSAC vocoder dynamically produces packets of many different sizes between 21 bytes and 166 bytes and yields a true VBR stream being significantly different from an on/off process on packet level. However, the underlying pattern of silence and talk in the speech sample is still recognizable in the resultung packet trace. Fig. 8(b) shows a packet trace of the iSAC codec. A silence phase appears as a cluster of relatively small packets from e.g., $t = 22$ s to $t = 45$ s and a talkspurt for the remaining time. Table IV summarizes for both codecs information about individual packet sizes and periods at which they are sent.

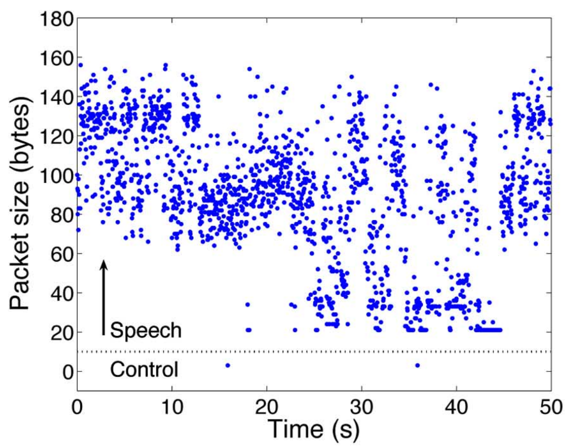
As on/off processes cannot model time series of different packet sizes, we use a memory Markov chain (MMC) [44] for that objective. An MMC is a Markov chain with two-dimensional states (m_i^s, m_i^a). The variables m_i^s and m_i^a can take n_s and n_a different values s_j and a_j, respectively. We use the following serialization of the two-dimensional state space: $((s_0,a_0),\ldots,(s_{n_s-1},a_0),\ldots,(s_0,a_{n_a-1}),\ldots,(s_{n_s-1},a_{n_a-1}))$. This equivalent conventional one-dimensional Markov chain has a $(n_s \cdot n_a) \times (n_s \cdot n_a)$ transition matrix. In our context, the m_i^s are packet sizes and the m_i^a correspond to the average of the last W_a packets. Thus, the m_i^s-projection of the MMC-state yields a synthetic trace of packet sizes.

The MMC can model time series X_i with strong positive correlations and a recipe is given in [44]. The X_i are discretized into n_s different values s_j and the corresponding moving averages $\bar{X}_i = \frac{1}{W_a} \cdot \sum_{0<k\leq W_a} X_{i-k}$ are discretized into n_a different values a_j. Thus, the tuples $(X_i, \bar{X}_i)$ are discretized into tuples $(X_i^d, \bar{X}_i^d)$. The empirical transition probabilities of the discretized process $(X_i^d, \bar{X}_i^d)$ are taken as the entries in the transition matrix of the MMC. In the following, we characterize memory Markov chains MMC(n_s, n_a, W_a) by the values of their parameters n_s, n_a, and W_a.

We tested different parameter settings to model the vocoder output by an appropriate MMC. The search for optimal parameters was performed until a good accordance of the ACFs of the original traces and the MMC was achieved. Removing iSAC's control traffic leads to better results. The parameters for both

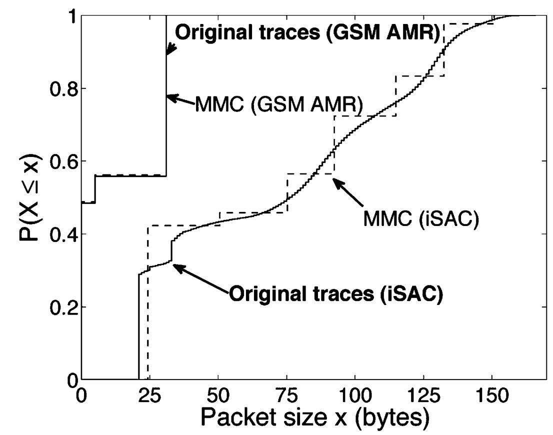


(a) GSM AMR.

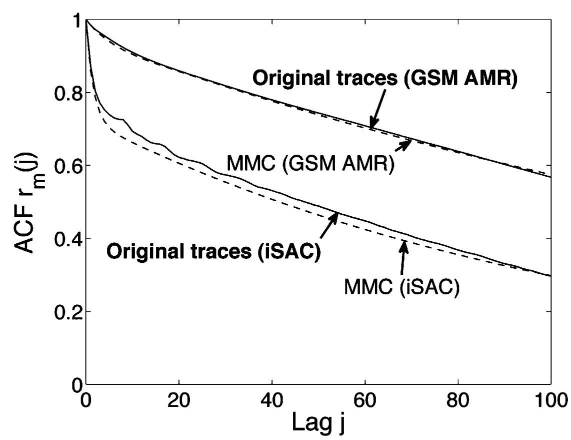


(b) iSAC.

Fig. 8. Time series of packet sizes.



(a) CDF of the packet sizes X.



(b) ACFs of consecutive packet sizes.

Fig. 9. Comparison of original traces and MMC output for the GSM AMR and the iSAC codec.

codecs are given in Table IV. Due to the lack of space we omit the presentation of the discretized packet sizes s_j and the transition matrices of the MMCs, but provide them for download from [45] or upon email request.

We validate the MMC models for the GSM AMR and the iSAC codec by comparing the statistical properties of their synthetic packet traces to those of the original traces. Table IV shows that the corresponding mean values and coefficients of variation hardly differ. Fig. 9(a) compares the analytically derived CDFs of the packet sizes to those of the original traces. The CDF of the MMC(3,10,15) model for the GSM AMR coincides with the one of the empirical data since the original codec also outputs only three different packet sizes. The trace of the iSAC codec has a more stepless distribution of the packet sizes, but the seven discretization levels of the MMC(7,20,20) model approximate the empirical distribution quite well. More discretization levels lead to a better approximation, but in this case the tradeoff was made towards a simpler and faster computable MMC.

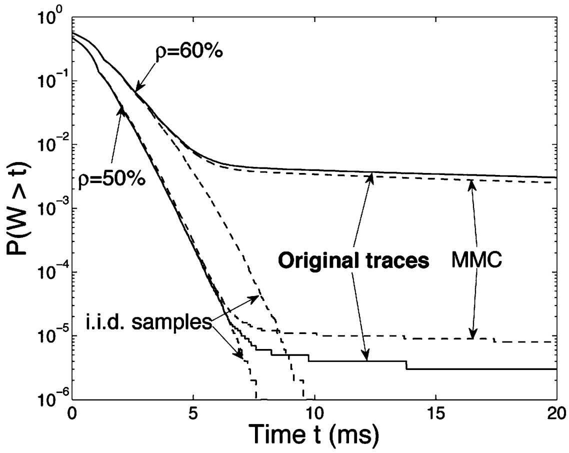
2) Autocorrelation of the VBR Traffic Models: Fig. 9(b) shows that the ACFs of the presented MMCs match those of the original sample traces very well for both codecs. We omitted the ACF for iid packet sizes that are generated based on the empirical distribution because they yield $r_m(j) = 0$ for all lags $j > 0$. Therefore, mere drawing from a packet size distribution cannot capture the ACF of the original codecs since we observe strong positive correlations in the original output traces.

3) Queueing Behavior of the VBR Traffic Models: To compare the queueing properties of the analytical models to those of the original traces, we feed their output to an $n \cdot D/G/1 - \infty$ queue like in Section III-B. We use $n = 20$ sources and choose the link bandwidth such that the system operates at different utilizations. Figs. 10(a) and (b) show the CCDF of the obtained packet waiting times for the GSM AMR and for the iSAC codec. The CCDFs for the original traces and the MMC match quite well for different load levels while the CCDFs of iid packet sizes sampled according to the empirical distribution function underestimate the waiting time of the sample traces significantly.

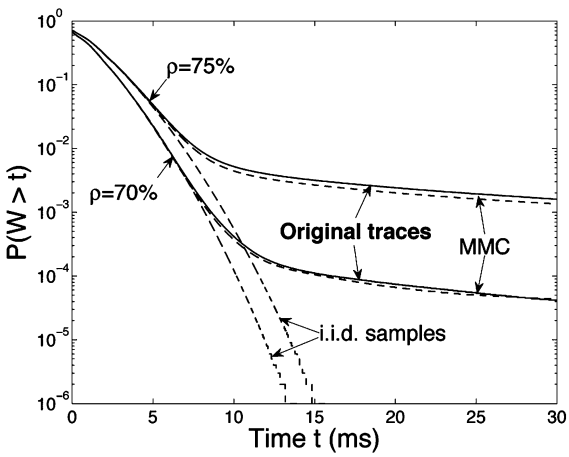
To summarize, the MMC-based model approximates quite well the CDF of the packet sizes, the ACFs, and the queueing properties of VBR voice sources while periodically sampled iid packet sizes fail to do so. The presented MMC is a special case of a Markovian arrival process (MAP) [46] and is accessible to analytical reasoning. The models are simple enough to be integrated in any simulation software and appropriate parameter sets are available at [45] for GSM AMR and iSAC.

IV. Conclusion

In this paper, we studied the output of fundamentally different voice codecs. We sampled a large set of standard telephone



(a) GSM AMR codec.



(b) iSAC codec.

Fig. 10. CCDFs of packet waiting times W of an $n \cdot D/G/1-\infty$ queue fed by original traces and synthetic traffic.

conversations [33] to produce vocoder output and analyzed the traces. We proposed stochastic models approximating the properties of the original traces and validated them. These models are useful for analytic and simulative performance studies.

G.711 and G.729.1 are constant bit rate codecs sending packets of fixed size in regular intervals. They differ in the length of these intervals, in the transmitted packet size, and the associated control information.

G.723.1 and iLBC are codecs with silence detection producing fixed packet sizes but in an on/off manner. Large on- and off-phases are interrupted by short breaks, therefore, their durations are difficult to determine. We measured the average on/off phase durations absolutely (W0) and by a two-window filtering approach (W2). We obtained values of about 10.4 s and 13.1 s with the more sophisticated W2 approach. We proposed new source traffic models based on geometrically (*Geom*) or negative-binomially (*NBin*) distributed on/off phases. The small breaks within the on-phases are either modelled (*IB*) or the phase durations are adapted (*APD*) to take account of the breaks in order to achieve the same voice activity factor as in the original traces. Our validation showed that the *NBin-IB-W2* model best approximates the complementary cumulative distribution functions (CCDFs) of on/off phase durations of the original traces, their autocorrelation function (ACF), and their queueing properties. In contrast to *NBin-IB-W2*, the *Geom-APD-W2* model is simple and its functional behavior is implemented in most simulation programs. Its characteristics are also acceptable if the average on/off phase durations are set to appropriate values.

However, the latter is not the case in practice. Most simulation studies use a source model for speech traffic with geometrically distributed on/off phases and average durations of 352 ms and 650 ms. Values in this order of magnitude are observed for non-digitized speech, but the characteristics for compressed digitized speech are different. When used for simulations, autocorrelations in the time series of generated or suppressed packets carrying compressed voice data must not be neglected, otherwise waiting times and packet loss are underestimated. Therefore, either *NBin-IB-W2* or *Geom-APD-W2* should be preferred to the widely spread model from literature in future simulation studies.

Variable bit rate (VBR) codecs such as the GSM AMR and iSAC also send data in regular intervals, but use variable packet sizes. We modelled the time series of consecutive packet sizes by a memory Markov chain (MMC). The synthetic output of the MMC matches the CDF of the packet sizes, the ACF of consecutive packet sizes, and the queueing properties of the original traces very well. IID packet sizes generated according to the empirical distribution function have a significantly different ACF and queueing properties. The full parametrization of the MMCs can be downloaded from our website [45]. To the best of our knowledge, this is the first study of VBR codecs at all.

The use of realistic models in simulations and analytical studies is especially important if waiting times and loss probabilities are of interest. Our proposed models are simple and easy to implement. Therefore, we hope that this paper leads to more realistic source models for speech traffic in future performance studies.

ACKNOWLEDGMENT

The authors would like to thank P. Kühn, D. Liu, D. Minder, O. Rose, K. Sriram, P. Tran-Gia, and W. Whitt for valuable pointers and fruitful discussions.

REFERENCES

[1] S. S. Wang and J. A. Silvester, "A discrete-time performance model for integrated service ATM multiplexers," in *Proc. IEEE Globecom*, 1993.

[2] S. Sibal, K. Parthasarathy, and K. S. Vastola, "Sensing the state of voice sources to improve multiplexer performance," in *Proc. IEEE INFOCOM*, 1995.

[3] S.-H. Jeong and J. A. Copeland, "Cell loss ratio and multiplexing gain of an ATM multiplexer for VBR voice sources," in *IEEE Conf. Local Computer Networks (LCN)*, 1998.

[4] C. Liu, S. Munir, R. Jain, and S. Dixit, "Packing density of voice trunking using AAL2," in *Proc. IEEE Globecom*, 1999.

[5] A. Bengt, A. Anders, H. Olof, and M. Ian, "Dimensioning Links for IP Telephony," in *Proc. Internet Telephony Workshop*, New York, Apr. 2001, pp. 14–24.

[6] H. Sze, S. Liew, J. Lee, and D. Yip, "A multiplexing scheme for H.323 voice-over-IP applications," *IEEE J. Sel. Areas Commun.*, vol. 20, 2002.

[7] M. K. Ranganathan and L. Kilmartin, "Performance analysis of secure session initiation protocol based VoIP networks," *Comput. Commun.*, vol. 26, pp. 552–565, 2003.

[8] A. Koubaa and Y.-Q. Song, "Loss-tolerant QoS using firm constraints in guaranteed rate networks," in *Proc. IEEE Real-Time and Embedded Technology and Applications Symp. (RTAS)*, Washington, DC, 2004.

[9] S. Obeidat and S. Gupta, "Towards voice over Ad Hoc networks: An adaptive scheme for packet voice communications over wireless links," in *Proc. IEEE Int. Conf. Wireless and Mobile Computing, Networking and Communications (WiMob)*, Aug. 2005.
[10] K. Sriram and W. Whitt, "Characterizing superposition arrival processes in packet multiplexers for voice and data," *IEEE J. Sel. Areas Commun.*, vol. 4, no. 6, pp. 833–846, Sep. 1986.
[11] C. J. May and T. J. Zebo, Private Work 1981.
[12] P. T. Brady, "A technique for investigating on-off patterns of speech," *Bell Syst. Tech. J.*, vol. 44, no. 1, pp. 1–22, Jan. 1965.
[13] P. T. Brady, "A statistical analysis of on-off patterns in 16 conversations," *Bell Syst. Tech. J.*, vol. 47, no. 1, pp. 73–91, Jan. 1968.
[14] P. T. Brady, "A model for generating on-off speech patterns in two-way conversations," *Bell Syst. Tech. J.*, vol. 48, no. 9, pp. 2445–2472, Sep. 1969.
[15] G. Wang, Y. Xia, and D. Harrison, An NS2 TCP Evaluation Tool. Apr. 2007 [Online]. Available: http://tools.ietf.org/id/draft-irtf-tmrg-ns2-tcp-tool-00.txt
[16] A. Adas, "Traffic models in broadband networks," *IEEE Commun. Mag.*, pp. 82–89, Jul. 1997.
[17] W. Jiang and H. Schulzrinne, "Analysis of on-off patterns in VoIP and their effect on voice traffic aggregation," in *Proc. IEEE Int. Conf. Computer Communications and Networks (ICCCN)*, 2000.
[18] A. C. Norwine and O. J. Murphy, "Characteristic time intervals in telephone conversation," *Bell Syst. Tech. J.*, vol. 17, no. 4, pp. 281–291, Apr. 1938.
[19] J. N. Daigle and J. D. Langford, "Models for analysis of packet voice communications systems," *IEEE J. Sel. Areas Commun.*, vol. 4, no. 6, pp. 847–855, Sep. 1986.
[20] H. P. Stern, S. A. Mahmoud, and K.-K. Wong, "Acomprehensive model for voice activity in conversational speech—Development and application to performance analysis of new-generation wireless communication systems," *Wireless Networks*, vol. 2, no. 4, pp. 359–367, Dec. 1996.
[21] Telephone Transmission Quality Objective Measuring Apparatus: Artificial Conversational Speech. ITU-T, P.59, Mar. 1993.
[22] J. G. Gruber, "A comparison of measured and calculated speech temporal parameters relevant to speech activity detection," *IEEE Trans. Commun.*, vol. 30, no. 4, pp. 728–738, Apr. 1982.
[23] H. H. Lee and C. K. Un, "A study of on-off characteristics of conversational speech," *IEEE Trans. Commun.*, vol. 34, no. 6, pp. 630–637, June 1986.
[24] S. Deng, "Traffic characteristics of packet voice," in *Proc. IEEE Int. Conf. Communications (ICC)*, Jun. 1995.
[25] C.-N. Chuah and R. H. Katz, "Characterizing packet audio streams from internet multimedia applications," in *Proc. IEEE Int. Conf. Communications (ICC)*, 2002.
[26] T. D. Dang, B. Sonkoly, and S. Molnár, "Fractal analysis and modeling of VoIP traffic," in *Proc. Int. Telecommunication Network Strategy and Planning Symp. (Networks)*, Vienna, Austria, Jun. 2004, pp. 217–222.
[27] M. Livny, B. Melamed, and A. K. Tsiolis, "The impact of autocorrelation on queueing systems," *Management Science*, vol. 39, no. 3, pp. 322–339, 1993.
[28] S.-Q. Li and J. W. Mark, "Traffic characterization for integrated services network," *IEEE Trans. Commun.*, vol. 38, no. 8, pp. 1231–1243, Aug. 1990.
[29] J. Beran, R. Sherman, M. S. Taqqu, and W. Willinger, "Long-range dependence in variable-bit-rate video traffic," *IEEE Trans. Commun.*, vol. 43, no. 2–4, pp. 1566–1579, 1995.
[30] O. Rose, "Simple and efficient models for variable bit rate MPEG video traffic," *Performance Evaluation*, vol. 30, no. 1–2, pp. 69–85, 1997.
[31] M. Krunz and A. M. Makowski, "A source model for VBR video traffic based on $M/G/\infty$ input processes," in *Proc. IEEE INFOCOM*, 1998.
[32] M. Menth, A. Binzenhöfer, and S. Mühleck, "A note on source models for speech traffic," *ACM SIGMETRICS Performance Evaluation Review (PER) (Short Paper)*, vol. 35, no. 3, pp. 22–24, Dec. 2007.
[33] Bavarian Archive for Speech Signals (BAS) Verbmobil 6.1. 1996 [Online]. Available: http://www.phonetik.uni-muenchen.de/Bas/BasHomedeu.html
[34] Pulse Code Modulation (PCM) of Voice Frequencies. ITU-T, G.711..
[35] Coding of Speech at 8 kbit/s Using Conjugate-Structure Algebraic-Code-Excited Linear Prediction (CS-ACELP) ITU-T, G.729.
[36] X-Lite 3.0. Counter Path Solutions, Inc. [Online]. Available: http://www.xten.com/, http://www.xten.com/
[37] Skype. [Online]. Available: http://www.skype.com
[38] Dual Rate Speech Coder for Multimedia Communications Transmitting at 5.3 And 6.3 kbit/s ITU-T, G.723.1.
[39] M. Vitez, Picophone. [Online]. Available: http://www.vitez.it/picophone/
[40] S. Andersen, A. Duric, H. Astrom, R. Hagen, W. Kleijn, and J. Linden, RFC3951: Internet Low Bit Rate Codec (iLBC). Dec. 2004.
[41] A. Erramilli, O. Narayan, and W. Willinger, "Experimental queueing analysis with long-range dependent packet traffic," *IEEE/ACM Trans. Networking*, vol. 4, no. 2, pp. 209–223, Apr. 1996.
[42] GSM AMR Speech Codec. 3rd Generation Partnership Project (3GPP) [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/26_series/26.073/
[43] iSAC. Global IP Sound [Online]. Available: http://www.globalipsound.com/datasheets/iSAC.pdf
[44] O. Rose, "A memory markov chain model for VBR traffic with strong positive correlations," in *Proc. 16th Int. Teletraffic Congr. (ITC)*, Edinburgh, Scotland, U.K., Jun. 1999, pp. 827–836.
[45] M. Menth, A. Binzenhöfer, and S. Mühleck, Transition Matrices and Parameters of the MMCs Modelling VBR Traffic. [Online]. Available: http://www3.informatik.uni-wuerzburg.de/TR/mmc
[46] D. M. Lucantoni, K. S. Meier-Hellstern, and M. F. Neuts, "A single-server queue with server vacations and a class of non-renewal arrival processes," *Adv. Appl. Probabil.*, vol. 22, pp. 676–705, 1990.

Michael Menth (M'06) studied computer science and mathematics at the University of Wuerzburg, Germany, the University of Texas at Austin, and the University of Ulm, Germany. He received the Diploma in 1998 and the Ph.D. degree in 2004 from the University of Wuerzburg.

Currently, he is an Assistant Professor and heading the research group Next Generation Networks at the Institute of Computer Science at the University of Wuerzburg. His special interests are performance analysis, optimization of communication networks, resource management, resilience issues, and future Internet. Dr. Menth holds numerous patent applications and has received various scientific awards for innovative work. He has been a member of the ACM since 2006.

Andreas Binzenhhöfer studied computer science and mathematics at the University of Wuerzburg, Germany, and the University of California at Los Angeles. He received the Diploma in 2002 and the Ph.D. degree in 2008 from the University of Wuerzburg.

Since 2002, he has been a research assistant at the Institute of Computer Science at the University of Wuerzburg. His current research area is performance analysis of structured P2P systems and their application to distributed network management.

Stefan Mühleck studied computer science and mathematics at the University of Wuerzburg, Germany, and the State University of New York at Albany. He received the Diploma in 2007 from the University of Wuerzburg.

His special interests are future Internet routing, performance analysis, and optimization of communication networks.