Waveform Codecs

- Waveform codecs provide a digital representation of the sound's waveform
- Waveform codecs are relatively simple, they have good quality, but bit rate is relatively large

The material of this section is mainly drawn from: Sigcomm '01: "Chord: A Scalable Peertopeer Lookup Service for Internet Applications", by Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, Hari Balakrishnan

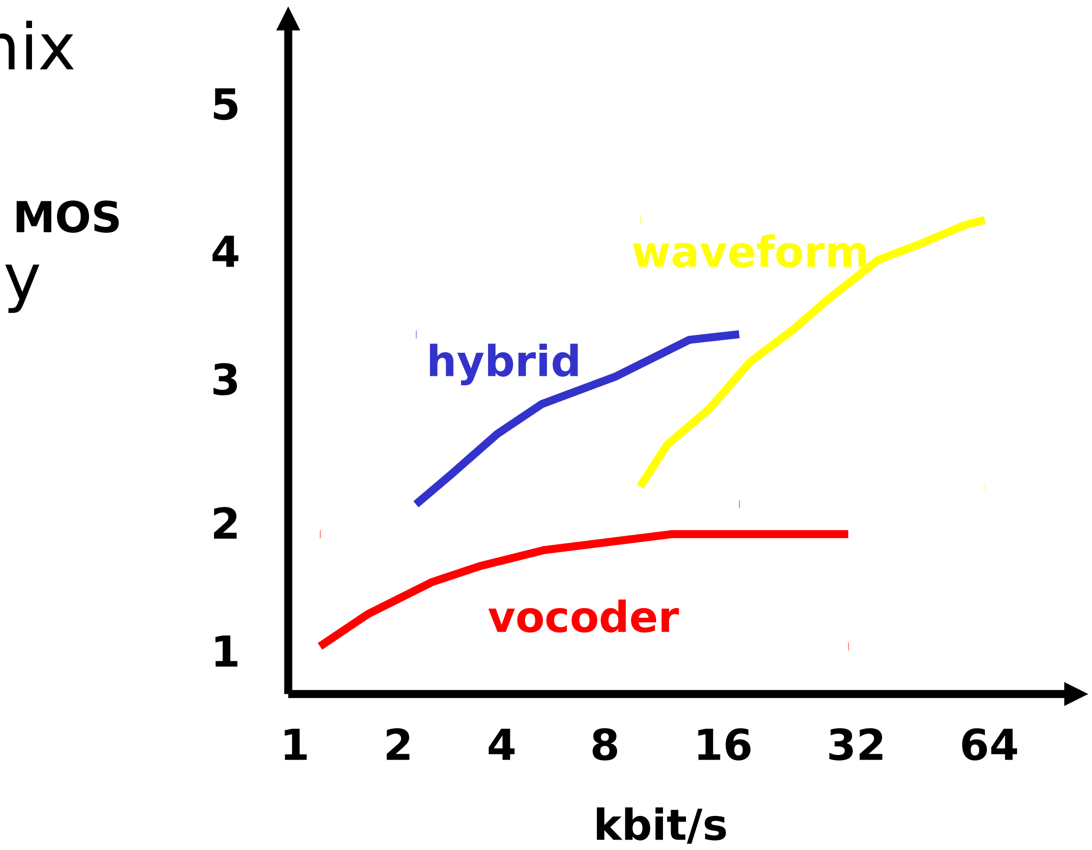
Vocoders

- A vocoder implements a digital filter model of the human system to produce sounds
- Usually linear digital filters are used, and also a system of filter excitations are provided
- The information sent by vocoders through the network are the filter parameters together with the description of the filter's inputs
- Vocoders consume less bandwidth than wavefor codecs, but quality in general is not satisfactory

The material of this section is mainly drawn from: Sigcomm '01: "Chord: A Scalable Peertopeer Lookup Service for Internet Applications", by Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, Hari Balakrishnan

Voice codecs

- Hybrid codecs mix the two approaches and strike a generally good tradeoff between bandwidth and quality



Voice quality

- The most important metric for voice quality is the MOS (mean opinion score)
- MOS ranges from 1 to 5:
  - Excellent: 5
  - Good: 4
  - Fair: 3
  - Poor: 2
  - Bad: 1
- MOS is a subjective metric

Voice quality

- The standard PCM codec @ 64 kbps has an intrinsic MOS of about 4.3
- When a coded is used for IP telephony, its intrinsic MOS is a reference performance figure, given a perfect network transport
- Since the network can delay and loose packets, the actual quality is lower than the reference intrinsic MOS of the adopted codec

The material of this section is mainly drawn from: Sigcomm '01: "Chord: A Scalable Peertopeer Lookup Service for Internet Applications", by Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, Hari Balakrishnan

Voice quality

- The main impairment factors for voice communication are:
  - Intrinsic MOS;
  - Degradation due to delay;
  - Degradation due to packet loss;
  - Degradation due to echo
  - ...

E-model

- Per prevedere, almeno approssimativamente, la qualita’ del segnale vocale su un collegamento IP, si puo’ fare ricorso al E-model
- Il E-model e’ standardizzato dalla ITU-T, e consente di eseguire una previsione approssimata del MOS che si potra’ ottenere su di un collegamento telefonico IP, dati il codec, il ritardo, la perdita di pacchetti ed altri eventuali fattori di degrado del segnale

E-model

- Utilizzare il E-model presenta il vantaggio di poter effettuare delle previsioni e quindi di condurre un progetto preliminare
- Come svantaggio si ha l'approssimazione della previsione, da verificarsi poi sul campo
- Il E-model, sebbene standardizzato dalla ITU-T, non e' ancora supportato dalla ITU-T stessa come metodo valido per la previsione della qualita' della telefonia su IP
- D'altra parte, nell'industria il E-model e' ampiamente utilizzato

E-model

- E-model is a standard algorithm to calculate the MOS level of a voice communication, given objective performance measures such as delay and loss
- A metric *R is defined, ranging from 1 to 100*
- The *R* quality is defined as:
  - $R = R_0 - Is - Id - Ie + A$

The material of this section is mainly drawn from: Sigcomm '01: "Chord: A Scalable Peertopeer Lookup Service for Internet Applications", by Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, Hari Balakrishnan

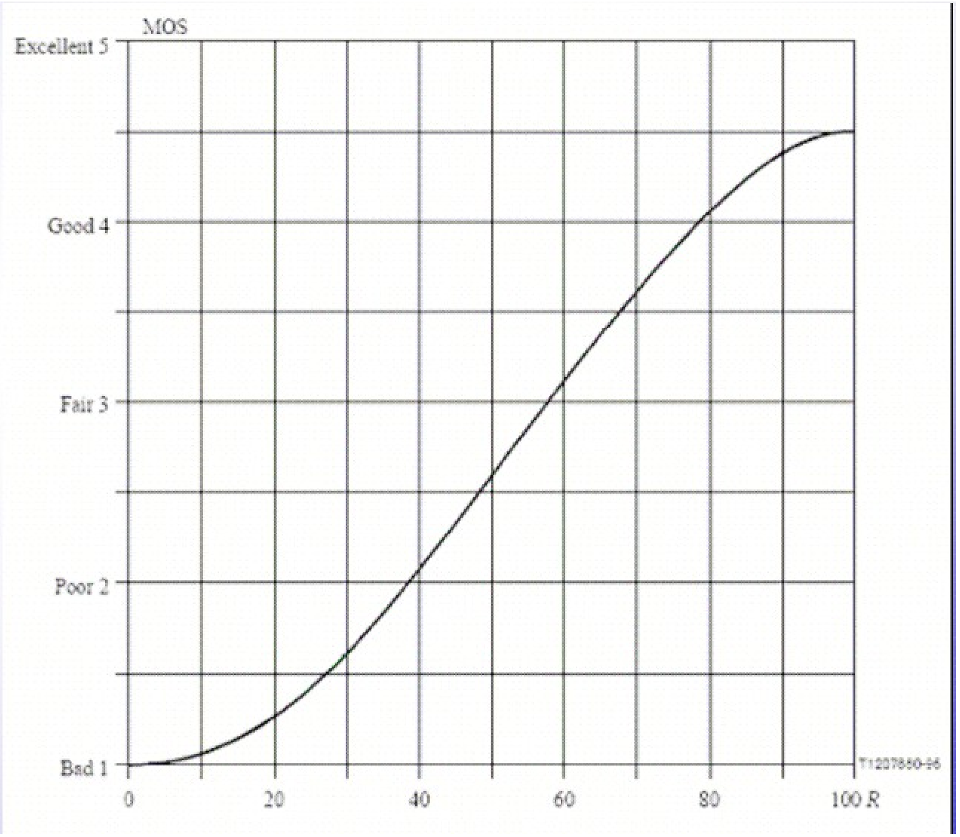
E-model

- where
  - R_0 is the base signal to noise ratio;
  - *Is* standard degradation factors due to voice coding;
  - *Id* impairments due to delay;
  - *Ie* impairments due to loss;
  - *A*: user's advantage factor (user is hard-to-reach locations is usually more willing to sustain performance degradations)

E-model

- Given *R*, MOS is obtained through the plotted curve

$$MOS = \begin{cases} R < 6{,}5 & 1 \\ 6{,}5 \le R \le 100 & 1 - \frac{7}{1000}R + \frac{7}{6250}R^2 - \frac{7}{1000000}R^3 \\ R > 100 & 4{,}5 \end{cases} \quad (4.9)$$



The material of this section is mainly drawn from: Sigcomm '01: "Chord: A Scalable Peertopeer Lookup Service for Internet Applications", by Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, Hari Balakrishnan

Some relevant voice codecs

Class	name	Speed (kbit/s)	Ie	R	MOS
ADPCM	G.726	40	2	91.34	4.37
	G.721(1988) G.726,G.727	32	7	86.34	4.24
	G.726,G.727	24	25	68.34	3.52
	G.726,G.727	16	50	43.34	2.23
LD-CELP	G.728	16	7	86.34	4.24
		12.8	20	73.34	3.75
CS-ACELP	G.729	8	10	83.34	4.14
	G.729-A+VAD	8	11	82.34	4.11
VSELP	IS-54	8	20	73.34	3.75
ACELP	IS-641	7.4	6	87.34	4.27
QCELP	IS-96 a	8	19	74.34	3.79
RCELP	IS-127	8	6	87.34	4.27
VSELP	Japanese PDC	6.7	24	69.34	3.57
RPE-LTP	GSM 06.10, Full rate	13	20	73.34	3.75
VSELP	GSM 06.20, Half rate	5.6	23	70.34	3.61
ACELP	GSM 06.60, Enhanced Full Rate	12.2	5	88.34	4.30
ACELP	G.723.1	5.3	19	74.34	3.79
MP-MLQ	G.723.1	6.3	15	78.34	3.96

The material of this section is mainly drawn from: Sigcomm ’01: “Chord: A Scalable Peertopeer Lookup Service for Internet Applications”, by Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, Hari Balakrishnan

G.711

- G.711 is the standard PCM codec
- It is a waveform codec, with a sampling rate of 8 kHz, non-linear quantization, 256 quantization levels, A or μ laws
- Bit rate: 64 kbit/s
- MOS: 4.3

ADPCM

- ADPCM: adaptive differential PCM
- The codec codes the difference between consecutive voice samples, instead of the absolute value of samples
- For example, G.721 and G.726, have bit rate of 16, 24, 32, 40 kbit/s
- At 32 kbit/s, MOS is about 4.0

CELP codecs

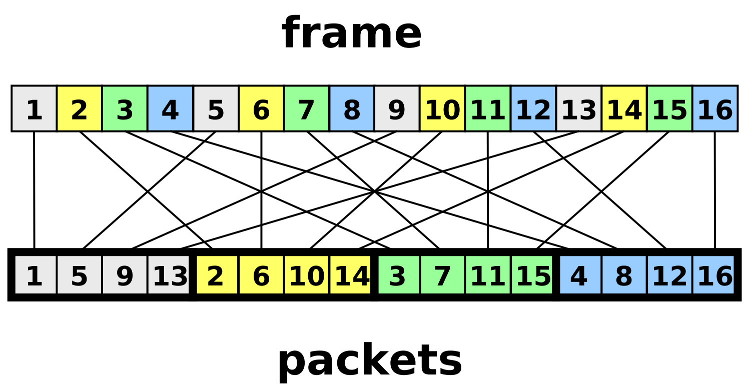
- CELP: code-excited linear predictive (usually they are hybrid codecs)
- Filter inputs are stored in a codebook
- G.728 operates on groups of 5 samples, it has a codebook of 1024 filter inputs
- G.728 works at 16 kbit/s

G.723.1 ACELP

- ACELP: algebraic CELP
- It can operate with small bandwidth (5.3 kbit/s)
- The intrinsic MOS is about 3.8
- Coding delay is about 37.5 ms

Packet loss concealment

- L'*interleaving* is used to reduce the impact of the loss of voice packets
- This is done by distributing on a longer time frame consecutive packet losses
- In the example, 16 frames are scattered in 4 packets
- The loss of one packet results is the lost of 4 dinstant frames, instead of 4 consecutive frames
- The impact of a packet loss is smaller
- However, delay is larger and the combined effect of minor impact and increased delay is not guaranteed to be positive



The material of this section is mainly drawn from: Sigcomm '01: "Chord: A Scalable Peertopeer Lookup Service for Internet Applications", by Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, Hari Balakrishnan

Forward Error Correction

- Forward Error Correction (FEC) is used to cope with voice packet losses
- FEC works by sending redundant packets by which the received can correct errors due to losses
- With parity encoding, every *n* voice packets, an additional packet containing the bitwise XOR of the previous n packets is transmitted
- The receiver can compensate the loss of 1 packet in each group of n packets
- With *n*=5, the consumed bandwidth increases by 20%
- However, delay increases, as time to recever errors must be allowed

The material of this section is mainly drawn from: Sigcomm '01: "Chord: A Scalable Peertopeer Lookup Service for Internet Applications", by Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, Hari Balakrishnan

Forward Error Correction

- With the piggybacking FEC (pFEC) the XOR packet is transmitted in the payload of the nth packet, rather then send it as an independent packet
- This decreases the required bandwidth (the header of VoIP packets is large, relatively to payload size)

The material of this section is mainly drawn from: Sigcomm '01: "Chord: A Scalable Peertopeer Lookup Service for Internet Applications", by Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, Hari Balakrishnan

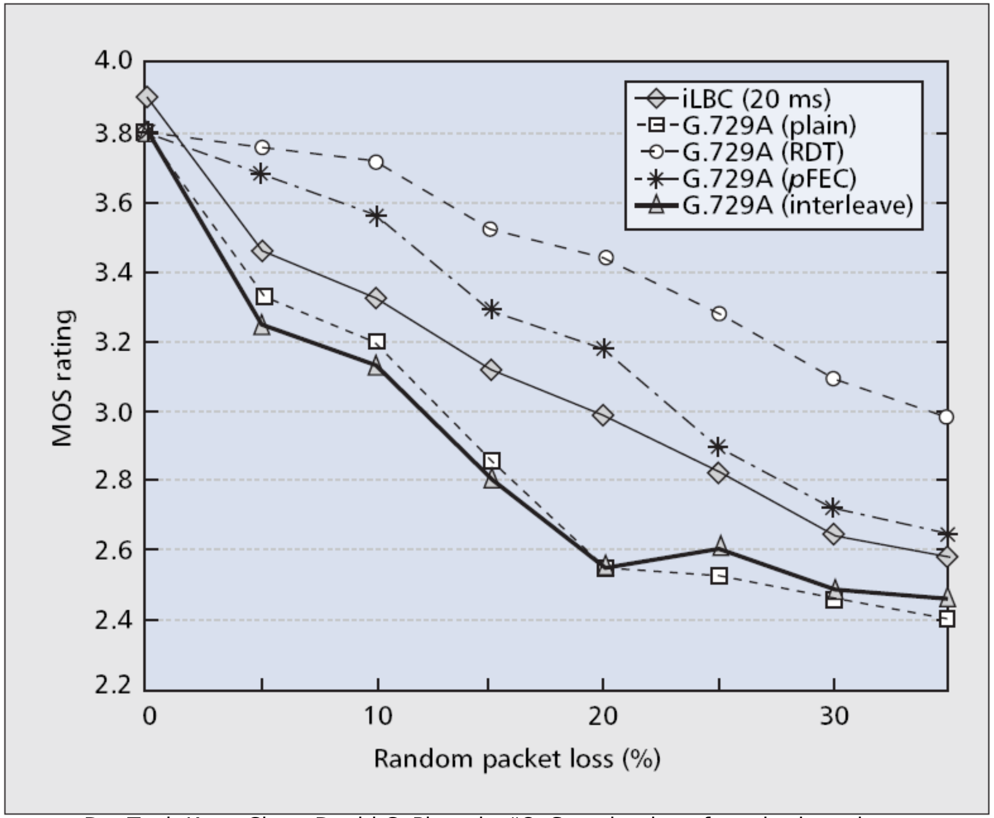
Redundant data transmission

- With redundant data transmission each packet carries the payload that would be transported by two packets in plain delivery
- For example, one packet could transmit frames 1 and 2, the next packets frames 2 and 3, and so on
- Clearly, bandwidth increases significantly
- With the Duplicate Packet approach, each packet is sent two times

The material of this section is mainly drawn from: Sigcomm '01: "Chord: A Scalable Peertopeer Lookup Service for Internet Applications", by Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, Hari Balakrishnan

Voice quality

- The figure plots experimental data on the quality of the voice signal with network packet loss, given different packet loss concealment techniques
- With the G.729 codec, interleaving does not provide positive effects
- However, other techniques provide better results (at the expense of higher bandwidth)



Da: Teck-Kuen Chua, David C. Pheanis, “QoS evaluation of sender-based loss-recovery techniques for VoIP”, *IEEE Network Magazine*, Vol. 20, NO. 6, Nov/Dec 2006, pp. 14-22