Effects of Fully Redundant Dispersity Routing on VoIP Quality

Stephan Bettermann and Yue Rong
Department of Electrical and Computer Engineering
Curtin University of Technology
Bentley, WA, 6102, Australia
P +61 8 9266 2600 F +61 8 9266 2584
stephan.bettermann@postgrad.curtin.edu.au, y.rong@curtin.edu.au

***Abstract*—Quality and reliability problems are characteristic of real-time services on the Internet such as Voice over IP (VoIP). In this paper the availability of diversity in the Internet is investigated to overcome these flaws. Fully redundant dispersity routing exploits diversity by routing complete copies of the data to be communicated along multiple paths. By actively replicating the data along multiple paths, the effect of a failure on one path may be reduced — or even masked completely — by other paths. This paper presents simulations of such a system, drawing on real VoIP traffic data for the loss, latency and jitter characteristics that data may experience while traversing a path. These simulations show that this form of dispersity routing reduces loss and mean loss burst length, has a de-jittering effect through competition among the paths, and that small numbers of paths already yield significant improvements in deliverable VoIP quality — from 84.12% of calls with which users are estimated as being 'very satisfied', to 99.86% using fully redundant dispersity routing with just two paths.**

Keywords: *Dispersity Routing, QoS, Telephony, VoIP*

I. INTRODUCTION

Users of Voice over IP (VoIP) services over the public Internet have most likely experienced quality and reliability problems. To solve these problems this paper considers a fully redundant dispersity routing system that exploits the path diversity available in the Internet to improve deliverable quality of VoIP services by transporting concurrently multiple instances of the data along multiple paths towards its destination. Assuming that paths have independent failure behaviors, actively replicating the data along multiple paths gives these paths the opportunity to reduce or even mask completely the effect of a failure on other paths. While users of real-time services such as VoIP may adapt to their needs infrastructure that they control, fully redundant dispersity routing allows these users to overcome failures in the public Internet over which they may have no control.

This paper presents simulation results using real VoIP traffic data measured in a commercial call center. The first simulation comprises of a single scenario using three paths, and illustrates salient effects of fully redundant dispersity routing on metrics relevant to VoIP. This is followed by two sets of simulations using two to six paths which demonstrate collectively the effect on deliverable VoIP quality, measured using Mean Opinion Score (MOS) [1] estimates computed with the E-model [2], of fully redundant dispersity routing using increasing numbers of paths.

Using these simulations, we show that: (1) fully redundant dispersity routing is effective in reducing loss and mean loss burst length; (2) competition among the paths has a de-jittering effect; and (3) significant improvements in VoIP quality are achievable with small numbers of paths. MOS estimates for telephone calls measured in a commercial call center show that 84.12% of calls have a MOS of at least 4.34. A MOS of 4.34 or greater may be interpreted as a level of quality with which users are 'very satisfied' [2]. Our simulations show that fully redundant dispersity routing with just two paths may be able to increase the proportion of calls with a 'very satisfied' quality rating to 99.86% in that commercial call center, providing a premium service that is more on par with traditional telephony than VoIP is currently.

While replicating the data of a real-time service such as VoIP along multiple paths increases Internet usage by that service, the Internet is a commodity and consumers pay to use it. Compared to services such as Internet television and on-demand video streaming the additional usage, and thus cost, is likely to be negligible. Similarly, the increased cost compared to that for traditional telephony is likely to be negligible. Each individual user may decide whether the gain in quality for a premium real-time service warrants the additional cost.

Recent work related to ours includes [3] and [4] which use non-redundant dispersity routing, and [5] which use path switching. Our work differs from [3]–[5] in that we consider fully redundant dispersity routing, being prepared to expend additional resources in return for a premium service that is on par with traditional telephony. Furthermore, we use real VoIP traffic data measured in a commercial call center in simulations to demonstrate the effects of fully redundant dispersity routing on VoIP quality.

The remainder of this paper is structured as follows: section II establishes the background summarizing alternative approaches to dispersity routing, before describing dispersity routing in its various forms, and introducing the MOS that is used in this paper to quantify quality. Section III describes the measuring of VoIP traffic data that both establishes a baseline of current performance for comparison and is used in the simulations to provide realistic path characteristics, section IV

This work was supported in part by the Western Australian Telecommunications Research Institute.

978-1-4577-1298-2/11/$26.00 ©2011 IEEE

outlines the dispersity routing simulator model, followed by the simulation results in section V and the conclusion in section VI.

II. Background

Firstly in this section alternative approaches to dispersity routing are summarized. Next, dispersity routing is described, including the form used in this paper. Finally, the MOS is discussed as a means of measuring quality.

A. Forward Error Correction

Forward Error Correction (FEC) techniques may mask failures by including redundant data in the stream [6], and then using that redundant data to recover lost data, at least in part. However, real-time communication by definition has time-constraints, which limits the usage of FEC techniques.

For VoIP, latencies of 150 ms or more are generally accepted as impeding on quality [7]. As loss tends to be bursty [8]–[9], FEC techniques may not be able to mask failures within these time constraints. A loss burst with a duration in excess of t cannot be masked within t when the redundant data for reconstructing that lost data is not available within t.

B. Path Switching

Path switching relies on identifying alternative paths preemptively and upon detecting or even just predicting better performance on another path, switching to that path [5][10]–[12]. The accuracy of the predictor determines the effectiveness of this scheme.

Being able to predict degradation on the current path allows the scheme to switch to a better path before the degradation occurs. In the absence of this ability, path switching would occur reactively only once the degradation has occurred, which would cause an interruption until the switch to another path has completed. Similarly, in addition to knowing which alternative path currently gives better performance, being able to predict which alternative path will offer better performance over long time scales allows the scheme to avoid switching to a path that is about to experience the same degradation or worse.

C. Dispersity Routing

Delivering data to its intended destination by transmitting it towards multiple nodes in the general direction of the destination was first considered as a routing technique called *selective flooding* [13]. Although it was dismissed as inefficient at the time, variations of this technique may show promise and are considered in this paper.

Maxemchuk [14]–[15] is generally credited to be the first to describe a system that employs multiple concurrent routes to transport data for the benefits that it brings to data transport. Called *dispersity routing*, Maxemchuk identifies three forms: (i) non-redundant, (ii) fully redundant, and (iii) partially redundant dispersity routing.

Non-redundant dispersity routing is used primarily for the performance gains possible when combining the capacity of multiple paths, effectively inverse multiplexing. In this approach, the data to be communicated is distributed among the paths, such that each path is given a subset of the data and each subset is given to one path only. Fully redundant dispersity routing is used primarily for the quality gains possible when a failure on one path may be masked by other paths. In this approach, the data to be communicated is given to all paths, such that each path is given all of the data. Partially redundant dispersity routing seeks to minimize bandwidth requirements by balancing performance and quality using techniques such as erasure codes. However, instead of a code covering a range of data along a given path, here a code may cover a block of data across the set or subset of paths, in various arrangements. For the remainder of this paper dispersity routing refers to the fully redundant form.

D. Mean Opinion Score

The Mean Opinion Score (MOS) is widely accepted as the standard method of evaluating perceived conversational voice quality [1]. In this method, human evaluators rate the call according to a standardized procedure with an opinion score in the range 1 (Bad) to 5 (Excellent).

However, the very subjective nature of this test which makes it so valuable — the perceived quality by humans — is also its weakness as this subjectiveness makes it difficult to be consistent, and the need for human evaluators makes it expensive and in this context impractical. Objective approaches such as Perceptual Evaluation of Speech Quality (PESQ) [16] and the E-model do not suffer these shortcomings.

PESQ compares the original signal (from the speaker) to the degraded signal (as observed by the listener) to yield a quality estimate. However, both the original and the degraded signal are needed; when the degraded signal only is available this method is unsuitable.

The E-model is a computational model that yields a rating factor from a set of telephone system characteristics, where the rating factor can be mapped to a MOS estimate. That set of characteristics comprises of observable properties such as latency, loss, burstiness, and codec. The audio for which the MOS estimate is being computed is not included in that set.

In this paper, all MOS estimates are computed using the E-model. While the E-model accepts 21 parameters, each equating to a telephone system characteristic, the MOS estimates in this paper are computed from 6 parameters derived from latency, codec, loss rate and burst ratio, using the default values recommended by the E-model for the remaining parameters.

III. Measuring VoIP Traffic Data

This section describes the measuring of VoIP traffic data used both to establish a baseline by estimating currently achieved quality without dispersity routing, and to provide realistic path characteristics for the simulations. In addition, some salient properties of the data collected are presented.

To measure real VoIP traffic data, the telephone system in a commercial call center was modified to measure a subset of its telephone calls. Within the organization, dedicated and isolated resources are allocated to the telephone system, with capacity utilization negligible. The telephone system connects to two VoIP service providers over the public Internet, each handling approximately half the calls. Between 20 August 2009 and

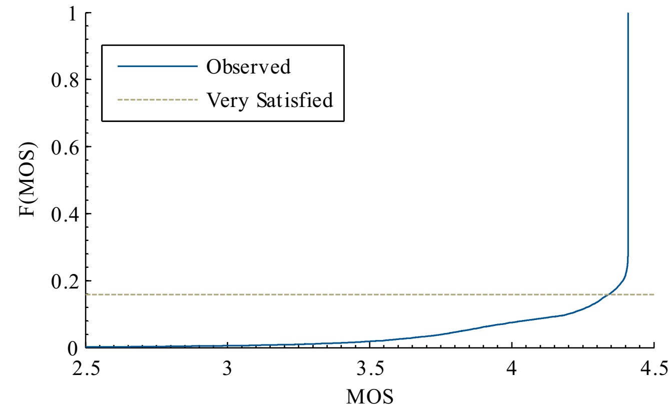

Figure 1. Cumulative distribution of MOS estimates for measured calls. A MOS of 4.34 or above may be interpreted as quality with which users are ‘very satisfied’. 84.12% of calls measured have a MOS of at least 4.34.

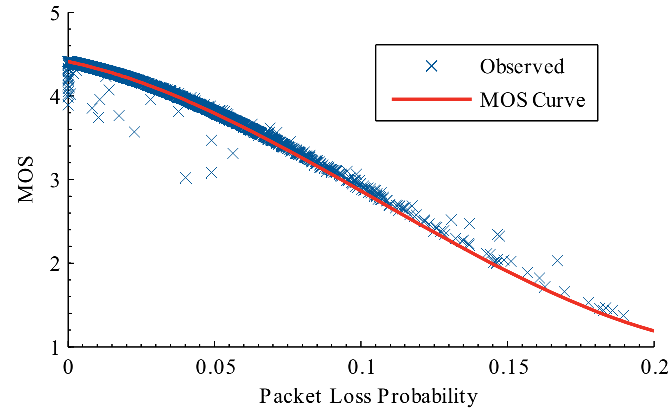

Figure 2. Scatter plot for measured calls of observed packet loss probability against estimated MOS. Trend follows MOS curve variant only on packet loss probability, with burst ratio and latency given as mean for all call profiles.

22 October 2010, 18 574 calls totaling to over 888 hours were measured encompassing 158 651 893 received packets.

For each call selected for measuring, the telephone system creates a file called a *call profile* that contains information about the call (including call date and time, a latency estimate, and the trunk identity) and each received packet (including packet type, sequence number, send timestamp, and receipt timestamps). All lost packets — identifiable using sequence numbers — are assumed to be voice packets.

Fig. 1 depicts the cumulative distribution of the MOS estimates (computed with the E-model) for these measured calls, and indicates the proportion of calls with a MOS estimate of at least 4.34, that is, calls with a very satisfied quality rating. The distribution depicted in Fig. 1 supports the perception that most of the time (approximately 84%) the performance of VoIP over the public Internet is acceptable, but that quality and reliability problems are characteristic of VoIP. Approximately 16% of measured calls have a quality rating indicating users are less than ‘very satisfied’.

To identify the cause of these quality and reliability problems, Fig. 2 plots the packet loss probability against the MOS estimate for each measured call. Fig. 2 also plots a MOS curve for packet loss probabilities in the range 0 to 0.2, where all parameters of this curve other than the packet loss probability are constant. As the estimated MOS correlates strongly with the observed packet loss probability, as shown in Fig. 2, along a MOS curve that is variant only on the packet loss probability, these quality and reliability problems are due mostly to packet loss. For the sake of completeness, the constant parameters to the MOS curve are as follows. The burst ratio and latency are the mean burst ratio and latency for all call profiles. Codec-derived parameters (that is, *Equipment Impairment Factor*, *Packet-loss Robustness Factor* and *Number of Quantization Distortion Units*) are adopted from the call profiles, which all use the same codec and therefore have the same values for these parameters. For the remaining parameters the default values defined by the E-model are used.

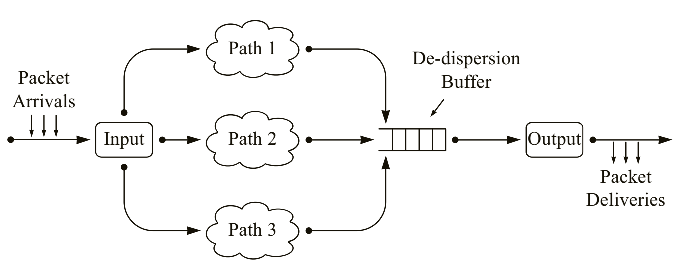

Figure 3. A dispersity routing system using three paths. Packets enter the system on the left, are delivered over multiple paths concurrently, and pass through a de-dispersion buffer before leaving the system on the right.

IV. Simulation Model for Dispersity Routing

The effects of dispersity routing may be observed by simulating a system as depicted in Fig. 3. Each packet entering the system is sent over all paths, three in this example. While traversing a path, packet instances may be lost or experience latency and some degree of jitter. Packet instances that are not lost while traversing a path enter the *de-dispersion buffer*, which discards all but the first received instance of each packet and schedules the packet for delivery from the system.

To simulate what happens to packet instances as they traverse a path, loss, latency and jitter are drawn from a measured call profile, the measuring of call profiles being described in section III above. This is similar to the approach taken by [17], except that we draw the loss, latency and jitter directly from call profiles, rather than using the measurements to create stochastic processes from which loss, latency and jitter are then drawn. While the call profiles may be described by models such as a Gilbert model for packet loss [6][8][11] and a shifted gamma distribution for jitter [18]–[19], in this paper we use call profiles directly to be as realistic as possible.

Sequence numbers are used to identify lost packets and to order packets. For a system with three paths, three unique call profiles are selected. As the shortest call profile constrains the duration of the simulation, where the length of a call profile is the number of received packets and identified lost packets, the call profiles are truncated to the length of the shortest call profile selected. Each truncated call profile is assigned to one path, such that each path has one call profile. The first packet instance traversing a path adopts the loss, latency and jitter observed for the first packet in the call profile assigned to that path, the second packet instance those for the second and so on.

Loss of a packet in the call profile is adopted as loss of the respective packet instance on the path. Latency and jitter are adopted by their inclusion in the computation of the arrival time of the non-lost packet instance into the de-dispersion buffer, and then in the delivery time of the first of each packet instance from the system.

Let $\mathcal{P}$ be the set of n received packets in a call profile assigned to a path, and $\mathcal{S}$ and $\mathcal{R}$ the set of send and receive times respectively for the packets in $\mathcal{P}$. Assuming that the clocks at the sender and receiver increment at the same rate, jitter j for packet i in relation to packet k, where $1 \leq i \leq n-1$, $0 \leq k \leq n-2$, and $k < i$, is computed as

$$j_{i,k} = (\mathcal{R}_i - \mathcal{R}_k) - (\mathcal{S}_i - \mathcal{S}_k). \tag{1}$$

The *cumulative jitter* a for received packet i, where $i > 0$ and $i \in \mathcal{P}$ is given by

$$a_i = \sum_{x=1}^{i} j_{x,x-1} = (\mathcal{R}_i - \mathcal{R}_0) - (\mathcal{S}_i - \mathcal{S}_0). \tag{2}$$

Jitter and cumulative jitter cannot be computed for packet 0, the first packet, since jitter is computed against a predecessor and the first packet has no predecessor. Furthermore, as a packet traversing a path with constant latency l by definition takes at least l to traverse that path, the minimum additional delay is 0. Thus, the delay experienced by the first packet in addition to latency is estimated as $-min(\mathcal{A})$ where $\mathcal{A} = \{a_x : x \in \{x \in \mathcal{P} : x > 0\}\}$, and to adopt the jitter observed for the packets in $\mathcal{P}$, the delay in addition to latency for received packet i is then given by

$$d_i = \begin{cases} -min(\mathcal{A}), i = 0 \\ d_0 + a_i, i > 0 \end{cases}. \tag{3}$$

Therefore, the arrival time $\mathcal{B}$ in the de-dispersion buffer of non-lost packet instance i received over a path with latency l is computed as

$$\mathcal{B}_i = \mathcal{S}_i + l + d_i. \tag{4}$$

While the de-dispersion buffer may delay packets similar to a de-jitter buffer, for illustration purposes no de-dispersion buffer delay is adopted in this paper. The first instance of each packet that arrives in the de-dispersion buffer is delivered from the system as soon as it arrives.

TABLE I. EFFECT ON MOS, LOSS, AND BURST LENGTH

Measuring Location	Latency (ms)	MOS	Total Loss (Packets)	Mean Burst (Packets)
Path 1	57.0	1.50	1419	33.44
Path 2	56.5	1.54	1409	31.38
Path 3	52.0	1.43	1465	34.02
Output	52.0	4.38	35	8.50

V. SIMULATION RESULTS

This section presents the results of three simulations. The first illustrates salient effects of dispersity routing, the second simulation quantifies the effectiveness of dispersity routing in improving quality for the overall range of conditions observed in section III, and the third quantifies the effectiveness of dispersity routing in improving quality for some of the most extreme conditions observed in section III.

A. Salient Effects of Dispersity Routing

Salient effects of dispersity routing may be illustrated by considering a dispersity routing system with three paths. Since in an actual system paths with comparable latencies would be chosen, let the three paths have latencies between 45 ms and 60 ms. Furthermore, let these three paths experience the loss, latency and jitter experienced by the three calls measured in section III with the *lowest* MOS estimates. Therefore, let the three call profiles selected for assignment to these three paths be those with (1) estimated latencies between 45 ms and 60 ms and (2) the lowest MOS estimates.

A simulation of this system demonstrates the effects of dispersity routing for a single call, using paths experiencing the worst conditions measured in section III for comparable latencies. Table I summarizes the latency, MOS, loss and mean burst length measured for the calls whose call profiles are adopted by the three paths, and at the output (see Fig. 3) of the simulation of this system.

As shown in Table I, the output MOS of 4.38 is much closer to the maximum possible MOS of 4.41 for a lossless output here than the MOS of the individual paths (averaging to 1.49). This improvement is due primarily to the reduction in packet loss and burstiness, as supported by the correlation observed in Fig. 2. Loss decreases from a mean of 1431 packets on the individual paths to 35 on the output, and burstiness from a mean (of the mean burst length of each path) of 32.95 packets to 8.5 packets at the output.

Fig. 4 depicts for each path and the output the cumulative jitter for non-lost packets in packet range 400 – 4400. Loss bursts are clearly visible as blocks of missing cumulative jitter. While in this example cumulative jitter was not considered in the MOS (by dropping packets that miss the de-jitter buffer and affecting the MOS in the form of loss), another effect of note is the reduction in cumulative jitter that occurs when low-jittered packets out-compete high-jittered packets [20]. In the example above, path 3 has the lowest latency and, therefore, its packets tend to arrive before packets from the other paths. However, while packet 722 experiences a peak in cumulative jitter on path 3, on the output that packet does not: another path out-competes path 3 and delivers that packet earlier.

Packets 4063 – 4268 illustrate this effect further. On path 3 these packets are lost, so paths 1 and 2 mask the loss. When packet 4182 on path 2 experiences high cumulative jitter, it is outcompeted by path 1. Indeed, the range of cumulative jitter on the output for that period is smaller than the range outside that period, because the latencies of path 1 and 2 are very close. Since cumulative jitter is heavy tailed and packets are more likely to be low jittered than high-jittered, competition has the opportunity to reduce cumulative jitter.

However, out-of-order packets may occur when a loss burst ends on a lower latency path. Packets traversing that path after the loss burst ends may arrive before earlier packets traversing a higher-latency path arrive and which were lost in the lost burst on the lower latency path. In addition, at the points that higher latency paths begin and end masking for a loss burst on a lower latency path, jitter is likely to occur.

The probabilities of causing jitter and out-of-order packets may be reduced by the de-dispersion buffer scheduling, similar to a de-jitter buffer except that late packets are not discarded, the first instance of each packet for delivery instead of delivering them as soon as they arrive. Given the set of path latencies $\mathcal{L}$, a de-dispersion buffer of size $max(\mathcal{L}) - min(\mathcal{L})$ compensates for the difference in latencies causing jitter and packet re-ordering.

B. *Effectiveness of Dispersity Routing for Overall Conditions*

Drawing on the full set of call profiles measured in section III, a simulation of dispersity routing systems using 2 – 6 paths quantifies the effectiveness of dispersity routing in improving deliverable quality for the full range of observed conditions. Let i be the number of paths in the range 2 – 6. For each value of i, select 10 000 random and different combinations of i call profiles, such that (1) each combination of i call profiles contains i different call profiles, and (2) no two combinations of i call profiles contain the same i call profiles. Fig. 5 depicts the probability that a call has a MOS of at least 4.34 for calls measured in section III (shown as the probability for a 1 path system as these calls did not use dispersity routing) and simulations of dispersity routing systems using 2 – 6 paths.

Of calls measured in section III, 84.12% have a rating of 4.34, a score indicating a quality with which users are 'very satisfied'. Going to dispersity routing with just two paths increases the proportion of calls with a MOS estimate of 4.34 or above to 99.86%. These simulations show that for the full range of conditions measured in section III, dispersity routing with just two paths provides a premium service that is more on par with traditional telephony than VoIP is currently.

C. *Effectiveness of Dispersity Routing for Extreme Conditions*

A simulation of dispersity routing systems using 2 – 6 paths, drawing on the call profiles measured in section III with the lowest 100 MOS estimates, quantifies the effectiveness of dispersity routing in improving quality for some of the most extreme conditions observed. Similarly to the previous simulation, let i be the number of paths in the range 2 – 6. For each value of i, select up to 10 000 random and different combinations of i call profiles, such that (1) each combination of i call profiles contains i different call profiles, and (2) no two combinations of i call profiles contain the same i call profiles. For 2 paths select all possible $\binom{100}{2} = 4950$ combinations, for the remainder select 10 000 combinations.

Fig. 6 depicts six cumulative distributions and their lowest 5th percentile. The percentiles are shown as vertical lines. From

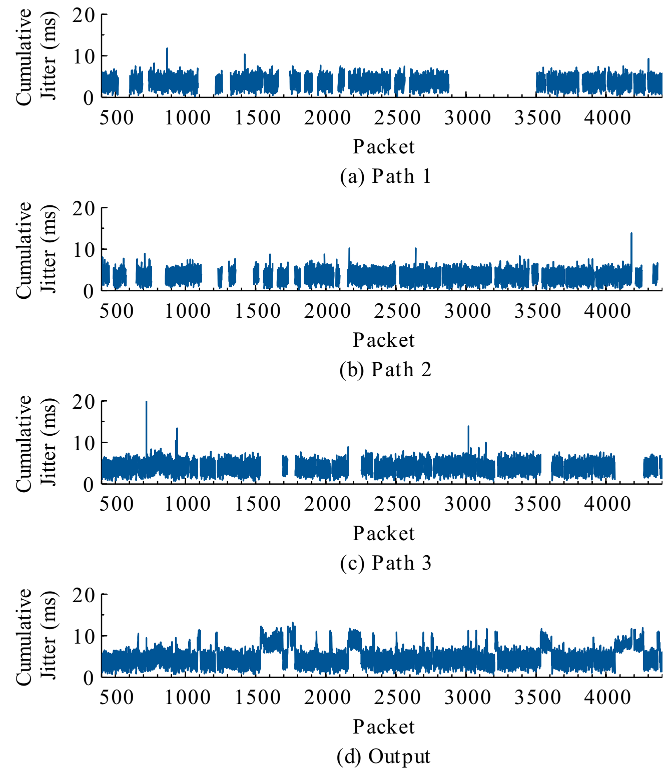

Figure 4. Fragment of simulation illustrating distortions experienced by packets traversing paths 1 – 3 shown by (a) – (c) respectively, and observed at output (d). Jitter shown as *cumulative jitter*, loss as *missing* cumulative jitter.

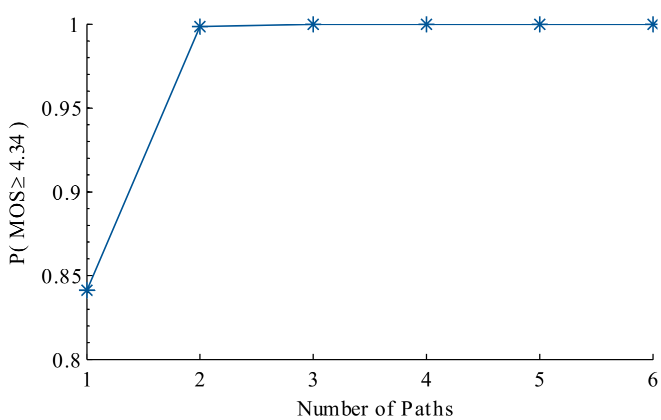

Figure 5. Probabilities of MOS being at least 4.34 for overall observed conditions with 1 path (no dispersity routing) and 2 – 6 paths (with dispersity routing). Two paths already yield significant improvements.

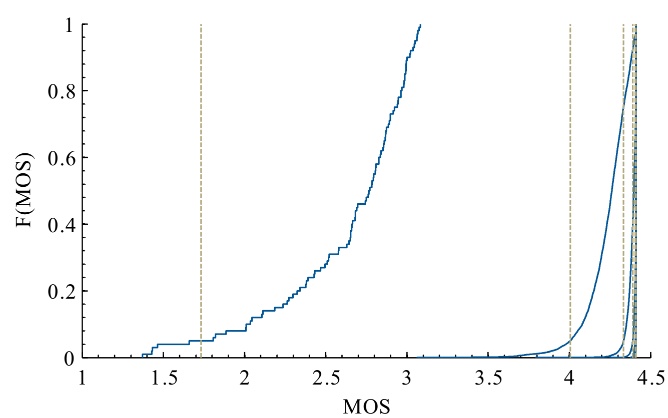

Figure 6. Cumulative distributions of output MOS estimates in extreme conditions for, from left to right, systems with 1 – 6 paths. Also shown is the lowest 5th percentile of these distributions, also from left to right.

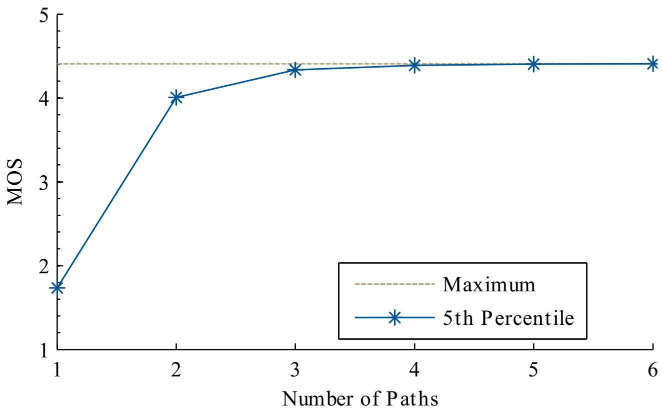

Figure 7. 5th percentiles of output MOS estimates under extreme conditions. Additional paths yield diminishing returns, with biggest gain returned by going from no dispersity routing to dispersity routing with 2 paths.

left to right, the first 5th percentile is for the first cumulative distribution, the second 5th percentile is for the second cumulative distribution and so on. The left-most cumulative distribution shown in Fig. 6 is for the MOS estimates of the call profiles measured in section III with the lowest 100 MOS estimates. This is the deliverable quality for these call profiles in the absence of dispersity routing, that is, a system with 1 path. The remaining five cumulative distributions are, from left to right, the output MOS estimates from the simulations of dispersity routing systems using 2 – 6 paths respectively. Note that the cumulative distribution and 5th percentile for 5 paths are obscured by the almost identical results for 6 paths.

The impact of dispersity routing on deliverable quality is illustrated in Fig. 6 by the shift in the output MOS estimate cumulative distribution towards the maximum possible MOS of 4.4094 for a lossless output here. Fig. 7 summarizes that shift by plotting the 5th percentiles from Fig. 6, highlighting a trend of diminishing returns. However, most importantly this shows that dispersity routing can improve quality even in extreme conditions.

VI. Conclusion

This paper shows that fully redundant dispersity routing can exploit the redundancy inherent in the Internet to increase deliverable VoIP quality as measured by MOS estimates computed using the E-model, reduces loss, has a de-jittering effect through competition among the paths, and reduces mean burst length. Simulations using actual VoIP traffic demonstrate that a small number of paths already yield significant returns. Even using the worst 0.54% of call profiles on record for example, 98.07% of the scenarios simulated with 6 paths drawing path characteristics from those call profiles yield the maximum possible MOS. In typical conditions observed in a commercial call center, where 84.12% of calls may be interpreted as being of a quality that is very satisfactory without dispersity routing, using dispersity routing with just two paths increases that percentage to 99.86%, delivering a premium service that is more on par with traditional telephony than VoIP is currently.

Acknowledgment

The authors would like to thank John Siliquini for his valuable assistance in our many discussions and in the preparation of this paper.

References

[1] ITU-T, "Methods for objective and subjective assessment of quality," Recommendation P.800, 1996.

[2] ITU-T, "The E-model, a computational model for use in transmission planning," Recommendation G.107, 2005.

[3] Y. Cai, "A Thin-Layer Protocol for Utilizing Multiple Paths," in *Proc. 2009 IEEE/RSJ Int. Conf. Intelligent Robots and Syst.*, St. Louis, USA, 2009, pp. 1049–1054.

[4] H. Levy and H. Zlatokrilov, "The Effect of Packet Dispersion on Voice Applications in IP Networks," *IEEE/ACM Trans. Netw.*, vol. 14, pp. 277–288, April 2006.

[5] S. Tao, K. Xu, A. Estepa, T. Fei, L. Gao, R. Guérin, J. Kurose, D. Towsley, and Z.-L. Zhang, "Improving VoIP Quality Through Path Switching," in *Proc. IEEE INFOCOM 2005 24th Annu. Joint Conf. IEEE Comp. Commun. Societies*, Miami, USA, 2005, pp. 2268–2278.

[6] J.-C. Bolot, S. Fosse-Parisis, and D. Towsley, "Adaptive FEC-Based Error Control for Internet Telephony," in *Proc. 18th Annu. Joint Conf. IEEE Comput. and Commun. Soc.*, New York, USA, 1999, pp. 1453–1460.

[7] ITU-T, "One-way transmission time," Recommendation G.114, 2003.

[8] E. N. Gilbert, "Capacity of a Burst-Noise Channel," *The Bell System Technical Journal*, vol. 39, pp. 1253–1265, Sep. 1960.

[9] E. O. Elliott, "Estimates of Error Rates for Codes on Burst-Noise Channels," *The Bell System Technical Journal*, vol. 42, pp. 1977–1997, Sep. 1963.

[10] S. Tao, K. Xu, Y. Xu, T. Fei, L. Gao, R. Guérin, J. Kurose, D. Towsley, and Z.-L. Zhang, "Exploring the Performance Benefits of End-to-End Path Switching," in *Proc. 12th IEEE Int. Conf. Network Protocols (ICNP 2004)*, Berlin, Germany, 2004, pp. 304–315.

[11] S. Tao, "Improving the Quality of Real-Time Applications through Path Switching," Ph.D. dissertation, Univ. Pennsylvania, USA, 2005.

[12] T. Fei, S. Tao, L. Gao, and R. Guérin, "How to Select a Good Alternate Path in Large Peer-to-Peer Systems?," in *Proc. INFOCOM 2006 25th IEEE Int. Conf. Comp. Commun.*, Barcelona, Spain, 2006, pp. 1–13.

[13] B. Boehm and R. Mobley, "Adaptive Routing Techniques for Distributed Communications Systems," *IEEE Trans. Commun.*, vol. 17, pp. 340–349, Jun. 1969.

[14] N. Maxemchuk, "Dispersity Routing in Store-and-Forward Networks," Ph.D. dissertation, Univ. Pennsylvania, USA, 1975.

[15] E. Gustafsson and G. Karlsson, "A Literature Survey on Traffic Dispersion," *IEEE Network*, vol. 11, no. 2, pp. 28–36, 1997.

[16] ITU-T, "Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs," Recommendation P.862, 2001.

[17] J. Sliwinski, A. Beben, and P. Krawiec, "EmPath: Tool to Emulate Packet Transfer Characteristics in IP Network," in *Proc. 2nd Int. Workshop on Traffic Monitoring and Anal. (TMA'10)*, Zürich, Switzerland, 2010.

[18] A. Mukherjee, "On the Dynamics and Significance of Low Frequency Components of Internet Load," Univ. Pennsylvania, USA, Rep. MS-CIS-92-83/DSL-12, 1992.

[19] J.-C. Bolot, "End-to-End Packet Delay and Loss Behavior in the Internet," SIGCOMM Comput. Commun. Rev., vol. 23, no. 4, pp. 289–298, Oct. 1993.

[20] Y. J. Liang, E. G. Steinbach, and B. Girod, "Real-time Voice Communication over the Internet Using Packet Path Diversity," in *Proc. 9th ACM Int. Conf. on Multimedia*, 2001, pp. 431–440.