Dimensioning of a dejittering buffer for variable bit rate video streams

DIMENSIONING OF A DEJITTERING BUFFER FOR VARIABLE BIT RATE VIDEO STREAMS

Kathleen Spaey
Chris Blondia
University of Antwerp
IBBT
Performance Analysis of Telecommunication Systems Research Group
Middelheimlaan 1, B-2020 Antwerp, Belgium
E-mail: {kathleen.spaey,chris.blondia}@ua.ac.be

KEYWORDS

Dejittering buffer, Markov model, underflow probability, overflow probability.

ABSTRACT

In this paper a Markov model is developed that, given a simple description of the network delay and the packet interarrival times of a variable bit rate video stream, can be used to dimension the two important parameters of the dejittering process: the capacity Q the dejittering buffer should have and the dejittering delay T_{jit} the dejittering process should wait before starting the playout of the first packet. The correct setting of these parameters is important to keep both the packet loss due to dejittering buffer overflow and/or underflow and the introduced dejittering delay below some acceptable limits.

INTRODUCTION

In the domain of access networks there is a tendency towards increasing the capacity of the access bit pipes. This provides opportunities to offer beyond the usual data transmission the Internet was originally designed for, also bandwidth-hungry services like high-quality digital television. A major problem with the real-time transmission of audio-visual information over the Internet is however that the end-to-end delay is variable, depending on the load of the network. These delay variations (also called jitter) need to be reduced at the receiving end, or the decoder will not operate correctly (Apostolopoulos et al., 2002).

The task of reducing the jitter is typically performed by placing a dejittering buffer before the decoder. In this buffer, the first packet of a video stream is retained for some time (the *dejittering delay*) before it is offered to the decoder. From then on, packets are played out with the same cadence as with which the encoder generated the video packet stream. Packets that arrive before their playout instant are temporarily stored in the dejittering buffer, packets that arrive too late are useless and considered lost. For each packet that arrives too late, the dejittering buffer underflows (i.e., runs empty) at the moment such a packet is supposed to be played out.

For a given dejittering delay and dejittering buffer capacity, there is a maximal amount of network jitter that can be accommodated without buffer underflow or overflow, either of which would result in a poor video quality. But while a dejittering buffer can compensate for jitter, the dejittering mechanism also introduces additional delay. Specifying the dejittering delay and the size of the dejittering buffer thus always involves a trade-off between packets being lost to the video application and their delay.

The aim of this paper is to develop a mathematical model that, given a simple description of the network delay and the packet interarrival times of a variable bit rate video stream, can be used to dimension the two important parameters of the dejittering process: the capacity Q the dejittering buffer should have and the dejittering delay T_{jit} the dejittering process should wait before starting the playout of the first packet. The correct setting of these parameters is important to keep both the packet loss due to dejittering buffer overflow and/or underflow and the introduced dejittering delay below some acceptable limits.

SYSTEM DESCRIPTION AND ASSUMPTIONS

Consider a video source that sends its traffic over a packet-based network towards a video player, at a rate that is not necessarily constant (i.e., a variable bit rate source). Because of varying (queueing) delays within the routers in the network, the end-to-end delay between source and receiver can vary from packet to packet. To compensate for this delay jitter introduced by the network, the video player uses a dejittering buffer.

Assume that the time is divided into fixed-length intervals, referred to as slots, and that the variable character of the interarrival time (IAT) of successive packets at the network is characterized by the probability distribution

$$b_i = \mathrm{P}\{\mathrm{IAT} = i \text{ slots}\}, \quad 0 < iat_{\min} \le i \le iat_{\max}.$$

Remark that taking $iat_{\min}$ equal to $iat_{\max}$ results in the characterization of constant bit rate traffic.

The network delays experienced by the packets are assumed to be identically distributed random variables (not necessarily independent) with probability distribution

$$d_i = \mathrm{P}\{\text{network delay} = i \text{ slots}\}, \qquad 0 < delay_{\min} \le i \le delay_{\max}.$$

For ease of notation later on, we assume that $d_i \ne 0$, $\forall i \in \{delay_{\min}, \ldots, delay_{\max}\}$. This is however not an essential requirement.

A delay distribution could be obtained from delay measurements in the considered network, or from the service level specification of the network provider. Assuming that the delay of each packet is described by the same probability distribution is only valid if the network conditions (e.g., routes, loads, etc.) do not change drastically over the time that the video flow exists.

Independency of the delay of succeeding packets can however not be assumed, since packet delays in general exhibit a temporal dependency, particularly for high delays (Jiang and Schulzrinne, 2000). The reason for this dependency is intuitively explained as follows. A high delay for packet k indicates a non-empty router buffer in the network. Since it takes some time for such a buffer to drain, if the inter-packet gap at the sender is small, the buffer depth may not have changed much upon arrival of packet $k+1$ at that buffer, and this packet will also experience a high delay. We will take this phenomenon into account by assuming that the probability with which the delay $D(k+1)$ of a packet $k+1$ takes a certain value, given the delay $D(k)$ experienced by packet k and the interarrival time $\mathrm{IAT}(k, k+1)$ between the two packets at the network, has the form of a conditional distribution:

$$\mathrm{P}\{D(k+1) = j \mid D(k) = i \wedge \mathrm{IAT}(k,k+1) = l\} = \begin{cases} 0, & \text{if } j \le i - l, \\ f_j \Big/ \left(1 - \sum_{m=delay_{\min}}^{i-l} f_m\right), & \text{otherwise,} \end{cases}$$

with $delay_{\min} \le i, j \le delay_{\max}$, $iat_{\min} \le l \le iat_{\max}$, all $f_j \ge 0$ and $\sum_{j=delay_{\min}}^{delay_{\max}} f_j = 1$, where the exact values of the f_j's are still unknown at this point.

Since $\mathrm{P}\{D(k) = i\} = d_i$ and $\mathrm{P}\{D(k+1) = j\} = d_j$, the theorem on total probability gives for all $j \in \{delay_{\min}, \ldots, delay_{\max}\}$ that

$$\begin{aligned} d_j &= \mathrm{P}\{D(k+1) = j\} \\ &= \sum_{i=delay_{\min}}^{delay_{\max}} \sum_{l=iat_{\min}}^{iat_{\max}} \mathrm{P}\{D(k+1) = j \mid D(k) = i \wedge \mathrm{IAT}(k,k+1) = l\} \\ &\quad \mathrm{P}\{\mathrm{IAT}(k,k+1) = l \mid D(k) = i\}\, \mathrm{P}\{D(k) = i\} \\ &= \sum_{i=delay_{\min}}^{delay_{\max}} \sum_{l=\max\{iat_{\min}, i-j+1\}}^{iat_{\max}} f_j\, d_i \left(\frac{b_l}{1 - \sum_{m=delay_{\min}}^{i-l} f_m}\right). \end{aligned} \quad (1)$$

By solving this set of $delay_{\max} - delay_{\min} + 1$ equations, the $delay_{\max} - delay_{\min} + 1$ unknown f_j's are obtained. Note that solving this set of equations is not difficult, since $i - l < j$ in Equation (1). So the first equation corresponding to $j = delay_{\min}$ only contains the unknown $f_{delay_{\min}}$. The second equation contains $f_{delay_{\min}+1}$ and $f_{delay_{\min}}$, where $f_{delay_{\min}}$ can be substituted from the solution of the first equation, and so on.

Denote by $\mathrm{IDT}(k, k+1)$ the interdeparture time from the network between packet k and packet $k+1$ (i.e., their interarrival time at the dejittering buffer). Then

$$\mathrm{IDT}(k,k+1) = \mathrm{IAT}(k,k+1) + D(k+1) - D(k). \quad (2)$$

Define for notational convenience later on the matrices $\mathbf{U}_s$, with elements $(\mathbf{U}_s)_{i,j}$, ($delay_{\min} \le i, j \le delay_{\max}$) equal to

$$\begin{aligned} (\mathbf{U}_s)_{i,j} &= \mathrm{P}\{\mathrm{IDT}(k,k+1) = s+1 \wedge D(k+1) = j \mid D(k) = i\} \\ &= \mathrm{P}\{\mathrm{IAT}(k,k+1) = s+1-j+i \wedge D(k+1) = j \mid D(k) = i\} \\ &= \mathrm{P}\{D(k+1) = j \mid \mathrm{IAT}(k,k+1) = s+1-j+i \wedge D(k) = i\} \\ &\quad \mathrm{P}\{\mathrm{IAT}(k,k+1) = s+1-j+i \mid D(k) = i\} \\ &= \begin{cases} \frac{f_j\, b_{s+1-j+i}}{1-\sum_{m=delay_{\min}}^{j-s-1} f_m}, & \text{if } s \ge 0 \text{ and } iat_{\min} \le s+1-j+i \le iat_{\max}, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Remark that for $s \notin \{0, \ldots, S\}$, with $S = iat_{\max} + delay_{\max} - delay_{\min} - 1$, $\mathbf{U}_s = \mathbf{0}$.

DEJITTERING PROCESS

For every slot n, define the states (l_n, t_n, j_n) with $l_n > 0$, and (l_n, t_n, w_n, j_n) with $l_n = 0$, where the different variables have the following interpretation:

- l_n: for $l_n > 0$, the number of slots ago the next packet for playout arrived, and for $l_n = 0$, the fact that the next packet for playout did not yet arrive, or arrives in the current slot n,

- t_n: the number of slots to go until the next playout should start, where $t_n = 0$ means that the next packet should be played out in the current slot n,

- w_n (only defined if $l_n = 0$): the number of slots to go until the next packet for playout arrives, where $w_n = 0$ means that it will arrive in the current slot n,

- j_n: the network delay experienced by the next packet for playout.

Remark that states with $l_n = 0$, $t_n = 0$ and $w_n > 0$ correspond to a packet being late for playout. Note further that given the state of the dejittering process in slot n, it is possible to calculate with which probability the process will be in which state in slot $n+1$. So the states defined above describe a discrete-time Markov chain. The transition probabilities of going from one state to another state will be derived now.

First of all, transitions from a state with $t_n > 0$ are trivial, since the next packet for playout is the same packet in slot $n+1$ as in slot n. With probability one the system will evolve

- from a state $(l_n > 0, t_n > 0, j_n)$ in slot n to the state $(l_{n+1} = l_n + 1, t_{n+1} = t_n - 1, j_{n+1} = j_n)$ in slot $n+1$,

- from a state $(l_n = 0, t_n > 0, w_n > 0, j_n)$ in slot n to the state $(l_{n+1} = 0, t_{n+1} = t_n - 1, w_{n+1} = w_n - 1, j_{n+1} = j_n)$ in slot $n+1$, and

- from a state $(l_n = 0, t_n > 0, w_n = 0, j_n)$ in slot n to the state $(l_{n+1} = 1, t_{n+1} = t_n - 1, j_{n+1} = j_n)$ in slot $n+1$.

If the system is in slot n in a state with $t_n = 0$, i.e., the next packet should be played out in this slot, then with probability $(\mathbf{U}_s)_{i,j}$, $s \in \{0, \ldots, S\}$ and $delay_{\min} \leq i, j \leq delay_{\max}$, a transition is made

- from a state $(l_n > 0, t_n = 0, j_n = i)$ in slot n to the state $(l_{n+1} = l_n - s, t_{n+1} = s + i - j, j_{n+1} = j)$ in slot $n+1$ if $s - l_n < 0$, and to the state $(l_{n+1} = 0, t_{n+1} = s + i - j, w_{n+1} = s - l_n, j_{n+1} = j)$ if $s - l_n \geq 0$,

- from a state $(l_n = 0, t_n = 0, w_n, j_n = i)$ in slot n to the state $(l_{n+1} = 0, t_{n+1} = s + i - j, w_{n+1} = w_n + s, j_{n+1} = j)$.

This is because when a packet k should be played out in slot n, and packet $k+1$ arrives $s+1$ slots later than packet k at the player,

- packet $k+1$ should be played out $\mathrm{IAT}(k, k+1)$ slots later than packet k, i.e., in slot $n + s + 1 - j + i$ (by Equation (2)),

- packet k arrived in slot $n - l_n$ if $l_n > 0$, and thus packet $k+1$ arrives in slot $n - l_n + s + 1$, which is before or later than slot $n+1$ depending on if $s - l_n$ is positive or negative,

- packet k arrives in slot $n + w_n$ if $l_n = 0$, and thus packet $k+1$ arrives in slot $n + w_n + s + 1$.

Denote $\mathcal{D} = \{delay_{\min}, \ldots, delay_{\max}\}$. Because the receiver waits T_{jit} slots after the arrival of the first packet before starting the playout of the video, the evolution of the Markov chain defined above will start upon arrival of the first packet in a state $(0, T_{jit}, 0, d^*)$, where $d^* \in \mathcal{D}$ is the network delay experienced by the first packet.

Define the set of states $\mathcal{C}(T_{jit} + d^*)$ as

$$\begin{aligned}
&\mathcal{C}(T_{jit} + d^*) = \\
&\quad \big\{(0, t, w, j) | j \in \mathcal{D}, w = t + j - T_{jit} - d^* \text{ and} \\
&\quad t \in \{\max(0, T_{jit} + d^* - j), \ldots, iat_{\max} - 1\}\big\} \\
&\quad \bigcup \big\{(l, t, j) | j \in \mathcal{D}, l = T_{jit} + d^* - t - j \text{ and} \\
&\quad t \in \{0, \ldots, \min(T_{jit} + d^* - j, iat_{\max}) - 1\}\big\}. \quad (3)
\end{aligned}$$

The states of $\mathcal{C}(T_{jit} + d^*)$ form a Markov chain with exactly one irreducible closed set of states (see the appendix for details). States not in this closed set are transient. As a consequence, this Markov chain has a unique stationary distribution $\boldsymbol{\pi}$.

If $T_{jit} \leq iat_{\max} - 1$, the start state $(0, T_{jit}, 0, d^*)$ belongs to the set $\mathcal{C}(T_{jit} + d^*)$. If $T_{jit} > iat_{\max} - 1$, $(0, T_{jit}, 0, d^*)$ is not an element of $\mathcal{C}(T_{jit} + d^*)$, but after T_{jit} transitions with probability 1, the system reaches the state $(T_{jit}, 0, d^*)$, which does belong to $\mathcal{C}(T_{jit} + d^*)$. So when the dejittering process starts in the state $(0, T_{jit}, 0, d^*)$, the steady-state behavior of this process is described by the Markov chain with the elements of $\mathcal{C}(T_{jit} + d^*)$ as states.

To calculate the stationary distribution $\boldsymbol{\pi}$, the set of equations

$$\boldsymbol{\pi}\mathbf{Q} = \boldsymbol{\pi}, \quad \boldsymbol{\pi}\mathbf{e} = 1, \qquad (4)$$

should be solved, where $\mathbf{Q}$ is the transition matrix describing the transition probabilities among the states of $\mathcal{C}(T_{jit} + d^*)$. By dividing $\mathbf{Q}$ into blocks corresponding to the first index of the states, i.e.,

$$\mathbf{Q} = \begin{pmatrix}
\mathbf{B}_1^{(0)} & \mathbf{B}_0^{(0)} & \mathbf{0} & \ldots & \mathbf{0} \\
\mathbf{B}_2^{(1)} & \mathbf{A}_1^{(1)} & \mathbf{A}_0^{(1)} & \ldots & \mathbf{0} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\mathbf{B}_{l_{\max}}^{(l_{\max}-1)} & \mathbf{A}_{l_{\max}-1}^{(l_{\max}-1)} & \mathbf{A}_{l_{\max}-2}^{(l_{\max}-1)} & \ldots & \mathbf{A}_0^{(l_{\max}-1)} \\
\mathbf{B}_{l_{\max}+1}^{(l_{\max})} & \mathbf{A}_{l_{\max}}^{(l_{\max})} & \mathbf{A}_{l_{\max}-1}^{(l_{\max})} & \ldots & \mathbf{A}_1^{(l_{\max})}
\end{pmatrix},$$

the lower block-Hessenberg form of this matrix becomes clear, which means that equation (4) can be solved using for example the algorithm described in (Van Houdt et al., 2003), which is an adaptation of the algorithm of Latouche, Jacobs and Gaver (Latouche et al., 1984). The diagonal blocks of $\mathbf{Q}$ are square matrices, the other blocks are rectangular with appropriate dimensions. Once $\boldsymbol{\pi}$ is calculated, several performance measures of the dejittering process can be calculated.

PERFORMANCE MEASURES

This section explains how to calculate the underflow and overflow probabilities from the presented model. Remember from before that for each different value of $T_{jit} + d^*$, a Markov chain with a different set of states $\mathcal{C}(T_{jit} + d^*)$, and consequently a different stationary distribution should be considered. Despite this, the simple notation '$\boldsymbol{\pi}$' is used for such a stationary distribution, since adding a reference to the value $T_{jit} + d^*$ will just overload notations, while the stationary distributions corresponding to different values for $T_{jit} + d^*$ will never be used together in a single formula.

Underflow Probability

The probability that the dejittering buffer underflows, i.e., that a packet is not in the dejittering buffer on its playout time, given that the receiver waits T_{jit} slots after the arrival of the first packet to start the playout, and given that the first packet experiences a network delay of d^* slots, is

$$P_{late}(T_{jit}, d^*) = \frac{\sum_{w>0} \sum_{j\in\mathcal{D}} \boldsymbol{\pi}_{(0,0,w,j)}}{\sum_{w} \sum_{j\in\mathcal{D}} \boldsymbol{\pi}_{(0,0,w,j)} + \sum_{l>0} \sum_{j\in\mathcal{D}} \boldsymbol{\pi}_{(l,0,j)}}, \quad (5)$$

i.e., the fraction of all slots in which a packet should be played out, but the packet did not yet arrive, to all slots in which a packet should be played out.

Remark that when $T_{jit} + d^* \geq delay_{\max}$, all packets are always on time for playout, since there does not exist a state of the form $(0, 0, w, j) \in \mathcal{C}(T_{jit} + d^*)$ with $w > 0$ and $j \in \mathcal{D}$. As a consequence, the numerator of (5) equals zero. However, the delay d^* experienced by the first packet will not be known to the receiver at the moment it sets the parameter T_{jit}. Choosing $T_{jit} \geq delay_{\max} - delay_{\min}$ ensures that all packets will always be on time for playout, although the introduction of an extra delay of this amount will not always be acceptable if the difference between the maximum and the minimum delay becomes large.

The average underflow probability given that the receiver waits T_{jit} slots after the arrival of the first packet to start the playout is obtained by considering all the different values d^* can take:

$$P_{late}(T_{jit}) = \sum_{k\in\mathcal{D}} P_{late}(T_{jit}, d^* = k)\, d_k.$$

Overflow Probability

Packets that arrive too late for playout will not enter the dejittering buffer, since they are of no use to the application anymore. Packets that arrive on time could however still get lost when upon their arrival the dejittering buffer is full. The probability that this happens when the dejittering buffer has a dimension of Q packets, will be approximated by the probability that a timely arriving packet sees upon its arrival Q or more packets present in an infinite capacity buffer.

Consider a random timely arriving packet, and suppose that the playout instant corresponding to this packet is slot n. Then the Markov chain describing the dejittering process will be in a state of the following types in slot n: $(l_n > 0, t_n = 0, j_n \in \mathcal{D})$ or $(l_n = 0, t_n = 0, w_n = 0, j_n \in \mathcal{D})$. This means that the packet arrived in slot $n - l_n$, and the packet saw at that moment as many packets present in the dejittering buffer as will be played out in the time interval $[n - l_n; n[$. So the probability that a random timely arriving packet that arrives $l \geq 0$ slots before its playout time finds q packets in the dejittering buffer upon its arrival, equals the probability that q succeeding interarrival times of the variable bit rate stream last l or less slots, while $q + 1$ succeeding interarrival times last more than l slots.

Denote by $X(q)$ the duration of q succeeding interarrival times. Then

$$\begin{aligned}
&\mathrm{P}\{X(q) \leq l \wedge X(q+1) > l\} \\
&= \sum_{i=\max(q.iat_{min},\, l-iat_{\max}+1)}^{\min(l,\, q.iat_{\max})} \mathrm{P}\{X(q) = i \wedge X(q+1) > l\} \\
&= \sum_{i=\max(q.iat_{min},\, l-iat_{\max}+1)}^{\min(l,\, q.iat_{\max})} \mathrm{P}\{X(q+1) > l \,|\, X(q) = i\} \\
&\qquad\qquad\qquad\qquad \mathrm{P}\{X(q) = i\} \\
&= \sum_{i=\max(q.iat_{min},\, l-iat_{\max}+1)}^{\min(l,\, q.iat_{\max})} \; \sum_{j=\max(l-i+1,\, iat_{\min})}^{iat_{\max}} b_j \;\, I_q^i,
\end{aligned}$$

where I_q^i denotes the probability that q succeeding interarrival times of the variable bit rate stream last i slots. This probability is calculated using the following recursive relation:

$$I_0^i = \begin{cases} 1, & \text{if } i = 0, \\ 0, & \text{otherwise,} \end{cases}$$

$$I_1^i = \begin{cases} b_i, & \text{if } iat_{\min} \leq i \leq iat_{\max}, \\ 0, & \text{otherwise,} \end{cases}$$

and for $q > 1$, $I_q^i =$

$$\begin{cases} 0, & \text{if } i < q.iat_{\min} \text{ or } i > q.iat_{\max}, \\ \sum_{j=\max(i-(q-1).iat_{\max}, iat_{\min})}^{\min(iat_{\max}, i-(q-1).iat_{\min})} I_1^j \, I_{q-1}^{i-j}, & \text{otherwise.} \end{cases}$$

So the probability that a timely arriving packet sees upon arrival q packets in the dejittering buffer, given T_{jit} and d^*, is:

$$\mathrm{P}\{Q_a = q \,|\, T_{jit}, d^*\} = \frac{\sum_{j\in\mathcal{D}} \boldsymbol{\pi}_{(0,0,0,j)} 1_{\{q=0\}}}{\sum_{j\in\mathcal{D}} \boldsymbol{\pi}_{(0,0,0,j)} + \sum_{l>0}\sum_{j\in\mathcal{D}} \boldsymbol{\pi}_{(l,0,j)}} +$$

$$\frac{\sum_{l>0}\sum_{j\in\mathcal{D}} \boldsymbol{\pi}_{(l,0,j)} \sum_{i=\max(q.iat_{min}, l-iat_{\max}+1)}^{\min(l, q.iat_{\max})} \sum_{k=\max(l-i+1, iat_{\min})}^{iat_{\max}} b_k \, I_q^i}{\sum_{j\in\mathcal{D}} \boldsymbol{\pi}_{(0,0,0,j)} + \sum_{l>0}\sum_{j\in\mathcal{D}} \boldsymbol{\pi}_{(l,0,j)}},$$

and given only T_{jit}, this becomes

$$\mathrm{P}\{Q_a = q \,|\, T_{jit}\} = \sum_{k\in\mathcal{D}} \mathrm{P}\{Q_a = q \,|\, T_{jit}, d^* = k\} \, d_k.$$

Then the overflow probability given T_{jit} and d^* for a dejittering buffer of size Q is approximated by

$$P_Q(T_{jit}, d^*) \approx \mathrm{P}\{Q_a \geq Q \,|\, T_{jit}, d^*\},$$

and given T_{jit} only by

$$P_Q(T_{jit}) \approx \mathrm{P}\{Q_a \geq Q \,|\, T_{jit}\}.$$

Remark that the maximal value l can take in a state with $t = 0$ of $\mathcal{C}(T_{jit} + d^*)$, is $l_{\max} = \max(T_{jit} + d^* - delay_{\min}, 0)$. Because there will be at most one playout per $iat_{\min}$ slots, a timely arriving packet sees never more than $\lfloor \frac{l_{\max}}{iat_{\min}} \rfloor$ packets in the dejittering buffer upon arrival. So when the dejittering buffer size Q is chosen equal to or larger than $\lfloor \frac{T_{jit}+delay_{\max}-delay_{\min}}{iat_{\min}} \rfloor + 1$, no packets will get lost due to buffer overflow. Setting $Q = \lfloor \frac{2(delay_{\max}-delay_{\min})}{iat_{min}} \rfloor + 1$ and $T_{jit} = delay_{\max} - delay_{\min}$ corresponds to the optimal choice for T_{jit} and Q for which no packets get lost due to the dejittering process, i.e., not because of buffer underflow and not because of buffer overflow.

NUMERICAL EXAMPLE

Consider a network delay distribution with the shape of a shifted gamma distribution, as shown in Figure 1. Studies (Corlett et al., 2002; Bovy et al., 2002) into one-way end-to-end packet delays have found that the vast majority of these delays is well approximated by a shifted gamma distribution, and that excessive delays are in reality quite rare. Figure 2 shows the packet interarrival time distribution that will be considered in this example.

Figures 3 until 6 present performance results obtained with the mathematical model developed in this paper, when the data shown in the Figures 1 and 2 is used as input to the model.

Figure 3 shows the underflow probability for different values of $T_{jit} + d^*$. $T_{jit} + d^*$ is always larger than or equal to $delay_{\min}$, which is 751 slots in this example. For $T_{jit} + d^*$ larger or equal to $delay_{\max}$, which is 800 slots, the underflow probability is zero. A curve as shown in Figure 3 can be used to dimension the parameter T_{jit} of the dejittering process. This parameter should be set as small as possible to reduce the extra delay introduced by the dejittering process, but large enough to keep the packet loss due to underflow below some target value. For a certain target

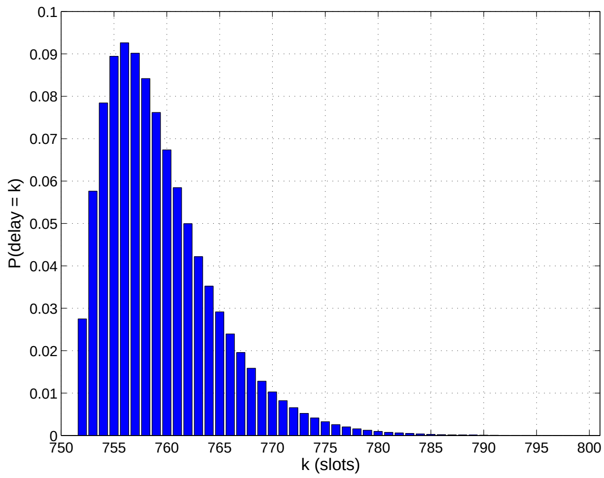

Figure 1: Network delay distribution

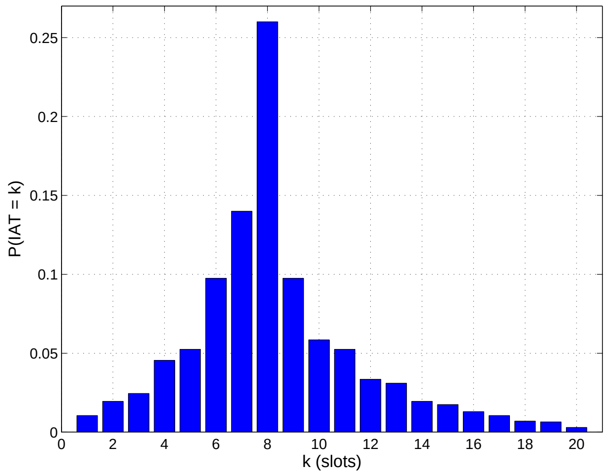

Figure 2: Packet interarrival time distribution of the variable bit rate stream

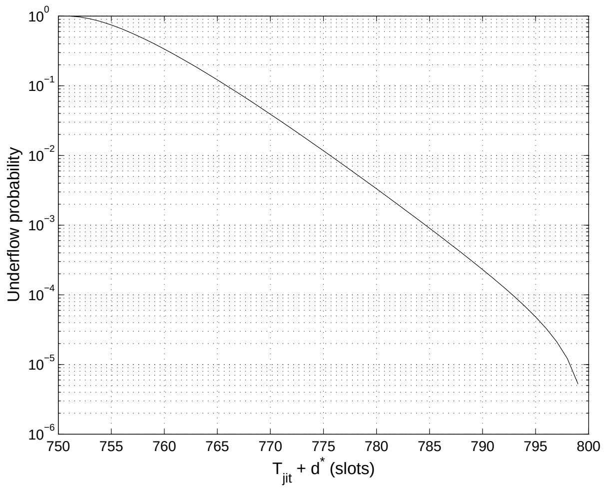

Figure 3: Underflow probabilities given T_{jit} and d^*

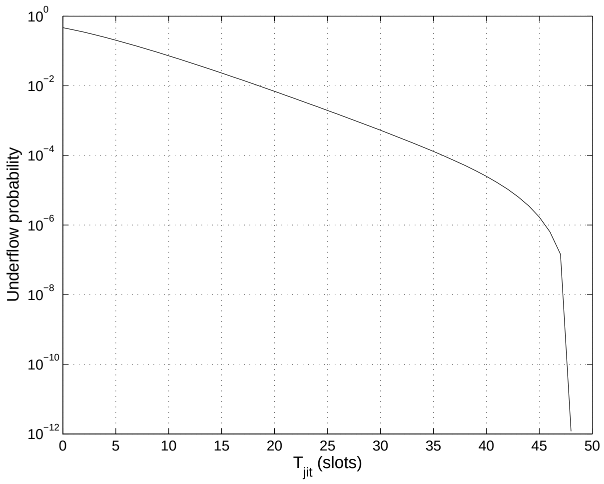

Figure 4: Average underflow probabilities given T_{jit}

loss limit, the corresponding value for $T_{jit}+d^*$ is read from the curve. Remark that in general the delay d^* experienced by the first packet of the video stream is not known at the moment a value for T_{jit} should be set. So under the worst case assumption that the first packet of the video stream experiences a delay $d^* = delay_{\min}$, the value for T_{jit} to be used is thus found.

The curve shown in Figure 4 gives the average underflow probability that would be obtained in a network with certain delay characteristics, when the same value of T_{jit} would be used in the dejittering process of all users receiving a video stream with the considered characteristics. Users for which the first packet of the video stream experiences a delay larger than $delay_{\min}$ will get a packet loss probability due to underflow smaller than the guaranteed one. As a consequence, the average underflow probability will be smaller than the guaranteed one.

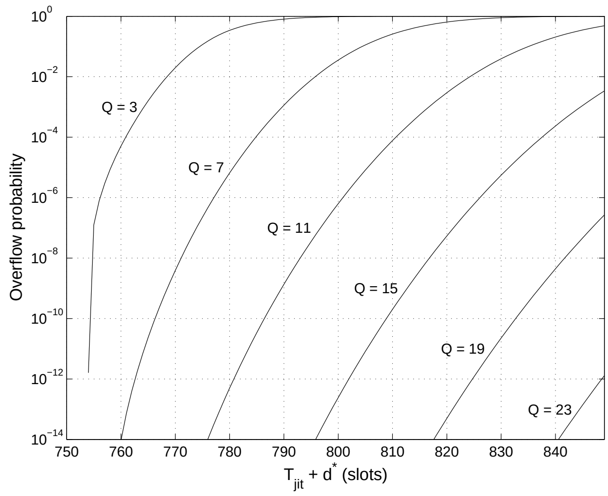

Figure 5: Overflow probabilities given T_{jit} and d^*, for different values of the dejittering buffer size Q

Figure 5 and 6 depict the overflow probabilities obtained for some values of Q, the size of the dejittering buffer. Figure 5 shows overflow probabilities given $T_{jit} + d^*$, while in Figure 6 average overflow probabilities given T_{jit} are seen.

Again the results given $T_{jit} + d^*$ can be used for dimensioning purposes, this time of the capacity Q the dejittering buffer should have to ensure that the packet loss probability due to overflow of the dejittering buffer stays below a guaranteed limit. As can be seen on Figure 5, for a target limit on the overflow probability and under the worst case assumption that d^* equals $delay_{\max}$, several valid combinations (T_{jit}, Q) are found. In practice however, one wants to guarantee a limit on both the packet loss due to overflow and due to underflow, which makes that only one combination (T_{jit}, Q) is considered as optimal, namely the one with the smallest value for T_{jit} that guarantees the limit on the underflow probability. Remark further that a guarantee ϵ for the underflow probability and a guarantee δ for the overflow probability together give a guarantee $\epsilon + \delta(1-\epsilon)$ on the general packet loss probability introduced by the dejittering process. So to guarantee a general loss probability of for example 1e-4, both ϵ and δ should be chosen smaller than 1e-4, in such a way that $\epsilon + \delta(1-\epsilon) =$ 1e-4.

For a fixed value of Q, overflow probability results as shown in Figure 6 are the average overflow probabilities that are obtained in a network when a certain value for T_{jit} is used in the dejittering process of all users. Users for which the first packet of the video stream experiences a delay smaller than $delay_{\max}$ will get an overflow probability smaller than the guaranteed one. Remark that for typical delay distributions, the probability that a (first) packet will ex-

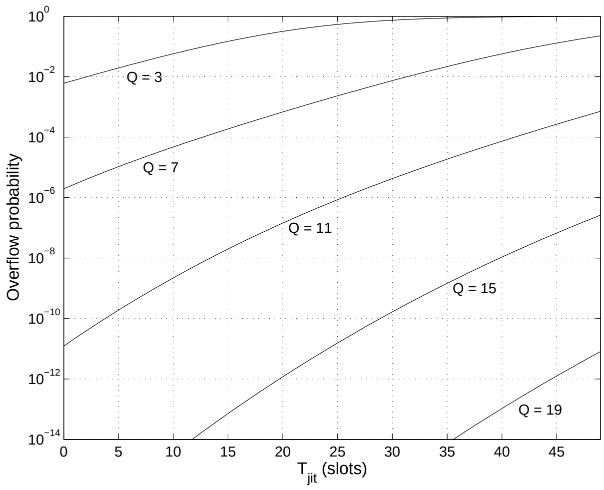

Figure 6: Average overflow probabilities given T_{jit}, for different values of the dejittering buffer size Q

perience the maximal delay is very small, so the average overflow probability will be considerably smaller than the guaranteed one which is based on the worst case delay the first packet of a video stream can experience.

CONCLUSIONS

In this paper a Markov model was developed that can be used to dimension the two important parameters of the dejittering process of variable bit rate video streams: the dejittering delay T_{jit} the dejittering process should wait before starting the playout of the first packet, and the capacity Q of the dejittering buffer. The model takes two probability distributions as input: that of the interarrival time of packets of the variable bit rate stream at the network, and that of the network delay. For every chosen value of $T_{jit} + d^*$, where d^* is the network delay that is experienced by the first packet of the video stream, a Markov chain with a different set of states $\mathcal{C}(T_{jit}+d^*)$ was defined. From the stationary distribution of such a Markov chain, the dejittering buffer underflow and overflow probability corresponding to $T_{jit}+d^*$ is then calculated. By weighting the results with the probability of all possible values d^* can take, underflow and overflow probabilities in function of T_{jit} are derived. To illustrate its calculability, the model was applied to a numerical example, where the delay distribution used in this example is rather typical.

In the video source considered in this paper, no temporal dependency among the interarrival times at the network of succeeding packets was assumed. As was shown, the overflow probability results depend on the probability of having more than x interarrival times in a time interval of a certain length. As a consequence, it is to be expected that a larger buffer capacity Q will be needed with bursty video sources, because then there will be periods with low arrival rates (large interarrival times) followed by periods with higher arrival rates (small interarrival times), and the buffer size Q should be large enough to capture a possible burst of early arrivals. The exact implications of bursty arrival processes on the dimensioning of the dejittering buffer are for further study.

Another area for further study is the use and modelling of dynamic or adaptive dejittering techiques (ETSI, 2003; Narbutt and Murphy, 2001). Dejittering considered in this paper is static dejittering, since once the decision about the value of the dejittering delay T_{jit} is taken at the establishment of the connection, it is never reviewed during the lifetime of the video stream. Dynamic dejittering techniques adjust the choice about the dejittering delay from time to time. These techniques become more common for voice streams, where the dejittering delay is typically changed at a new talkspurt, to limit the audible distortion. For video the instants at which such a change could be made are less clear. In (Laoutaris et al., 2004) dynamic dejittering for constant bit rate video traffic is considered. For variable bit rate video it is however more complicated to estimate on time which part of the variability in the arrival rate of the packets at the receiver is due to network delay variations, and which part is due to the variable character of the video stream.

REFERENCES

Apostolopoulos, J., Tan, W., and Wee, S. (2002). Video streaming: Concepts, algorithms and systems. Technical Report HPL-2002-260, HP Laboratories Palo Alto.

Bovy, C., Mertodimedjo, H., van Mieghem, P., and Uijterwaal, H. (2002). Analysis of end-to-end delay measurements in the internet. RIPE41.

Corlett, A., Pullin, D., and Sargood, S. (2002). Statistics of one-way internet packet delay. 53rd IETF.

ETSI (2003). Telecommunications and Internet protocol harmonization over networks (TIPHON) release 4; end-to-end quality of service in TIPHON systems; part 7: Design guide for elements of a TIPHON connection from an end-to-end speech transmission performance point of view. Technical Report ETSI TR 102 024-7 V4.1.1.

Jiang, W. and Schulzrinne, H. (2000). Modeling of packet loss and delay and their effect on real-time multimedia service quality. In *Proceedings of the International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV)*, Chapel Hill, NC.

Laoutaris, N., Van Houdt, B., and Stavrakakis, I. (2004). Optimization of a packet video receiver under different levels of delay jitter: an analytical approach. *Performance Evaluation*, 55:251–275.

Latouche, G., Jacobs, P., and Gaver, D. (1984). Finite Markov chain models skip-free in one direction. *Naval Research Logistics Quarterly*, 31:571–588.

Narbutt, M. and Murphy, L. (2001). Adaptive playout buffering for audio/video transmission. In *Proceedings of the IEE 17th UK Teletraffic Symposium*, Dublin, Ireland.

Van Houdt, B., Lenin, R., and Blondia, C. (2003). The delay distribution of (im)patient customers in a discrete-time DMAP/PH/1 queue with age dependent service times. *Queueing Systems and its Applications (QUESTA)*, 45(1):59–73.

APPENDIX

The set of states $\mathcal{C}(T_{jit} + d^*)$, defined in (3), has the following properties:

1. $\mathcal{C}(T_{jit} + d^*)$ contains $\#\mathcal{D}.iat_{\max}$ elements.

2. From a state in $\mathcal{C}(T_{jit} + d^*)$ it is impossible to reach a state outside $\mathcal{C}(T_{jit} + d^*)$.

3. $\mathcal{C}(T_{jit} + d^*)$ always contains at least one state with $j = delay_{\max}$, namely the state $(0, iat_{\max} - 1, iat_{\max} - 1 + delay_{\max} - T_{jit} - d^*, delay_{\max})$ if $T_{jit} + d^* \leq iat_{\max} - 1 + delay_{\max}$, and the state $(T_{jit} + d^* - iat_{\max} + 1 - delay_{\max}, iat_{\max} - 1, delay_{\max})$ otherwise.

4. From all states in $\mathcal{C}(T_{jit} + d^*)$ it is possible to reach the state mentioned in property 3.

Namely,

1. $\mathcal{C}(T_{jit} + d^*)$ contains exactly one element for every combination (j, t), with $j \in \mathcal{D}$ and $t \in \{0, \ldots, iat_{\max} - 1\}$.

2. This is because all states reachable with one transition from a random state in $\mathcal{C}(T_{jit}+d^*)$ are again elements of $\mathcal{C}(T_{jit} + d^*)$, as is easily seen by considering all one-step transitions described in the dejittering process Section.

3. Trivial.

4. Consider a random state of $\mathcal{C}(T_{jit} + d^*)$. Such a state has one of the following forms: (a) $(0, t, t + j - T_{jit} - d^*, j)$ with $T_{jit} + d^* - j > 0$, (b) $(0, t, t+j-T_{jit}-d^*, j)$ with $T_{jit}+d^*-j \leq 0$, or (c) $(T_{jit} + d^* - t - j, t, j)$. After a finite number of transitions with probability 1, the following states are reached: $(T_{jit} + d^* - j, 0, j)$ in cases (a) and (c), and $(0, 0, j - T_{jit} - d^*, j)$ in case (b). Then with probability $(\mathbf{U}_s)_{j,delay_{\max}} > 0$, where $s = iat_{\max} - 1 - j + delay_{\max}$, the state mentioned in property 3 is reached.

Combining all these properties leads to the property that the $\#\mathcal{D}.iat_{\max}$ states of $\mathcal{C}(T_{jit} + d^*)$ form a Markov chain with exactly one irreducible closed set of states. This set contains all states of $\mathcal{C}(T_{jit} + d^*)$ that are reachable from the state mentioned in property 3. States not in this closed set are transient.

ACKNOWLEDGEMENTS

This work was carried out within the framework of the IWT (Flemish Institute for the promotion of Scientific and Technological Research in the Industry) project CHAMP.

AUTHOR BIOGRAPHIES

KATHLEEN SPAEY received her M.Sc. degree in mathematics and computer science and a PhD in science from the University of Antwerp (Belgium) in 1995 and 2002, respectively, and a post-graduate degree in Telecommunications in 1999. Since 1995 she is a member of the "Performance Analysis of Telecommunication Systems" (PATS) research group at the University of Anwerp, where she has been involved in many research projects. Her main research interest goes to mathematical models for the performance evaluation of traffic control mechanisms and Quality of Service measures in telecommunication systems.

CHRIS BLONDIA obtained his M.Sc. and PhD in Mathematics, both from the University of Ghent (Belgium) in 1977 and 1982 respectively. In 1983 he joined Philips Belgium, where he was a researcher between 1986 and 1991 in the Philips Research Laboratory Belgium (PRLB). Between August 1991 and end 1994 he was an Associate Professor in the Computer Science Department of the University of Nijmegen (The Netherlands). In 1995 he joined the Department of Mathematics and Computer Science of the University of Antwerp (UA), where he is currently a Professor and head of the research group "Performance Analysis of Telecommunication Systems" (PATS - www.pats.ua.ac.be). He is lecturing Telecommunication and Performance Evaluation courses. His main research interests are mathematical models for performance evaluation of computer and communication systems. The systems that are studied are related to traffic management in multi-service networks, mobility management in IP networks, access control in access networks, wired and wireless including satellite networks, routing in ad hoc networks and optical networks. He has published a substantial number of papers in international journals and conferences. He has been and is currently involved in many National and European Research Programs. He is editor of the Journal of Network and Computer Applications and of Computer Networks. The PATS research group is member of the Interdisciplinary institute for BroadBand Technology (IBBT) (www.ibbt.be).