Modelling the Arrival Process for Packet Audio

Ingemar Kaj[1] and Ian Marsh[2]

[1] Dept. of Mathematics, Uppsala University, Sweden
`ikaj@math.uu.se`
[2] Ian Marsh, CNA Lab, SICS, Sweden
`ianm@sics.se`

Abstract. Packets in an audio stream can be distorted relative to one another during the traversal of a packet switched network. This distortion can be mainly attributed to queues in routers between the source and the destination. The queues can consist of packets either from our own flow, or from other flows. The contribution of this work is a Markov model for the time delay variation of packet audio in this scenario. Our model is extensible, and show this by including sender silence suppression and packet loss into the model. By comparing the model to wide area traffic traces we show the possibility to generate an audio arrival process similar to those created by real conditions. This is done by comparing the probability density functions of our model to the real captured data.

Keywords:: Packet delay, VoIP, Markov chain, Steady state

1 Introduction

Modelling the arrival process for audio packets that have passed through a series of routers is the problem we will address. Figure 1 illustrates this situation: Packets containing audio samples are sent at a constant rate from a sender, shown in step one. The

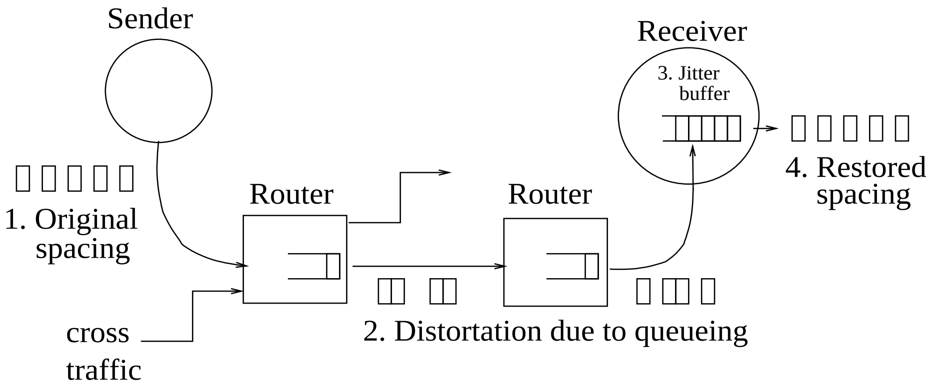

Fig. 1. The networks effect on packet audio spacing

spacing between packets is compressed and elongated relative to each other. This is due to the buffering in intermediate routers and mixing with cross-traffic, shown in step two. In order to replay the packets with their original spacing, a buffer is introduced at the receiver, commonly referred to as a *jitter buffer* shown in step three. The objective of the buffer is to absorb the variance in the inter-packet spacing introduced by the delays due to cross traffic, and (potentially) its own data. In step four, using information coded into the header of each packet, the packets are replayed with their original timing restored.

The motivation for this work derives from the inability of using known arrival processes to approximate the packet arrival process at the receiver. Using a known arrival process, even a complex one, is not always realistic as the model does not include characteristics that real audio streams experience. For example the use of silence suppression or the delay/jitter contribution of cross traffic. One alternative is to use real traffic traces. Although they produce accurate and representative arrival processes, they are inherently static and do not offer much in the way of flexibility. For example, observing the affect of different packet sizes without re-running the experiments. When testing the performance of jitter buffer playout algorithms, for example, this inflexibility is undesirable. Thus, an important contribution of this paper is to address the deficiencies of these approaches by *combining* the advantages of both a model of the process, with data from real traces.

This paper presents in a descriptive manner, a packet delay model, based on the main assumption that packets are subjected to independent transmission delays. It is intended that readers not completely familiar with Markovian theory can follow the description. We assume no prior knowledge of the model as it is built from first principles starting in section 2. We give results for the mean arrival and interarrival times of audio packets in this section. We add silence suppression to the model in section 3 and packet loss in the next section, 4. Real data is incorporated in section 5, related work follows in section 6 and we customarily round off with some conclusions in section 7.

2 The packet delay model

There are two causes of delay for packet audio streams. Firstly, the delay caused by our own traffic, i.e. packets queued up behind ones from the same flow, this we refer to as the *sequential* delay. Secondly, the delay contributed by cross traffic, usually TCP Web traffic, which we call *transmission* delay in this paper. It is important to state we consider these two delays as separate, but study their combined interaction on the observed delays and interarrivals. Propagation and scheduling delay are not modelled as part of this work.

In this model packets are transmitted periodically using a packetisation time of 20 milliseconds. For convenience, the packetisation interval is used as the time unit for the model. Saying that a packet is sent at time k signifies that this particular packet is sent at clock time $20k$ ms into the data stream. The first packet is sent at time 0.

We begin with the transmission delay of a packet. Suppose that packet k could be sent isolated from the rest of the audio stream and let

$$Y_k = \text{transmission delay of packet } k \text{ (no. of 20 ms periods).}$$

To see the impact of the sequential delay, let

$$T_k = \text{the arrival time of packet } k \text{ at the jitter buffer}, \quad k \geq 1.$$

The model used in this paper is shown in Figure 2. The figure shows packets being transmitted from a sender at regular intervals. They traverse the network, where as stated, their original spacing is distorted. Packet k arrives at time T_k at the receiver. The difference in time between when it departed and arrived we call the *observed delay*, which we denote

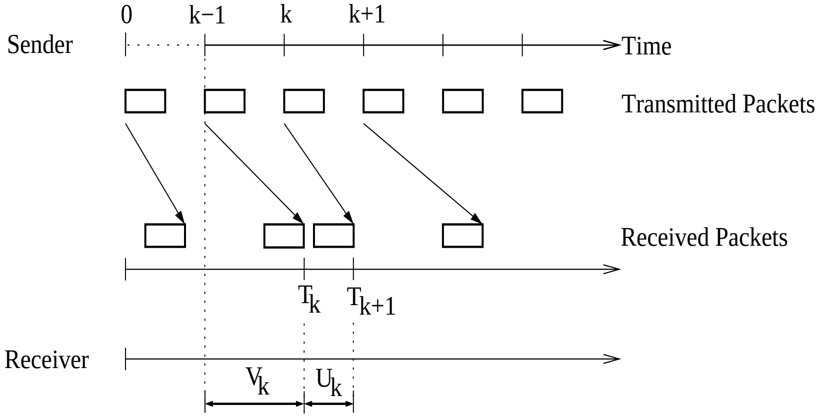

Fig. 2. T_k arrival times before playout, V_k observed delays, U_k observed interarrival times

$$V_k = \text{arrival time} - \text{departure time} = T_k - k + 1 \quad k \geq 1.$$

The time when the next packet (numbered $k + 1$) arrives is T_{k+1} and so the *observed interarrival times* are obtained as the differences between T_{k+1} and T_k, denoted

$$U_k = T_{k+1} - T_k.$$

A packet k, sent at time $k-1$, requires time Y_k to propagate through the network and arrives therefore at $T_k = k - 1 + Y_k$, as long as it is not delayed further by other audio packets (which we call sequential delays). It may however catch up to audio packets transmitted earlier ($1 \to k - 1$). This packet is forced to wait before being stored in the playout buffer. This shows that the actual arrival times satisfy:

$$\begin{aligned} T_1 &= Y_1 \\ T_k &= \max(T_{k-1},\, k - 1 + Y_k), \quad k \geq 2. \end{aligned} \tag{1}$$

Since T_{k-1} and Y_k are independent, we conclude from the relation above (1) that T_k forms a transient Markov chain. Moreover, the interarrival times satisfy

$$U_k = T_k - T_{k-1} = \max(0,\, k - 1 + Y_k - T_{k-1}) \quad k \geq 2. \tag{2}$$

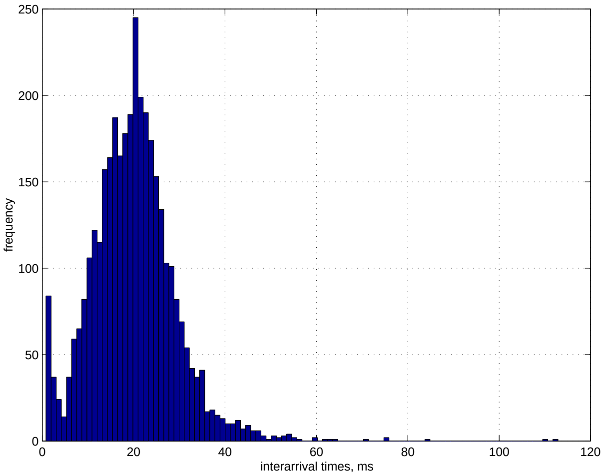

Fig. 3. Histogram of the interarrival times (U_k)

The arrival times (T_k), interarrival times (U_k) and observed delays (V_k) can be easily observed from traffic traces. As an example, Figure 3 shows the histogram for an empirical sequence of interarrival times. The data is from a recording of a Voice over IP session between Argentina and Sweden, more details of the traffic traces are given in section 5.1. The transmission delay sequence (Y_k) should be on the other hand considered as non-observable. The approach in this study is to consider (Y_k) having a general (unknown) distribution and investigate the resulting properties of the observed delay (V_k) and interarrival times (U_k). Since the latter sequences can be empirically observed, this leads to the question to whether the transmission delay distribution can be reconstructed using statistical inference. In this direction we will indicate some methods that could be used to compare the theoretical results with the gathered empirical data.

To carry out the study, we assume from this point the sequence (Y_k) is independent and identically distributed, with distribution function

$$F(x) = P(Y_k \leq x), \quad k \geq 1,$$

and finite mean transmission delay $\nu = \int_0^\infty (1 - F(x))\, dx < \infty$. For the data in our study, typical values of ν are 20-40, i.e. 400-800 ms. We consider these assumptions justified for the purpose of studying a reference model, obviously it would be desirable to allow dependence over time.

2.1 Mean arrival and interarrival times

It is intuitively clear that in the long run $E(U_k) \approx 1$ as on average packets arrive with 20 ms spacing, which we will now verify for the model. The representation (1) for T_k can be written

$$T_k = \max(Y_1, 1 + Y_2, \ldots, k - 1 + Y_k) \quad k \geq 1,$$

which gives the alternative representation

$$T_k = \max(Y_1, 1 + T'_{k-1}), \quad k \geq 2 \tag{3}$$

where on the right side

$$T'_{k-1} = \max(Y_2, 1 + Y_3, \ldots, k - 2 + Y_k)$$

has the same marginal distribution as T_{k-1} but is independent of Y_1. From (3) follows that we can write $\{T_k > t\}$ as a union of two disjoint events, as

$$\{T_k > t\} = \{1 + T'_{k-1} > t\} \cup \{Y_1 > t, 1 + T'_{k-1} \leq t\}.$$

Hence, using the independence of T'_{k-1} and Y_1,

$$\begin{aligned} P(T_k > t) &= P(1 + T'_{k-1} > t) + P(Y_1 > t, 1 + T'_{k-1} \leq t) \\ &= P(1 + T_{k-1} > t) + P(Y_1 > t) P(1 + T_{k-1} \leq t) \end{aligned}$$

and so

$$\begin{aligned} E(T_k) &= \int_0^\infty P(T_k > t)\, dt \\ &= E(1 + T_{k-1}) + \int_1^\infty P(Y_1 > t) P(T_{k-1} \leq t - 1)\, dt. \end{aligned} \tag{4}$$

Therefore

$$E(U_k) = 1 + \int_1^\infty P(Y_1 > t) P(T_{k-1} \leq t - 1)\, dt \to 1, \quad k \to \infty \tag{5}$$

(since $\nu = \int_0^\infty P(Y_1 > t)\, dt < \infty$ and $T_k \to \infty$, the dominated convergence theorem applies forcing the integral to vanish in the limit).

A further consequence of (4) is obtained by iteration,

$$E(T_k) = k - 1 + E(Y_1) + \int_1^\infty P(Y_1 > t) \sum_{i=1}^{k-1} P(T_i \leq t - 1)\, dt.$$

If we introduce

$N(t) =$ the number of arriving packets in the time interval $(0, t]$,

so that $\{N(t) \geq n\} = \{T_n \leq t\}$, this can be written

$$E(V_k) = E(Y_1) + \int_1^\infty P(Y_1 > t) \sum_{i=1}^{k-1} P(N(t-1) \geq i)\, dt, \tag{6}$$

which, as $k \to \infty$, gives an asymptotic representation for the average observed delay as

$$E(V_k) \to \nu + \int_1^\infty P(Y_1 > t) E(N(t-1)))\, dt. \tag{7}$$

2.2 Steady state distributions

By (1),

$$P(T_k \leq x) = \prod_{i=1}^{k} P(i + Y_i \leq x + 1) = \prod_{i=0}^{k-1} F(x - i),$$

and therefore the sequence (V_k), which we defined by $V_k = T_k - k + 1$, $k \geq 1$, satisfies

$$P(V_k \leq x) = \prod_{i=0}^{k-1} F(x + k - 1 - i) = \prod_{i=0}^{k-1} F(x + i) \quad x \geq 0.$$

This shows that (V_k) is a Markov chain with state space the positive real line and asymptotic distribution given by

$$P(V_\infty \leq x) = \prod_{i=0}^{\infty} F(x + i) \quad x \geq 0. \tag{8}$$

Furthermore, for $x \geq 0$

$$\begin{aligned} P(U_k \geq x) &= P(k - 1 + Y_k - T_{k-1} \geq x) = P(V_{k-1} \leq Y_k + 1 - x) \\ &= \int_0^\infty P(V_{k-1} \leq y + 1 - x)\, dF(y), \end{aligned}$$

where in the step of conditioning over Y_k we use the independence of Y_k and V_{k-1}. Therefore the sequence (U_k) has the asymptotic distribution

$$P(U_\infty \leq x) = 1 - \int_0^\infty \prod_{i=1}^{\infty} F(y - x + i)\, dF(y) \quad x \geq 0, \tag{9}$$

in particular a point mass in zero of size

$$P(U_\infty = 0) = 1 - \int_0^\infty \prod_{i=1}^{\infty} F(y + i)\, dF(y). \tag{10}$$

This distribution has the property that $E(U_\infty) = 1$ for any given distribution F of Y with $\nu = E(Y) < \infty$. In fact, this follows from 5 under a slightly stronger assumption on Y (uniform integrability), but can also be verified directly by integrating (9). Figure 4 shows numeric approximations of the (non-normalised) density function $\frac{d}{dx}P(U_\infty \leq x)$ of (9) for three choices of F. All three distributions show a characteristic peak close to time 1 corresponding to the bulk of packets arriving with more or less correct spacing of 20 ms. A fraction of the probability mass is fixed at $x = 0$ in accordance with (10), but not shown explicitly in the figure. These features of the density functions can be compared with the shape of the histogram in Figure 3 with its peak at the 20 ms spacing. Also, close to the origin is a small peak which corresponds to packets arriving back-to-back usually arriving as a burst, probably due to a delayed packet ahead of them. In Figure 4 the density function with the highest peak close to 1 time unit is a Gaussian distribution with arbitrarily selected parameters mean 5 and variance 0.2. Of the two exponential distributions, the one with the higher variance (Exp(3)) has a lower peak and more mass at zero compared with an exponential with smaller variance (Exp(2)).

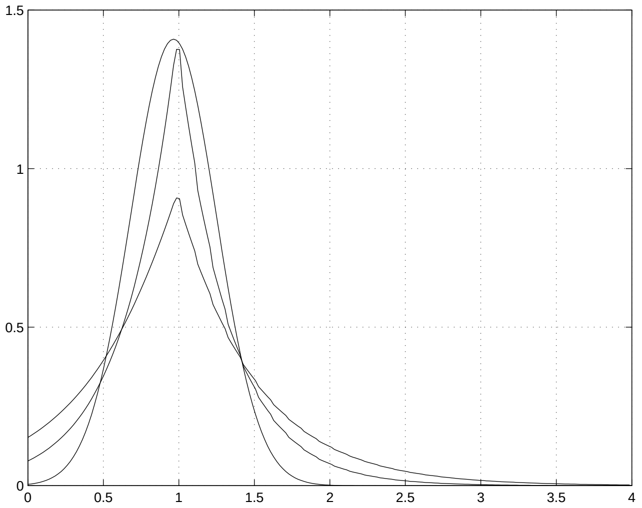

Fig. 4. Density functions of U for N(5,0.2), Exp(2) and Exp(3)

3 Silence suppression mechanism

In this section we incorporate an additional source of random delays due to silence suppression into the model. Silence suppression is employed at the sender so as not to transmit packets when there is no speech activity. During a normal conversation this accounts for about half of the total number of packets, considerably reducing the load on the network. Assign to packet number k the quantity

$$X_k = \text{duration of silent period between packets } k-1 \text{ and } k.$$

A silent period is the time interval during which the silence suppression mechanism is in effect. We assume that the silence suppression intervals are independent of $(Y_k)_{k\geq 1}$ and are given by a sequence of independent random variables $X_1, X_2, \ldots$, such that

$$G(x) = P(X_k \leq x), \quad 1-\alpha = G(0) = P(X_k = 0) > 0, \quad \mu = E(X_k) < \infty.$$

The (small) probability $\alpha = P(X_k > 0)$ represents the case where silence suppression is activated just after packet $k-1$ is transmitted from the sender. Note that

$$S_k = \sum_{i=1}^{k} X_i = \text{total time of silence suppression affecting packet } k,$$

which implies that the delivery of packet k from the sending unit now starts at time $k-1+S_k$. The representation (1) takes the form

$$T_1 = S_1 + Y_1, \quad T_k = \max(T_{k-1},\, k-1+S_k+Y_k), \quad k \geq 2, \tag{11}$$

hence

$$U_k = T_k - T_{k-1} = \max(0,\, k-1+S_k+Y_k-T_{k-1}) \quad k \geq 2. \tag{12}$$

Similarly,

$$V_k = \text{arrival time} - \text{departure time} = T_k - k + 1 - S_k \quad k \geq 1.$$

The alternative representation (3) is

$$T_k = X_1 + \max(Y_1, 1 + T'_{k-1}), \tag{13}$$

where

$$T'_{k-1} = \max(Y_2 + S_2 - X_1, 1 + Y_2 + S_2 - X_1, \ldots, k-2+Y_k+S_k-X_1)$$

has the same marginal distribution as T_{k-1} but is independent of X_1 and Y_1. In analogy with the calculation of the previous section leading up to (4), this relation gives

$$E(T_k) = E(X_1 + 1 + T_{k-1}) + \int_1^\infty P(X_1 + Y_1 > t, X_1 + T'_{k-1} \leq t-1)\,dt. \tag{14}$$

Exchanging the operations of integration and expectation shows that the last integral can be written

$$E\left[\int_{1+X_1}^\infty \mathbf{1}\{Y_1 > t - X_1, T'_{k-1} > t - X_1 - 1\}\,dt\right]$$

where we have also used that the integrand vanishes on the set $\{t \leq 1 + X_1\}$. Apply the change-of-variables $t \to t - X_1$ to get $E\left[\int_1^\infty \mathbf{1}\{Y_1 > t, T'_{k-1} > t-1\}\,dt\right]$. Then shift integration and expectation again to obtain from (14) the relations

$$E(T_k) = 1 + E(X_1) + E(T_{k-1}) + \int_0^\infty P(Y_1 > t) P(T_{k-1} \leq t-1)\,dt$$

and

$$E(U_k) = 1 + E(X_1) + \int_1^\infty P(Y_1 > t) P(T_{k-1} \leq t-1)\,dt.$$

Hence with silence suppression, as $k \to \infty$,

$$E(U_k) \to 1 + \mu, \quad E(V_k) \to \nu + \int_1^\infty P(Y_1 > t) E(N(t-1))\,dt, \tag{15}$$

using the same arguments as in the simpler case of the previous section.

4 Including packet loss in the model

We return to the original model without silence suppression but consider instead the effect of lost packets. Suppose that each IP packet is subject to loss with probability p, independently of other packet losses and of the transmission delays. Lost packets are unaccounted for at the receiver and hence, in this section, the sequence (T_k) records arrival times of non-lost packets only. To keep track of their delivery times from the sender introduce

$$K_k = \text{number of attempts required between} \\ \text{successful packets } k-1 \text{ and } k, \quad k \geq 1,$$

which gives a sequence $(K_k)_{k\geq 1}$ of independent, identically distributed random variables with the geometric distribution

$$P(K_k = j) = (1-p)p^j, \quad j \geq 0.$$

Moreover,

$$\begin{aligned} L_k &= K_1 + \ldots + K_k \\ &= \text{number of attempts required for } k \text{ successful packets} \end{aligned}$$

is a sequence of random variables with a negative binomial distribution. The arrival times of packets are now given by

$$T_1 = K_1 - 1 + Y_{K_1}, \quad T_k = \max(T_{k-1}, L_k - 1 + Y_{L_k}), \quad k \geq 2.$$

Due to the independence we may re-index the sequence of Y_{L_k}'s to obtain

$$T_1 = K_1 - 1 + Y_1, \quad T_k = \max(T_{k-1}, L_k - 1 + Y_k), \quad k \geq 2.$$

and thus

$$T_k = K_1 - 1 + \max(Y_1, 1 + T'_{k-1}), \quad k \geq 2 \tag{16}$$

with K_1, Y_1 and T'_{k-1} all independent, and again T_{k-1} and T'_{k-1} identically distributed. This is the same relation as (13) with X_1 replaced by $K_1 - 1$ and hence, as in (15), $E(U_k) \to 1 + E(K_1 - 1) = \frac{1}{1-p}$, $k \to \infty$, which provides a simple method to estimate packet loss based on observed interarrival times. Similarly, combining silence suppression and packet loss,

$$E(U_k) \to 1 + E(X) + E(K_1 - 1) = \mu + \frac{1}{1-p}, \quad k \to \infty, \tag{17}$$

5 Incorporating Real Data

5.1 Trace data

We give a brief description of the experiments we performed in order to obtain estimates for the parameters in the model. Pulse Code Modulated (PCM) packet audio streams were sent from a site in Buenos Aires, Argentina to Stockholm, Sweden over a number of weeks[1]. The streams are sent with a 64kbits/sec rate in 160 byte payloads. This implies the packets leave the sender with a inter-packet spacing of 20 ms. The remote site is approximately 12,000 kilometres, 25 Internet hops and four time zones from our receiver. The tool is capable of silence suppression, in which packets are not sent when the speaker is silent. Without silence suppression, 3563 packets are sent during 70 seconds and with suppression 2064 are sent. We record the absolute times the packets leave the sender and the absolute arrival times at the receiver. This gives an observed

---

[1] Available from http://www.sics.se/~ianm/COST263/cost263.html

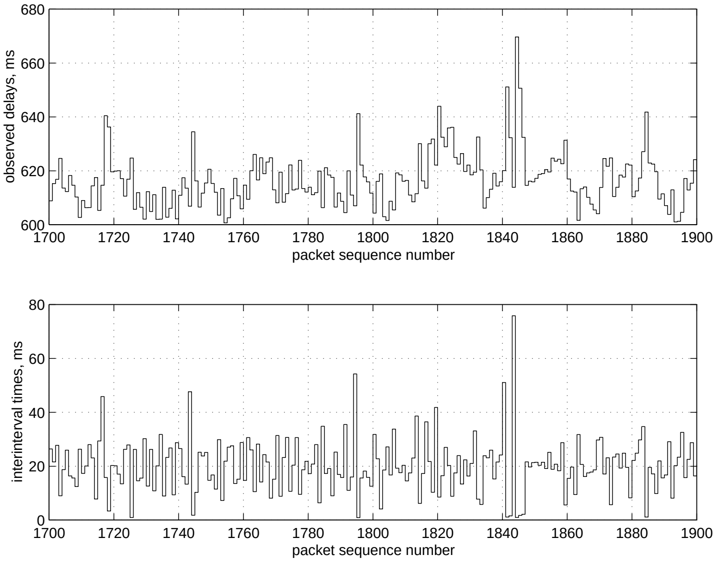

Fig. 5. Four second audio packet traces: a)delays b)interarrival times

sequence

$$v_k = \text{arrival time no } k - \text{departure time no } k$$

of the Markov chain (V_k). In particular, the sample mean $\bar{v}$ is an estimate of the one-way delay. Similarly,

$$u_k = \text{arrival time no } k - \text{arrival time no } (k-1)$$

is a sample of the interarrival time sequence (U_k).

A typical sequence of trace data used in this study *without* silence suppression is shown in Figure 5, which shows (v_k) and (u_k) for a small sequence of 200 packets ($1700 \leq k < 1900$), corresponding to four seconds of audio. To further illustrate such trace data, Figure 6 shows a histogram of the delays (v_k) and Figure 3 a histogram for the interarrival times (u_k). It can be noted that large values of interarrival times are sometimes followed by very small ones, manifesting that a severely delayed packet forces subsequent packets to arrive back-to-back. The fraction of packets arriving in this manner corresponds to the height of the leftmost peak in the histogram of Figure 3.

Returning to traces with silence suppression, Figure 7 gives the statistics of the recorded voice signal used. The upper part shows a histogram of the talkspurts and the lower part the corresponding histogram for the non-zero part of the distribution G of the silence intervals X discussed in section 3. The probability $\alpha = P(X = 0)$ and the expected value $\mu = E(X)$ were estimated to

$$\alpha^* = 0.0456 \qquad \mu^* = 25.7171.$$

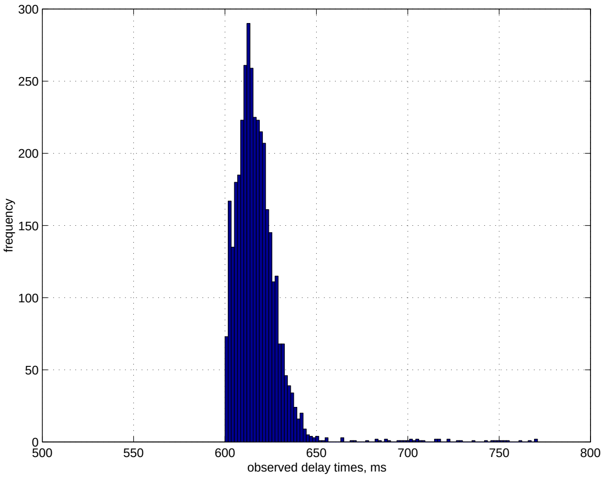

Fig. 6. Histogram of the observed delays (V_k)

5.2 Numerical estimates

In this section we indicate a few simple numerical techniques that give parameter estimates based on trace data. In principle such methods based on the model presented here can be used for systematic studies of the delays and losses and for comparison of traces sampled in different environments.

Considering first the case of no silence suppression, it was pointed out in section 4 that given an observed realization $(u_k)_{k=1}^n$ of (U_k), a point estimate of the packet loss probability p is obtained from (17) (with $\mu = 0$), using

$$p^* = 1 - \frac{20}{\bar{u}}, \qquad \bar{u} = \frac{1}{n}\sum_{k=1}^{n} u_k \text{ ms.}$$

Our measurements gave consistently $\bar{u} \approx 20.002 - 20.005$ ms, indicating loss probabilities of the order 10^{-4}.

Next we look at an experiment where the pre-recorded voice is transmitted at seven different times using silence suppression, and the interarrival times measured at the receiver during each transmission. Table 1 shows the expected silence interval $E(X)$ and the estimated μ from the trace files. The obtained estimates indicate a systematic bias of the order 0.5 milliseconds in the mean values of the silence suppression intervals. Packet losses do not seem to explain fully the observed deviation. A more comprehensive statistical analysis might reveal the source of this slight mismatch. For the present preliminary investigation we find the numerical estimates satisfactory.

We now consider the problem of estimating the distribution F of packet delays Y given a fixed length sample observation (v_k) of the Markov chain (V_k) for observed delays. One method for this can be based on the steady state analysis in section 2.2.

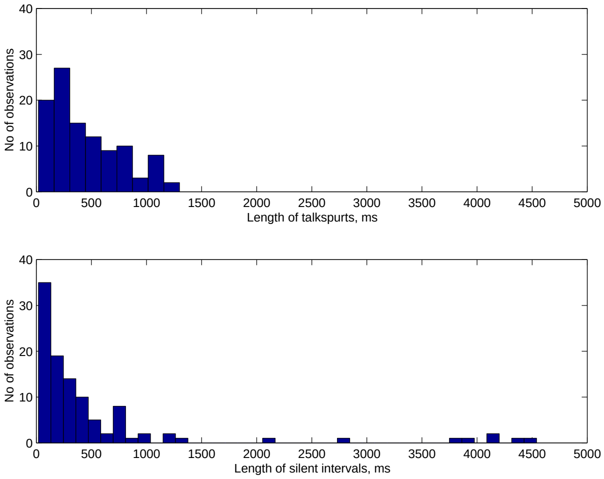

Fig. 7. Lengths of talkspurts and silence periods

Table 1. Silence Interval Parameters

Trace	**E(X)**	**E(X)-μ^***
trace 1	25.7492	0.0321
trace 2	26.2204	0.4639
trace 3	26.2284	0.5113
trace 4	26.2164	0.4993
trace 5	26.2186	0.5015
trace 6	26.2124	0.4953
trace 7	26.2209	0.5038

Indeed, rewriting (8) as the simple relation

$$P(V_\infty \leq x) = F(x)\prod_{i=1}^{\infty} F(x+i) = F(x)\, P(V_\infty \leq x+1)$$

shows that if we let $\bar{F}_V$ denote an empirical distribution function of V, then we obtain an estimate $\bar{F}$ of F by taking

$$\bar{F}(x) = \frac{\bar{F}_V(x)}{\bar{F}_V(x+1)} \quad x \geq 0, \tag{18}$$

where we recall that the variable x is measured in units of 20 ms intervals. An application of this numerical algorithm to the trace data of the previous figures (5 and 6) yields an estimated density function for Y as in Figure 8. The numerical scheme is sensitive for small changes in the data, so it is difficult to draw conclusions on the finer details of the distribution of F. As expected the graph is very similar to that of the observed delays, Figure 6, but with certain differences due to the Markovian dependence structure

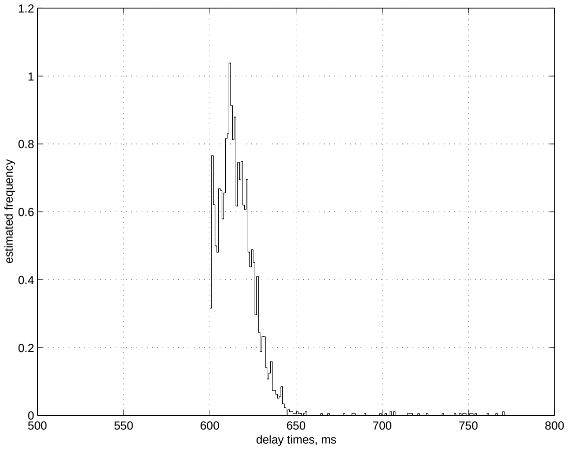

Fig. 8. Estimated density of Y

in the sequence (V_k) as opposed to the independence in (Y_k). The main difference is the shift towards smaller values for Y in comparison to those of V. This corresponds to the inequality $\bar{F}(x) \geq \bar{F}_V(x)$ valid for all x, which is obvious from (18).

6 Related Work

Many researchers have looked at the needs in terms of buffer size for packet streams characterised by Markov (semi or modulated) behaviour especially in the case of multiplexed sources. Their goal was to derive the waiting time of packets spent in the buffer shown as probability density function of the waiting times. Relatively few, however, have looked at the arrival process using a stage of buffers and identifying embedded Markov chains from a single source. Additionally we concentrate on this scenario, including both streams with and without silence suppression. Additionally as far as we know, no-one has used real trace data to enhance and verify their models to the level we show.

Some early analytical work on the buffer size requirements for packetised voice is summarised by Gopal *et al.* [1]. One often cited piece of work is Barberis [2]. As part of this work he assumes the delays experienced by packets of the same talkspurt are i.i.d according to an exponential distribution $p(t) = \lambda e^{-\lambda T}$ where $1/\lambda$ is the average transmission delay and standard deviation. M.K. Mehmet Ali *et al.* in their work of buffer requirements [3] model the arrival process as a Bernoulli trial with probability $[1 - F(j, n - j + 1)]$ of the event “no arrival yet” at each interval up to its arrival. The outcome of the trial is represented by the random variable $k(j, n)$:

$$k(j,n) = \begin{cases} 1 \text{ if packet j has arrived at or before time n} \\ 0 \text{ otherwise.} \end{cases}$$

Ferrandiz and Lazar in [4] look at the analysis of a real time packet session over a single channel node and compute its performance parameters as a function of their model primitives. They do not use any Markovian assumptions, rather an approach which uses a series of overload and under-load periods. During overload packets are discarded. They derive an admission control scheme based on an average of the packet arrival rate. Van Der Wal *et al.* derive a model for the end to end delay for voice packets in large scale IP networks [5]. Their model includes different factors contributing to the delay but not the arrival process of audio packets per se. The mathematical model described here is also discussed in the book [6].

7 Conclusions

We have addressed the problem of modelling the arrival process of a single packet audio stream. The model can be used to produce packet audio streams with characteristics, at least, quite similar to the particular traces we have obtained. The model is suitable for generating streams both with and without silence suppression applied at the source, in addition the case where packets are lost has been included.

The work can be generally applied to research where modelling arriving packet audio streams needs to be performed. A natural next step is to use the arrival model presented here for evaluation of jitter buffer performance, such as investigating waiting times and possible packet loss in the jitter buffer. We observed from our model that the interarrival times are negatively correlated (as mentioned Section 5.1). This will have an impact on the dynamics and performance of a jitter buffer. With an accurate model, based on real data measurements, a realistic traffic generator can be written. In separate work we have gathered nearly 25,000 VoIP traces from ten globally dispersed sites which we can utilise for ’parameterising’ the model, depending on the desired scenario.

References

1. Prabandham M. Gopal, J. W. Wong, and J. C. Majithia, “Analysis of playout strategies for voice transmission using packet switching techniques,” *Performance Evaluation*, vol. 4, no. 1, pp. 11–18, Feb. 1984.
2. Giulio Barberis, “Buffer sizing of a packet-voice receiver,” *IEEE Transactions on Communications*, vol. COM-29, no. 2, pp. 152–156, Feb. 1981.
3. Mehmet M. K. Ali and C. M. Woodside, “Analysis of re-assembly buffer requirements in a packet voice network,” in *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, Bal Harbour (Miami), Florida, Apr. 1986, IEEE, pp. 233–238.
4. Josep M. Ferrandiz and Aurel A. Lazar, “Modeling and admission control of real time packet traffic,” Technical Report Technical Report Number 119-88-47, Center for Telecommunications Research, Columbia University, New York 10027, 1988.
5. Walm van der Wal, Mandjes Michel, and Rob Kooij, “End-to-end delay models for interactive services on a large-scale IP network,” in *IFIP*, June 1999.
6. Ingemar Kaj, *Stochastic Modeling for Broadband Communications Systems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002, Prepared with LATEX.