COMPARISONS OF FEC AND CODEC ROBUSTNESS ON VOIP QUALITY AND BANDWIDTH EFFICIENCY

WENYU JIANG AND HENNING SCHULZRINNE

Columbia University, Department of Computer Science
1214 Amsterdam Ave, Mail Code 0401, New York, NY 10027, USA
Email: {wenyu,hgs}@cs.columbia.edu

Packet loss and delay in the Internet degrades the quality of Voice over IP (VoIP). Forward error correction (FEC) reduces packet loss at the cost of higher delay and bandwidth. An alternative is a loss-robust voice codec, but also at the expense of higher bit-rates. We compare the perceived quality of both schemes under the same bandwidth requirement. We evaluate two conventional low bit-rate codecs, G.729 and G.723.1, and one robust codec, iLBC. G.729 with FEC generally prevails, but when considering delay impairment, iLBC is better under low packet loss. If iLBC also uses FEC, then its quality can sometimes be better than G.729 plus FEC with the same gross bit-rate. The comparison with G.723.1 is similar, except that G.723.1 is less advantageous for low packet loss because its intrinsic quality is slightly lower than iLBC. Finally, given a target bit-rate, we compute the maximum achievable MOS using any of these codecs. It exhibits an upper limit however large the target bit-rate is, due to the inherent delay impairment of FEC.

1 Introduction

Voice over IP uses low bit-rate codecs to save on bandwidth and communication costs. However, packet loss can result in lower voice quality. Because the Internet is still a best-effort network without general quality of service support, users must rely on end-to-end techniques to improve quality.
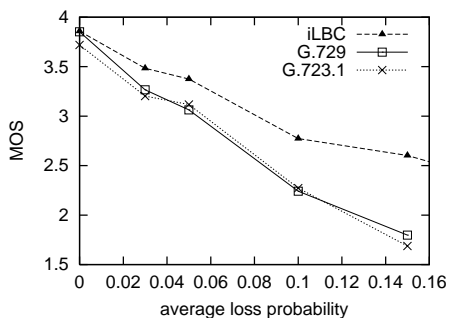
Forward error correction (FEC) [1] recovers lost packets by transmitting redundant data. A commonly used FEC code is the Reed-Solomon [2] code. Its notation is (n, k), where n and k are the number of total and payload units in a block, respectively. It can recover the lost packets if and only if at least k out of n packets in a block are received. FEC's main disadvantage is the additional recovery delay and bandwidth overhead. An (n, k) code introduces a delay of $n \cdot T$ (including packetization delay), where T is the packet interval.

Alternatively we can design a more robust codec. It should be frame-independent, as otherwise one frame loss would cause many subsequent voice frames to be decoded incorrectly. The recent iLBC codec [3] is the only such candidate, [a] which runs at 14 kb/s with a 30 ms frame length.

---

[a] G.711 is also frame-independent, but its high bit-rate (64 kb/s) makes it less suitable for wide area network telephony, where bandwidth can be scarce during congestion

Given the common objective of FEC and robust codec design, it is valuable to know which scheme gives better quality. The perceived quality in VoIP is measured by the Mean Opinion Score (MOS). [4] MOS rates the audio from 5 to 1, with 5 being Excellent, 4 Good, 3 Fair, 2 Poor, and 1 Bad, respectively. MOS is obtained by listening tests, where each listener rates an audio clip from 5 to 1, and the resulting average is the MOS.

Figure 1(a) shows the MOS performance of iLBC, G.729 [5] (8 kb/s) and G.723.1 [6] (6.3 kb/s) under packet loss. The MOS of iLBC at 15% loss is close to G.729 under 8% loss. Such robustness comes at a slightly higher bit-rate, because iLBC encodes every frame independently as opposed to incrementally.

	codec	media pkt_size	r_{media}	r_{IP}
no FEC	iLBC	54 bytes	14.4kb/s	25.1kb/s
	G.729	30 bytes	8kb/s	18.7kb/s
	G.723.1	24 bytes	6.4kb/s	17.1kb/s
(4,2) FEC	iLBC	108 bytes	28.8kb/s	39.5kb/s
	G.729	60 bytes	16kb/s	26.7kb/s
	G.723.1	48 bytes	12.8kb/s	23.5kb/s

(a) MOS performance without FEC [7]   (b) Bandwidth overhead of FEC

Figure 1. Example codec MOS performance and FEC overhead; packet interval T=30 ms

The overhead of FEC is not necessarily high once we consider the typical 40 byte IP/UDP/RTP [8] packet header. Assuming piggy-backing [9] in FEC, the media and IP bit-rates for an (n, k) code are then calculated as:

$$r_{media} = r_{codec} \cdot \frac{n}{k}, \qquad r_{IP} = r_{media} + \frac{40 bytes}{T} \qquad (1)$$

where r_{codec} is the codec bit-rate, T is the packet interval. Figure 1(b) shows some typical packet sizes and bit-rates FEC introduces.

Our main contributions include an original study on the quality trade-off between FEC and codec robustness under packet loss with the same bandwidth requirement, evaluating both listening and *conversational* quality, where delay is an important factor. We also investigate the best quality attainable under packet loss utilizing both schemes within a target bit-rate.

The rest of this paper is organized as follows. Section 2 shows how to predict MOS quality of FEC. Section 3 compares FEC and codec robustness. Section 4 discusses the effect of target bit-rate on achievable quality. Section 5 lists related work and section 6 concludes the paper.

2 Prediction of MOS Quality under FEC

2.1 The E-model

MOS listening tests are time consuming. An alternative is the E-model, [10] an analytical model for predicting MOS. Cole and Rosenbluth [11] discusses its application in a VoIP environment by simplifying the model into two parameters, loss and delay. The procedures are summarized below:

1. Translate loss and delay into impairment scores, denoted as I_e and I_d, respectively, as in Figure 2. Mapping for I_e is derived from MOS curves.

2. Compute a score called the R value as follows:

$$R = 94.2 - I_d - I_e \qquad (2)$$

3. Map the R value to MOS with a fixed, reversible mapping function. [11]

$$MOS = 1 + 0.035 \cdot R + 7 \times 10^{-6} \cdot R \cdot (R - 60) \cdot (100 - R) \qquad (3)$$

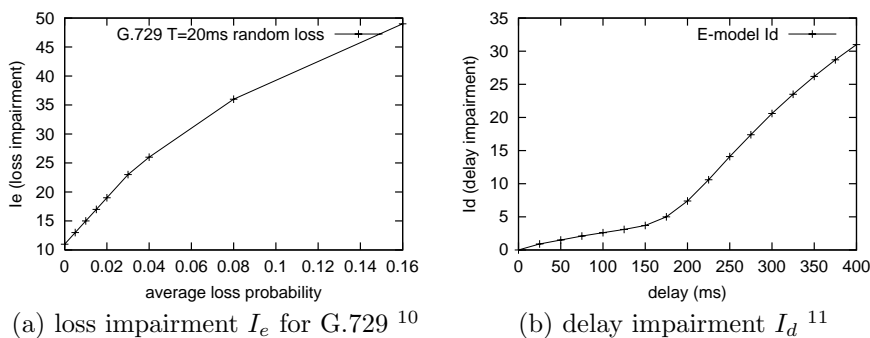(a) loss impairment I_e for G.729 [10]   (b) delay impairment I_d [11]

Figure 2. Effect of loss and delay in the E-model

In an earlier paper [9] we present a method for predicting FEC MOS based on the E-model. First we reuse an existing MOS curve (e.g., Figure 1) and reverse map MOS to R value by inverting equation (3). Then R is mapped to I_e, where $I_e = 94.2 - R$, a reverse application of equation (2). This way we establish a loss to I_e mapping just as in Figure 2(a). Second, we compute the final loss probability (denoted as p_f) after applying FEC. The method of calculating p_f is either analytical for simple FEC codes, [9] or by simulation for complex FEC codes. Then, we plug p_f into our reverse-engineered loss to I_e mapping and obtain I_e. The delay of an (n, k) code is $n \cdot T$, which is translated into I_d using Figure 2(b). From there we can compute the R value and obtain the FEC MOS. We have compared this analytical method with our real FEC MOS test results [9], and the analytical prediction is fairly accurate.

3 Quality Evaluation of FEC vs. Codec Robustness

We compare iLBC, G.729 and G.723.1 [b] and our evaluation comprises three steps. First, only G.729 uses FEC. Next, we add FEC to both iLBC and G.729 while keeping the same gross bit-rate between them. Finally, we repeat these steps for G.723.1 and iLBC. We examine both *listening* MOS and *conversational* MOS (MOS_c), with the latter considering delay impairment.

Packet loss in the Internet is often bursty, [12] so we use the Gilbert model [9] to simulate bursty loss. In our evaluation, loss burstiness, expressed in conditional loss probability, is set to 30% at 30 ms packet interval.

3.1 G.729 with FEC vs. iLBC without FEC

Figure 3 compares the plain iLBC and G.729 with a (5,3) FEC code. Thus G.729 has a gross bit-rate of $8 \times 5/3 = 13.3$ kb/s, about the same as iLBC (14 kb/s). From Figure 3 through 9, sub-figure (a) and (b) compare the *listening* and *conversational* MOS, respectively.
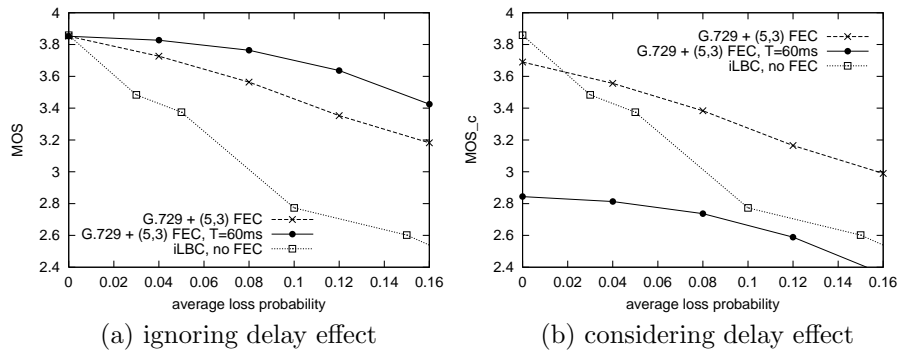

(a) ignoring delay effect   (b) considering delay effect

Figure 3. G.729 + (5,3) FEC code vs. iLBC without FEC; T=30 ms by default

In Figure 3(a) MOS is highest for G.729 with a 60 ms interval, because FEC is more effective with a larger packet interval under bursty loss. [9] Here both G.729 curves out-perform iLBC. But once we consider delay impairment, iLBC then prevails for low loss ($< 2\%$), as in Figure 3(b). The G.729 curve starts lower than iLBC, because its FEC delay ($5 \times 30 = 150$ ms) attributes to a small impairment. However, beyond 2% loss, G.729 becomes increasingly better than iLBC due to its low packet loss after FEC. Finally, G.729 at

---

[b] We did not choose GSM EFR, AMR FR, or G.728 because their bit-rates are close to or even higher than iLBC, making the comparison moot or uninteresting.

60 ms now performs quite poorly, because its FEC delay is too high (5×60 = 300 ms). In Figure 2(b) delay impairment rises drastically beyond 150 ms. Consequently, the gain in quality by using a larger packet interval is usually not enough to offset the excessive delay impairment. Therefore a rule of thumb is to keep FEC delay within 150 ms, a requirement for an average interactive conversation. If a task or conversation is highly interactive, a lower threshold such as 100 ms may be used, and all we need to do is to adjust the delay impairment mapping in Figure 2(b) accordingly.

Our first result indicates G.729 with FEC is better than iLBC for a similar gross bit-rate, except when delay is considered and loss is low ($< 2\%$).

3.2 Adding FEC to Both G.729 and iLBC, Keeping Same Gross Bit-rate

If we can afford more than 14 kb/s, then iLBC may use FEC as well, as in Figure 4, where iLBC uses a (3,2) code, resulting in a 21 kb/s gross bit-rate. G.729 uses a (5,2) code, giving a comparable 20 kb/s.
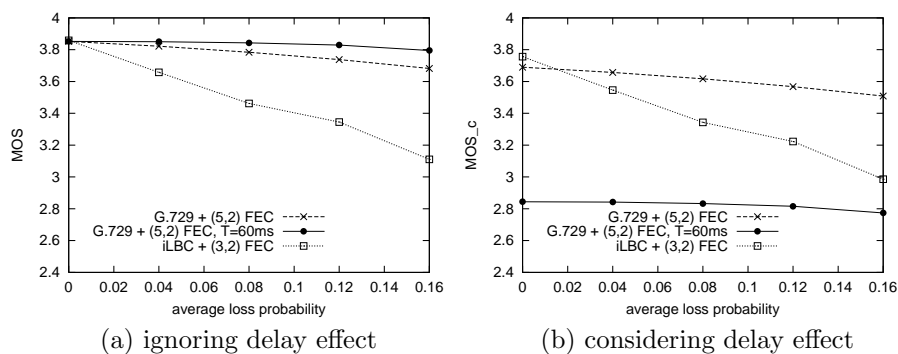

(a) ignoring delay effect   (b) considering delay effect

Figure 4. G.729 + (5,2) FEC code vs. iLBC + (3,2) FEC code; T=30 ms by default

In Figure 4(a), both G.729 curves are quite similar, because the FEC codes are already so redundant that virtually all losses are recovered. The quality of iLBC is much improved with FEC, but still lower than G.729. When considering delay, similar to Figure 3(b), iLBC is better when loss is low ($< 2\%$), but G.729 prevails afterwards.

In Figure 5, iLBC employs a (4,2) code, giving a 28 kb/s gross bit-rate. G.729 uses a long (7,2) code, resulting in 28 kb/s as well. However, we do not show the G.729 60 ms curve, because its difference with G.729 30 ms curve is minimal when ignoring delay but too big otherwise. In Figure 5(a), iLBC still falls short of G.729. However, in Figure 5(b), iLBC is noticeably better

than G.729 for loss rate as high as 14%, because G.729's FEC delay is 7×30 = 210 ms, 60 ms larger than the 150 ms threshold in Figure 2(b). This is the case where iLBC dominates.

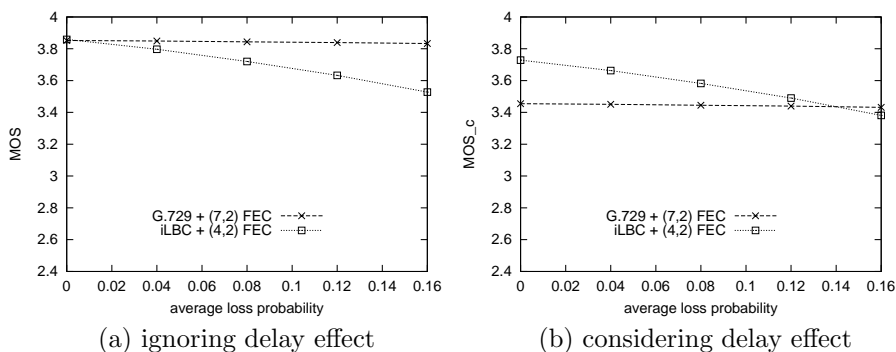(a) ignoring delay effect (b) considering delay effect

Figure 5. G.729 + (7,2) FEC code vs. iLBC + (4,2) FEC code; T=30 ms by default

The lower MOS of G.729 in Figure 5(b) is due to long FEC delay. By using a shorter FEC code, even with some concession in the bit-rate, higher conversational MOS may be possible. This is illustrated in Figure 6, where iLBC still uses a (4,2) code, but G.729 changes to a (3,1) code, using 24 kb/s instead of 28 kb/s. The result: in both Figure 6(a) and (b), iLBC and G.729 perform very alike. Therefore sacrificing some FEC redundancy or using shorter FEC blocks may actually result in better conversational MOS.
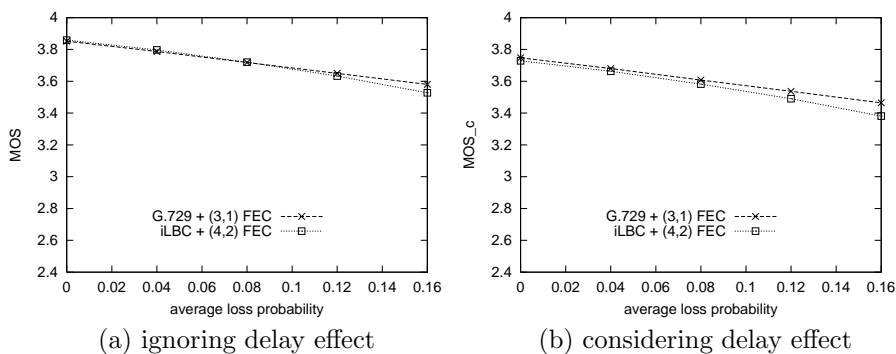

(a) ignoring delay effect (b) considering delay effect

Figure 6. G.729 + (3,1) FEC code vs. iLBC + (4,2) FEC code; T=30 ms by default

3.3 Comparison with G.723.1

Due to its lower bit-rate, G.723.1 has a lower *intrinsic* MOS at 0% loss than
G.729, but this also allows more FEC redundancy. In Figure 7, G.723.1 uses
a (2,1) FEC code (12.6 kb/s), and iLBC without FEC. Similar to Figure 3(b),
iLBC is better under low loss, but its dominant loss range is wider when
considering delay. In Figure 7(b), G.723.1 performs better at 60 ms than at
30 ms, because its FEC delay is only $2 \times 60 = 120$ ms < 150 ms threshold.

(a) ignoring delay effect   (b) considering delay effect
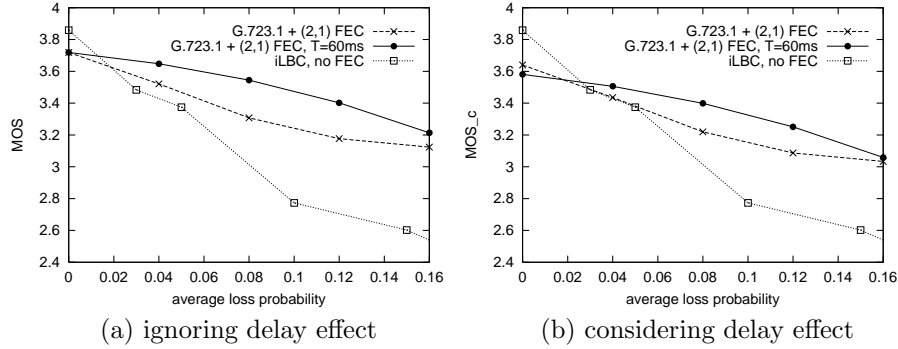
Figure 7. G.723.1 + (2,1) FEC code vs. iLBC without FEC; T=30 ms by default

Figure 8 compares iLBC with a (3,2) code (21 kb/s) and G.723.1 with a
(3,1) code (18.9 kb/s). Again, iLBC is better for low loss (< 3-6%).
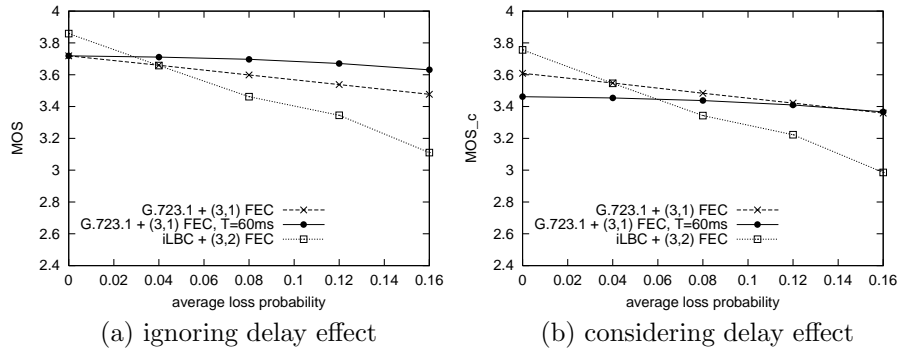

(a) ignoring delay effect   (b) considering delay effect

Figure 8. G.723.1 + (3,1) FEC code vs. iLBC + (3,2) FEC code; T=30 ms by default

Finally, Figure 9 compares iLBC with (4,2) code (28 kb/s) and G.723.1
with (4,1) code (25.2 kb/s). In both Figure 9(a) and (b), iLBC is the best for
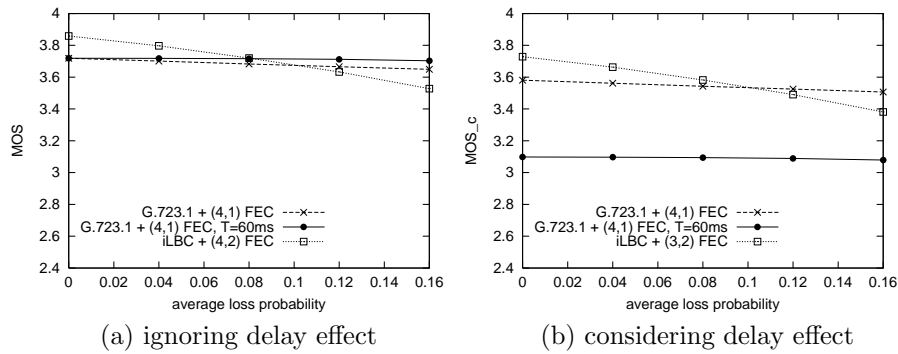at least up to 8% loss, because G.723.1 has a lower intrinsic MOS.

(a) ignoring delay effect   (b) considering delay effect

Figure 9. G.723.1 + (4,1) FEC code vs. iLBC + (4,2) FEC code; T=30 ms by default

4   Discussion: The Best of Both Worlds

In section 3, a common observation when considering delay is that iLBC prevails under low loss, but G.729 or G.723.1 is better for medium to high loss. Since none of them dominate solely, it is only beneficial to unite the best of both schemes. Therefore, we compute for a particular target bit-rate the maximum MOS these codecs can achieve. We consider conversational MOS (MOS_c) only, as it is realistic to VoIP.

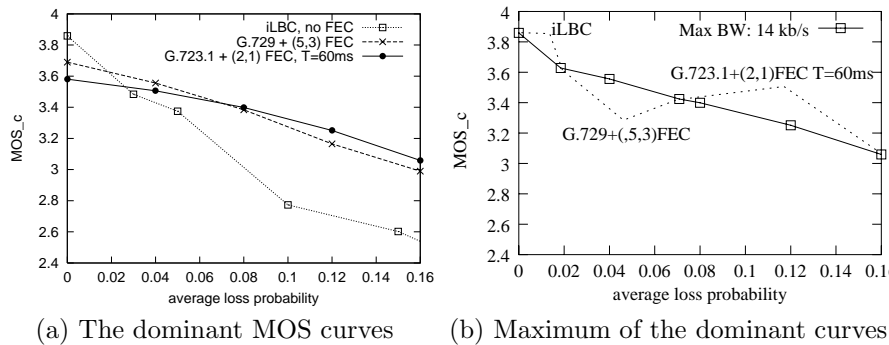(a) The dominant MOS curves   (b) Maximum of the dominant curves

Figure 10. Maximum conversational MOS achievable within 14 kb/s; T=30 ms by default

The first target bit-rate we evaluate is 14 kb/s. Based on Figure 3(b) and 7(b), we find the highest MOS_c curves, [c] as shown in Figure 10(a). Then

---

[c]With a 10 ms frame, G.729 can use 40 ms, 50 ms or even 20 ms packet intervals too, but these cases turn out not to correspond to the dominant curves

we take their maximum and arrive at the curve in Figure 10(b). This max MOS_c curve has three segments, the first portion (low loss) corresponds to iLBC without FEC, the second part (medium loss) corresponds to G.729 + (5,3) FEC. The last segment (high loss) is dominated by G.723.1 + (2,1) FEC with a 60 ms packet interval, because G.723.1's lower bit-rate allows more FEC redundancy and its FEC delay (120 ms) is not too high.
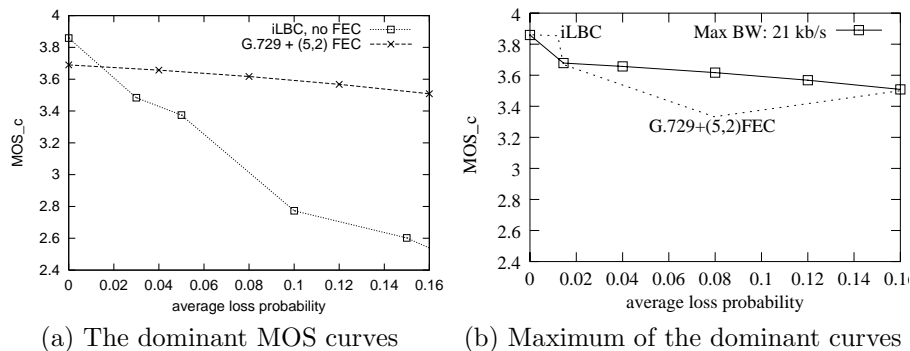

(a) The dominant MOS curves   (b) Maximum of the dominant curves

Figure 11. Maximum conversational MOS achievable within 21 kb/s; T=30 ms by default

Figure 11 and 12 demonstrate the results for higher target bit-rates, 21 kb/s and 28 kb/s, respectively. Figure 11(b) has two segments, the first one (low loss) again corresponds to iLBC with no FEC, and the second is G.729 + (5,2) FEC. G.723.1 + (3,1) FEC is also within 21 kb/s, but its MOS curve is lower than G.729 + (5,2) FEC, therefore not listed for brevity.
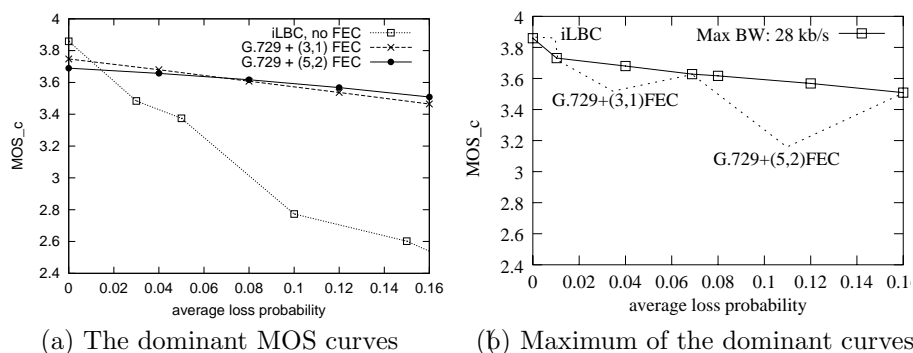

(a) The dominant MOS curves   (b) Maximum of the dominant curves

Figure 12. Maximum conversational MOS achievable within 28 kb/s; T=30 ms by default

Figure 12(b) has three segments. The first one (low loss) again belongs to iLBC without FEC, the second part is G.729 + (3,1) FEC. The third segment is G.729 + (5,2) FEC, the same as in Figure 11(b). Although a (3,1) code has more redundancy, its shorter block can sustain only 3-1 = 2 consecutive losses. A (5,2) code can sustain 5-2 = 3 consecutive losses and its FEC delay is within the 150 ms threshold in Figure 2(b).

Finally, we plot in Figure 13 a comparison of Figure 10(b), 11(b), and 12(b). Figure 13 demonstrates the quality realizable for a certain target bit-rate with the current state-of-the-art VoIP technology. The gain in MOS_c is quite big from 14 kb/s to 21 kb/s, but the improvement is marginal from 21 kb/s to 28 kb/s. Although a higher target bit-rate allows more FEC redundancy, due to FEC delay impairment, there is an inherent limit to MOS_c.
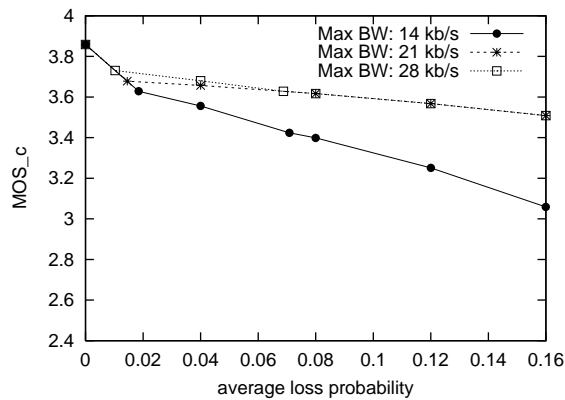

Figure 13. Effect of target bit-rate on maximum achievable conversational MOS

5   Related Work

Bolot [13] discusses adaptation of FEC to maximize quality. However, the FEC used there is really low bit-rate redundancy (LBR), [14] which is less general than FEC and it cannot recover lost audio with 100% accuracy. The perceived quality of LBR is also lower than FEC under the same gross bit-rate. [9] The utility functions used by Bolot are hypothetical and do not correspond to real quality measures, whereas our results are based on real MOS curves. Finally we evaluate the effect of both loss and delay, instead of just loss.

We are not aware of other work that addresses the trade off between robust codec design and FEC or that derives the maximum attainable perceived quality with respect to packet loss within a target bit-rate.

6 Conclusions and Future Work

We evaluate the perceived quality of conventional vs. robust low bit-rate codecs under the same bandwidth requirement, using the Mean Opinion Score (MOS) metric. When ignoring delay, the conventional codec with FEC is generally better. When considering delay impairment, the robust codec (iLBC) is better under low loss conditions. In some cases with high FEC redundancy, iLBC can dominate under most loss conditions.

We compute for a given target bit-rate, the quality achievable using either FEC or robust codec. We find that from $14\,\mathrm{kb/s}$ to $21\,\mathrm{kb/s}$ target bit-rate, there is a significant quality improvement. Yet from $21\,\mathrm{kb/s}$ to $28\,\mathrm{kb/s}$, the gain is marginal especially beyond 1% packet loss, limited inherently by increasing delay impairment introduced by FEC. Such limitation is unavoidable unless multiple communication paths are available, albeit at the expense of sending more packets and higher packet header overhead. Our evaluation reveals the quality realizable with the current state-of-the-art VoIP technology.

MOS results are subjective and require listening tests. However, the MOS performance of a given codec need only be measured once, and the remaining prediction of FEC MOS is done analytically, as outlined in Section 2. Our near-term plan is to implement the MOS prediction procedure in a VoIP application for real-time decision support on network adaptation. Another plan is to evaluate the effect of jitter on VoIP quality. Despite being effective against loss, FEC cannot reduce jitter unless out-of-order packets are common in the Internet. In contrast, packet loss concealment in the robust codec incurs a much lower delay. So we conjecture that a robust codec is more useful for battling jitter.

References

1. J. Rosenberg and H. Schulzrinne. An RTP payload format for generic forward error correction. RFC 2733, Internet Engineering Task Force, December 1999.
2. L. Rizzo. Effective erasure codes for reliable computer communication protocols. *ACM Computer Communication Review*, 27(2):24–36, April 1997.

3. S. Andersen et al. Internet low bit rate codec. Internet Draft, Internet Engineering Task Force, February 2002. Work in progress.

4. International Telecommunication Union. Subjective performance assessment of telephone-band and wideband digital codecs. Recommendation P.830, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, February 1996.

5. International Telecommunication Union. Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear-prediction. Recommendation G.729, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, March 1996.

6. International Telecommunication Union. Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s. Recommendation G.723.1, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, March 1996.

7. Soren Andersen and Alan Duric. IETF AVT WG draft-andersen-ilbc-00. Presentation, Global IP Sound, March 2002. Available at http://www.ietf.org/proceedings/02mar/slides/avt-6.pdf.

8. H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson. RTP: a transport protocol for real-time applications. RFC 1889, Internet Engineering Task Force, January 1996.

9. Wenyu Jiang and Henning Schulzrinne. Comparison and optimization of packet loss repair methods on voip perceived quality under bursty loss. In *Proc. International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV)*, May 2002.

10. International Telecommunication Union. Application of the e-model: A planning guide. Recommendation G.108, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, September 1999.

11. R. G. Cole and J. Rosenbluth. Voice over IP performance monitoring. *ACM Computer Communication Review*, 4(3), 2001.

12. Maya Yajnik, Sue Moon, Jim Kurose, and Don Towsley. Measurement and modelling of the temporal dependence in packet loss. In *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, New York, March 1999.

13. Jean-Chrysostome Bolot, Sacha Fosse-Parisis, and Don Towsley. Adaptive FEC-Based error control for interactive audio in the Internet. In *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, New York, March 1999.

14. Vicky Hardman, Angela Sasse, Mark Handley, and Anna Watson. Reliable audio for use over the Internet. In *Proc. of INET'95*, Honolulu, Hawaii, June 1995.