Modeling the effects of Burst Packet Loss and Recency on Subjective Voice Quality

Alan Clark

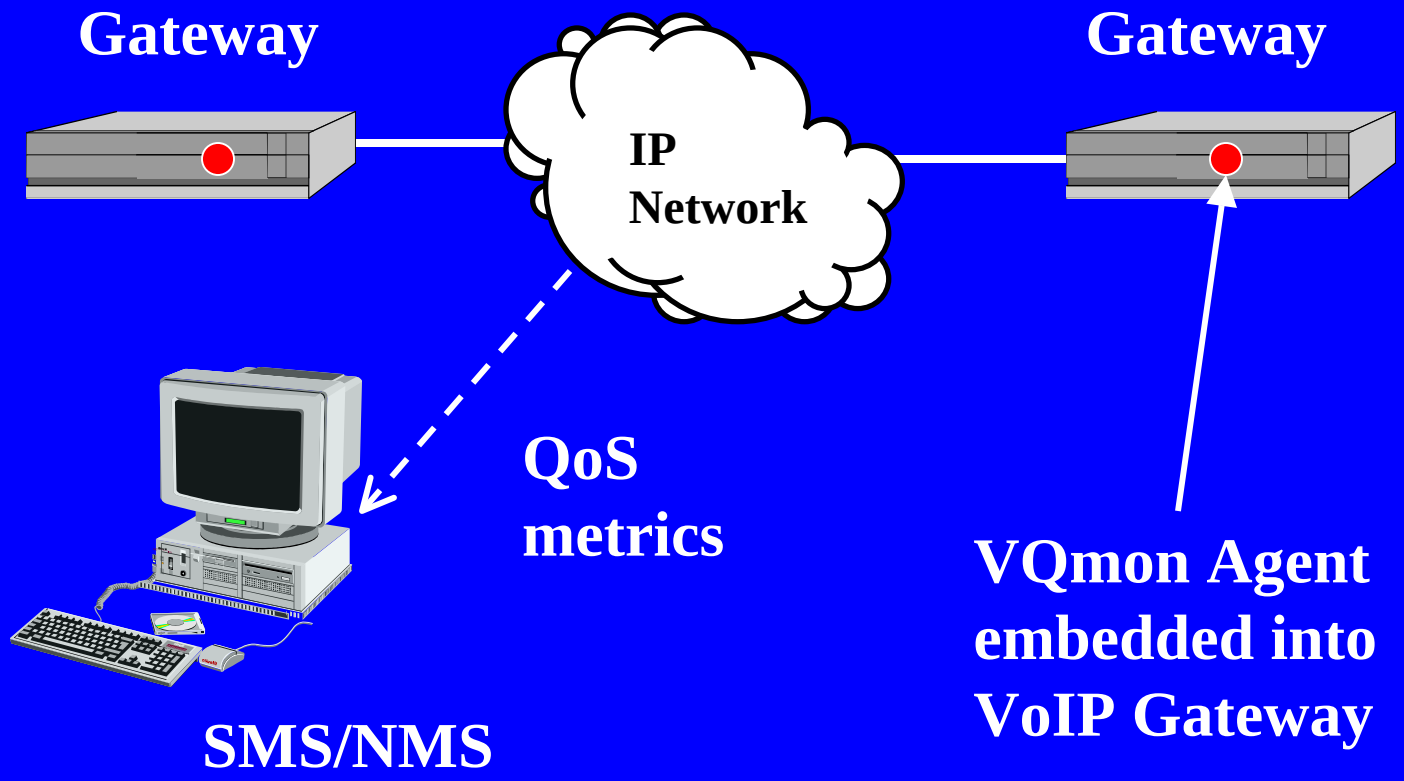
Telchemy

alan@telchemy.com

Embedded Monitoring

- Need to monitor QoS to provide feedback on network performance / impact on subjective quality
- Desirable to provide monitoring in the form of a lightweight software agent
- Focus on time varying impairments – burst packet loss and “recency”

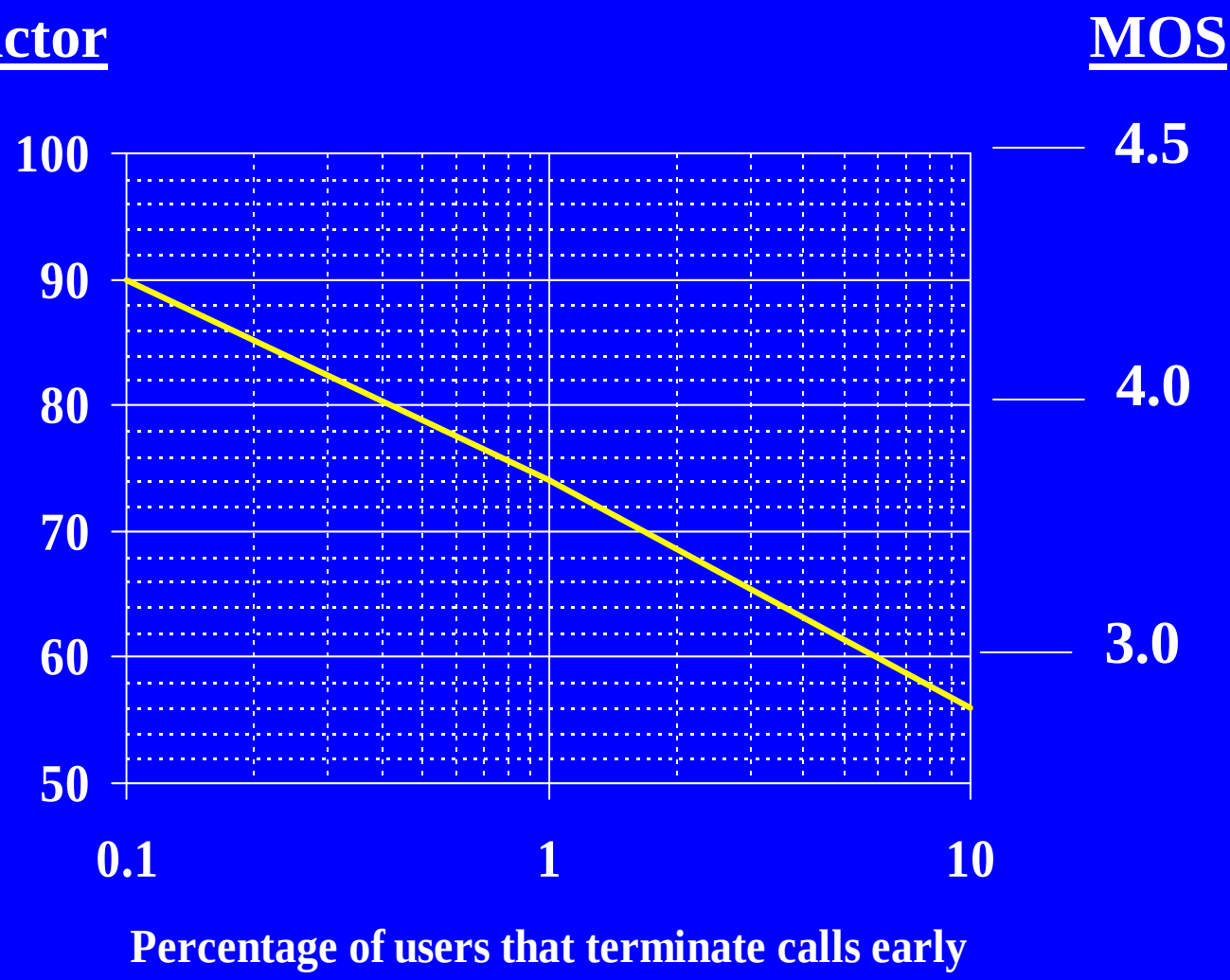
Embedded Monitoring



The E Model

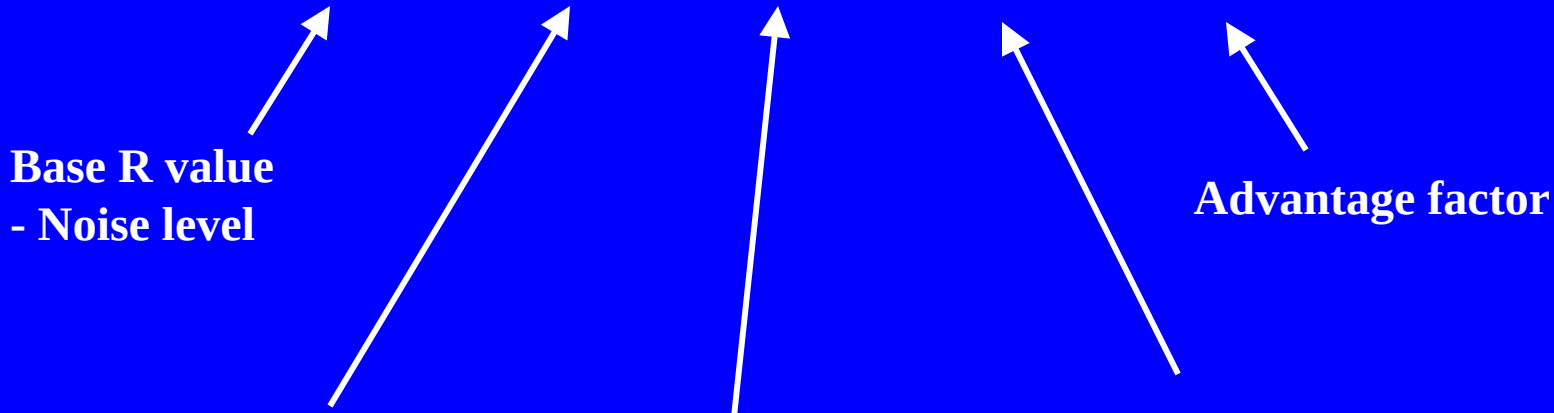
- “Mouth to ear” transmission quality measurement
- Produces an “R” factor typically in the range 50 (bad) -95 (good)
- R factor can be related to MOS score, Terminate Early (TME) etc.
- ITU G.107/ G.108 and ETSI ETR250

R Factor vs MOS



E Model

R = Ro - Is - Id - Ie + A



Base R value
- Noise level

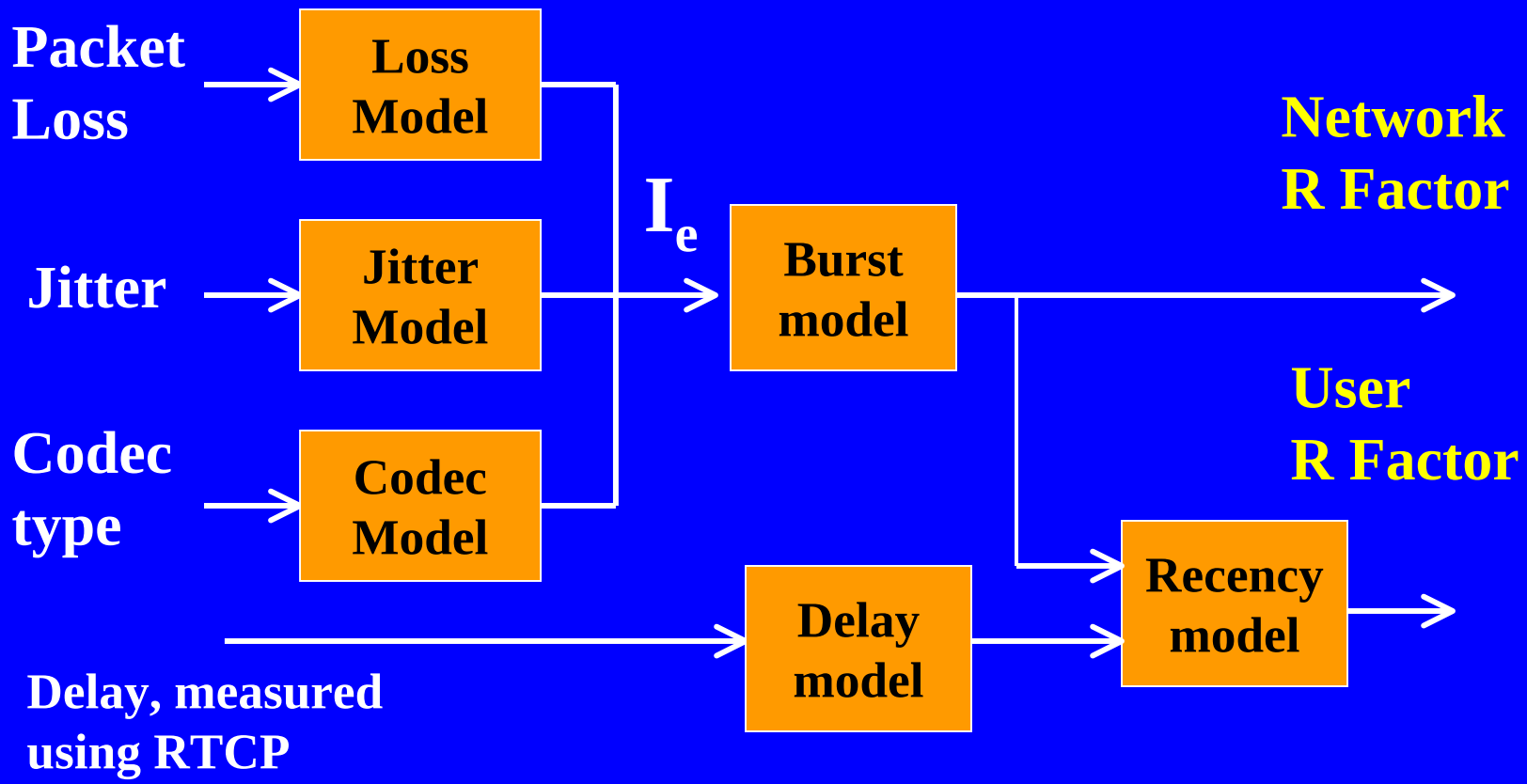
Advantage factor

Impairments that occur simultaneously with speech
- received speech level
- sidetone level
- quantization noise

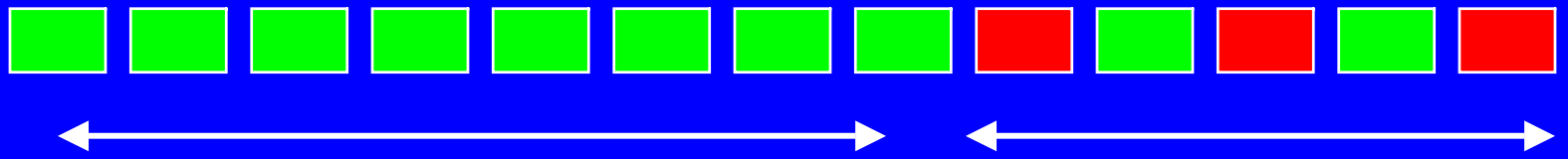
Impairments that are delayed with respect to speech
- talker echo
- listener echo
- round trip delay

Equipment Impairment Factor
- CODEC
- multiplexing effects

Extended E Model

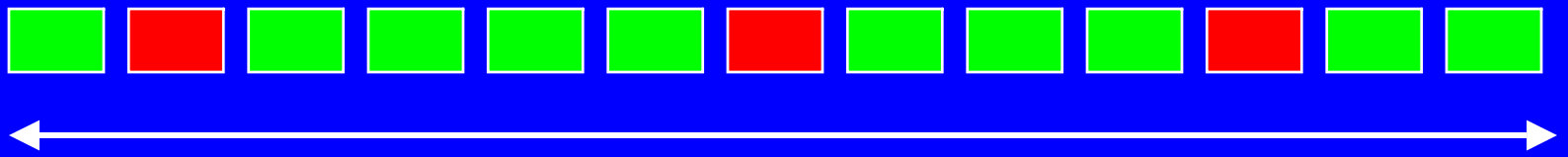


Burst vs average loss



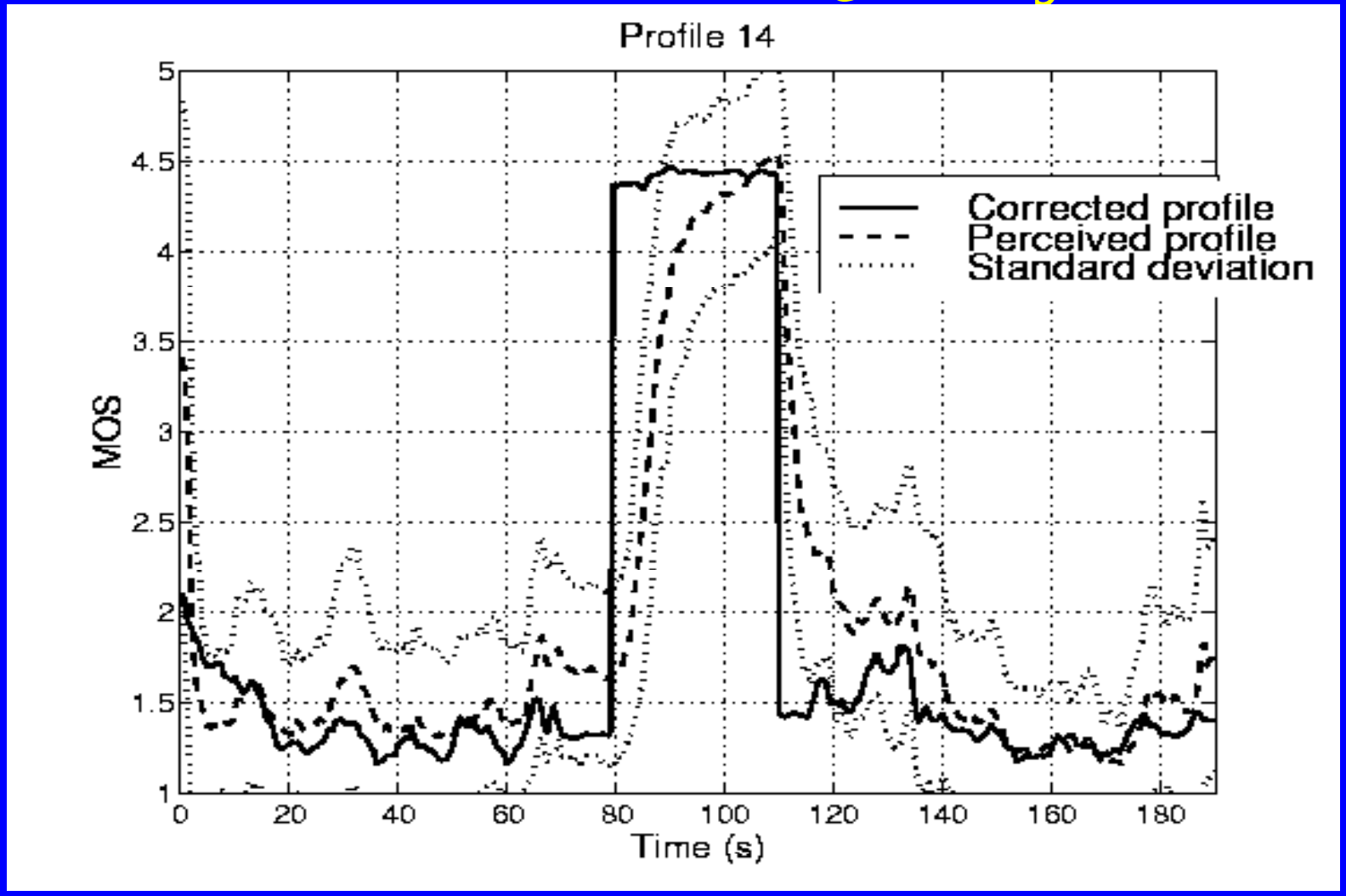
Zero packet loss

Burst of packet loss



Non-bursty packet loss

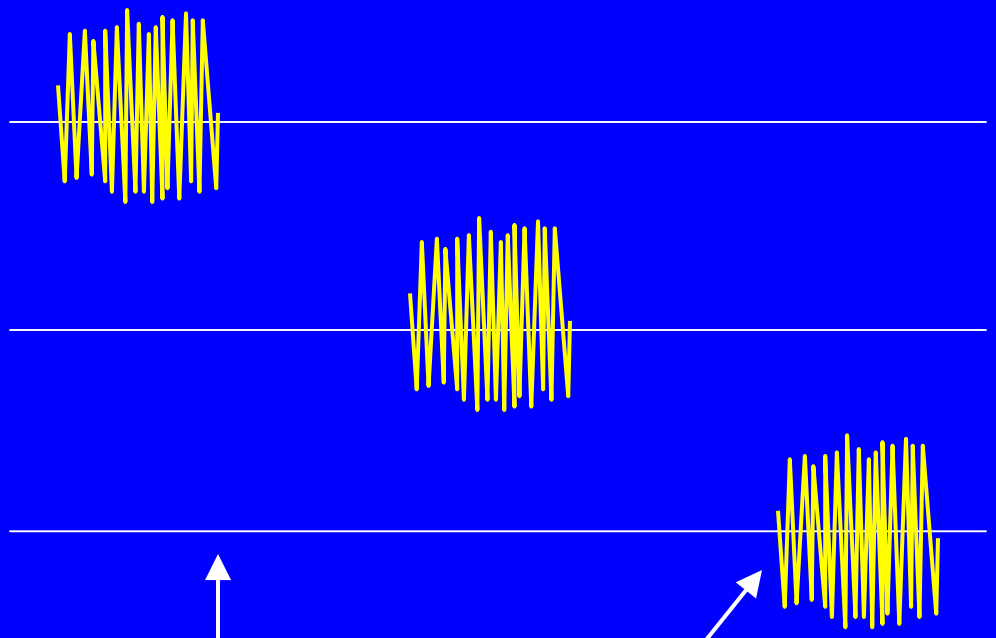
Instantaneous Quality



Source France Telecom

“Recency” effect

60 second call



MOS **3.82**

MOS **3.28**

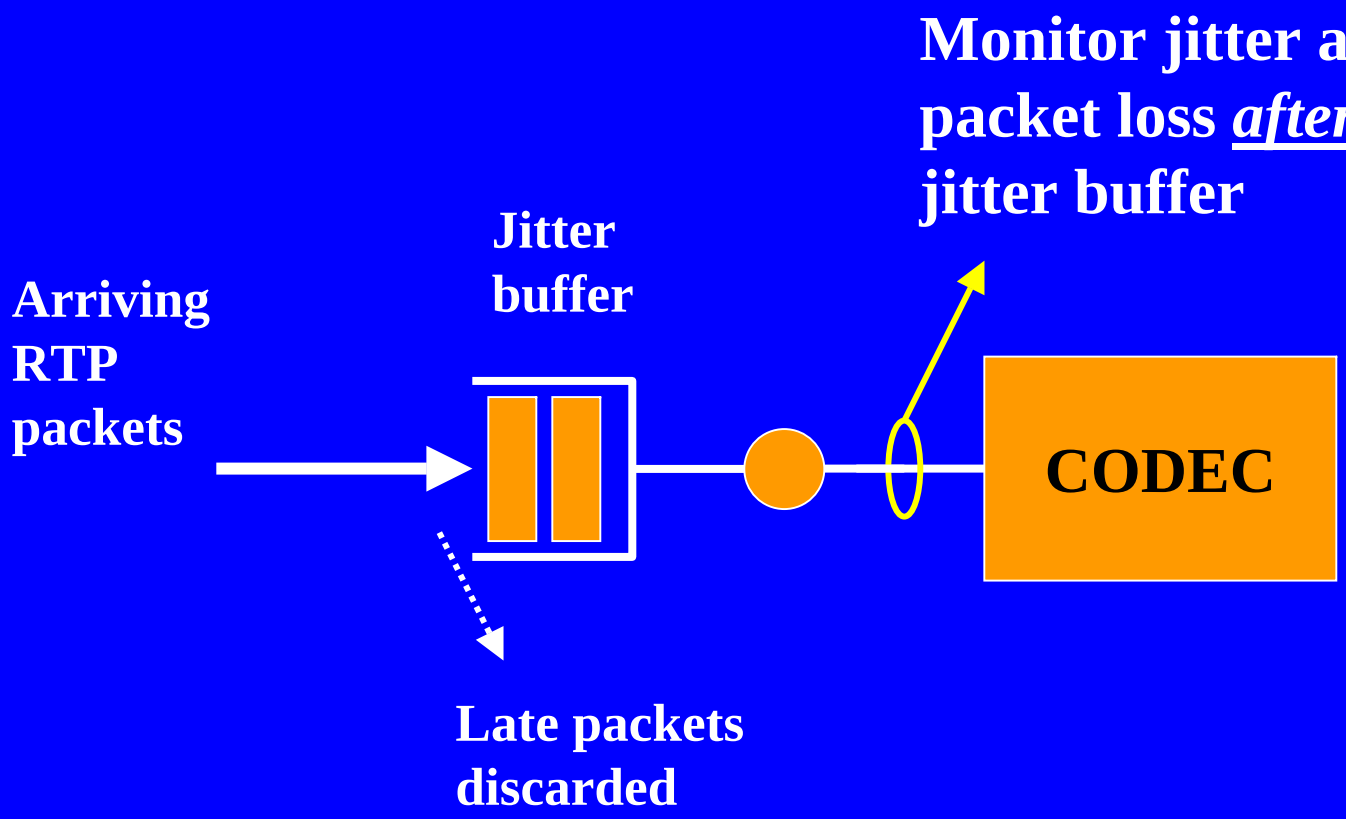
MOS **3.18**

“Good” 4.3MOS

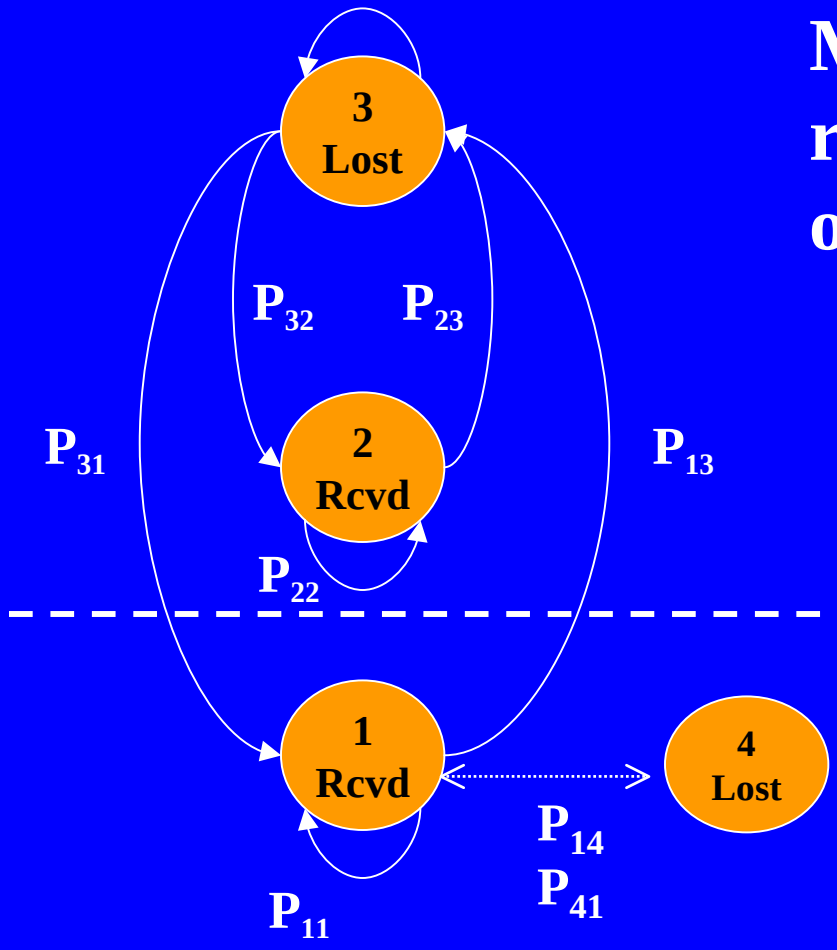
“Bad” 1.8 MOS
(3dB SNR)

Source AT&T
T1A1.7/98-031

Jitter and Packet Loss



Loss Model - Markov model

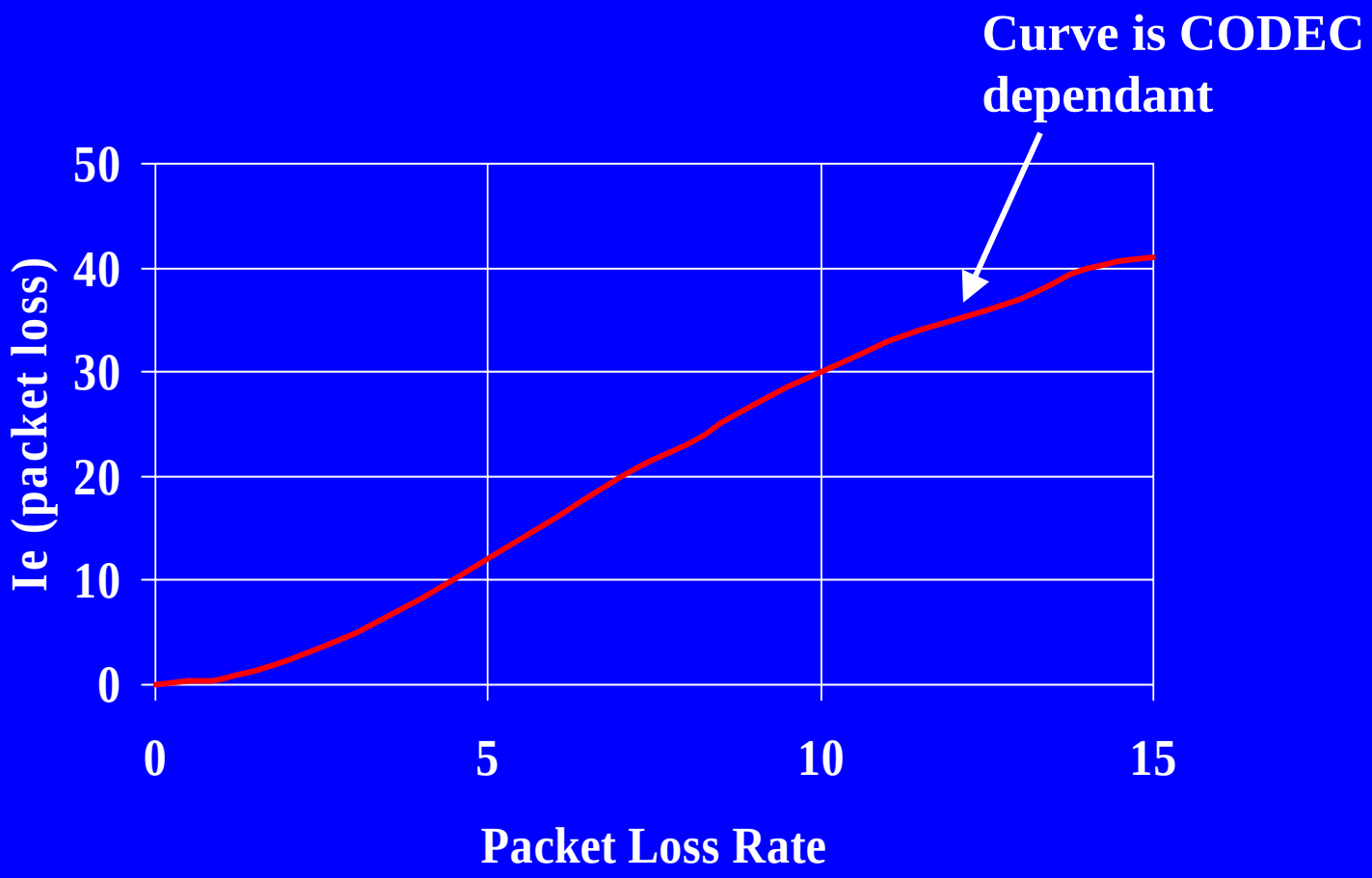


Model parameters reconstructed at end of call

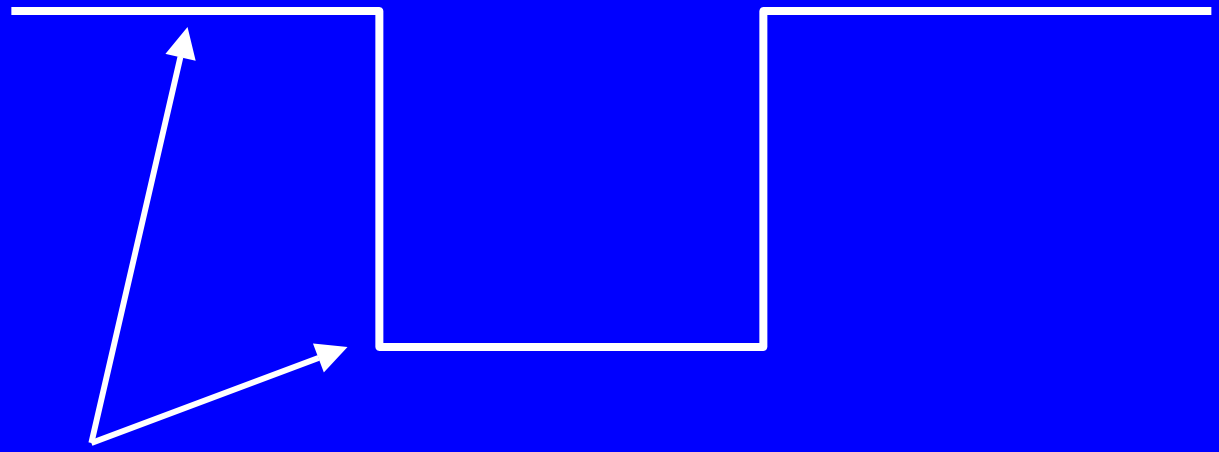
Burst state

Gap state

Loss Model - mapping loss to I_e

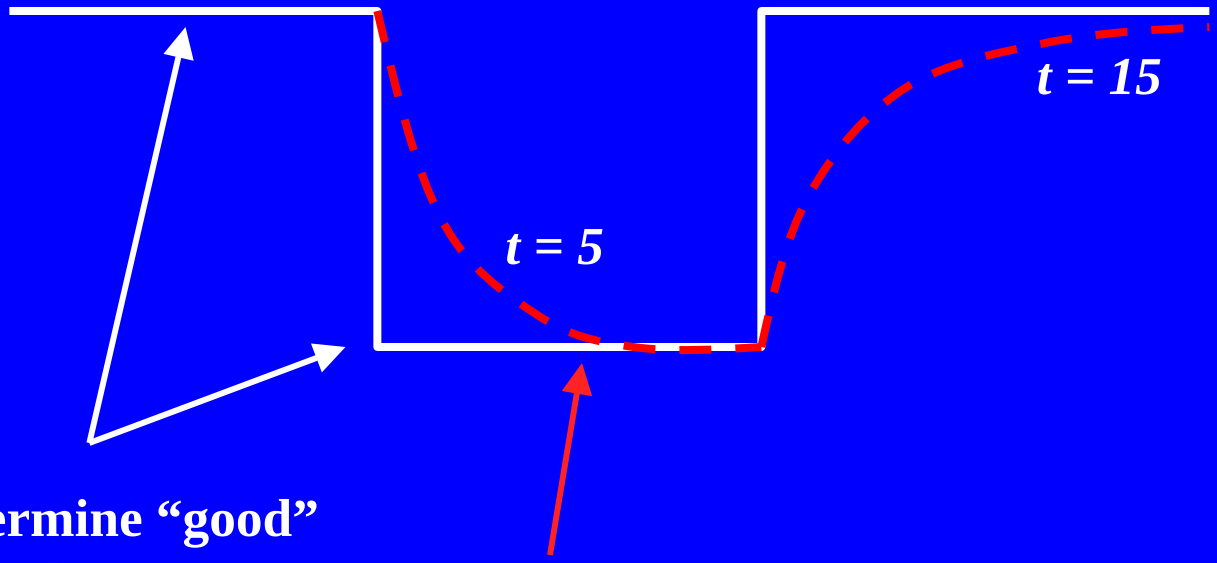


Determining QoS metrics



1. Determine “good” and “bad” state
Ie Factor

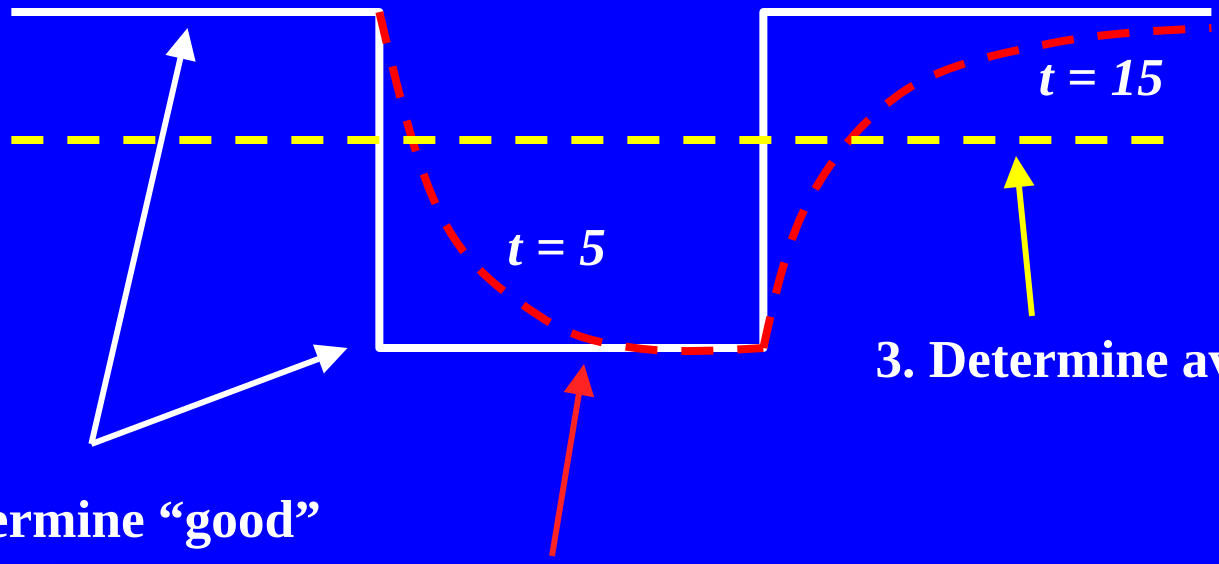
Determining QoS metrics



1. Determine “good” and “bad” state Ie Factor

2. Estimate Instantaneous R Factor for each state

Determining QoS metrics

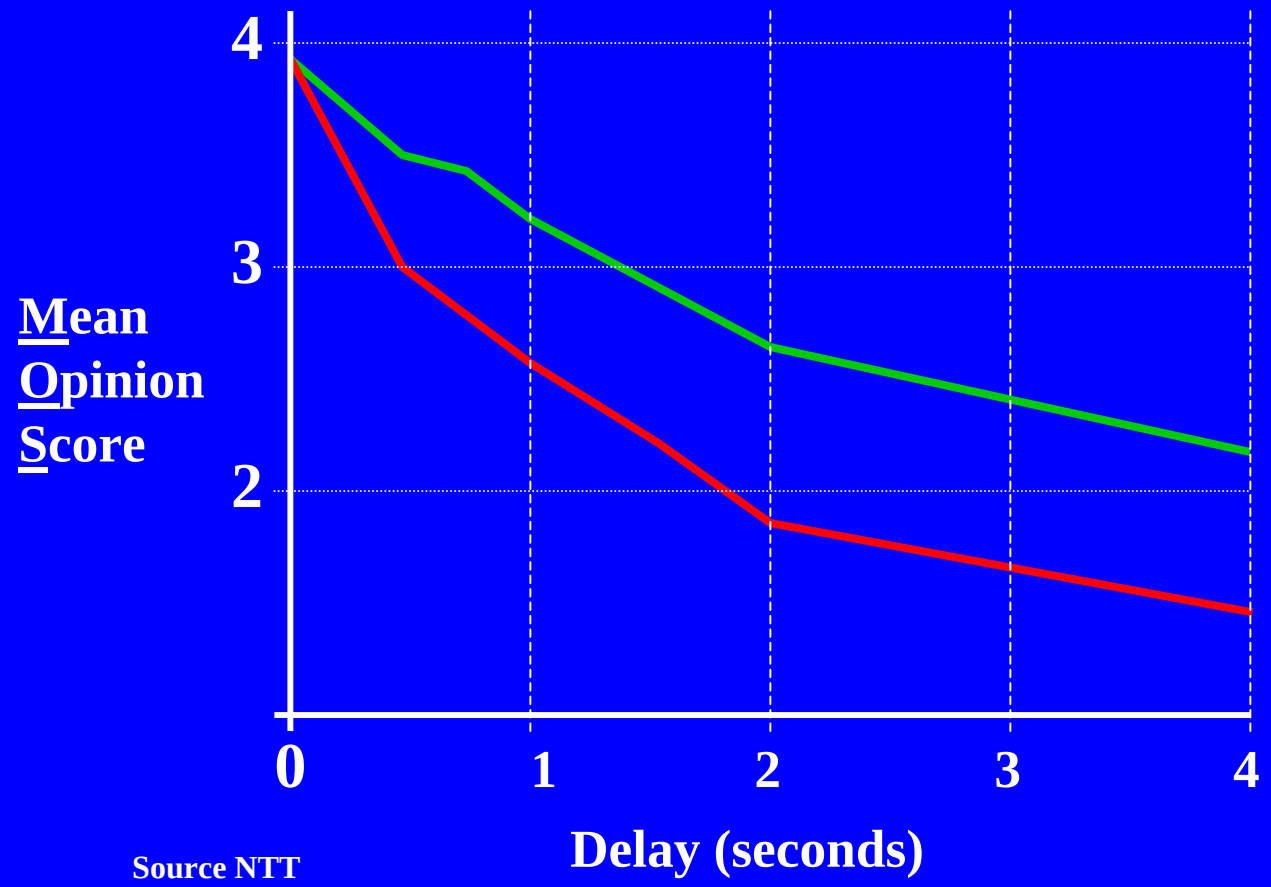


1. Determine “good” and “bad” state Ie Factor

2. Estimate Instantaneous R Factor for each state

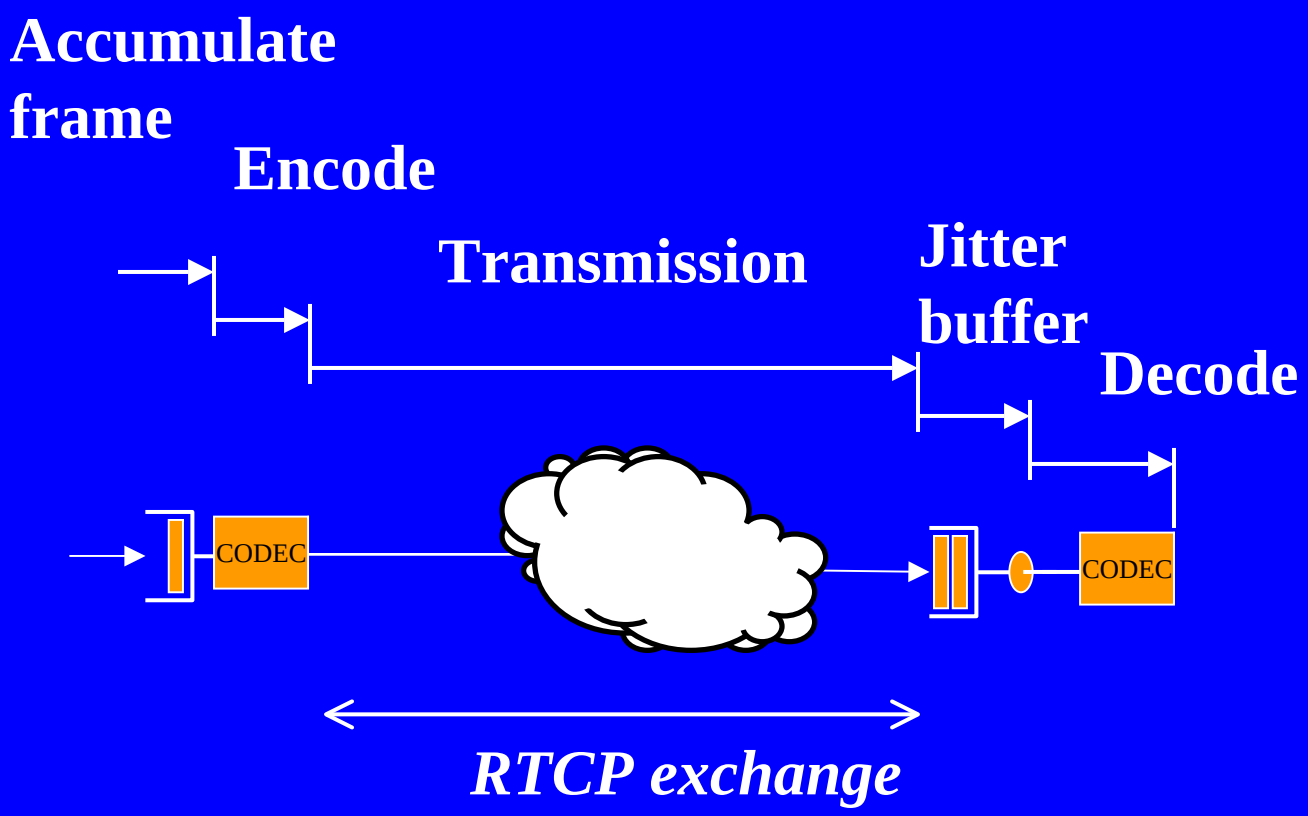
3. Determine average Ie

Delay effects

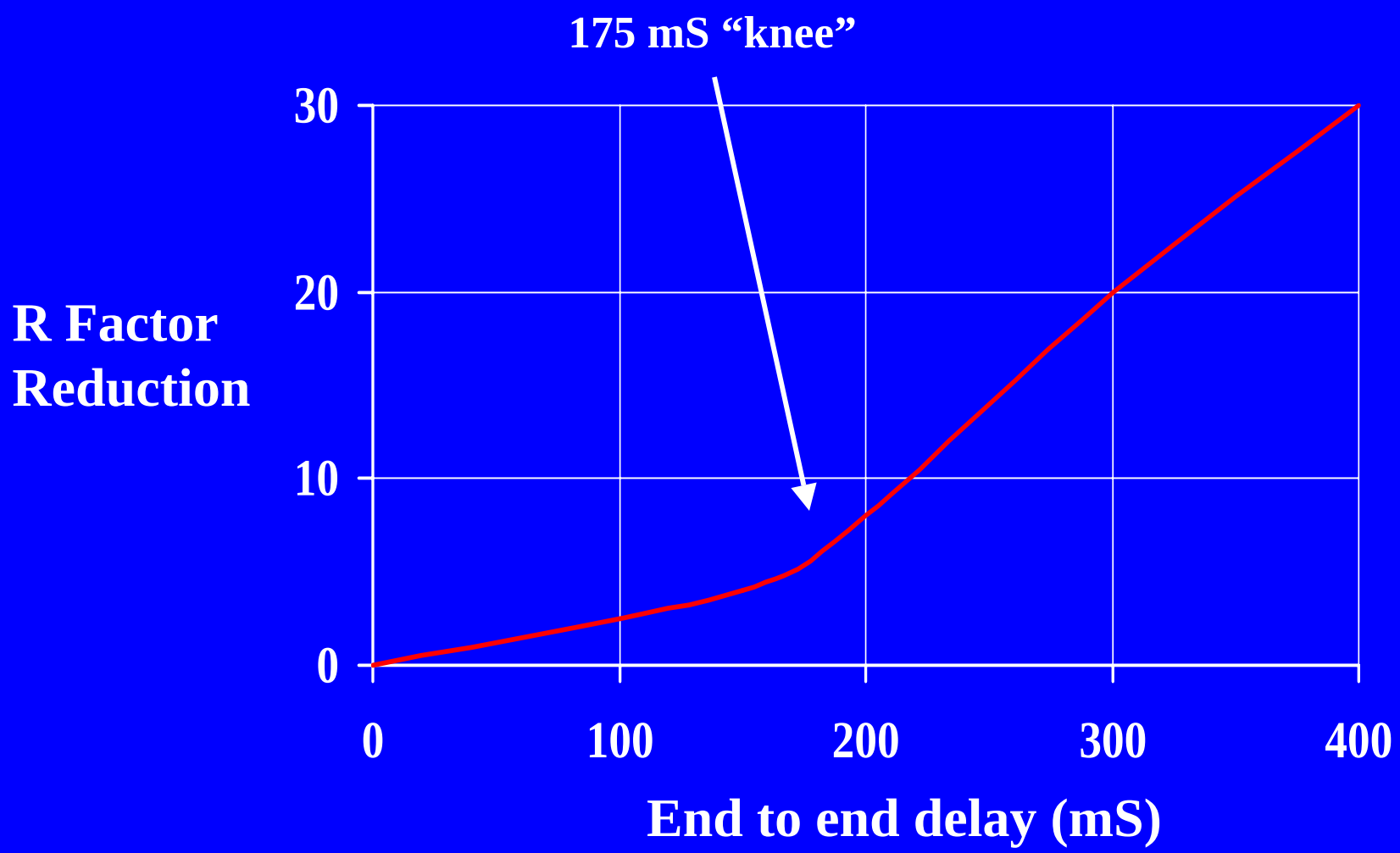


Source NTT

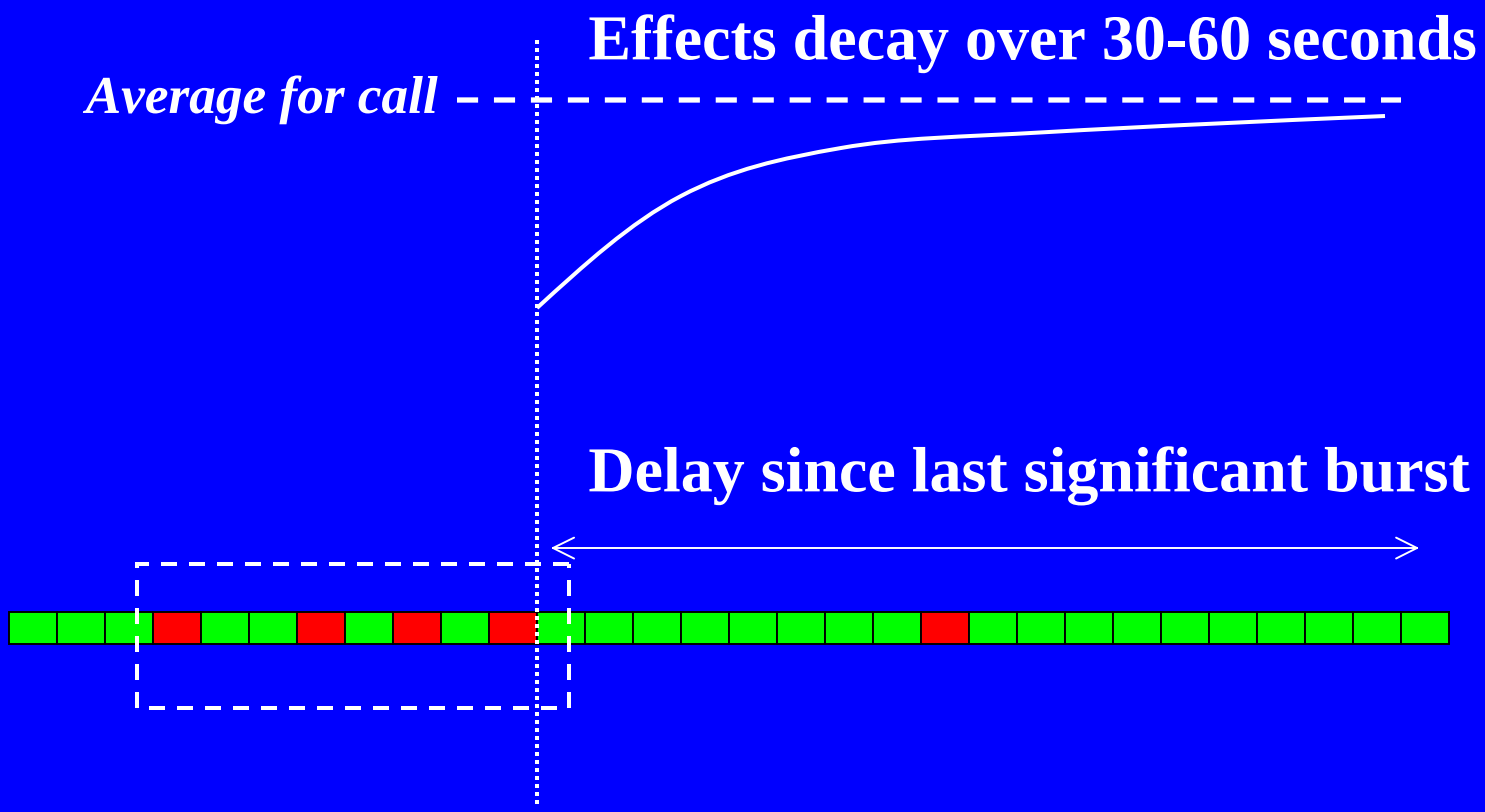
Measuring Delay



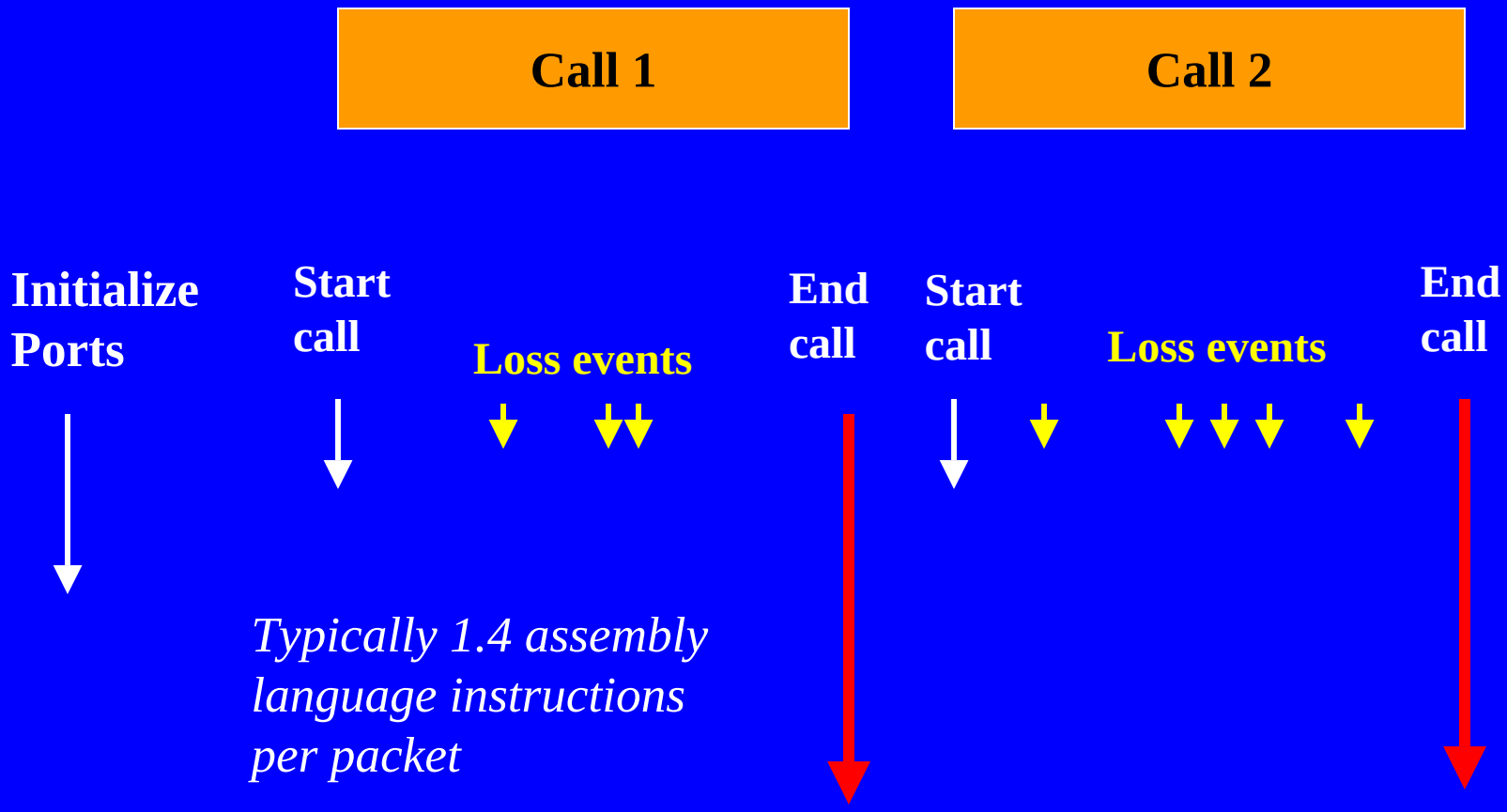
Delay Model


175 mS “knee”
R Factor Reduction
30
20
10
0
0
100
200
300
400
End to end delay (mS)

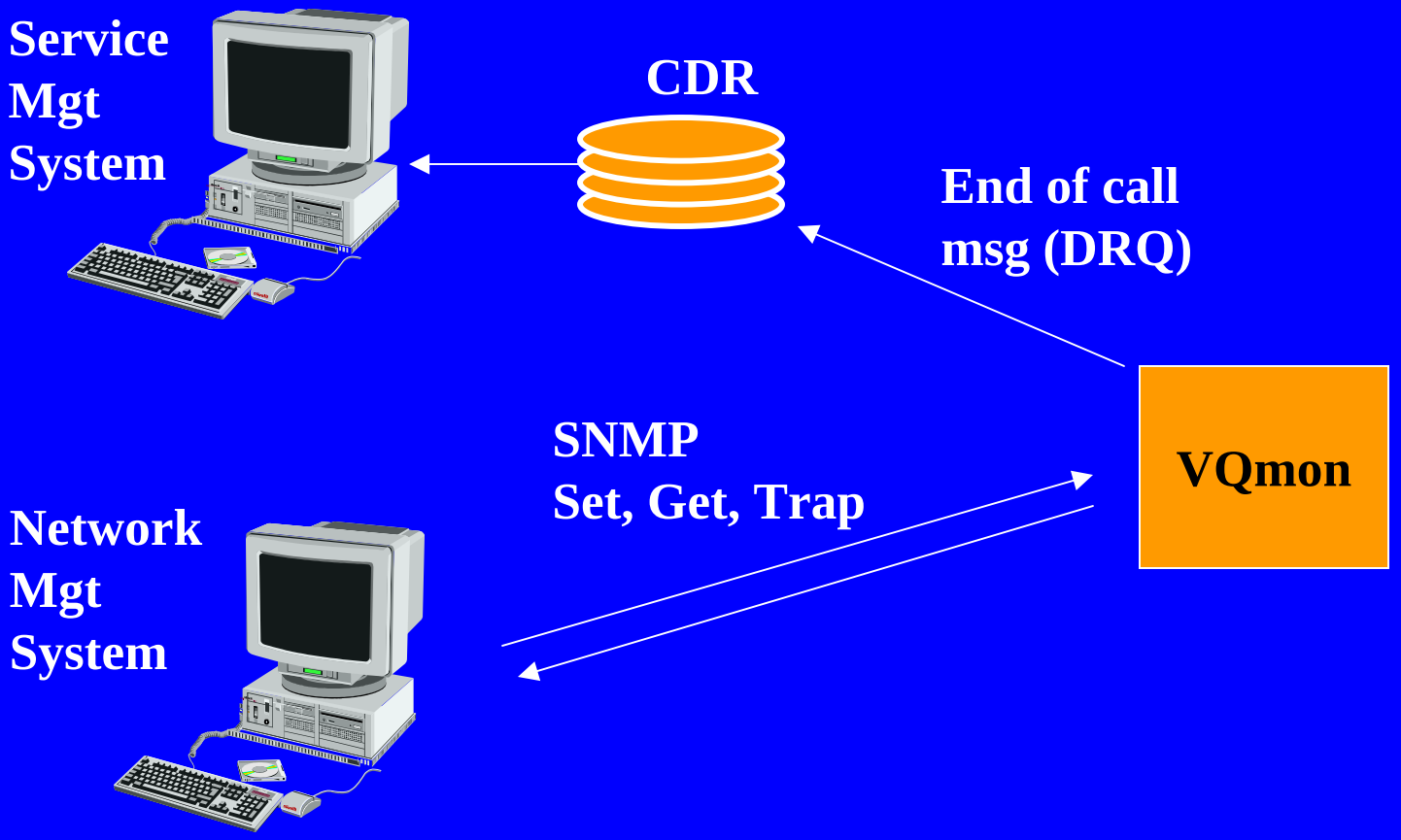
Estimation of recency effect



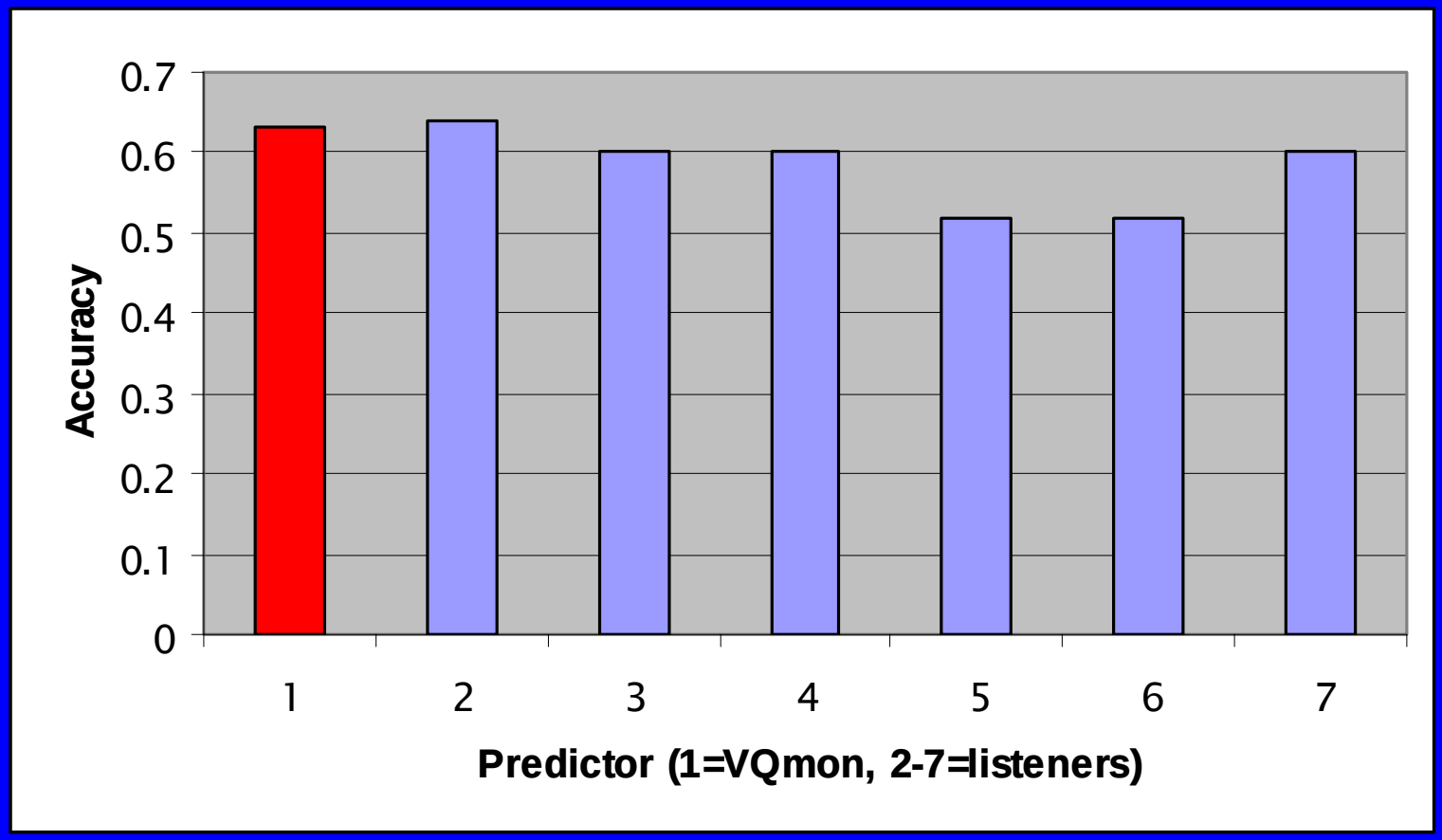
Execution model



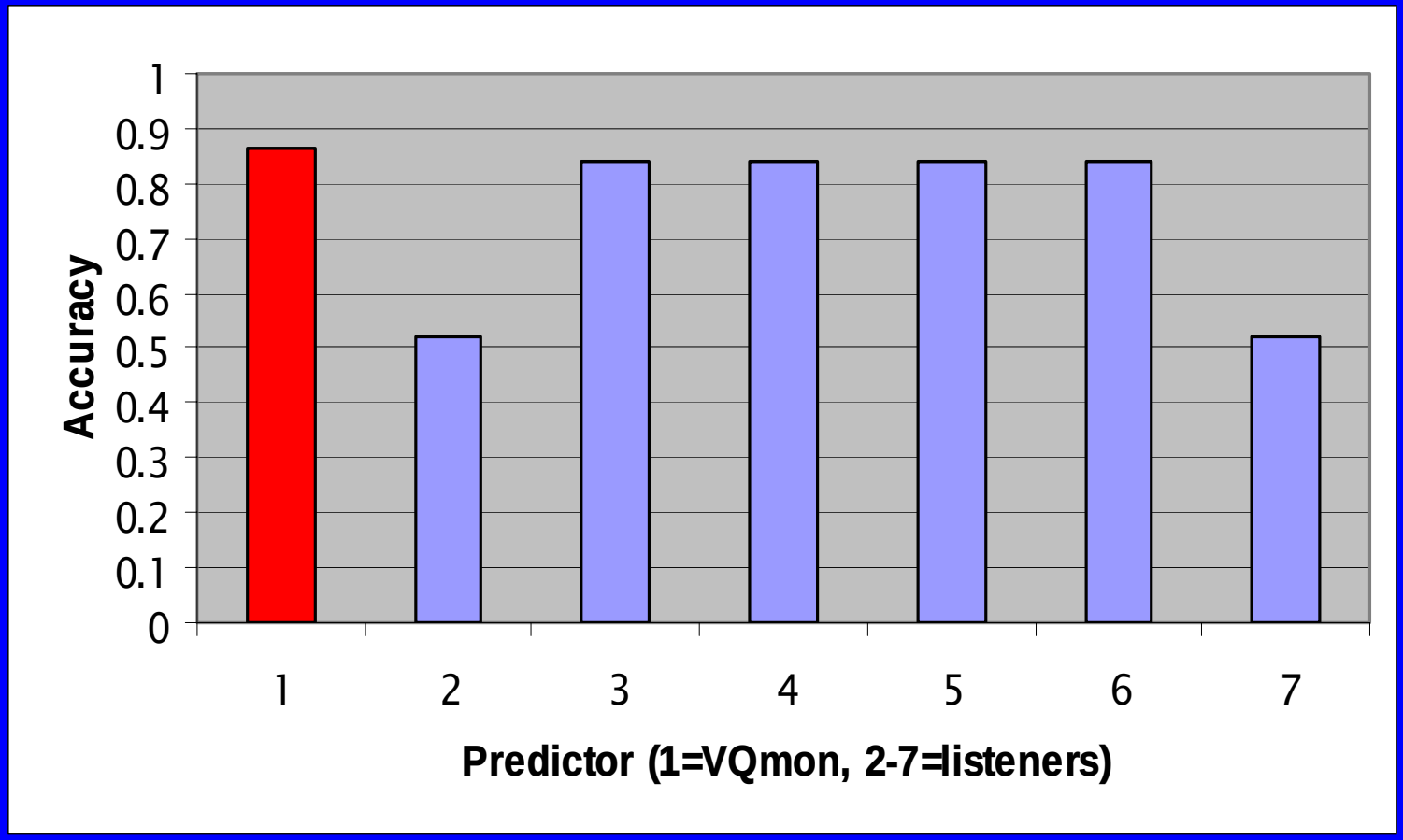
Integration with VoIP SMS



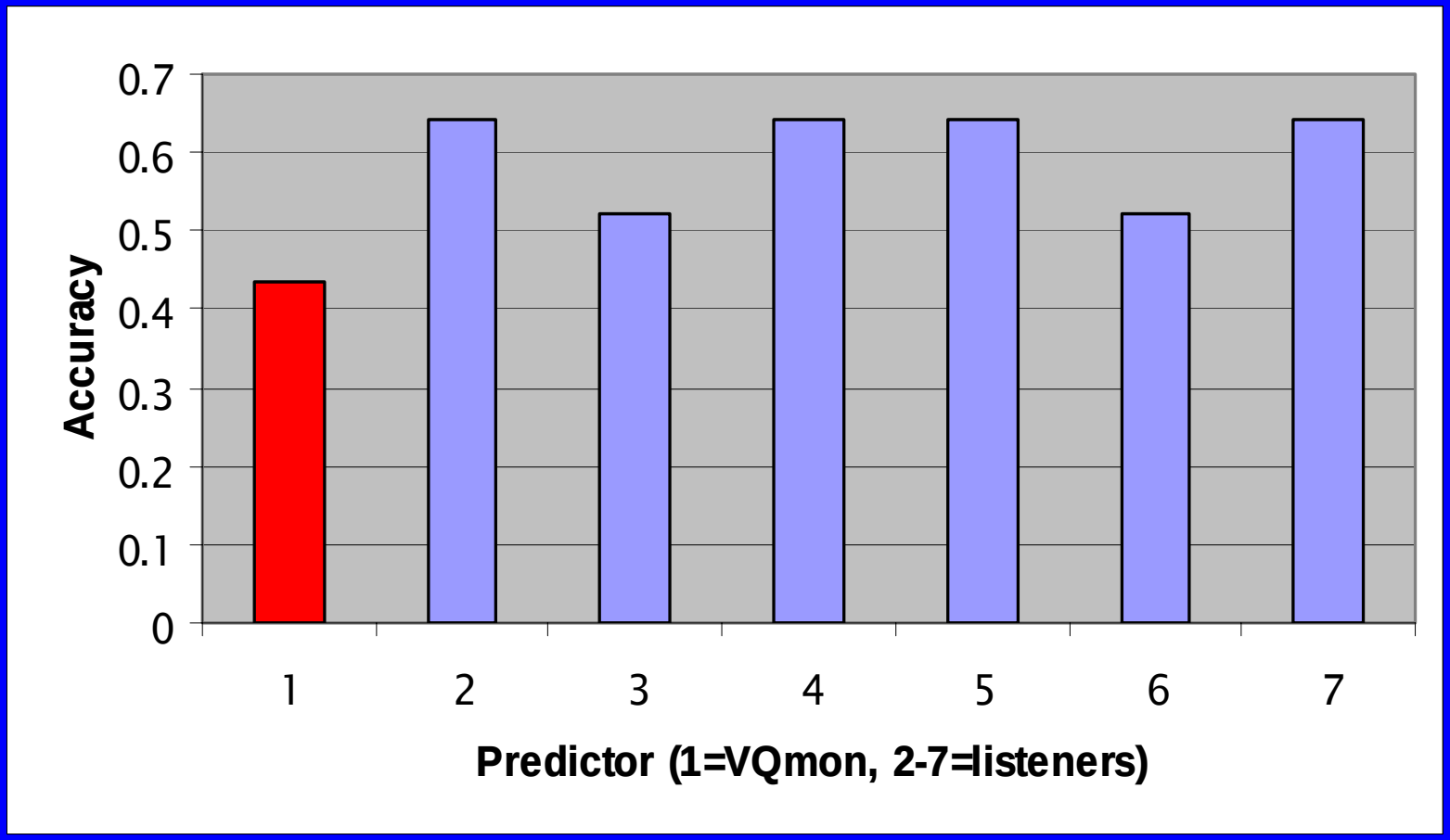
Ranking accuracy – Set 1



Ranking Accuracy – Set 2



Ranking Accuracy – Set 4



Conclusions

- Computational model meets design goals
- Ranking accuracy is comparable to human listener
- But need -
  - Systematic comparison with PSQM/ PESQ
  - Increased level of subjective testing
  - Add support for VAD, non-PLC
  - Improved accuracy requires some information on voice frame content

Further work areas

- Use CODEC generated frame loss event – indicate presence of speech energy
- Additional subjective testing
  - Wider variety of audio sources
  - Force listeners to focus on call content
  - Design impairments to isolate recency effect, burst characteristics, masking effects