The Effect of Packet Dispersion on Voice Applications in IP Networks

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

***Abstract*- Delivery of real time streaming applications, such as voice and video over IP, in packet switched networks is based on dividing the stream into packets and shipping each of the packets on an individual basis to the destination through the network. The basic implicit assumption on these applications is that shipping all the packets of an application is done, most of the time, over a single path along the network. In this study we present a model in which packets of a certain session are dispersed over multiple paths, in contrast to the traditional approach. The dispersion may be performed by network nodes from various reasons such as load-balancing, or implemented as a mechanism to improve quality, as will be presented in this work. To study the effect of packet dispersion on the quality of Voice over IP (VoIP) applications , we focus on the effect of the network loss on the applications, where we propose to use the *Noticeable Loss Rate* (*NLR*) as a measure (negatively) 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 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[1][2].**

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

1 Introduction

Traditionally, packet switched networks, such as IP networks, are based on a fundamental principle of directing the traffic over the "best single path" (single shortest path). This means that all the traffic to a certain destination follows a single, supposedly the best, path. This principle is used by real-time streaming applications, such as voice and video over IP, where the application divides the stream into packets at the source, and ships them toward the destination.

Violating this fundamental principle, packet dispersion[3] 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. Clearly, packet dispersion may have a critical effect on application's quality of service. While under some circumstances it may degrade quality, in some others it may improve it significantly.

The aim of this study is to examine whether packet dispersion can be used as a machinery to improve QoS of VoIP applications. Further we aim at evaluating the effect of dispersion on VoIP quality as a function of the statistical network properties.

Realization of packet dispersion can be done in several methods by the source node or by routers in the network. Technically, packet dispersion can be achieved by source routing, multi-homing devices, Content Delivery Network companies (such as Akamai, that uses edge architecture to achieve load-balancing and improved network utilization) and other methods.

Dispersing traffic over multiple parallel paths is a known mechanism in several technologies. Reference [6] provides a literature survey on traffic dispersion, mostly for issues in telecommunication networks. In [7] [29] traffic dispersion in IP networks is suggested to reduce traffic burstiness and therefore achieve higher resource utilization. In CDMA radio networks, traffic dispersion (also called frequency-hopping) is used for security reasons and in order to statistically multiplex noises. Another reason for implementing traffic dispersion, proposed in [13], 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.

The quality of VoIP applications is affected by two major factors: a) The underlying network behavior, and b) 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 increase, quality degrades. The acceptable delay for a bi-directional VoIP session (such as a phone call), is usually limited by values of 200-250 milliseconds. This limit may change dynamically in advanced playout techniques that adjust the deadline for speech packet [22][24]. VoIP application packet exceeding these margins will be

[1] Part of these results appear in [30] and [31].

[2] This work was supported in part by a grant from the Israeli Science Foundation.

[3] Since we focus our discussion on IP networks, we will use the term Packet Dispersion instead of the general term Traffic Dispersion.

counted as lost. Thus, both delay and jitter can be roughly translated, physically and mathematically, into a loss measure. 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).

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 VoIP 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 [8][28] (the impact of packet loss on VoIP quality is also studied in [13][14]). Supporting this method, perceptual studies of applications such as IP phones have shown that user dissatisfaction rises dramatically in presence of bursty losses. 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 [15] as the basic ingredient in computing the perceived quality of VoIP applications. The *NLR* metrics counts losses of 'close' packets and ignores losses of distant packets. Based on [28] 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[4]. In some scenarios the assumption taken in our analysis, that paths tend to experience similar delay characteristics, may be problematic. 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 2 we discuss the modeling considerations of this work, present the underlying assumptions of our model, and introduce the Noticeable Loss Rate model adopted from [15]. We then turn into mathematical analysis of packet dispersion strategies under the Bernoulli loss model (Section 3) and under the Gilbert model (section 4). 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.

[4] In [2] it is 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.

2 MODELING ASSUMPTIONS, MODEL AND NOTATIONS

2.1 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 [11] and PSQM [10]. Another non-intrusive monitoring technique for VoIP, incorporating the effects of time varying packets loss and "recency", based on the E-model [8] is suggested in [28].

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 [28] propose that given the codec type and other application parameters, loss (I_e) and delay (I_d) 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 affect 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 with 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 [28], 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 lost 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 presented in Figure 1, taken from [13]:

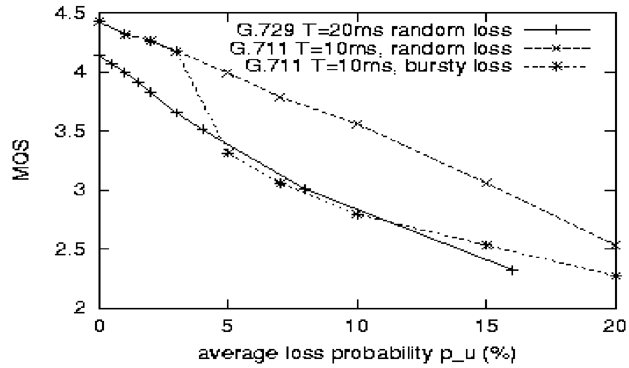

Figure 1. Call Quality (MOS) as a function of average loss probability

Perceptual studies, such as those referenced in [5], also support the fact that bursty losses may dramatically reduce perceptual quality, especially for audio. Other studies [20] have used real Internet measurements, bursty traffic model evaluation and the E-model to evaluate the quality of VoIP on various links of the Internet.

Common VoIP manipulation techniques also increase the importance of bursty losses. First, in modern codecs internal *Packet Loss Concealment* (*PLC*, see [9]) 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 [24]) mechanisms are also used to compensate for lost packets by appending the information of 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.

2.1.1 Noticeable Loss Rate (NLR)

The *IP Performance Metrics* (ippm) working group in the IETF has proposed a set of metrics for packet loss [15]. 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[5]. 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*.

2.1.1.1 A Definition of NLR

The *Loss Distance* is defined (as in [15]) as the difference in sequence numbers between two successively lost packets. The loss event of a packet is defined to be "a δ 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 δ, where δ, a positive integer, is the *loss constraint*. In order to measure how 'noticeable' a loss might be for quality purposes, the *loss distance* δ may be selected to be equal to $g_{\min}$ (the parameter used in [28], typically $g_{\min} = 16$), determining whether a packet belongs to a *burst.* Alternatively, small values of δ 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 [15]. Where necessary we will associate the parameter δ with the notion of noticeable loss rate, reading δ - *noticeable loss rate*, or $NLR^{(\delta)}$.

The *loss indicator function* for a certain flow reflects the loss event of packet *t:*

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

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

$$NL_i^{(\delta)}(k) = \begin{cases} 1 & l(k) = 1 \text{ and } \exists t \in [\text{k} - \delta, \text{k} - 1] \text{ where } l(t) = 1 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

The noticeable loss rate for session *i* with loss constraint δ, and for a sequence of *K* packets, is then given by:

$$NLR_i^{(\delta)}(K) = \frac{1}{K}\sum_{s=1}^{K} NL_i^{(\delta)}(s). \quad (3)$$

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

$$NL_i^{(\delta)}(k) = \begin{cases} 1 & l(k) = 1 \text{ and } \exists t \in [\text{k} + 1, \text{k} + \delta] \text{ where } l(t) = 1 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

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 our model we assume that VoIP sessions are small in comparison to path capacity and therefore dispersing traffic does not influence the losses on the links. If dispersion is conducted in mass volume, we conjecture that due to load-balancing effects it will reduce the losses on the links. This is conjectured to increase the benefits of dispersion. A model that combines this factor with the analysis of this paper is a subject for further study.

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

$$NLR_i^{(\delta)} = \lim_{M \to \infty} \frac{1}{M}\sum_{k=1}^{M} NL_i^{(\delta)}(k). \quad (5)$$

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 *NLR* taken over the *N* sessions, denoted $\overline{NLR}^{(\delta)}$, $\overline{NLR}^{(\delta)} = \frac{1}{N}\sum_{i=1}^{N} NLR_i^{(\delta)}$.

2.2 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 [27][6]. 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

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

[6] The construction of independent parallel paths might be problematic in the Internet, but feasible in managed networks.

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.

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.

2.3 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 the source of the loss (network loss or network delay[7]). 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[8]. 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. This model is limited to regular voice packet stream without accounting for compression aspects such as MDC [1][18] and other techniques used in VoIP applications such as FEC, PLC and VAD. 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 3) and the Gilbert loss model to model bursty losses (Section 4). The Gilbert loss model is used in many studies, such as [1][3][5][7] [19][21], to model network bursty loss behavior. This bursty loss behavior has been shown to arise from the drop-tail queuing disciplines implemented in many Internet routers. Measurement studies of Internet loss behavior, such as [4][26], show that loss modeling over Internet segments can be modeled by a Bernoulli model for some segments, and by a 2-state Markov chain (Gilbert-Elliot model) or by k-th order Markov chains for others. Other networks may be subject to either bursty or non-bursty traffic, depending on the network. Note also that satellite and radio links, that are becoming more and more popular, are known to be bursty.

[7] Roughly, we may say the packets delayed beyond 250ms are considered lost.

[8] Usually in duration of 10 to 30 milliseconds in VoIP applications.

2.4 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) 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[9].

3 PERFORMANCE ANALYSIS – NLR UNDER BERNOULLI 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 now use the variables $NL_i^{(\delta)}(k)$ as random variables and evaluate the

[9] This assumption is not mandatory since *random dispersion* or *periodic dispersion* of all packets, regardless of the application, will lead in many cases to the same results.

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 $s_1 \cdots s_N$, and possibly are routed over *P* parallel paths, denoted $p_1 \cdots p_P$.

3.1 The NLR under No-Dispersion

From *Proposition* 1 and the definition of noticeable loss in Eq. (4), the probability for packet *k* to be counted as a noticeable loss is given by:

$$\begin{aligned}&\Pr[NL_i^{(\delta)}(k)=1]\\&=\Pr[l(k)=1]-\Pr[l(k)=1,l(k+1)=0,...,l(k+\delta)=0]\end{aligned} \quad (6)$$

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

$$NLR_i^{(\delta)}=\Pr[NL_i^{(\delta)}=1]\ , \quad (7)$$

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_i is:

$$NLR_i^{(\delta)}=\Pr[NL_i^{(\delta)}=1]=L_i-L_i(1-L_i)^\delta . \quad (8)$$

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

3.2 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 dispersion* policy *Q*, with period length *K*. The policy is defined by $Q(k)$ ($Q(k)\in\{1...P\}$ and ($k=1...K$)), 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 $p_{Q((t \bmod K)+1)}$. The *NLR* for session s_i, starting at an arbitrary location of the period is then:

$$NLR_i^{(\delta)}=\frac{1}{K}\sum_{k=1}^{K}L_k\left(1-\prod_{j=1}^{\delta}(1-L_{Q((k+j \bmod K)+1)})\right), \quad (9)$$

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 $c_{i,j}$ ($\sum_{j=1}^{P}c_{i,j}=1\ \ \forall i$ 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, $j=1...P$. The *NLR* experienced by session s_i is:

$$NLR_i^{(\delta)}=(\sum_{j=1}^{P}c_{i,j}L_j)(1-\prod_{k=1}^{P}(1-L_k)^{c_{i,k}\delta})\ . \quad (10)$$

Note that the *NLR* experience 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:

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

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.

3.3 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 $\rho_{i,j}$ ($\sum_{j=1}^{P}\rho_{i,j}=1$) 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 $(t+1)^{st}$ packet is independent of the loss of the t^{th} packet. The average *NLR* over all sessions is then:

$$\overline{NLR}^{(\delta)} = \frac{1}{N}\sum_{i=1}^{N}\left(\widehat{L}_i(1-(1-\widehat{L}_i)^{\delta}\right). \qquad (13)$$

where $\widehat{L}_i = \sum_{k=1}^{P} \rho_{i,k} L_k$.

3.3.1 The NLR Under Random Dispersion with Limited Resources

Consider *random dispersion* where the system resources are limited. Under these conditions the use of independent Bernoulli decisions for each packet may cause traffic to exceed the link capacity. To this end an alternative random dispersion is needed, which is discussed next.

Consider the case of *N* sessions and *P* paths having together the capacity to carry exactly *N* sessions. For simplicity assume that $P \le N$. 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: $\sum_{j=1}^{P} \rho_{i,j} = 1$ and $\sum_{i=1}^{N} \rho_{i,j}$ equals to 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_2 and s_2 over p_1. 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 $\rho_{1,1}$.

3.4 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.

3.4.1 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.

3.4.2 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, $\widehat{L}_i$. 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).

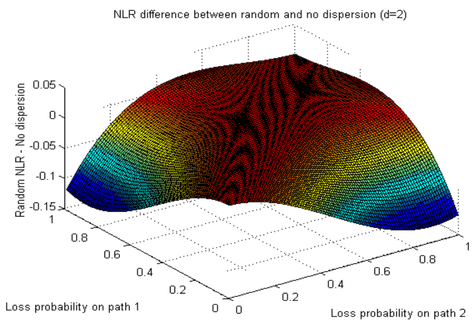

Figure 2. *NLR* difference between *random dispersion* and *no-dispersion* for $\delta = 2$

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 2. *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 3.

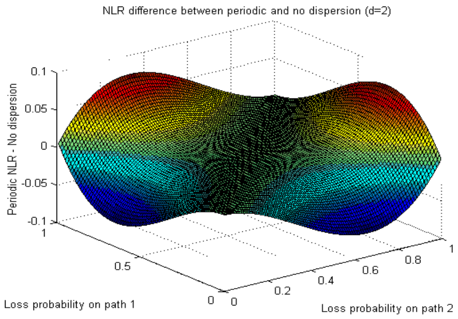

Figure 3. *NLR* difference between *periodic round robin dispersion* and *no-dispersion* for $\delta = 2$

Under the same conditions (two sessions to be sent over two paths with limited resources) we present the following question: Under what values of δ, *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 δ for which d*eterministic round robin dispersion* is superior to *no-dispersion*. This result, as function of the path loss rates, is given by:

$$\begin{aligned} \delta &< 2\frac{\log(L_1/L_2)}{\log(1-L_2/1-L_1)} \quad \text{for } L_1 \le L_2 \\ \delta &> 2\frac{\log(L_1/L_2)}{\log(1-L_2/1-L_1)} \quad \text{for } L_2 \le L_1 \end{aligned} \tag{14}$$

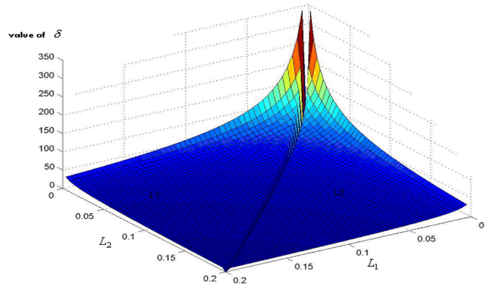

Figure 4. Comparison of *round robin packet* dispersion and *no-dispersion*: Above plane *no-dispersion* is superior; below plane dispersion is superior

In Figure 4 the region above the plane represents the values of δ 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 $1 \le \delta \le 32$. 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.

For two paths the gain of *periodic* and *random dispersion* over *no-dispersion* decreases once δ becomes larger (e.g. δ=10). However, for such values of δ the gain may again increase if the number of paths increases. Figures of these results are provided in [30].

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*.

3.4.3 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:

$$\prod_{k=1}^{P} \left(\overline{L}_k\right)^{c_{i,k}\delta} \le \left(\sum_{k=1}^{P} c_{i,k}\overline{L}_k\right)^{\delta} \tag{15}$$

where $\overline{L}_k = 1 - L_k$. Note that (15) holds since the arithmetic weighed average is always greater than the geometric weighted average when $\sum_{j=1}^{P} c_{i,j} = 1$ (see [22]).

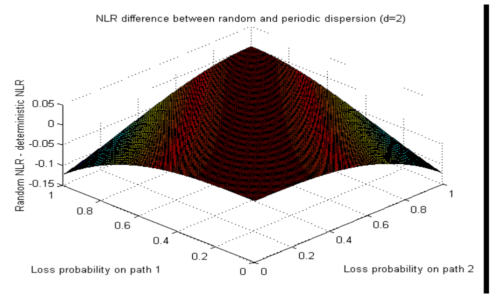

Figure 5. *NLR* difference between *random dispersion* and *periodic dispersion* for $\delta = 2$

Figures 5 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.

4 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][12].

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 Good. The states of the path, $S_i(t)$ are governed by a Markov chain depicted in Figure 6:

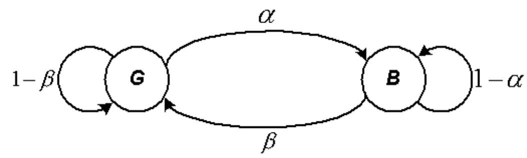

Figure 6. The Gilbert channel loss model

When the path is in state *G*(*B*) it is subject to Bernoulli loss at rate P_G (P_B)[10]. Considering path p_i we have:

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

$$P_{B_i} \stackrel{\Delta}{=} \Pr[\text{packet } t \text{ is lost over } p_i \mid S_i(t) = B]. \qquad (16)$$

Clearly $P_{G_i} < P_{B_i}$.

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 π_i denote the steady state probability vector, of path p_i.

Let B_i^1 be a vector representing the loss probability conditioned on the path state, that is $B_i^1 = \begin{bmatrix} P_{Gi} \\ P_{B_i} \end{bmatrix}$. Also let $\mathbf{1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $B_i^0 = \mathbf{1} - B_i^1$.

Note that the Bernoulli loss model can be represented by special cases of this model, such as $P_{G_i} = P_{B_i}$.

4.1 Dispersion Strategies: Analysis of the NLR

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 ϕ denotes either loss or success ("don't

[10] In many studies, such as [12], the values P_G=0 and P_B=1 are used, which leads to modeling bursts of consecutive losses.

care")[11]. Let $E_i(t,n)=(L_i(t),\cdots,L_i(t+n-1))$. For a particular event sequence $(l_i(t),\cdots,l_i(t+n-1))$ we want to compute $\Pr[E_i(t,n)=(l_i(t),\cdots,l_i(t+n-1))]$, which is done in the next theorem.

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

$$\Pr[E_i(t,n)=(l_i(t),\cdots,l_i(t+n-1))] = \pi_i^T(t-1)\left(\prod_{j=0}^{n-1}\widehat{A}_i^{l_i(t+j)}\right)\mathbf{1} \quad (17)$$

where $\widehat{A}_i^{l_i(t+j)}=\begin{cases}A_i\widehat{B}_i^1 & \text{if } l_i(t+j)='1'\\ A_i\widehat{B}_i^0 & \text{if } l_i(t+j)='0'\\ A_i & \text{if } l_i(t+j)='\phi'\end{cases}$,

$A_i=\begin{bmatrix}1-\alpha_i & \alpha_i\\ \beta_i & 1-\beta_i\end{bmatrix}$, $\widehat{B}_i^1=\begin{bmatrix}P_{Gi} & 0\\ 0 & P_{B_i}\end{bmatrix}$, $\widehat{B}_i^0=I-\widehat{B}_i^1$, and where $\pi_i^T(t-1)$ denotes the transpose of the state probability vector at time *t-1*. The proof is given in Appendix A. Note that $\widehat{A}_i^{l_i(k)}$ denotes the matrix of probabilities where:

$\widehat{A}_i^{l_i(k)}(m,n)=\Pr[L_i(k)=l_i(k)\wedge S_i(k)=n\mid S_i(k-1)=m]$. That is, the $(m,n)^{th}$ 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*.

4.1.1 The NLR under the no-dispersion

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

$$\begin{aligned}&\Pr[NL_i^{(\delta)}(t)=1]\\ &=\Pr[l_i(t)=1]-\Pr[l_i(t)=1,l_i(t+1)=0,\cdots,l_i(t+\delta)=0]\end{aligned}$$

$$=\pi_i^T(t-1)B_i^1-\pi_i^T(t-1)\left(\widehat{A}_i^1\left(\widehat{A}_i^0\right)^{\delta}\right)\mathbf{1} \quad (18)$$

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

$$NLR_i^{(\delta)}=\Pr[NL_i^{(\delta)}=1]=\pi_i^{\,T}B_i^1-\pi_i^{\,T}\left(\widehat{A}_i^1\left(\widehat{A}_i^0\right)^{\delta}\right)\mathbf{1}, \quad (19)$$

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

4.1.2 The NLR Under Periodic Packet Dispersion

Consider a *periodic dispersion* policy *Q*, with period length *K*. The policy is defined by $Q(k)$ ($Q(k)\in\{1...P\}$ and ($k=1...K$)), meaning that packet *k* in the period will always be sent on $p_{Q(k)}$ periodically. For packet *t*, define $k(t)=(t \bmod K)+1$. Thus, the path taken for packet *t* is $p_{Q(k(t))}$.

To calculate the *NLR* we examine every path individually. In a matrix notation, when packet *t* is sent over $p_{Q(k(t))}$, its loss/success probabilities and state transition are obtained by multiplying the state vector of path $p_{Q(k(t))}$ at *t-1*, by $\widehat{A}_{Q(k(t))}^{l_{Q(k(t))}(t)}$ (where $\widehat{A}_{Q(k(t))}^{l_{Q(k(t))}(t)}$ is as defined in *Theorem* 1 above).

Now, consider packet *t* routed over $p_{Q(k(t))}$ and consider path $p_j\ne p_{Q(k(t))}$. For this path one must account for the state transition at time *t* but not for the probability of loss/success. We thus introduce the transition matrix T_i^Q for *periodic dispersion* policy *Q*, over path p_i:

$$T_i^Q[t]=\begin{cases}\widehat{A}_i^0 & \text{if } p_{Q(\mathrm{k(t)})}=p_i\\ A_i & \text{otherwise}\end{cases} \quad (20)$$

First we want to compute the *NLR* for a packet transmitted at time *t*, sent over path $p_j=p_{Q(k(t))}$. The *NLR* for this packet is obtained by calculating the probability that the packet is lost and subtracting the probability that the packet is lost and all subsequent δ packets arrive. The latter probability is obtained by deriving the probability of the following event combination:

1) On path p_j there is:
   i) A loss of packet *t*.
   ii) No loss on all packets *t+l* $(l=1...\delta)$, obeying $Q(k(t+l))=j$.

[11] The actual event of cause is either '0' or '1'. The 'ϕ' event is modeled for cases where we do not care for the actual outcome of $l_i(t)$.

iii) Don't care on all other packets.

2) On path $p_r \neq p_j$ there is:
   i) Don't care at packet *t*.
   ii) No loss on all packets *t+l* $(l = 1...\delta)$, obeying *Q*(*k*(*t+l*)) = *r*.
   iii) Don't care on all other packets.

This yields:

$$\Pr[NL_i^{(\delta)}(t) = 1] = \pi_{Q(k(t))}^T B_{Q(k(t))}^1 - \left(\pi_{Q(k(t))}^T \widehat{\mathrm{A}}_{Q(k(t))}^1 \left(\prod_{l=1}^{\delta} \mathrm{T}_{Q(k(t))}^Q[t+l]\right)\mathbf{1}\right) \prod_{\substack{r=1 \\ r \neq Q(k(t))}}^{P} \left(\pi_r^T \prod_{l=1}^{\delta} \mathrm{T}_r^Q[t+l]\right)\mathbf{1} \quad (21)$$

Note that since we assume steady state we consider $t \to \infty$ and thus $\pi_{Q(k(t))}^T$ is the steady state distribution of the path state (and thus we do not use $\pi_{Q(k(t))}^T(t)$ in the equation). However, the index *t* is kept as to properly account for the periodicity of $Q(k(t))$.

By accounting for all possible starting positions in the period, the *NLR* for session s_i sent in a *periodic packet dispersion* policy *Q*, is then:

$$NLR_i^{(\delta)} = \frac{1}{K}\sum_{t=1}^{K} \Pr[NL_i^{(\delta)}(t) = 1], \quad (22)$$

where $\Pr[NL_i^{(\delta)}(t) = 1]$ is calculated from (21).

Remark 2: 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: $O(P \cdot K \cdot \delta)$.

To calculate the average *NLR* for *N* sessions using the *periodic dispersion* strategy, note that the *NLR* experienced by s*i* is not affected by session s*j*. Therefore the expected $\overline{NLR}^{(\delta)}$ for *N* sessions over *P* paths is simply calculated by averaging the *NLR* observed by the sessions.

For the sake of presentation, we demonstrate the methodology on the special case of r*ound-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 probability vectors on the paths, $\pi_1(t-1)$ and $\pi_2(t-1)$, we have:

$$\begin{aligned}
&\Pr[NL_i^{(\delta)}(t) = 1] = \Pr[\text{session starts at } p_1]\cdot \\
&\begin{pmatrix}\Pr[l_1(t) = 1] - \\ \Pr[E_1(t,\delta+1) = (1,(\phi,0)^{\delta/2}) \wedge \Pr[E_2(t,\delta+1) = ((\phi,0)^{\delta/2},\phi)]\end{pmatrix} \quad (23)\\
&+ \Pr[\text{session starts at } p_2]\cdot \\
&\begin{pmatrix}\Pr[l_2(t) = 1] - \\ \Pr[E_2(t,\delta+1) = (1,(\phi,0)^{\delta/2}) \wedge \Pr[E_1(t,\delta+1) = ((\phi,0)^{\delta/2},\phi)]\end{pmatrix}
\end{aligned}$$

where ϕ 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 δ, is then:

$$\begin{aligned}
NLR^{(\delta)} = &\frac{1}{2}\left(\pi_1^T B_1^1 - \left(\pi_1^T\left(\widehat{\mathrm{A}}_1^1\left(\mathrm{A}_1\widehat{\mathrm{A}}_1^0\right)^{\delta/2}\right)\mathbf{1}\right)\left(\pi_2^T\left(\left(\mathrm{A}_2\widehat{\mathrm{A}}_2^0\right)^{\delta/2}\right)\mathbf{1}\right)\right) \\
&+\frac{1}{2}\left(\pi_2^T B_2^1 - \left(\pi_2^T\left(\widehat{\mathrm{A}}_2^1\left(\mathrm{A}_2\widehat{\mathrm{A}}_2^0\right)^{\delta/2}\right)\mathbf{1}\right)\left(\pi_1^T\left(\left(\mathrm{A}_1\widehat{\mathrm{A}}_1^0\right)^{\delta/2}\right)\mathbf{1}\right)\right)
\end{aligned} \quad (24)$$

Note that the events $E_1()$ and $E_2()$ 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. (24). The derivation for odd δ is similar.

4.1.3 The NLR Under Random Packet Dispersion

In our analysis we assume that the loss models over the paths are independent, meaning that the state ($S_i(t)$) on path *pi* 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. We will denote the "equivalent path" by p', and we construct it next.

First we calculate the characteristics of p' as observed by the session. The "equivalent path" is characterized by 2^P states, resulting from the cross product of the states of the individual paths, $S' = \{S_1 \times S_2 \times \cdots \times S_P\}$. For the sake of clarity one may enumerate and index the states in S' as $1,2,...2^P$. Let $\Gamma(2^P \times 2^P)$ denote the transition matrix for p' where the $(m,n)^{th}$ entry is the probability that p' moves at time *t+1* to state *n,* given that is was at state *m* at time *t.* Matrix Γ is a transition matrix for a Markov chain with 2^P states. This Markov chain has a steady state probability vector denoted by $\widehat{\pi}$, and can be easily computed from the steady states over the individual paths by noting that: $P[(S_1(t),...,S_P(t)) = (x_1,...,x_P)] = \prod_{i=1}^{P} P[S_i(t) = x_i]$, where

$P[S_i(t)=x_i]$ is the steady state probability of being in state x_i, on path p_i.

In *random dispersion* the decision regarding over which path to send packet *t* of session s_q, is done in a random fashion. Let $\rho_{q,i}(\sum_{i=1}^{P}\rho_{q,i}=1)$ denote the probability that packets of s_q are sent on p_i. Similar to the expression received for the *no-dispersion* (see (21), given the initial state vector $\bar{\pi}(t-1)$, we get the following expression for *NLR* under *random dispersion* for session s_q:

$$NLR^{(\delta)} = \bar{\pi}^T(t-1)\hat{B}\mathbf{1} - \bar{\pi}^T(t-1)\left(\Gamma\hat{B}^1\left(\Gamma\hat{B}^0\right)^{\delta}\right)\mathbf{1} \quad (25)$$

where, (as in *Theorem* 1) Γ denotes the matrix of transition probabilities:

$\Gamma(m,n)=\Pr[S_i'(t)=n \mid S_i'(t-1)=m]$, and where $\hat{B}^1$ is a $2^P x 2^P$ diagonal matrix, whose $(m,m)^{\text{th}}$ entry ($1 \le m \le 2^P$), is the loss probability in state *m*, as follows: Assuming that state *m* is given by $(S_1,...,S_P)$ where $S_i \in \{G,B\}$, then this entry is given by $\sum_{i=1}^{P}\rho_{q,i}(P_{B_i}I(S_i=B)+P_{G_i}I(S_i=G))$, where $I(\ldots)$ is the indicator function. Similarly, the $(m,m)^{\text{th}}$ entry of $\hat{B}^0$ is given by $\sum_{i=1}^{P}\rho_{q,i}((1-P_{B_i})I(S_i=B)+(1-P_{G_i})I(S_i=G))$, that is $\hat{B}^0=\mathbf{I}-\hat{B}^1$, where $\mathbf{I}(2^P x 2^P)$ is a unit matrix.

Remark 3: Note that the computation complexity is exponential in the number of paths: $O(2^P \cdot \delta)$.

4.2 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 7-17 plots comparing ratios and differences between the strategies. 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 [30], 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 7-10; in these figures all the *NLR* ratios are smaller than 1, implying full superiority of dispersion. In Figures 16-17 we present a set of values where *no-dispersion* is superior in a small range of parameters. 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 11-12. The reason for this phenomenon can be understood by observing Figures 13-15. In these figures one can see that in *periodic dispersion* all packets suffering high loss will be from a single session and thus losses are not spread between them to achieve better overall average performance. All the losses will be counted as Noticeable Losses in one session, concluding in higher *NLR* as can be seen in Figure 15.

Remark 4: In the comparisons we can see that the largest differences between the strategies are when $2 \le \delta \le 10$. The reason for that is that we compare the strategies using two paths only. Clearly, if more paths are used for dispersion, the range of δ will grow, and thus have greater impact on quality.

5 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. The results of this study can be used by load-balancing devices to decide which particular scheduling policy to use.

REFERENCES

[1] D. G. Apostolopoulos. "Reliable video communication over lossy packet networks using multiple state encoding and path diversity. In *Proceeding Visual Communication and Image Processing, pages 392-409,* Jan. 2001.

[2] E. Altman, K. Avrachenkov, C. Barakat, "TCP in Presence of Bursty Losses", Proceedings of ACM SIGMETRICS, Santa Clara, California, June 2000.

[3] Y. Birk and N. Bloch, "Improving network performance with prioritized dispersal", IEEE INFOCOM 2000, vol. 3, pp. 817 -1826.

[4] J-C. Bolot, "End-to-end packet delay and loss behavior in the Internet", In Proceedings of ACM SIGCOMM '93 Conference on Communications Architectures, Protocols and Applications, pages 289-298, San Francisco, CA, USA,September 1993.

[5] Hung Ngo, S. Varadarajan and J. Srivastava, "Error-Spreading: Reducing Bursty Error in Continuous Media streaming", in Proceedings of *IEEE Multimedia Systems '99* (ICMCS), vol. 1 1999: 314-319.

[6] E. Gustafsson and G. Karlson, "A literature survey on traffic dispersion," *IEEE Network*, vol. 11, pp. 28–36, March/April 1997.

[7] F. Ishizaki, "Reduction effect of effective bandwidth and blocking rate with diespersion", Proceedings of Communications, Internet, & Information Technology (CIIT) 2003, pp.485-490, 2003.

[8] ITU-T, "The E-model, a computational model for use in transmission planning", recommendation G.107.

[9] ITU-T, "A high quality low complexity algorithm for packet loss concealment with G.711", recommendation G.711 Appendix I (1999).

[10] ITU-T, "Objective quality assessment of telephone band speech codecs (PSQM)", recommendation P.861.

[11] ITU-T, "Perceprual Evaluation of Speech Quality (PESQ)", Recommendation P.862.

[12] W. Jiang and H. Schulzrinne, "Modeling of Packet Loss and Delay and Their Effect on Real-Time Multimedia Service Quality", NOSSDAV 2000 Paper #27.

[13] W. Jiang and H. Schulzrinne. "Comparison and Optimization of Packet Loss Repair Methods on VoIP Perceived Quality under Bursty Loss", NOSSDAV 2002.

[14] W. Jiang, H. Schulzrinne, "Perceived Quality of Packet Audio under Bursty Losses", In procedings of IEEE INFOCOM 2002, June 2002.

[15] R. Koodli, R. Ravikanth "One-Way Loss Pattern Sample Metrics", ietf, draft-ietf-ippm-loss-pattern-05.txt, 1999.

[16] Y. J. Liang, N. Färber, and B. Girod, "Adaptive playout scheduling and loss concealment for voice communication over IP networks," IEEE Transactions on Multimedia, Dec. 2003.

[17] Yi J. Liang, E-G. Steinbach and B. Girod, "Multi-Stream Voice Over IP Using Packet Path Diversity", Proceedings IEEE Fourth Workshop on Multimedia Signal Processing, pp. 555-560, Cannes, France, Oct. 2001.

[18] Yi J. Liang, E-G. Steinbach and B. Girod, "Real-time Voice Communication over the Internet Using Packet Path Diversity", ACM Multimedia '01, Ottawa, Canada, Oct. 2001.

[19] V. Markovski, "Simulation and Analysis of Loss in IP networks", Thesis, University "Sts. Cyril and Methodius", Macedonia, October 2000.

[20] A. Markopoulou, F. Tobagi, and M. Karam. ,"Assesment of VoIP Quality over Internet Backbone", In Proceedings of IEEE INFOCOM 2002, June 2002.

[21] Y. Nebat and M. Sidi, "Resequencing Considerations in Parallel Downloads", In Proceedings of IEEE INFOCOM 2002, June 2002.

[22] S. B. Moon, J. Kurose, and D. Towsley, "Packet audio playout delay adjustment: Performance bounds and algorithms," Multimedia Systems, vol. 6, no. 1, pp. 17-28, Jan. 1998.

[23] F. Qi, J. Q. Mei, D. Xia, S. Xu, "New proofs of weighted power mean inequalities and monotonicity for generalized weighted mean values", Mathematical Inequalities and Applications 3 (2000), no.3, 377--383. RGMIA Research Report Collection 2 (1999), no.1, Article 10, 99--102; 2 (1999), no.4, Article 6, 469--472.

[24] Ramjee R, Kurose J, Towsley D, Schulzrinne H, "Adaptive playout mechanisms for packetized audio applications in wide-area networks. In: Proceedings of IEEE INFOCOM 94 Montreal, Canada.

[25] J. Rosenberg, L. Qiu, H. Schulzrinne, "Integrating Packet FEC into Adaptive Voice Playout Buffer Algorithms on the Internet", INFOCOM 2000.

[26] M. Yajnik, S. B. Moon, J. Kurose, D. Towsley. "Measurement and Modelling of the Temporal Dependence in Packet Loss," Proceedings of 1999 IEEE INFOCOM, New York, NY, March 1999.

[27] D. Thaler, C. Hopps, "Multipath Issues in Unicast and Multicast Next-Hop selection", RFC2991, Novemver 2000.

[28] TIA, "Telecommunications IP Telephony Equipment Voice Quality Recommendations for IP Telephony" – PN-4689 (TIA/EIA/TSB116) TIA Engineering Committee TR-41.1.

[29] S-J. Yang, "Traffic Dispersion For Bursty Traffic On Heterogeneous Networks", Master Thesis Report, 1998, UT Austin.

[30] H.Zlatokrilov, H. Levy, "Packet Dispersion and the Quality of Voice over IP Applications in IP networks", To appear in IEEE INFOCOM 2004, March 2004.

[31] H.Zlatokrilov "Packet Dispersion and the Quality of Voice over IP Applications in IP networks", Master thesis, School of Computer Science, Tel-Aviv Univeristy, May 2003.

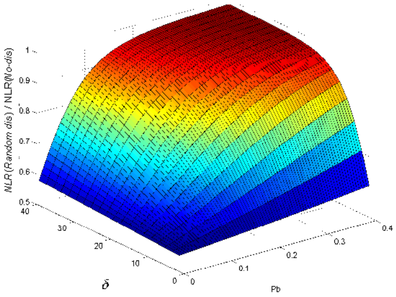

Figure 7. *NLR* ratio between *random* and *no-dispersion* for T_g=1000 and T_b=100

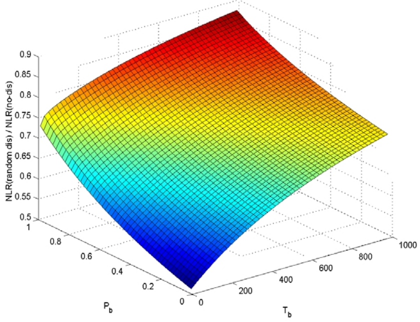

Figure 8. *NLR* ratio between *random* and *no-dispersion* for T_g=1000 and $\delta = 2$

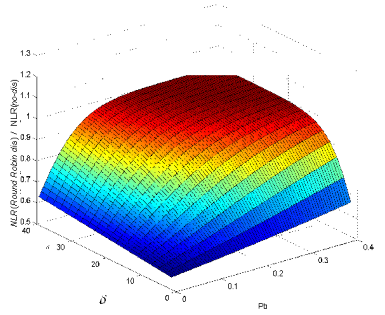

Figure 9. *NLR* ratio between *round robin* and *no-dispersion* for T_G=1000, T_B=100

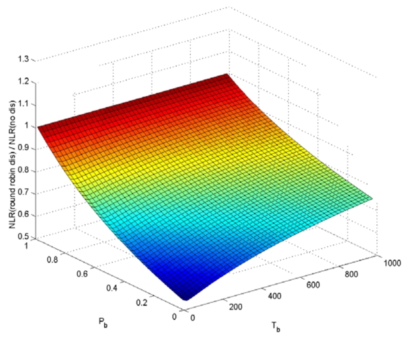

Figure 10. *NLR* ratio between *round robin* and *no-dispersion* for T_G=1000, $\delta = 2$

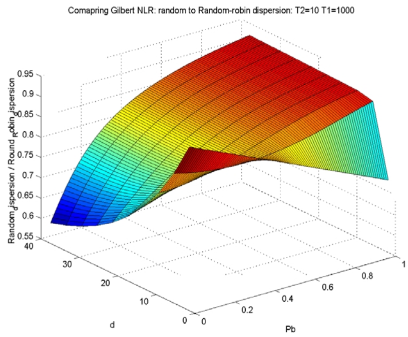

Figure 11. *NLR* ratio between *random* and *round robin dispersion* for T_G=1000 and T_B=10

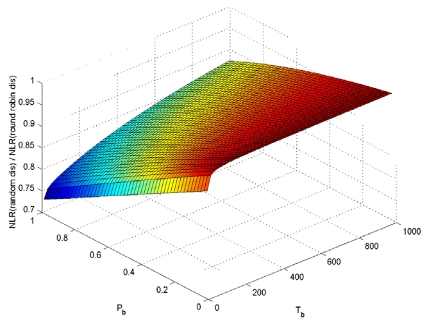

Figure 12. *NLR* ratio between *random* and *round robin dispersion* for T_G=1000, $\delta = 2$

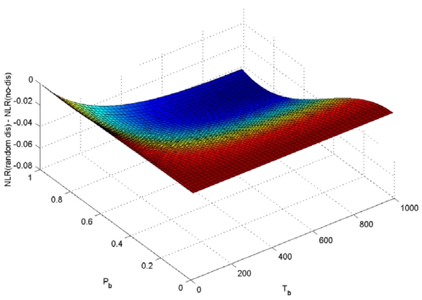

Figure 13. *NLR* difference between *random dispersion* and *no-dispersion* for T_G=1000, $\delta = 2$

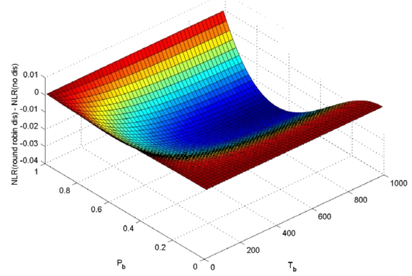

Figure 14. *NLR* difference between *round robin dispersion* and *no-dispersion* for T_G=1000, $\delta = 2$

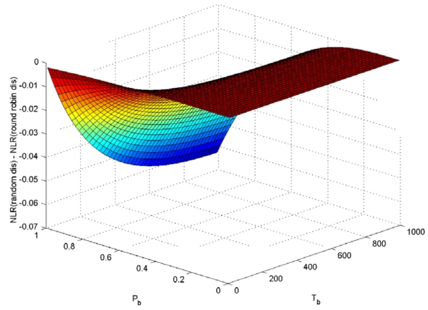

Figure 15. *NLR* difference between *random dispersion* and *round robin dispersion* for T_G=1000, $\delta = 2$

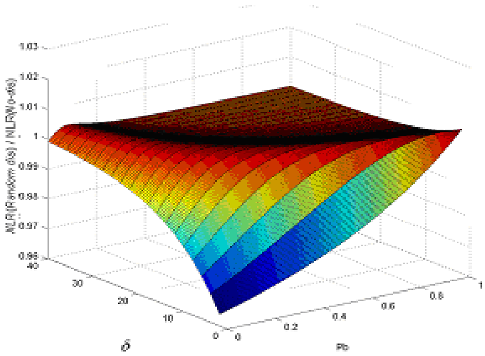

Figure 16. *NLR* ratio between *random dispersion* and *no-dispersion* for T_G=10 and T_B=100

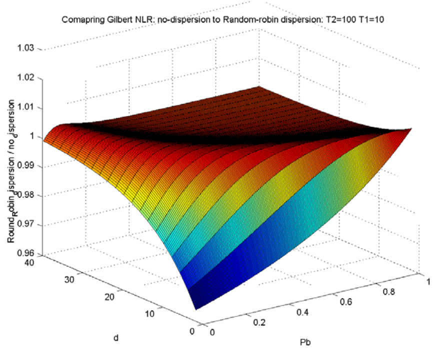

Figure 17. *NLR* ratio between *round robin dispersion* and *no-dispersion* for T_G=10 and T_B=100

Appendix A

Proof of Theorem:

We will prove the following two claims for $n \geq 1$.

1) $$\Pr[E_i(t,n) = (l_i(t),\cdots,l_i(t+n-1)) \wedge S_i(t+n-1) = G] = \pi_i^T(t-1)\left(\prod_{j=0}^{n-1} \widehat{A}_i^{l_i(t+j)}\right)[1] \quad (26)$$

2) $$\Pr[E_i(t,n) = (l_i(t),\cdots,l_i(t+n-1)) \wedge S_i(t+n-1) = B] = \pi_i^T(t-1)\left(\prod_{j=0}^{n-1} \widehat{A}_i^{l_i(t+j)}\right)[2] \quad (27)$$

The proof of these claims is carried out by induction on both claims concurrently.

Induction basis (*n=1*): the proof of the basis is immediate from the definitions:

Note that starting at *t-1*, a state transition is performed and based on the state at *t*, the loss/success event occurs.

$$\Pr[l_i(t) = '1' \wedge S_i(t) = B] = (\Pr[S_i(t-1) = G]\cdot\alpha + \Pr[S_i(t-1) = B]\cdot(1-\beta))\cdot P_{B_i}$$
$$\Pr[l_i(t) = '1' \wedge S_i(t) = G] = (\Pr[S_i(t-1) = G]\cdot(1-\alpha) + \Pr[S_i(t-1) = B]\cdot\beta)\cdot P_{G_i}$$

$$\Pr[l_i(t) = '0' \wedge S_i(t) = B] = (\Pr[S_i(t-1) = G]\cdot\alpha + \Pr[S_i(t-1) = B]\cdot(1-\beta))\cdot(1-P_{B_i})$$
$$\Pr[l_i(t) = '0' \wedge S_i(t) = G] = (\Pr[S_i(t-1) = G]\cdot(1-\alpha) + \Pr[S_i(t-1) = B]\cdot\beta)\cdot(1-P_{G_i})$$

$$\Pr[l_i(t) = '\phi' \wedge S_i(t) = B] = \Pr[S_i(t-1) = G]\cdot\alpha + \Pr[S_i(t-1) = B]\cdot(1-\beta)$$
$$\Pr[l_i(t) = '\phi' \wedge S_i(t) = G] = \Pr[S_i(t-1) = G]\cdot(1-\alpha) + \Pr[S_i(t-1) = B]\cdot\beta$$

Induction step: Assuming correctness of both claims for *n-1*, we prove them for *n*. Using conditional probabilities we get:

$$\Pr[E_i(t,n) = (l_i(t),\cdots,l_i(t+n-1)) \wedge S_i(t+n-1) = G] =$$
$$= \Pr[E_i(t,n-1) = (l_i(t),\cdots,l_i(t+n-2)) \wedge S_i(t+n-2) = G]\cdot$$

$$(\Pr[L_i(t+n-1)=l_i(t+n-1)) \wedge S_i(t+n-1)=G \mid E_i(t,n-1)=(l_i(t),\cdots,l_i(t+n-2)) \wedge S_i(t+n-2)=G]+$$
$$\Pr[L_i(t+n-1)=l_i(t+n-1)) \wedge S_i(t+n-1)=G \mid E_i(t,n-1)=(l_i(t),\cdots,l_i(t+n-2)) \wedge S_i(t+n-2)=B])$$

Since $L_i(t+n-1)$ and $S_i(t+n-1)$ depend on $E_i(t,n-1)$ and $S_i(t+n-2)$ only through $S_i(t+n-2)$, then the conditional probability is equal to:

$$\Pr[L_i(t+n-1)=l_i(t+n-1)) \wedge S_i(t+n-1)=G \mid S_i(t+n-2)=G]= \\ =\begin{cases}(1-\alpha_i)\cdot(1-P_B) & \text{if } l_i(t+n-1)='1' \\ (1-\alpha_i)\cdot P_B & \text{if } l_i(t+n-1)='0' \\ (1-\alpha_i) & \text{if } l_i(t+n-1)='\phi'\end{cases}$$

$$(\Pr[L_i(t+n-1)=l_i(t+n-1)) \wedge S_i(t+n-1)=G \mid S_i(t+n-2)=B]= \\ =\begin{cases}\beta_i\cdot(1-P_B) & \text{if } l_i(t+n-1)='1' \\ \beta_i\cdot P_B & \text{if } l_i(t+n-1)='0' \\ \beta_i & \text{if } l_i(t+n-1)='\phi'\end{cases}$$

$$(\Pr[L_i(t+n-1)=l_i(t+n-1)) \wedge S_i(t+n-1)=B \mid S_i(t+n-2)=G]= \\ =\begin{cases}\alpha_i\cdot(1-P_B) & \text{if } l_i(t+n-1)='1' \\ \alpha_i\cdot P_B & \text{if } l_i(t+n-1)='0' \\ \alpha_i & \text{if } l_i(t+n-1)='\phi'\end{cases}$$

$$(\Pr[L_i(t+n-1)=l_i(t+n-1)) \wedge S_i(t+n-1)=B \mid S_i(t+n-2)=B]= \\ =\begin{cases}(1-\alpha_i)\cdot(1-P_B) & \text{if } l_i(t+n-1)='1' \\ (1-\alpha_i)\cdot P_B & \text{if } l_i(t+n-1)='0' \\ (1-\alpha_i) & \text{if } l_i(t+n-1)='\phi'\end{cases}$$

Now from the inductive assumption we have:

$$\Pr[E_i(t,n-1)=(l_i(t),\cdots,l_i(t+n-2)) \wedge S_i(t+n-2)=G]=\pi_i^T(t-1)\left(\prod_{j=0}^{n-2}\widehat{\mathrm{A}}_i^{l_i(t+j)}\right)[1]$$

$$\Pr[E_i(t,n-1)=(l_i(t),\cdots,l_i(t+n-2)) \wedge S_i(t+n-2)=B]=\pi_i^T(t-1)\left(\prod_{j=0}^{n-2}\widehat{\mathrm{A}}_i^{l_i(t+j)}\right)[2]$$

Putting all these together, the two claims are proved for n (based on $n-1$). Using the induction we complete the proof of the two claims.

Finally, *theorem* 1 follows immediately from (25) and (26).

Note that the proof for any N state Markov chain is similar.