Packet Dispersion and the Quality of Voice over IP Applications in IP networks

Haim Zlatokrilov, Hanoch Levy
School of Computer Science
Tel Aviv University
Tel Aviv, Israel

Abstract - Next Generation Networks (NGN) and the migration towards IP networks is likely to make the IP technology the main vehicle for carrying voice and video calls on modern networks. Packet dispersion is a mechanism by which the packets of a certain session are dispersed over multiple paths, in contrast to the traditional approach by which they follow a single path most of the time. In this work we examine the quality of Voice over IP (VoIP) applications and the effects of packet dispersion on it. We focus on the effect of the network loss on the applications, where we propose to use Noticeable Loss Rate (NLR) as a measure correlated with the voice quality. We analyze the NLR for various packet dispersion strategies over paths experiencing memory-less (Bernoulli) or bursty (Gilbert model) losses, and compare them to each other. Our analysis reveals, that in many situations, in particular for most cases where losses are bursty, the use of packet dispersion reduces the NLR and thus improves session quality. The results suggest that the use of packet dispersion can be quite beneficial for these applications.

Keywords -- Stochastic processes; packet dispersion; noticeable loss rate; voice over IP quality; bursty losses

I. INTRODUCTION

Next Generation Networks and today’s migration towards IP based networks is likely to make these networks the main infrastructure for carrying Voice and Video applications. A major issue that needs to be solved to make this migration successful is that of the required quality of the applications over the “best effort” IP network.

Packet dispersion in IP networks is a mechanism in which application packets are dispersed between parallel paths leading from the source to the destination, based on a predefined dispersion strategy. Packet dispersion can be implemented by the source application (e.g. by using source routing or other techniques) or by nodes in the network (e.g. multi-homing devices or Content Delivery Network companies, such as Akamai, that use edge architecture to achieve load-balancing and improved network utilization). In this study we examine the quality of real time stream-oriented applications, in particular VoIP, in light of packet dispersion. The aim of this study is to examine whether packet dispersion can be used as a machinery to improve QoS of VoIP applications under known network statistical characteristics and to examine the effect of dispersion conducted in multi-homing architectures on VoIP quality.

Traffic dispersion techniques are used in many technologies, for a variety of reasons. In CDMA radio networks, traffic dispersion (also called frequency-hopping) is used for security reasons and in order to statistically multiplex noises. In [6][23] traffic dispersion in IP networks is suggested to reduce traffic burstiness and therefore achieve higher resource utilization. Another idea, proposed in [12], suggests using traffic dispersion as a better method to Forward Error Correction (FEC) technique for voice over IP (VoIP) applications. Traffic dispersion is implemented de-facto in IP networks for load balancing purposes.

Several factors affect the quality of VoIP applications. One can divide them into two classes, the underlying network behavior and the technology built-in mechanisms, such as codec type, Packet Loss Concealment (PLC) mechanisms and Forward Error Correction (FEC). Our focus is on the network behavior, which is usually measured in three measures: Packet loss, delay and jitter (delay variance). Clearly, as these measures grow, quality degrades. However, the acceptable delay, for bi-directional real-time streaming applications, is usually limited by values of 200-250 milli-seconds. For this reason, both delay and jitter can be roughly translated, physically and mathematically, into a loss measure, since late packets arriving at the destination are not useable and can be counted as lost. We therefore will concentrate in this study on the packet loss experienced by a session, regardless of the cause of the loss (whether a real network loss or a dropped packet due to late arrival).

Perceptual studies of applications such as IP phones have shown that user dissatisfaction rises dramatically in presence of bursty losses. Average packet loss rate property, as shown in many studies, is not enough to capture the effect of network behavior on VoIP applications. For better quality evaluation one should also take into account loss burstiness and recency effects. Taking these together with the technology built-in mechanisms can lead to a good estimation of VoIP application quality, as suggested in the E-model [7][22]. One intrinsic property shown in these studies is that bursty losses degrade voice quality. Due to these properties we conclude that in many situations the packet Loss-Rate measure should be replaced by the Noticeable-Loss-Rate (NLR) measure [14] as the basic ingredient in computing the perceived quality of

1 This work was partially supported by Israeli Science Foundation grant 235/03.
2 Since we focus our discussion on IP networks, we will use the term Packet Dispersion instead of the general term Traffic Dispersion.
VoIP applications. The NLR metrics counts losses of ‘close’ packets and ignores losses of distant packets. Based on [22] we value the NLR as a metrics well correlated with perceived voice quality (the lower the NLR the better the quality). Therefore, in this work we focus on the NLR experienced by VoIP sessions.

The analysis in this work is based on assuming that the losses experienced in the network are either memory-less (Bernoulli) or bursty (following the Gilbert loss model). Though the Bernoulli loss model is a special case of the Gilbert model, we start the analysis with the Bernoulli model, as to simplify the exposition. Our analysis provides the mathematical machinery needed for computing the NLR experienced by the sessions in these systems. Despite the fact that the dimension of problem addressed is very large (exponential state space) the results are formulated in expressions whose computational complexity is very small (linear). Thus, using our analysis, one can easily compute the NLR of a given network scenario.

Examining several common data-driven packet dispersion strategies using the Bernoulli loss model, we demonstrate that packet dispersion reduces the NLR in many practical cases. An examination of the NLR under bursty losses leads to the conclusion that in many cases packets dispersion can highly reduce NLR, though in some other cases, depending on path characteristics, there are opposite results. The formulae derived as well as the cases examined in the paper can be used in the process of network design and traffic engineering where dispersion is applied.

Though the results show that packet dispersion is beneficial in many cases for VoIP, one should be aware of the fact that packet dispersion may have some side effects and may cause other network problems (e.g. out of order packets), which may harm other applications. Thus, technologies implementing packet dispersion should take into consideration the specific application requirements, network conditions over the routes and the dispersion strategies for overall enhanced network performance. It is worthwhile to mention that traffic dispersion can also be used for QoS differentiation and enhanced network utilization purposes over asymmetric paths.

The structure of this work is as follows: In Section II we discuss the modeling considerations of this work, present the underlying assumptions of our model, and introduce the Noticeable Loss Rate model adopted from [14]. We then turn into mathematical analysis of packet dispersion strategies under the Bernoulli loss model (Section III) and under the Gilbert model (section IV). For both loss models, we first analyze the NLR experienced by a session traversing a single path (no-dispersion), as is typically the case in traditional networks. We then turn to analyze the NLR as experienced in multi-path environment, and examine two typical packet dispersion scheduling policies: i) The memory-less random packet scheduling, in which the paths taken for the packets of a stream are chosen using a memory-less probabilistic mechanism (selection from a predefined set of paths), and ii) The periodic packet scheduling in which the paths taken for a packet stream are selected according to some periodic order; a common special case of the latter scheduling is the Round-Robin scheduling. Having analyzed these systems we then compare them to each other and bring numerical results to support our findings.

II. MODELING ASSUMPTIONS, MODEL AND NOTATIONS

A. Voice quality, the factors affecting it and its evaluation

Traditionally, voice perceived quality is measured by the Mean Opinion Score (MOS) or by mechanical techniques such as PESQ [10] and PSQM [9]. Another non-intrusive monitoring technique for VoIP, incorporating the effects of time varying packets loss and “recency”, based on the ES-model [7] is suggested in [22].

There are many factors affecting voice quality in VoIP applications. In general, one can divide these factors into application factors (e.g. codec type, jitter buffer implementation, etc.) and network performance factors: delay, jitter and loss. The techniques suggested in [22] propose that given the codec type and other application parameters, loss ($L_d$) and delay ($L_R$) impairments are the main factors affecting voice quality. From these impairments one can compute the gross score, called $R$ value, which can be mapped to MOS. The delay impairment causes relatively small effect as long as it is bounded within certain constraints (usually up to 250ms). Roughly speaking, this factor can be used to translate network delay into network loss by counting all the packets whose delay exceeds a certain threshold as lost packets. This results in network loss being the major network performance parameter affecting voice quality.

The average packet loss rate metrics alone is not enough to determine voice quality. The other factors, mentioned in [22], are the recency effect (the relative location of the lost frames, e.g. losses occurring at the end of the session significantly degrades perceived quality in comparison to losses occurring at the session beginning) and the loss burstiness (a packet is considered to be in a burst if less than $g_{min}$ packets have arrived since the previous packet was lost). Loss burstiness, having the greater impact, can reduce MOS in more than one grade (out of five) as shown in [12]. Perceptual studies, such as those referenced in [5], also support the fact that bursty losses may dramatically reduce perceptual quality, especially for audio.

Common VoIP manipulation techniques also increase the importance of bursty losses. First, in modern codecs internal Packet Loss Concealment (PLC, see [8]) algorithms are used to reduce the effect of packet loss on perceived quality. When a loss occurs the decoder derives the data of the lost frame from previous frames to conceal losses. A simple example of a PLC mechanism would be to use the last (properly arrived) packet to replace a lost packet. Some codec concealment mechanisms may be effective for a single lost packet, but not for consecutive losses or bursts of losses. Second, Forward Error Correction (FEC, see [20]) mechanisms are also used to compensate for lost packets by appending the information of

3 In [2] it claimed that given the loss rate, the performance of TCP applications improves when losses tend to appear in bursts. Meaning that the same effect of reducing burstiness that is beneficial for VoIP is bad for TCP.
previous voice frames to packet payload. Clearly, for this technique sequential losses decrease FEC efficiency and reduce voice quality.

We thus conclude that the loss rate and loss burstiness are the major network performance factors affecting voice quality and we focus on their performance. Next we define and discuss the Noticeable Loss Rate (NLR) as a measure for loss burstiness that is well correlated with voice perceived quality.

B. Noticeable Loss Rate (NLR)

The IP Performance Metrics (ippm) working group in the IETF has proposed a set of metrics for packet loss [14]. This includes loss constraint distance (i.e. the threshold for distance between two losses) and the Noticeable Loss Rate (NLR) metrics, which is the percentage of lost packets with loss distance smaller than the loss constraint distance\(^4\). In VoIP applications the loss constraint distance is usually related to the convergence time of the decoder. Clearly, the perceived voice quality decreases with the NLR.

1) A Definition of NLR

The Loss Distance is defined (as in [14]) as the difference in sequence numbers between two successively lost packets. The loss event of a packet is defined to be “a \(\delta\) noticeable loss” event (and is denoted as \(NL^{(\delta)}\)), if the loss distance between the lost packet and the previously lost packet is no greater than \(\delta\), where \(\delta\), a positive integer, is the loss constraint. In order to measure how ‘noticeable’ a loss might be for quality purposes, the loss distance \(\delta\) may be selected to be equal to \(g_{\text{min}}\) (the parameter used in [22], typically \(g_{\text{min}} = 16\)), determining whether a packet belongs to a burst. Alternatively, small values of \(\delta\) can be used when FEC or PLC mechanisms are enabled.

Below we will define the Noticeable Loss Rate (NLR) as the fraction of all packets, which are noticeable loss packets. This definition agrees with, but slightly deviates from, the NLR metrics ‘Type-P-one-Way-Loss-Distance-Stream’ defined in [14]. Where necessary we will associate the parameter \(\delta\) with the notion of noticeable loss rate, reading \(\delta\) - noticeable loss rate, or \(NL^{(\delta)}\).

The loss indicator function for a certain flow reflects the loss event of packet \(t\):

\[
l(t) = \begin{cases} 
1 & \text{if packet } t \text{ is lost} \\
0 & \text{Otherwise} 
\end{cases} 
\]  

(1)

The event that packet \(k\) in session \(i\) is a noticeable loss with loss constraint \(\delta\), is denoted by indicator function \(NL^{(\delta)}_{i}(k)\):

\[
NL^{(\delta)}_{i}(k) = \begin{cases} 
1 & \text{if } l(k) = 1 \text{ and } \exists t \in [k - \delta, k - 1] \text{ where } l(t) = 1 \\
0 & \text{otherwise} 
\end{cases} 
\]  

(2)

The noticeable loss rate for session \(i\) with loss constraint \(\delta\), and for a sequence of \(K\) packets, is then given by:

\[
NL^{(\delta)}_{i}(K) = \frac{1}{K} \sum_{k=1}^{K} NL^{(\delta)}_{i}(k) . 
\]  

(3)

Next we propose an alternative definition to that given in Eq. (2) for the noticeable loss event (\(NL^{(\delta)}(k)\)):

\[
NL^{(\delta)}(k) = \begin{cases} 
1 & l(k) = 1 \text{ and } \exists t \in [k + \delta, k + 1] \text{ where } l(t) = 1 \\
0 & \text{otherwise} 
\end{cases} 
\]  

(4)

Proposition 1: For any sequence of loss events, the number of noticeable loss events under the definitions (2) and (4) are identical to each other.

The proposition is proven by counting, under both definitions, the number of losses that are not noticeable and subtract them from the total number of losses.

In the analysis we analyze the system under the assumption of steady state. Thus, for a session of \(M\) packets we have:

\[
NL^{(\delta)}_{i} = \lim_{M \to \infty} \frac{1}{M} \sum_{k=1}^{M} NL^{(\delta)}_{i}(k) . 
\]  

(5)

That is, the NLR equals the steady state probability that a packet is a ‘noticeable loss’. In order to conduct a meaningful comparison in scenarios where multiple sessions are involved, we will evaluate the average \(NL^{(\delta)}\) taken over the \(N\) sessions, denoted \(\overline{NL^{(\delta)}}\), \(\overline{NL^{(\delta)}} = \frac{1}{N} \sum_{i=1}^{N} NL^{(\delta)}_{i}\).

C. Independent Multiple-Paths over packet switched networks

The construction of parallel paths can be achieved by using parallel paths in MPLS networks, using Source Routing, constructing static parallel routes in the IP network or any other way, as discussed in [1] and [21]\(^5\). Moreover parallel paths exist de-facto in today’s networks via the multi-homing connectivity approach, where load-balancing devices disperse traffic to parallel routes.

We will assume that the losses on the different paths are independent of each other. This is likely to occur if the paths are fully disjoint or if at least the “noisy”, in terms of loss and delay, components of the different paths are disjoint. Theoretically speaking, this assumption can hold in a multi-homing environment in the Internet as well. Packets in the Internet usually cross only a few managed networks on the way to destination. Hence, it might be enough for the first domain to disperse the packets between two different managed networks to achieve the effect of dispersion over independent parallel paths.

---

\(^4\) Note that the Consecutive Loss Factor (CLF), mentioned in [5], is a special case of the NLR metrics.

\(^5\) The construction of independent parallel paths might be problematic in the Internet, but feasible in managed networks.
The destination endpoint, in VoIP applications, must be able to receive and synchronize packets arriving from parallel paths and manage the jitter-buffer optimally in order to reduce delay to minimum and handle out-of-order packets (which may be very common if the paths are not of equal delay). In our model we assume that parallel paths have small delay differences in comparison to the allowed buffering delay. This assumption can hold for many network scenarios. In applications where large buffering is allowed, such as one-way video or voice streaming, the gap in delay may be unimportant and compensated for by increased jitter-buffer. For interactive applications that demand quick response (e.g. phone-call) only small buffering is allowed, up to few tenths of milliseconds, and choosing eligible set of paths is crucial.

D. Modeling Path Loss

Losses at the application level are caused both by the IP network losses and by network delays. In this study, we model the application loss, regardless of the source of the loss (network loss or network delay\(^7\)). Internet loss models have been studied in many studies, such as [3][4]. Here we are focusing on modeling the losses experienced by VoIP applications. For this matter we look at these applications as constant packet rate applications. We assume that time is divided into time slots\(^6\). At each time slice \(t\), a packet is sent by the application. For clarity, in the analysis we refer to the packet sent at time slice \(t\) as packet \(t\). Thus the loss model, expresses the loss experienced by the application. We also assume that the traffic itself does not affect the loss model over the paths.

We will focus on a Bernoulli loss model to model memory-less losses (Section III) and the Gilbert loss model to model bursty losses (Section IV). The Gilbert loss model is used in many studies to model the bursty loss behavior in the Internet. This bursty loss behavior has been shown to arise from the drop-tail queuing disciplines implemented in many Internet routers.

E. Dispersion strategies

Packet dispersion can be implemented through a variety of strategies, of which we focus the following:

1. Deterministic scheduling dispersion
   a. Periodic dispersion – session packets are dispersed in a periodic schedule manner over the routes repeatedly. For example, if the schedule is \((i, i, i, j, j, j)\) then in every cycle 3 packets in a row are sent over path \(p_i\), and then the following two packets are sent over path \(p_j\), where this schedule repeats cyclically.
   b. Deterministic round robin dispersion – a special case of periodic dispersion where packets are sent in a round robin fashion (cyclic schedule) over the paths.

2. Random packet dispersion – for each packet of the session, the dispersing device picks randomly one of the paths leading to the destination and sends the packet over it.

The traditional delivery of packets over a single path is referred to as a no-dispersion strategy. We will assume that the packet dispersion strategies are executed in session context\(^8\).

III. PERFORMANCE ANALYSIS – NLR UNDER BERNOLLI LOSS MODEL

The aim of this section is to evaluate the effect that packet dispersion has on application performance, where the network paths experience Bernoulli (memory-less) losses, that is, each packet \(t\) shipped over path \(i\), has the probability of \(L_i\) to be lost. To this end we evaluate the NLR for sessions traversing a single or multiple paths, for a variety of packet dispersion strategies. We will consider situations, which possibly consist of \(N\) streams, denoted \(s_1, \ldots, s_N\), and possibly are routed over \(P\) parallel paths, denoted \(p_1, \ldots, p_P\).

A. The NLR under No-Dispersion

From the definition of noticeable loss in Eq. (4), the probability for packet \(k\) to be counted as a noticeable loss is given by:

\[
Pr[N_{L_i}^{(\delta)}(k) = 1] = Pr(l(k) = 1, l(k + 1) = \ldots, l(k + \delta - 1) = 0) \quad (6)
\]

As we do the analysis under the Bernoulli (memory-less) loss model: \(Pr[N_{L_i}^{(\delta)}(k) = x] = Pr[N_{L_i}^{(\delta)}(k + t) = x] \quad \forall t, x \in \{0,1\}\). Thus, under steady state we may define \(N_{L_i}^{(\delta)}\) as the limit \(N_{L_i}^{(\delta)} = \lim_{t \to \infty} N_{L_i}^{(\delta)}(k)\) and Eq. (5) translates to:

\[
NLR_i^{(\delta)} = Pr[N_{L_i}^{(\delta)} = 1] \quad (7)
\]

which is the probability for an arbitrary packet in the session to be counted as ‘noticeable loss’. Below we assume that each session is directed over a single path (no-dispersion strategy). Based on (7), the NLR, when the system is under steady state, experienced by session \(s_i\), sent over \(p_j\) is:

\[
NLR_{ij}^{(\delta)} = Pr[N_{L_{ij}}^{(\delta)} = 1] = Li - Li(1 - Li)^{\delta} \quad (8)
\]

Now, assuming that each session takes a single path, the expected network NLR for the \(N\) sessions, \(NLR_{\delta}^{(\delta)}\), is then simply calculated by averaging the \(N\) sessions.

B. The NLR Under Periodic Packet Dispersion

In periodic dispersion, packets of session \(s_i\) are dispersed over the paths according to a fixed policy. Consider a periodic

---

\(^{6}\) Roughly, we may say the packets delayed beyond 250ms are considered lost.

\(^{7}\) Usually in duration of 10 to 30 milliseconds in VoIP applications.
dispersion policy \( Q \), with period length \( K \). The policy is defined by \( Q(k), \quad (k \in \{1,2,\ldots,P\} \), meaning
that packet \( k \) in the period will always be sent on \( P_{Q(k)} \) periodically. Thus, the path taken for packet \( t \), without loss of
generality, is \( P_{Q((t \mod K)+1)} \). The NLR for session \( s_j \), starting at
an arbitrary location of the period is then:

\[
NLR_j^{(\delta)} = \frac{1}{K} \sum_{k=1}^{K} L_k \left( 1 - \sum_{j=1}^{\delta} (1 - L_{Q((k+j \mod K)+1)}) \right), \quad (9)
\]

where \( L_k \) is the loss probability over the path \( p_k \) taken by
the session.

For simplicity of presentation consider periodic dispersion
where the period length is a whole multiple of \((\delta + 1)\). Given
the periodic dispersion selected, let \( c_{i,j} \left( \sum_{j=1}^{\delta} c_{i,j} = 1 \quad \forall i \right) \) and
assume \( c_{i,j} \delta \) is an integer) denote the fraction of packets
belonging to session \( s_i \) that are sent on path \( p_j, \quad j = 1,...,P \).
The NLR experienced by session \( s_i \) is:

\[
NLR_i^{(\delta)} = \left( \sum_{j=1}^{P} c_{i,j} L_j \right) (1 - \sum_{k=1}^{\delta} (1 - L_k)^{c_{i,j} \delta}) . \quad (10)
\]

Note that the NLR experienced by \( s_i \) is not affected by
session \( s_j \). Therefore, the expected average NLR for \( N \) sessions over \( P \) routes is then given by:

\[
\overline{\text{NLR}}^{(\delta)} = \frac{1}{N} \sum_{i=1}^{N} \left( \sum_{j=1}^{P} c_{i,j} L_j \right) (1 - \sum_{k=1}^{\delta} (1 - L_k)^{c_{i,j} \delta}) . \quad (11)
\]

Under limited resources (e.g. the total capacity of paths
equals or approximately equals to the required sessions’
payload), periodic dispersion can be used for QoS purposes by
spreading the sessions in a way that as many sessions as
possible will meet their QoS requirements. Finding the optimal periodic dispersion assignment is a problem left for
further study.

C. The NLR Under Random Dispersion

In random dispersion the decision regarding over which
path to send packet \( t \) of session \( s_i \), is done in a random
fashion. Let \( \rho_{i,j} \left( \sum_{j=1}^{P} \rho_{i,j} = 1 \right) \) denote the probability that
packets of \( s_i \) are sent on path \( p_j \). The NLR experienced by
session \( s_i \) is then given by:

\[
NLR_i^{(\delta)} = \sum_{j=1}^{P} \rho_{i,j} L_j \left( 1 - \left( 1 - \sum_{k=1}^{P} \rho_{i,k} L_k \right)^{\delta} \right), \quad (12)
\]

Under the random dispersion strategy we assume that the
path selection of one session is independent of that of another
session. Under this setting the loss experienced by the \( t \)-th
packet of \( s_i \) is independent of the loss experienced by the \( t \)-th
packet of \( s_j \). Further, the loss of the \((t+1)^{\text{th}}\) packet is
independent of the loss of the \( t \)-th packet. The average NLR
over all sessions is then:

\[
\overline{\text{NLR}}^{(\delta)} = \frac{1}{N} \sum_{i=1}^{N} \left( \sum_{j=1}^{P} \rho_{i,j} L_j \right) (1 - (1 - \tilde{L}_j)^{\delta}) . \quad (13)
\]

1) The NLR Under Random Dispersion with Limited
Resources

Consider random dispersion where the system resources
are limited. That is, the combined paths capacity equals, or
approximately equals, to the sessions’ payload. Thus, the NLR
of session \( s_i \) is dependent on the NLR of session \( s_j \)
through the sharing of the resources.

Consider the case of \( N \) sessions and \( P \) paths having
together the capacity to carry exactly \( N \) sessions. For
simplicity assume that \( P=N \). The source endpoint can choose
one of \( \binom{N}{P} \) possible dispersion combinations for assigning
sessions over the paths. The formulation is similar to that
given in Eq. (12) where:

\[
\sum_{j=1}^{P} \rho_{i,j} = 1 \quad \text{and} \quad \sum_{i=1}^{N} \rho_{i,j} = \text{the number of sessions within the capacity of path } p_j .
\]

The NLR observed by each session depends only on the loss
probabilities of the paths it travels over, and is similar to the
case of random dispersion. Note that the NLR of session \( s_i \)
depends on the NLR of session \( s_j \). But this dependency is
taken into account in the calculation of \( \rho_{i,j} \). Once \( \rho_{i,j} \) is set,
this model is completely similar to the NLR observed in the
random dispersion model without any path capacity limitations.

To demonstrate how the transmission probabilities can be
set, consider two sessions \( s_1 \) and \( s_2 \), and two parallel paths \( p_1 \)
and \( p_2 \) each with the capacity of one session. There are two
possible combinations for sending the packets: 1) Send \( s_1 \) over
\( p_1 \) and \( s_2 \) over \( p_2 \), and 2) Send \( s_1 \) over \( p_1 \) and \( s_2 \) over \( p_2 \). To meet
the objective of sending a fraction \( \rho_{1,1} \) packets of \( s_1 \) over \( p_1 \)
and \( 1 - \rho_{1,1} \) over \( p_2 \) (with complement probabilities for \( s_2 \)), the
first dispersion combination should be assigned probability of
\( \rho_{1,1} \).

D. Comparison of Dispersion Strategies under Bernoulli loss
model

Clearly, if there are no capacity limitations it would always
be better to send all the traffic over the best path using the no-
dispersion strategy. The comparison of strategies under the
Bernoulli loss model is thus significant under limited path resources and provides insight to the question of which dispersion strategy to implemented by load-balancing devices for VoIP sessions.

For the sake of presentation, we will present the tradeoffs between the strategies under the scenario of two sessions that need to be delivered over two parallel paths with limited resources (for simplicity consider capacity of single session on each path). We will compare the average $NLR$, $\overline{NLR}^{(\delta)}$, observed by the sessions.

1) Equal Quality paths

Corollary 1: For equal loss rate over the paths, $L_1 = L_2 = \cdots = L_N$, all dispersion strategies provide the same $NLR$.

This implies that under the Bernoulli loss model, dispersing packets over paths with similar random loss probabilities has no affect on the VoIP quality. From the practical point of view, under no capacity limitations, the use of packet dispersion in a multi-path environment is undesirable due to the possible effects of delay variation, packet out-of-order events, etc.

2) Random and Periodic Dispersion vs. No-Dispersion

The form of the expression of $NLR$ of a single session, under random dispersion is identical to that of $NLR$ under no-dispersion, where the loss parameter $L_i$ is replaced by the average loss experienced by session $s_i$, $\vec{L}_i$. This means that random dispersion in practice averages out the loss over all paths. For meaningful comparison one should compare the average $NLR$ (averaged over multiple sessions).

The difference in average $NLR$ (for the two session system – two path system) between random dispersion and no-dispersion, for $\delta = 2$, is presented in Figure 1. Random dispersion is superior, in this scenario, to no-dispersion if one of the paths experiences low loss rate while the other experiences very high loss rate and can significantly reduce the $NLR$ (in up to 13%). However, if the paths experience very high loss rate (non-identical) the no-dispersion strategy becomes superior. The reason is that dispersing the session increases the probability for losses over the ‘better’ path to be counted as noticeable.

Comparing the periodic round robin dispersion and no-dispersion brings to similar results as presented in Figure 2. Under the same conditions (two sessions to be sent over two paths with limited resources) we present the following question: Under what values of $\delta$, deterministic round robin packet dispersion is superior to no-dispersion. By comparing $\overline{NLR}^{(\delta)}$ under no-dispersion (calculated as the $NLR$ averaged over the sessions (9)) to (11), we may compute the values of $\delta$ for which deterministic round robin dispersion is superior to no-dispersion. This result, as function of the path loss rates, is given by:

$$
\delta < 2 \frac{\log(L_1 / L_2)}{\log(1 - L_2 / 1 - L_1)} \quad \text{for } L_1 \leq L_2
$$

$$
\delta > 2 \frac{\log(L_1 / L_2)}{\log(1 - L_2 / 1 - L_1)} \quad \text{for } L_2 \leq L_1
$$

In Figure 3 the region above the plane represents the values of $\delta$ for which no-dispersion is superior and the region below the plane represents superiority of round-robin dispersion. Note that for most practical situations, that is, if loss probabilities on both paths are lower than 5%, periodic dispersion is superior for all practical ranges of $1 \leq \delta \leq 32$. Further, periodic dispersion is superior also for loss probability between 5% and 20%, for any $\delta < 8$. The Figure also demonstrates (as mentioned in Corollary 1) that for equal paths the $NLR$ is equal.

Figure 1. $NLR$ difference between random and no dispersion for $\delta = 2$

Figure 2. $NLR$ difference between periodic round robin dispersion and no-dispersion for $\delta = 2$
For two paths the gain of periodic and random dispersion over no-dispersion decreases once $\delta$ becomes larger (e.g. $\delta = 10$). However, for such values of $\delta$ the gain may again increase if the number of paths increases. Figures of these results are provided in [24].

We thus conclude that both periodic and random dispersion can reduce the average NLR in many scenarios and thus improve quality in comparison to the traditional no-dispersion.

3) The Superiority of Random Dispersion over Periodic Dispersion

Corollary 2: Random dispersion results in lower NLR than periodic dispersion (where the period length is a multiple of $\delta + 1$) achieved under similar conditions.

Given a periodic dispersion one can always produce a random dispersion that results in lower NLR. Consider random dispersion and periodic dispersion where $c_{i,j} = \rho_{i,j}$. This means that the random dispersion sends on average the same fraction of packets belonging to session $s_i$ over path $p_j$. By comparing (11) to (13), random dispersion results in lower NLR since:

$$\prod_{k=1}^{P} (\bar{L}_k)^{c_{i,j}^\delta} < \left(\sum_{k=1}^{P} c_{i,j} \bar{L}_k\right)^\delta$$

where $\bar{L}_k = 1 - L_k$. Note that (15) holds since the arithmetic weighed average is always greater than the geometric weighted average when $\sum_{j=1}^{L} c_{i,j} = 1$ (see [19]).

Figure 4. demonstrates the reduction of NLR by random dispersion in comparison to periodic dispersion, when two sessions are sent over two paths and $\delta = 2$. The gain grows when the difference in loss rates between the paths increases.

IV. BURSTY LOSSES – THE NLR UNDER THE GILBERT LOSS MODEL

The aim of this section is to evaluate the effect that packet dispersion has on VoIP performance. To this end we evaluate the NLR for sessions traversing a single or multiple paths that are subject to bursty losses, for a variety of packet dispersion strategies. Intuitively speaking, packet dispersion can reduce NLR and thus improve voice quality, especially over paths suffering bursty losses, since dispersion is expected to spread the losses. We will use the Gilbert loss model to model the bursty losses over the paths. We will consider a general situation in which $N$ streams, denoted $s_1 \cdots s_N$, are possibly routed over $P$ parallel paths, denoted $p_1 \cdots p_P$.

A. The Gilbert loss Model – A Two States Markov Chain

The loss probability as expressed in the Bernoulli model, is a basic parameter that affects the performance of VoIP applications. However, it is insufficient in capturing loss burstiness which is highly important for these applications. The Gilbert model allows one to express history-dependent losses and thus to capture loss burstiness. This model has been used in many studies to characterize bursty loss in the Internet [3][11][15].

The model uses a two-state Markov chain to represent the packet losses. We consider a discrete time model where the time unit corresponds to packet transmission for path $p_i$. Let $S_i(t)$ denote the state of the path at time $t$. We assume that $t = 0, \cdots, \infty$, where $B$ stands for Bad and $G$ stands for Good. The states of the path, $S_i(t)$ are governed by a Markov chain depicted in Figure 5:

![Figure 5. The Gilbert channel loss model](image)
When the path is in state \( G(B) \) it is subject to Bernoulli loss at rate \( P_{gi} (P_{bj})^j \). Considering path \( p_i \) we have:

\[
P_{gi} = \Pr[\text{packet } t \text{ is lost over } p_i \mid S_i(t) = G],
\]

\[
P_{bj} = \Pr[\text{packet } t \text{ is lost over } p_i \mid S_i(t) = B].
\] (16)

Clearly \( P_{gi} < P_{bj} \).

To put this in matrix notation let state 1 represent \( G \) and state 2 represent \( B \), and let \( A_i \) be the state transition matrix for path \( p_i \), that is \( A_i(m,n) = \Pr[S_i(t) = n \mid S_i(t-1) = m] \).

Then we have: \( A_i = \begin{bmatrix} 1 - \alpha_i & \alpha_i \\ \beta_i & 1 - \beta_i \end{bmatrix} \). Let \( \pi_i \) denote the steady state probability vector, of path \( p_i \).

Let \( B_i^l \) be a vector representing the loss probability conditioned on the path state, that is \( B_i^l = \begin{bmatrix} P_{gi} \\ P_{bj} \end{bmatrix} \). Also let

\[
1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad B_i^0 = 1 - B_i^l.
\]

Note that the Bernoulli loss model can be represented by special cases of this model, such as \( P_{gi} = P_{bj} \).

B. The NLR Under Various Dispersion Strategies

We start our analysis by first studying the NLR as observed over a single path. Let \( L_i(t) \) be a random variable denoting the event of loss or success at time \( t \) on path \( p_i \). Let \( l_i(t) \) be the actual event occurring at \( t \) on \( p_i \), \( l_i(t) \in \{0,1,\phi\} \) where ‘1’ denotes loss, ‘0’ denotes success and \( \phi \) denotes either loss or success (a “don’t care”)\(^9\).

Let \( E_i(t,n) = (l_i(t), \ldots, l_i(t + n - 1)) \). For a particular event sequence \( (l_i(t), \ldots, l_i(t + n - 1)) \) we want to compute \( \Pr[E_i(t,n) = (l_i(t), \ldots, l_i(t + n - 1))] \) which is done in the next theorem.

**Theorem 1:** Let \( (l_i(t), \ldots, l_i(t + n - 1)) \) be an arbitrary success/loss sequence where \( l_i(j) \in \{0,1,\phi\} \). Assume that the state probabilities at \( t-1 \) are given by \( \pi_i(t-1) \). Then:

\[
\Pr[E_i(t,n) = (l_i(t), \ldots, l_i(t + n - 1))] = \pi_i^{T}(t-1)\prod_{j=0}^{n-1} (\Lambda_i^{j+1})^l \quad \text{(17)}
\]

where \( \Lambda_i^{j+1} = \begin{cases} A_i & \text{if } l_i(t+j) = '1' \\ \tilde{A}_i & \text{if } l_i(t+j) = '0' \\ \phi_i & \text{if } l_i(t+j) = '0' \end{cases} \), and where \( \pi_i(t-1) \) denotes the transpose of the state probability vector at time \( t-1 \).

For lack of space we omit the proof, which can be found in [24].

Note that \( \Lambda_i^{j+1} \) denotes the matrix of probabilities where:

\[
\Lambda_i^{j+1} (m,n) = \Pr[L_i(k) = l_i(k) \land S_i(k) = n \mid S_i(k-1) = m].
\]

That is, the \((m,n)\) entry is the probability for the Markov chain to transit from \( S_i(k-1) \) to \( S_i(k) \) and for packet \( k \) to be a loss/success/don’t care, based on the value of \( l_i(k) \).

**Remark 1:** One should note the low complexity for computing (Eq. (17)). Despite the fact that the number of possible sequences is exponential in \( n \), the special form of Eq. (17) allows one to compute the probability of \( E_i(t,n) \) in linear time in \( n \).

1) The NLR under the no-dispersion

Based on (6) and assuming that the state probability at \( t-1 \) is given by \( \pi_i(t-1) \), we may now compute the noticeable loss rate for session \( s_i \) delivered over path \( p_i \) (based on the definition in (4)):

\[
\Pr[NL_i^{(\delta)}(t) = 1] = \Pr[l_i(t) = 1] - \Pr[l_i(t) = l_i(t+1) = 0, \ldots, l_i(t+\delta) = 0] \\
= \pi_i^{T}(t-1)B_i^l \cdot \pi_i^{T}(t-1) \left( \Lambda_i \left( \Lambda_i^0 \right)^\delta \right) \quad \text{(18)}
\]

When the system is under steady state we substitute \( \pi_i \), by \( \pi_i = \lim_{t \to \infty} \pi_i(t) \). The noticeable loss rate, \( NLR^{(\delta)} \), is then given by:

\[
NLR_i^{(\delta)} = \Pr[NL_i^{(\delta)}(t) = 1] = \pi_i^{T}B_i^l \cdot \pi_i^{T} \left( \Lambda_i \left( \Lambda_i^0 \right)^\delta \right) \quad \text{(19)}
\]

from which the average over \( N \) sessions, \( \overline{NLR}^{(\delta)} \), readily follows.

2) The NLR Under Periodic Packet Dispersion

The analysis of the NLR under periodic dispersion policy is based on calculating the NLR of a session visiting paths according to the specific periodic dispersion policy. To

---

\(^9\) In many studies, such as [11], the values \( P_{gi}=0 \) and \( P_{bj}=1 \) are used, which leads to modeling bursts of consecutive losses.

\(^{10}\) The actual event of cause is either ‘0’ or ‘1’. The ‘\( \phi \)’ event is modeled for cases where we do not care for the actual outcome of \( l_i(t) \).
calculate this properly, the states in the Markov chain on each of the \( P \) paths must be accounted for. One should note that a straightforward analysis of the \( P \) path system may require using a \( P \) dimensional state space, with computational complexity exponential in \( P \). However, our analysis shows that the problem is decomposable and thus the computational complexity is only linear in \( P \). The overall computational complexity is only: \( O(P \cdot K \cdot \delta) \), where \( K \) is the deterministic period length. Full analysis of the NLR under periodic dispersion policy can be found in [24].

For the sake of presentation, we demonstrate the methodology described in [24], on the special case of round-robin dispersion. We assume a simple round robin dispersion policy conducted over two paths \( p_1 \) and \( p_2 \), in which the odd packets are sent over \( p_1 \) and the even packets are sent over \( p_2 \). Writing the probabilities implicitly, given the initial state \( \pi_1(t-1) \) and \( \pi_2(t-1) \), we have:

\[
\Pr[\{N(t) = 1\} = \Pr[\text{session starts at } p_1].
\]

\[
\Pr[\{tS_j(t) = 1\} = \begin{cases}
\Pr[E_1(t, \delta + 1) = (1, (\phi, 0)^{\delta/2}) \land \Pr[E_2(t, \delta + 1) = ((\phi, 0)^{\delta/2}, \phi)] \bigg] \quad (20) \\
\Pr[E_1(t, \delta + 1) = (1, (\phi, 0)^{\delta/2}) \land \Pr[E_2(t, \delta + 1) = ((\phi, 0)^{\delta/2}, \phi)]
\end{cases}
\]

where \( \phi \) stands for a ‘don’t care’ and \( (\phi, 0)^{\delta/2} \) stands for a sequence of \( \delta / 2 \) ‘don’t cares’ and packet arrivals. The NLR for the system, assuming steady state and even \( \delta \), is then:

\[
NLR^{(\delta)} = \frac{1}{2} \left( \pi_1^T B_1 \pi_1 \right) \left( \pi_2^T \left( A_1 A_2 \phi^{\delta/2} / 2 \right) \pi_2 \right) + \frac{1}{2} \left( \pi_2^T B_2 \pi_2 \right) \left( \pi_1^T \left( A_1 A_2 \phi^{\delta/2} / 2 \right) \pi_1 \right)
\]

(21)

Note that the events \( E_1(t) \) and \( E_2(t) \) in Eq. (20), reflect the behavior of the paths \( p_1 \) and \( p_2 \) respectively and are independent of each other (due to the independence of the path behavior). This leads to the product form in Eq. (21). The derivation for odd \( \delta \) is similar.

3) The NLR Under Random Packet Dispersion

In our analysis we assume that the loss models over the paths are independent, meaning that the state \( (S_j(t)) \) on path \( p_i \) is independent of the state \( (S_j(t)) \) on path \( p_j \) \( \forall j \neq i \), at time \( t \). A session dispersed over the paths using the random dispersion strategy, experiences losses as if it was delivered over a single path with the underlying loss model that is the combination of loss models over the paths. The loss model experienced by the session, is characterized by a \( 2^P \times 2^P \) Markov chain, and a matching set of loss probabilities on each state. The calculation of the NLR is then very similar to that of the no-dispersion calculation, Eq. (19). For the lack of space we omit the full analysis, which can be found in [24].

The computation complexity of this analysis is exponential in the number of paths, that is: \( O(2^P \cdot \delta) \).

C. Comparison of the Dispersion Strategies Under the Gilbert loss model

In this section we compare the NLR experienced by sessions sent using various dispersion strategies over paths experiencing bursty losses (following the Gilbert loss model). Since the loss model is affected by four parameters, it is difficult to present a thorough comparison. For simplicity we will compare paths with equal characteristics and will assume that in all paths \( P_G = 0 \). A numerical comparison of paths with different characteristics leads to similar conclusions.

For a better understanding of the results we present in Figures 6-11 plots comparing ratios between the strategies given. In the plots we present the Markov chain parameters in term of \( T_G \) and \( T_B \), which are the average duration time for the chain to be in states \( G \) and \( B \), respectively (\( T_G = 1/\alpha \), \( T_B = 1/\beta \)). The time duration in our model is actually measured in the number of packets sent in each state (i.e. \( T_G = 100 \) means that 100 packets are sent on average in state \( G \). For packetization periods of 30ms in codecs this would mean 3 seconds).

In a thorough examination we conducted [24], the cases we examined demonstrate that under a vast range of network conditions, packet dispersion, both via random and periodic dispersion, can highly reduce the NLR in comparison to the traditional no-dispersion strategy. Only in a very small set of parameter ranges the no-dispersion strategy is superior to dispersion. A sample of those cases is given in Figures 6-9; in these figures all the NLR ratios are smaller than 1, implying full superiority of dispersion. Similarly to the results under the Bernoulli loss model, Random dispersion is in many cases superior to periodic dispersion, as can be seen in Figures 10-11.

Figure 6. NLR ratio between random and no-dispersion for \( T_g=1000 \) and \( T_s=100 \)
Remark 2: In the comparisons we can see that the largest differences between the strategies are when $2 \leq \delta \leq 10$. The reason for that is that we compare the strategies using two paths only. Clearly, if more paths are used, dispersion will have greater impact on quality even for higher values of $\delta$.

V. DISCUSSION AND CONCLUSIONS

We addressed the factors affecting voice quality of VoIP and focused on packet loss. We proposed the noticeable loss rate (NLR) as a metrics well correlated with voice quality for VoIP applications. We studied the effect of packet dispersion strategies, as performed de-facto by load balancing (multi-homing) devices or can be implemented using other mechanisms, on the NLR. We conducted this analysis under the assumption of Bernoulli losses and the Gilbert loss model, over the network paths.

We showed that under the Bernoulli loss model, in many cases the discussed packet dispersion strategies could reduce NLR and thus improve voice quality. We showed that for identical paths all dispersion strategies and no-dispersion are equally good and thus packet dispersion is not recommended. We also showed that random dispersion is superior to periodic dispersion (under several assumptions) and as such preferred for VoIP applications.

We provided mathematical analysis of the NLR for sessions traveling over paths experiencing bursty loss model (Gilbert model). We provided low complexity expressions for the computation of the NLR under the dispersion strategies. We demonstrated, using numerical examples, that the effectiveness of the various packet dispersion strategies strongly depends on the model parameters, and that in many environments both periodic dispersion and random dispersion can highly reduce NLR in comparison to the traditional routing, where a single path is used. We observed that as the number of paths used for dispersion grows, the impact of packet dispersion increases and therefore is recommended in many scenarios.

The superiority of packet dispersion implies that this strategy can improve VoIP application quality, regardless of how dispersion is realized, whether by a multi-homing device located in the network or by a dedicated dispersing element.
intended to improve quality. Due to this improvement it might be worthwhile to place dispersing devices in the network. Such devices should be located on the path between the sender and the receiver and may take automatic dispersion decisions based on current network conditions or base on a-priori knowledge gathered by network management elements.

REFERENCES


