Les Cahiers du GERAD ISSN: 0711–2440

Mean delay variation applicability for jitter buffer

H. Dbira, A. Girard,
B. Sansò

G–2014–91

November 2014

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Mean delay variation applicability for jitter buffer

Hadhami Dbira [a]
André Girard [b]
Brunilde Sansò [a]

[a] *GERAD & Electrical Engineering Department, Polytechnique Montréal, Montréal (Québec) Canada, H3C 3A7*

[b] *GERAD & Polytechnique Montréal & INRS-Énergie Matériaux Télécommunications, Montréal (Québec) Canada, H5A 1K6*

hadhami.dbira@polymtl.ca
andre.girard@gerad.ca
brunilde.sanso@polymtl.ca

November 2014

Les Cahiers du GERAD

G–2014–91

Abstract: Jitter buffering is a key component in multimedia and real-time services. A jitter buffer is used at the receiver side to control packet rate and to compensate for the delay variation, commonly called jitter, by its function of tradeoff between delay and loss. The packet loss at the jitter buffer can be induced by either underflow and overflow, and we note that buffer underflow probability is an important performance metric for the Quality of user Experience (QoE). Since the estimation of this parameter is difficult for most buffer models, we propose in this paper to show that by controlling the mean delay variation in the buffer, we could control the underflow probability. Thus, we provide an analytical approach for jitter delay estimation in jitter buffer.

Key Words: Delay variation, jitter buffer, overflow, underflow.

1 Introduction

The fast growth of the Internet is driven by the wide spread of multimedia streaming applications and real-time services such as television over IP, video conferencing, video on demand, telemedicine, etc. For these services, the users' satisfaction is expressed by the quality of experience (QoE) which reflects the human perception of the service quality. The most important measure of QoE [1] seems to be service interruption, where the screen freezes to some fixed image for some time, but other measures have been proposed as well [2, 3] such as the start-up delay before actual transmission or simply poor image quality.

The need for good QoE can be met by implementing strict quality of service (QoS) requirements within the network to avoid impairments that affect the service quality like packet loss, delay and delay variation, also known as *delay jitter*. In particular, multimedia traffic is very sensitive to the jitter and an important solution to control jitter is to store the received packets in a buffer commonly called the *playout* or *jitter buffer*. In that context, a service interruption happens when the client buffer is empty because the network cannot deliver packets fast enough for some time. This is know as a buffer underflow, or buffer starvation. Another cause of QoE degradation is when the buffer is full and packets are dropped, which leads to picture degradation such as pixelation. This is called *buffer overflow* and is generally considered less annoying that an actual freeze of the media stream.

This is why the buffer over- and underflow probability are important parameters used to design and dimension jitter buffers [4]. Playout buffer control is usually based on the packet delay distribution, the end-to-end delay variance or on some form of perceptual evaluation. For instance, the work of [3, 5] has shown how minimizing the underflow probability can be used to manage the jitter buffer parameters. For this reason, calculating the probability of buffer underflow is an important issue. Unfortunately, this is very difficult and there exist relatively little work on this topic and then only for simple systems.

As an example, a full characterization of the buffer underflow distribution was provided by [3] when the jitter buffer is modeled by a M/D/1 queue. On the other hand, there is a large body of work [6, 7] showing that in many cases, the packet arrival process on the receiver side is definitely not Poisson so that modeling the jitter buffer with an M/D/1 queue may not be very accurate. Still, there does not seem to be any method to compute buffer underflow for more realistic arrival processes.

Given the difficulty of computing the underflow probability, it has been suggested [8] to use the packet jitter as the QoS measure to use for controlling the playback buffer. The main reason why this approach may be more tractable than computing the underflow and overflow probability is based on [9] where some fast and accurate approximations were presented for computing the jitter in a $G/G/1/\infty$ queue for some interesting traffic processes.

Still, there are two main problems with using this jitter model for buffer control. One is the fact that jitter buffers are finite so that the first question is to what extent we can approximate the jitter with a model designed for infinite buffer. The other issue is how the jitter is actually related to the underflow probability. Although it is intuitively clear that there is a connection, it is possible to give counter-examples in some special cases where there is very little correlation between the two.

This is why we need to look more closely to the relationship between the mean delay variation and underflow probability and evaluate to what extend the mean delay jitter could help for controlling the over- and underflow.

First we describe in Section 2 the jitter buffer model we use in this study. Next, we present some definitions of jitter in Section 3 and we recall the results of [9] for the fast calculation of jitter. The first set of results in Section 4 shows that the infinite buffer model can be quite accurate even for relatively small buffers. The main result of this paper is described in Section 5 where we present some simulation results showing that the mean delay jitter is well correlated to the underflow and overflow probability for realistic non-Poisson arrival processes. Finally, concluding remarks are discussed in Section 6.

2 Jitter buffer model

In this paper, we consider the jitter buffer of a particular application. We assume that the interarrival distribution of the packets arriving at the buffer follows a general distribution. Next, we consider that the packet service time is constant. This is not unrealistic for video traffic since the B blocks for a standard MP4 codec, for instance, are much larger than the standard 1500-byte MTU, so that most of the packets are of this length. Also, there is some evidence [10] that a large portion of internet traffic is made up of packets of constant sizes with only a few specific values. We also assume that there is room for K packets in the system, i.e., K denotes places used *both* for waiting and in-service packets. We also assume a FCFS service discipline so that the jitter buffer is modeled as a G/D/1/K queue.

Recall that we are considering the playback buffer for a single application. This means that the packet generation rate at the source codec must be equal to that of the destination server, at least in the long term. This is why we assume in what follows that the service rate of the queue μ is equal to the traffic arrival rate λ, so that the traffic load $\rho = \lambda/\mu = 1$.

From this model we can define the jitter buffer over- and underflow. An overflow occurs when we have a full buffer upon the arrival of a new packet, and an underflow when packet arrive to destination and finds the buffer empty.

3 Mean delay variation model

We now present some definitions that have been proposed for jitter and recall some previous results for fast numerical algorithms showing how we can compute jitter quickly for some queues of interest.

3.1 Some definitions of jitter

Delay variation is a measure of the variation of delay in a sequence of packets over time. It is one of the most important parameters characterizing network transmission and QoS. There are different reasons for jitter but the most prominent cause is queuing within the network.

The definition of jitter differs according to different standards organizations. The International Telecommunication Union—Telecommunication Standardization Sector (ITU-T) defines jitter as the variation of delay from some reference value like the minimum or mean delay [11, 12]. It is difference between the one-way delay of a packet i, T_i, and the local reference delay, a_i

$$J = E\left[|T_i - a_i|\right]. \tag{1}$$

Using this definition, it is possible to detect quickly an increase in packet delay and for this reason, it increasing, this definition is widely used for computing the jitter buffer size [8].

The Internet Engineering Task Force (IETF), on the other hand, defines delay variation by the difference in successive transit time between two measurement points [13]. Let T_i and T_{i+1} be the transit time of consecutive packets. The average delay jitter is then given by the expected absolute value of this random variable

$$J = E\left[|T_i - T_{i+1}|\right]. \tag{2}$$

This metric is suitable for design and QoS routing because there are accurate and fast computation models for estimating its value [14, 15, 9]. This is why we will be using it in this paper. The fact that there are an analytical model for jitter in some general type of queue presents an important aspect for studying the applicability of this parameter in jitter buffers and its effect on de-jittering performance.

3.2 No waiting

We can give the exact expressions for the jitter with finite buffer queues when $K = 1$, i.e., when there is no waiting space at all. In that case, arriving packets cannot wait so that the transit time of a packet is simply its service time so that $T_i = S_i$ and we get

$$\begin{aligned} J &= |T_i - T_{i-1}| \\ &= |S_i - S_{i-1}|. \end{aligned} \tag{3}$$

This is the same value that we get for the $K = \infty$ queue when $\rho \to 0$. Here, on the other hand, ρ can be arbitrarily large. This is not really surprising since in both cases, packets don't wait, although for different reasons.

3.3 Jitter for the G/D/1/∞ queue

We now recall the result of [9] for an approximate model to compute jitter in a G/D/1 queue. This is based on a very important proposition in [9] that could be applied to G/D/1 queue at $\rho \simeq 1$:

Proposition 1 *For a G/G/1 queue, in the high traffic limit, the jitter is given by the expectation of the absolute value of the difference between the service and inter-arrival times of a packet.*

To see this, we use the definitions

W_i	Waiting time of packet i
S_i	Service time of packet i
T_i	Transit time of packet i, $T_i = W_i + S_i$
R_i	Interarrival time of packet i
τ_i	Instant just before packet i arrives to the queue
$W_i(\tau_i)$	Workload at time τ_i

The assumption of heavy traffic in [9] is equivalent to the statement that for $\rho \to 1$, when packet i arrives to the queue, packet $i-1$ has not yet started service. We then have

$$W_i(\tau_i) = W_{i-1}(\tau_{i-1}) - R_i + S_{i-1}. \tag{4}$$

From which we get

$$\lim_{\rho \to 1} J = E\left[|S_i - R_i|\right]. \tag{5}$$

If we define the random variable $Z = S_i - R_i$, we get the average (5) as

$$E[|Z|] = \int_0^\infty z f_{|Z|}(z) dz \tag{6}$$

where $f_{|Z|}(z)$ is the probability density function of $|Z|$ and f_R and f_S that of R and S. We get

$$f_{|Z|}(z) = \int_0^\infty f_S(z+y) f_R(y) dy + \int_0^\infty f_S(y) f_R(z+y) dy. \tag{7}$$

We see that using (6–6), we can compute the jitter either with and analytic expression or by numerical integration.

Next, we can get a simplified expression for J for the G/D/1 queue. Here, the service time S follows a deterministic distribution

$$f_S(y) = \delta(y - \overline{S}) \tag{8}$$

where $\overline{S} = 1/\mu$ is the mean service time. From (7), we get

$$E\left[|Z|\right] = \int_0^\infty z f_R(z + \overline{S}) + \int_0^{\overline{S}} z f_R(\overline{S} - z) \tag{9}$$

which can be evaluated analytically in some cases and computed reasonably quickly by numerical integration when this is not possible

4 Accuracy for finite buffer

The queuing model of Section 3.3 is very close to the jitter buffer system described in Section 2 except that it has an infinite capacity. In this section, we want to see to what extent we can use the model given in [9] to approximate a finite buffer of size $K-1$. It is clear that the approximation will be good when K is large so that the real question is the accuracy when K becomes small.

We use simulations to compare the jitter for a G/D/1/K queue with results computed by (9) for an infinite queue. We present results for a gamma interarrival process in Figure 1 and for a lognormal in Figure 2. Analytic results, labeled `Analytic`, are computed for $\lambda \simeq \mu = 1$ and we vary the standard deviation of the interarrival distribution. The same conditions are applied for simulation results, labeled `Sim`, where we change in each simulation the buffer capacity $K-1$.

To summarize, we see that the G/D/1/∞ model becomes approximately close to the curve for a finite capacity when $K \approx 120$ so that we could set the capacity threshold $K-1$ at 120. Moreover, the results from Figures 1 and 2 show that the accuracy is quite good for values of K as low as 20. It is known [16] that jitter buffers usually have a large size so that we conclude that the model (9) is useful for estimating mean delay variation in a jitter buffer of limited capacity.

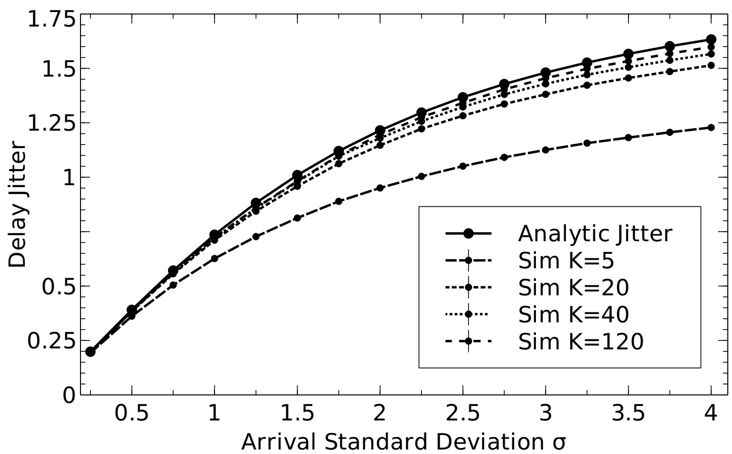

Figure 1: Mean delay variation: Gm/D/1 vs Gm/D/1/K

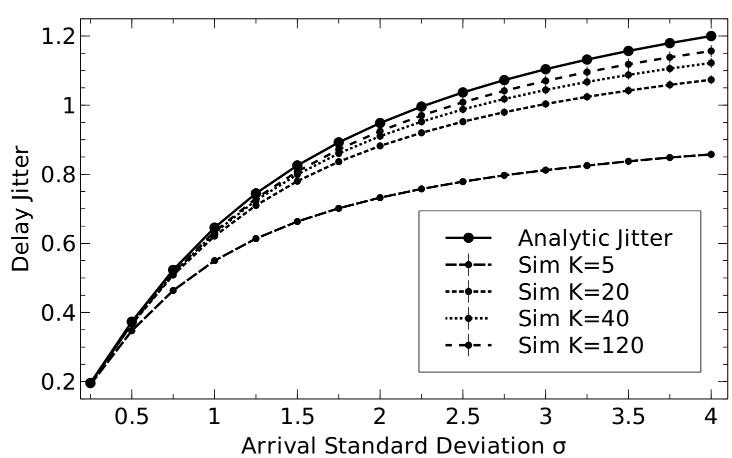

Figure 2: Mean delay variation: LogN/D/1 and LogN/D/1/K

5 Correlation with probabilities

In this section, we look into the relationship between the jitter and the buffer over- and underflow probability. Although it is intuitively clear that there should be some connection, it is possible to construct limit cases where there is actually none. For instance, in a D/D/1/K queue, the jitter is always 0 but one of the overflow or underflow probabilities will always be 1 depending on the actual value of ρ. In that case, jitter is not a good replacement for the probabilities. As another case, consider the M/M/1/K buffer operating at $\rho = 1$. It is known that in this case, the packet loss probability is given by

$$P_l = \frac{1}{K+1}. \tag{10}$$

However, we can see from the simulation results in Figure 3 that the jitter measured by (2) is constant and cannot be a good estimation for loss probabilities. Note also that the simulation results show that the jitter is actually the same as that for the M/M/1/∞ queue for the whole range of buffer sizes.

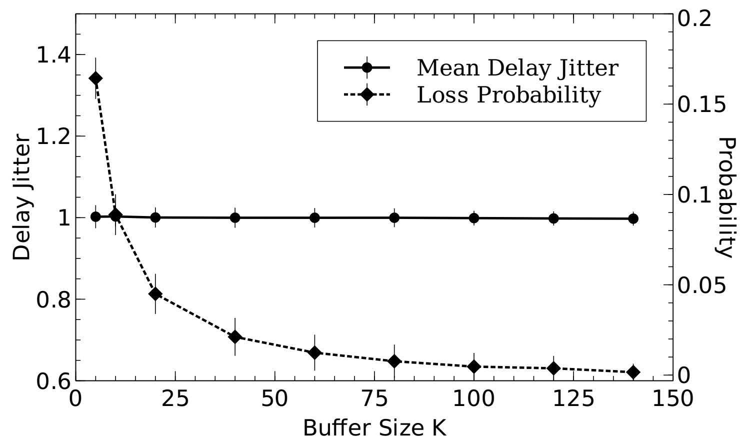

Figure 3: M/M/1/K buffer

In what follows, we present simulation results indicating that there is in fact a rather strong correlation between jitter and probabilities for realistic arrival processes. The underflow probability is defined as the ratio of the server idle periods to the total simulation time and the overflow probability is the ratio of dropped packets to the total number of generated packets.

5.1 Small buffer

We consider first a Gm/D/1/K queue where the interarrival time follows a Gamma distribution with $K = 5$ for a fixed packet size of 1500 bytes. This buffer operates at $\rho = 1$ with a single flow. Results are given in Figure 4 and expressed in service time unit. We present in this figure the jitter and over- and underflow probabilities produced by simulations. The 95% confidence interval are also shown in the figure. The x axis presents the standard deviation of the inter-arrival time distribution σ. The vertical axis on the left refers to the jitter values and the right one is for loss and overflow probabilities.

Similar results are shown in Figure 5 for a LogN/D/1/K queue where the interarrival variable follows a LogNormal distribution with the same simulation conditions. It can be seen from Figures 4 and 5 that mean delay jitter in this queue is correlated with the over- and underflow probabilities. The correlation coefficient for the Gm/D/1/K is about 0.9603 and 0.9525 for the LogN/D/1/K.

5.2 Large buffer

Actually, most real-time applications use a large jitter buffer. We repeat the simulations in Subsection 5.1 but for $K = 140$. Figures 6 and 7 show simulation results respectively for a Gm/D/1/140 and LogN/D/1/140

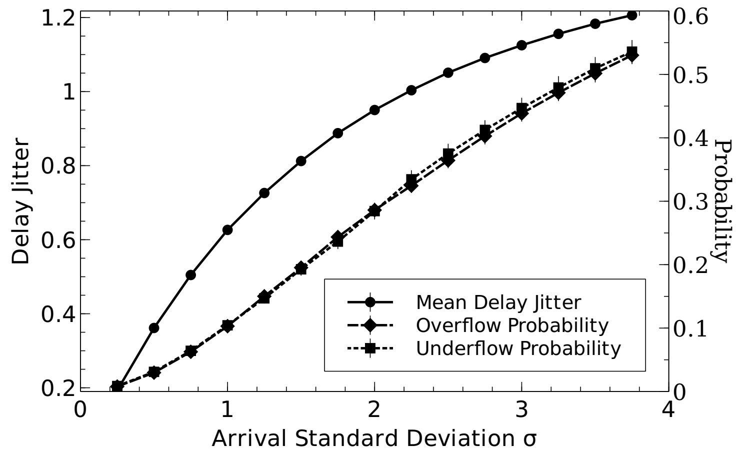

Figure 4: Gm/D/1/5-small buffer

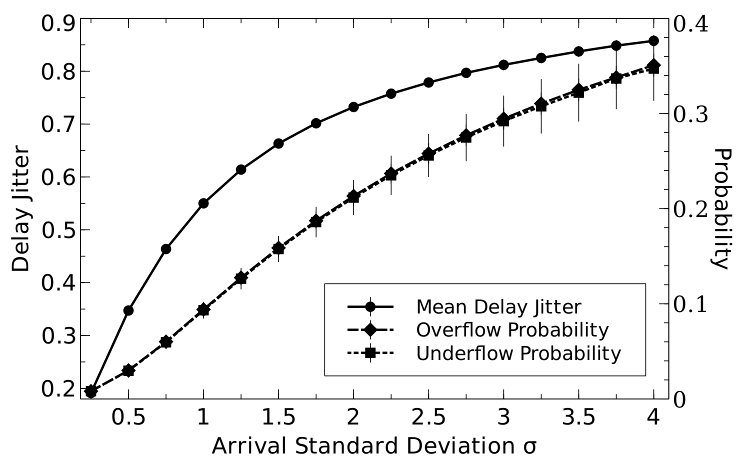

Figure 5: LogN/D/1/5-small buffer

buffers. Here too we find that the jitter is correlated to the probabilities. The correlation coefficient for the Gamma distribution is 0.9236 and 0.9323 for the Lognormal distribution.

To summarize, we see that the jitter is correlated with loss probabilities in a jitter buffer. The correlation is quite good for small buffer and somewhat less strong for a large buffer. More importantly, we see that the jitter presents an upper bound for the underflow probability. This facts make the control of underflow probability possible in a jitter buffer modeled by a G/D/1/K queue.

6 Conclusion

We found that we can approximate the jitter in a G/D/1/K queue by an infinite-buffer model for buffer sizes as low as 20, a value which is appropriate for a playout buffer. We have examined the relation between the loss probabilities in a jitter buffer modeled by a G/D/1/K queue with the jitter. We found that those parameters are reasonably well correlated, especially for small buffers. We are consequently able to control the underflow probability through the mean delay variation. Then, we conclude that estimating delay jitter analytically by means of (2) in a jitter buffer is a good solution for quantifying the underflow in the same buffer.

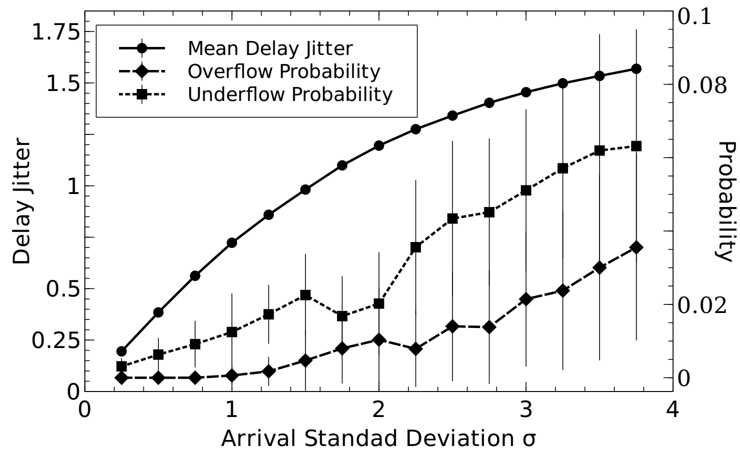

Figure 6: Gm/D/1/140-large buffer

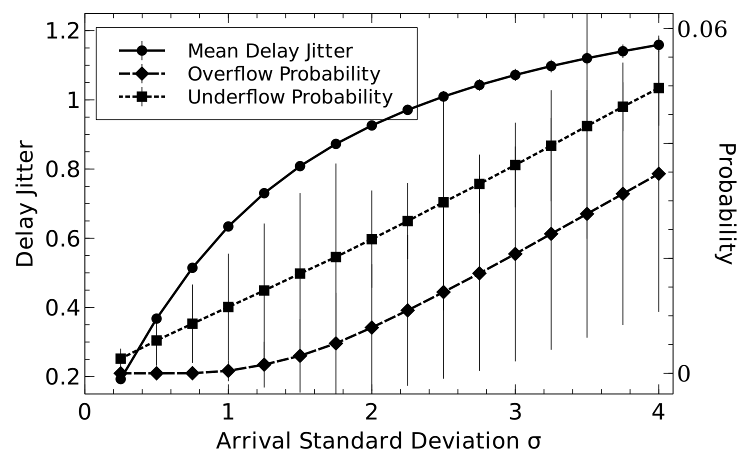

Figure 7: LogN/D/1/140-large buffer

References

[1] A. ParandehGheibi, M. Médard, A. Ozdaglar, and S. Shakkotai, Avoiding interruptions–a QoE reliability function for streaming media applications, IEEE Journal on Selected Areas in communications, 29(5), 1064–1074, May 2011.

[2] M. Li, QoE-Based performance evaluation for adaptive media playout systems, Advances in Multimedia, 2013, Jan. 2013. [Online]. Available: http://dx.doi.org/10.1155/2013/152359

[3] Y. Xu, E. Altman, R. El-Azouzi, M. Haddad, S. Elayoubi, and T. Jimenez, Analysis of buffer stravation with application to objective QoE optimization of streaming services, IEEE Transactions on Multimedia, 16, 813–827, Apr. 2014.

[4] B. Steyaert, K. Laevens, D.D. Vleeschauwer, and H. Bruneel, Analysis and design of a playout buffer for VBR streaming video, Annals of Operations Research, 162, 159–169, 2008.

[5] J. Seo, V. Leung, and H. Lee, Optimizing a playout buffer with queueing performance metrics for one-way streaming video, in Proc. Globecom. IEEE, Nov. 2008, 1–6.

[6] T. Luan, L. Cai, and X. Shen, Impact of network dynamics on user's video quality: Analytical framework and QoS provision, IEEE Transactions on Multimedia, 12(1), 64–78, 2010.

[7] M. Asefi, J.W. Mark, and X. Shen, A mobility-aware and quality-driven retransmission limit adaptation scheme for video streaming over VANETs, IEEE Transactions on Wireless Communications, 11(5), 1817–1827, May 2012.

[8] A. Morton and G. Claise, RFC 5481: Packet delay variation applicability statement, IETF, Mar. 2009.

[9] H. Dbira, A. Girard, and B. Sansò, Calculation of packet jitter for non-Poisson traffic, Submitted to Annals of Telecommunications, Mar. 2014.

[10] C. Fraleigh, S. Moon, B. Lyles, C. Cotton, M. Khan, D. Moll, R. Rockell, T. Seely, and S. Diot, Packet-level traffic measurements from the Sprint IP backbone, Network, 17(6), 6–16, 2003.

[11] Internet protocol data communication service–IP packet transfer and availability performance parameters, ITU-T Recommendation Y.1540, Mar. 2011. [Online]. Available: https://www.itu.int/rec/T-REC-Y.1540-201103-I/en

[12] A. Clark, Analysis, measurement and modeling of jitter, Telchemy Incorporated, USA, Tech. Rep., Jan. 2003. [Online]. Available: https://www.telchemy.com/reference/ITUSG12_JitterAnalysis.pdf

[13] C. Demichelis and P. Chimento, RFC 3393: IP packet delay variation metric for IP performance metrics (IPPM), IETF, Nov. 2002.

[14] H. Dahmouni, A. Girard, and B. Sansò, An analytical model for jitter in IP networks, Annals of Telecommunications, 67, 81–90, Jan. 2012. [Online]. Available: http://www.springerlink.com/content/5046104335872460/

[15] H. Dahmouni, A. Girard, M. Ouzineb, and B. Sansò, The impact of jitter on traffic flow optimization in communication networks, IEEE Transactions on Network and Service Management, 9(3), 279–292, Sep. 2012.

[16] C.-C. Wu, K.-T. Chen, C.-Y. Huang, and C.-L. Lei, An empirical evaluation of VoIP playout buffer dimensioning in Skype, Google Talk, and MSN Messenger, in Proceedings of the 18th International Workshop on Network and Operating Systems Support for Digital Audio and Video, ser. NOSSDAV '09. New York, NY, USA: ACM, 2009, 97–102. [Online]. Available: http://doi.acm.org/10.1145/1542245.1542268