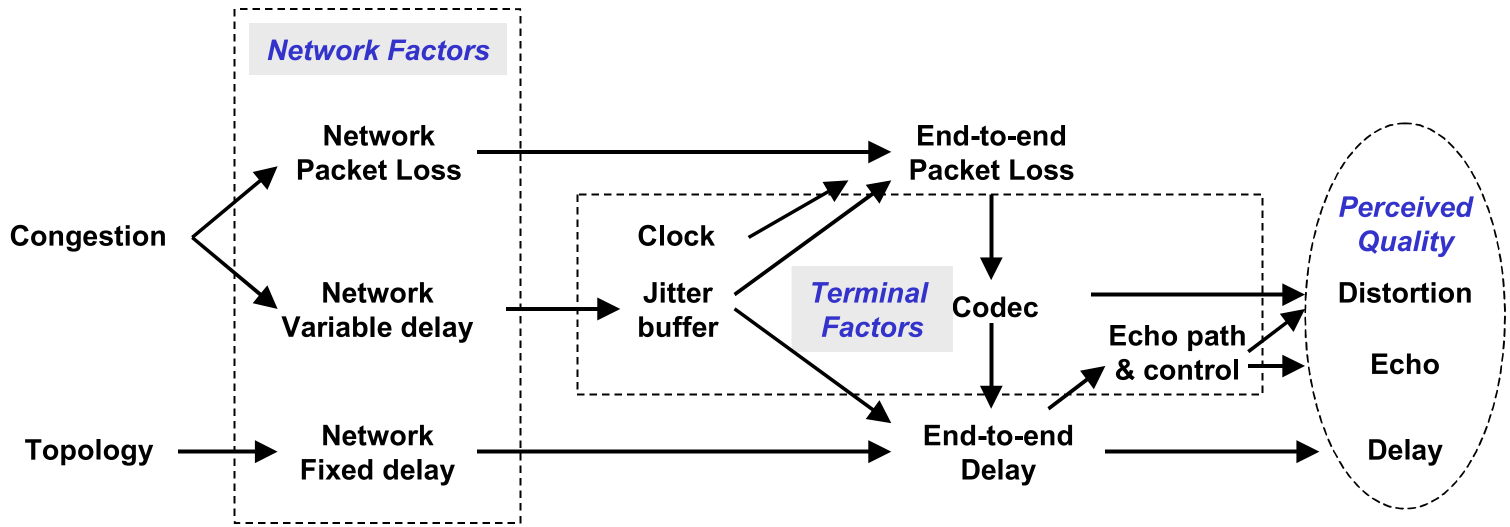
Assessing the quality of VoIP transmission affected by playout buffer scheme and encoding scheme

Miroslaw Narbutt, Mark Davis

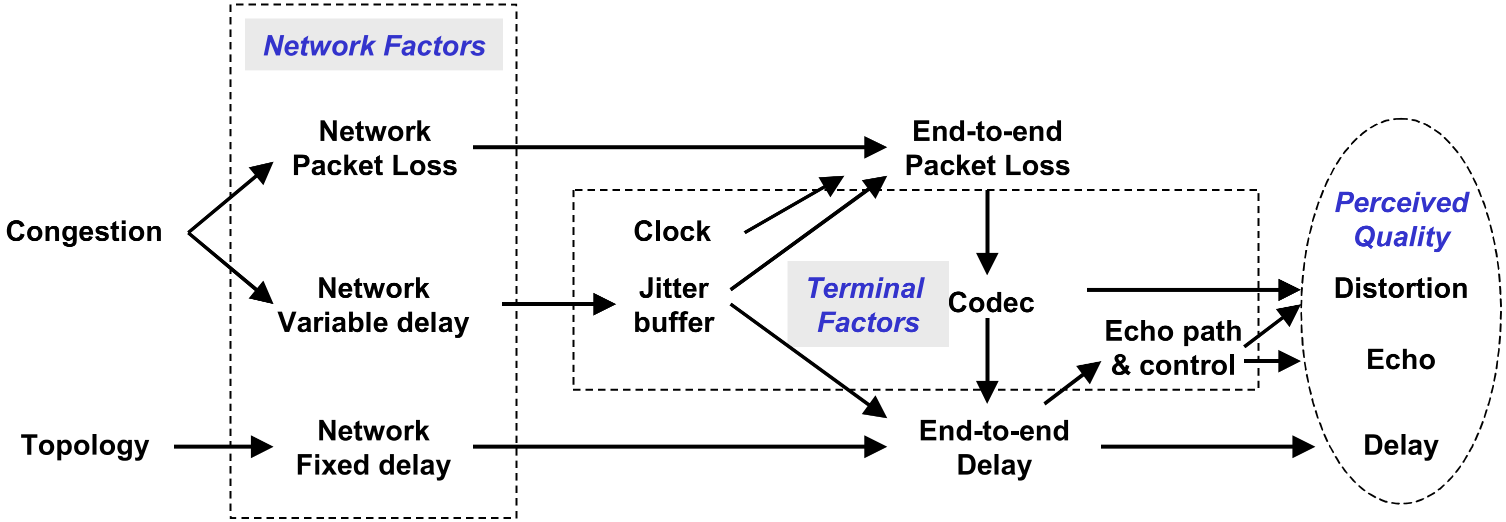
Communications Network Research Institute
Dublin Institute of Technology

Wireless Cluster Workshop, Maynooth
June 21/22nd 2005

Application mechanisms used to improve QoS

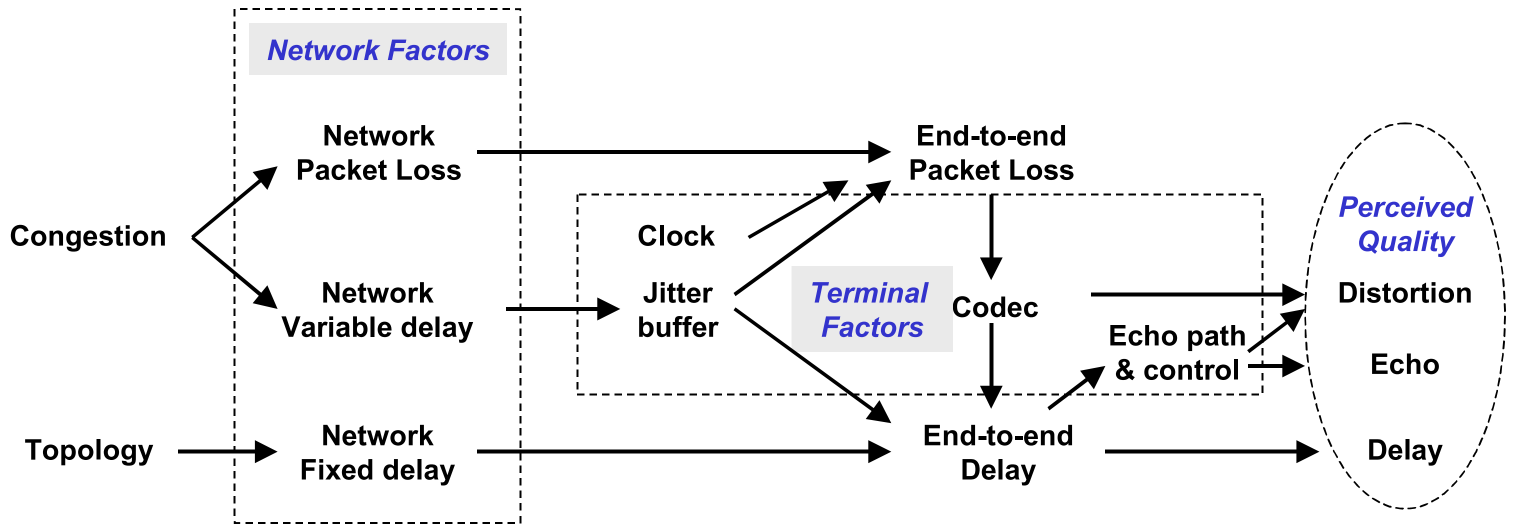


Application mechanisms used to improve QoS



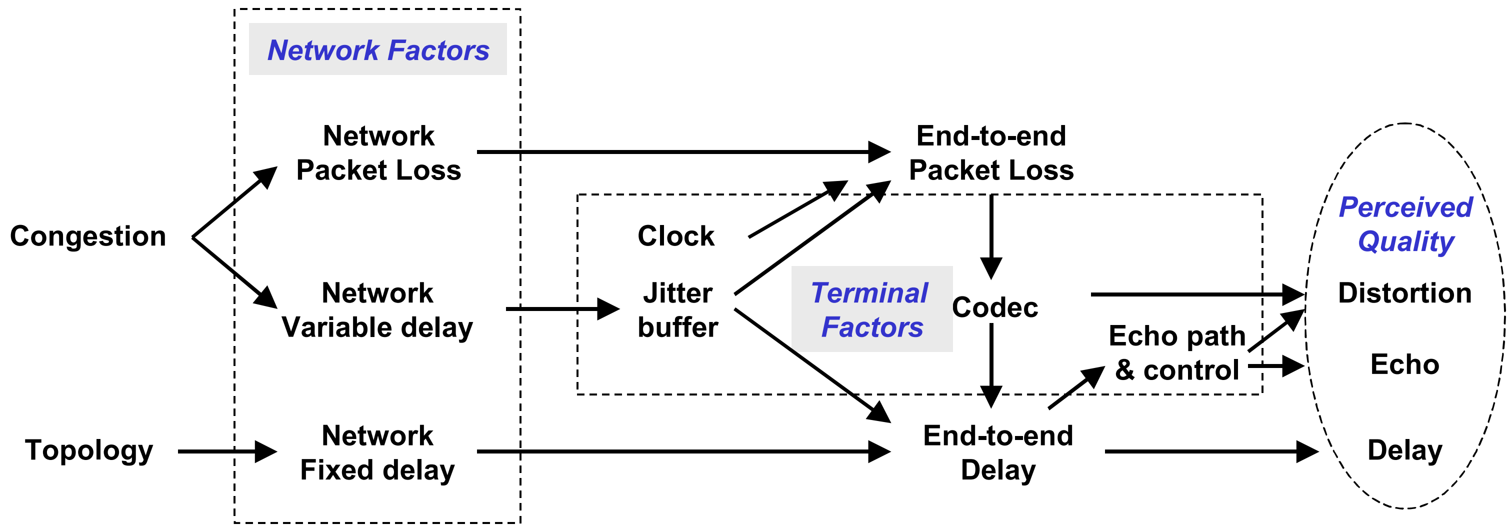
- **transmission impairments (delay, loss, jitter) affect conversational speech**

Application mechanisms used to improve QoS



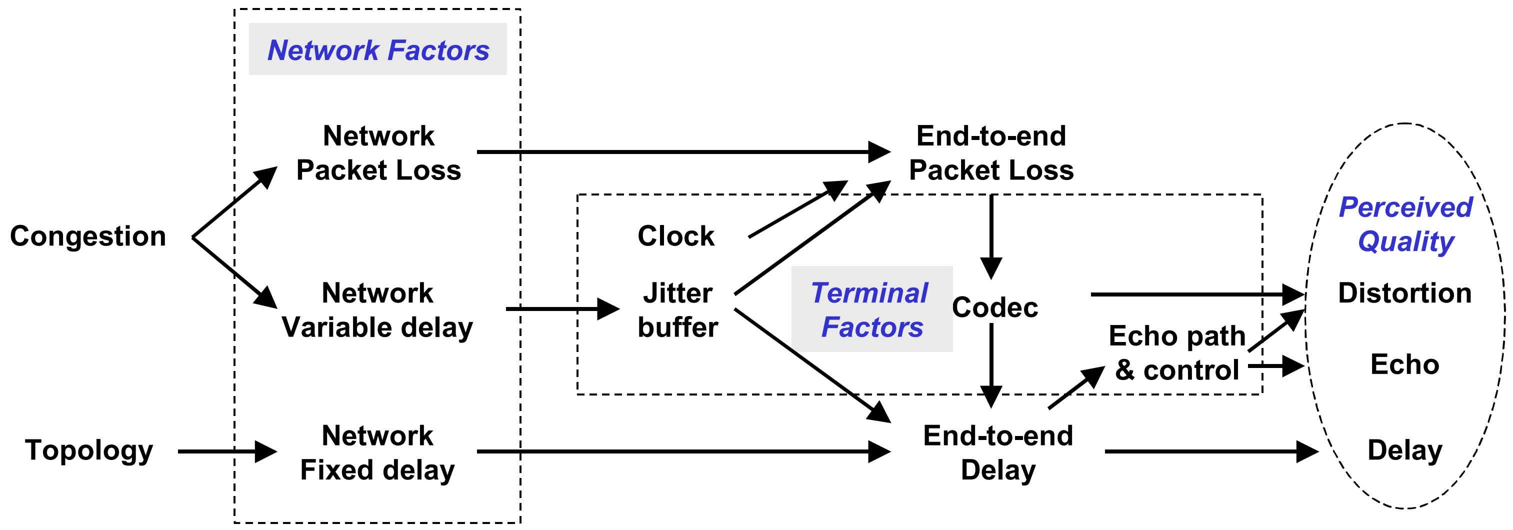
- transmission impairments (delay, loss, jitter) affect conversational speech
- actions at the end-systems can greatly mitigate their effect:

Application mechanisms used to improve QoS



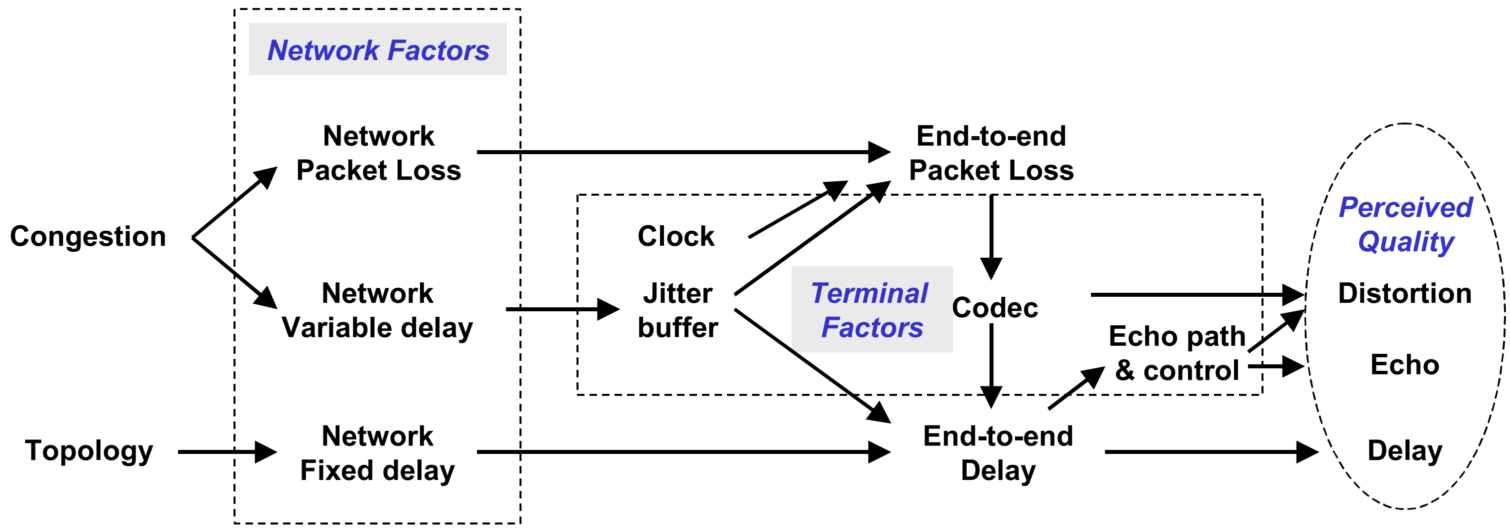
- **transmission impairments (delay, loss, jitter) affect conversational speech**
- **actions at the end-systems can greatly mitigate their effect:**
  - ***low bit-rate codecs***

Application mechanisms used to improve QoS



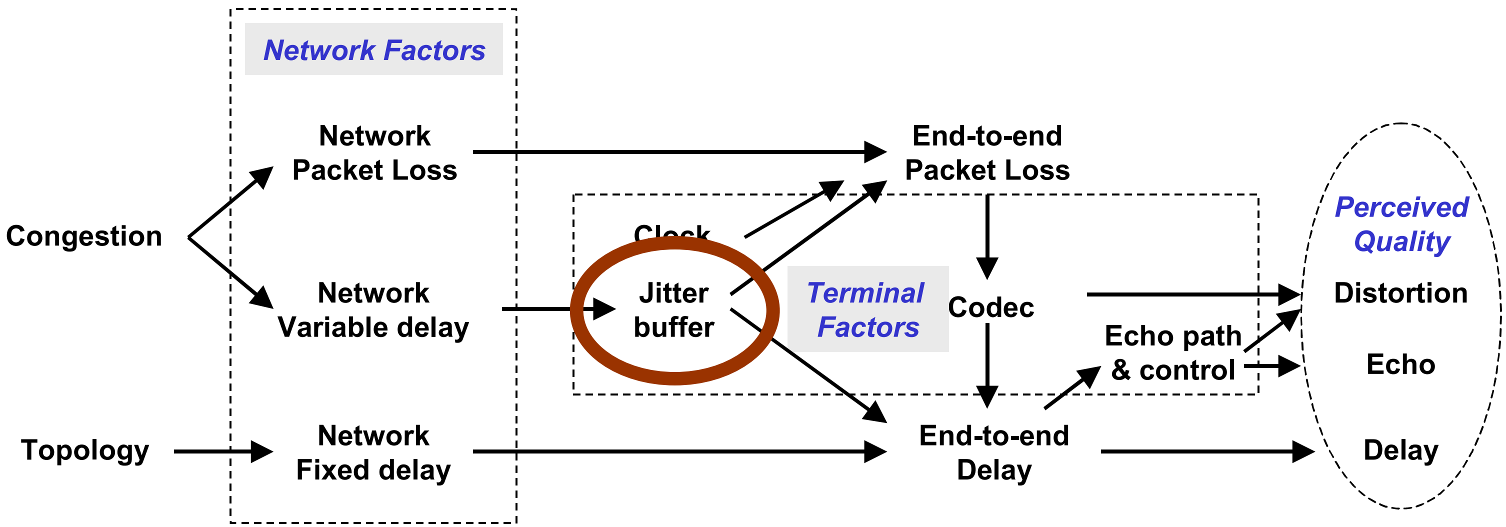
- **transmission impairments (delay, loss, jitter) affect conversational speech**
- **actions at the end-systems can greatly mitigate their effect:**
  - ***low bit-rate codecs***
  - ***packet loss concealment - PLC***

Application mechanisms used to improve QoS



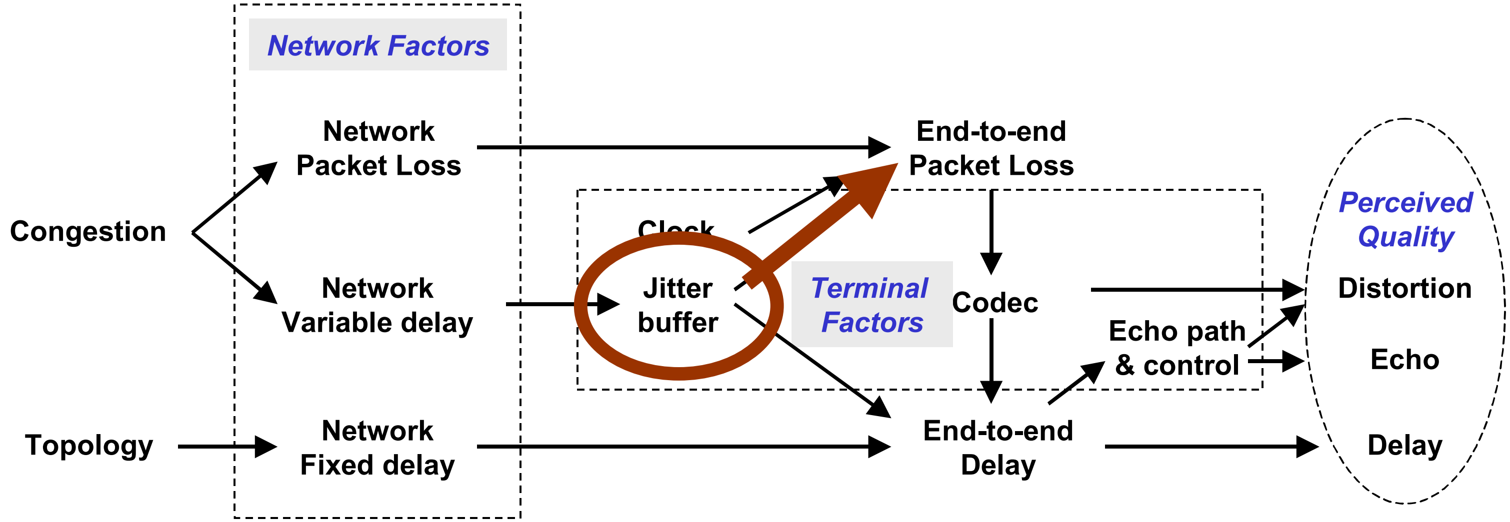
- **transmission impairments (delay, loss, jitter) affect conversational speech**
- **actions at the end-systems can greatly mitigate their effect:**
  - ***low bit-rate codecs***
  - ***packet loss concealment - PLC***
  - ***echo cancellation***

Application mechanisms used to improve QoS



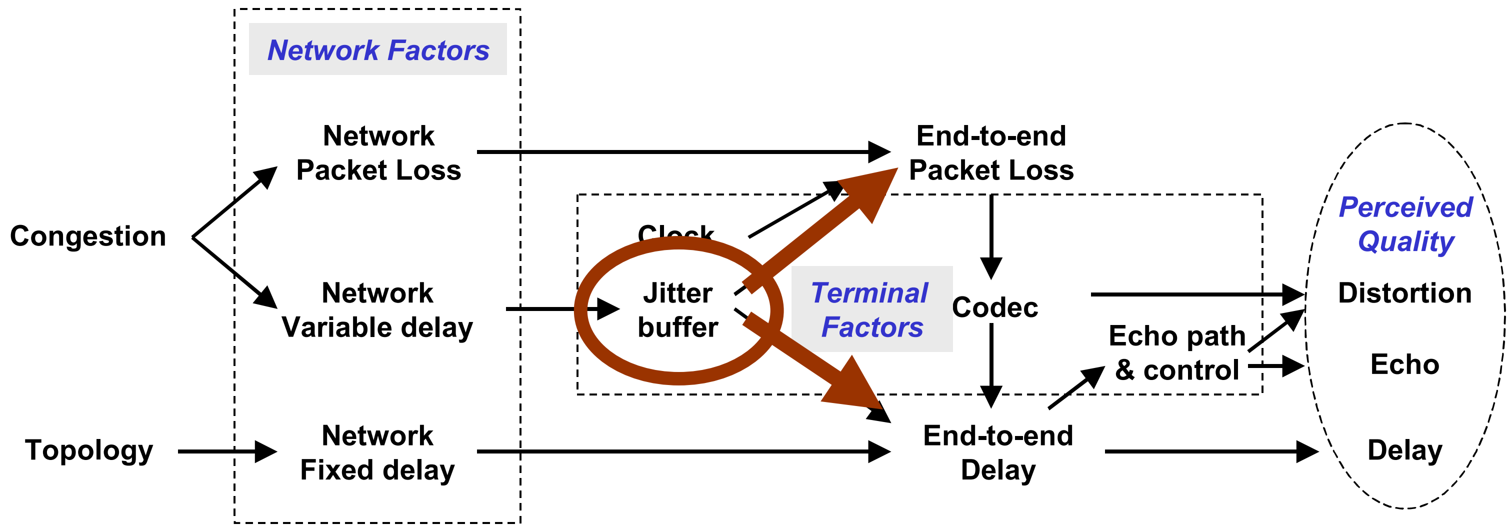
- **transmission impairments (delay, loss, jitter) affect conversational speech**
- **actions at the end-systems can greatly mitigate their effect:**
  - ***low bit-rate codecs***
  - ***packet loss concealment - PLC***
  - ***echo cancellation***
  - ***<u>de-jitter buffering</u>***

Application mechanisms used to improve QoS



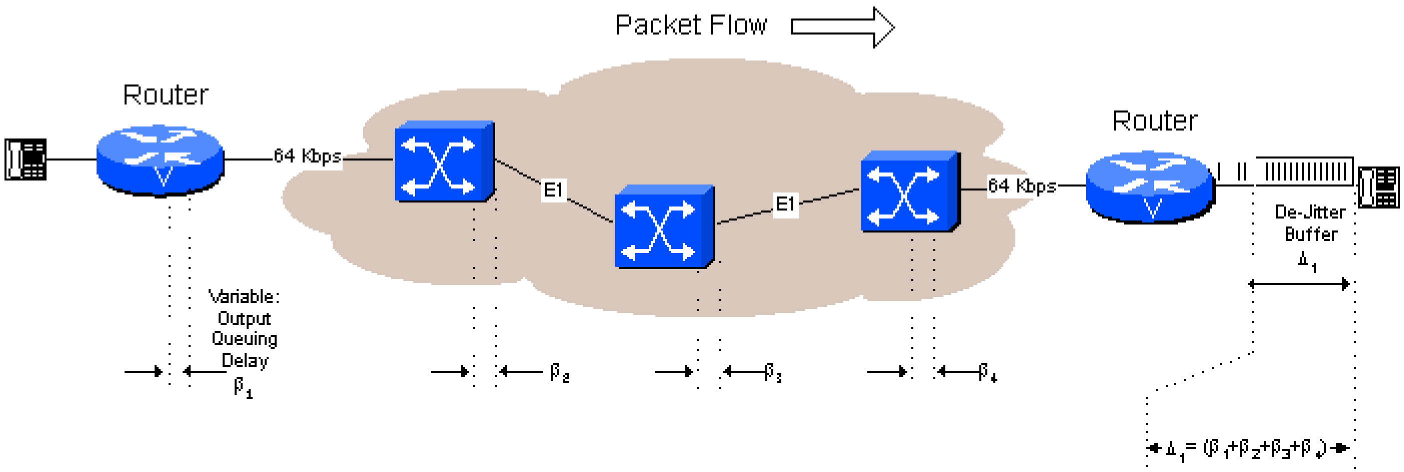
- **transmission impairments (delay, loss, jitter) affect conversational speech**
- **actions at the end-systems can greatly mitigate their effect:**
  - ***low bit-rate codecs***
  - ***packet loss concealment - PLC***
  - ***echo cancellation***
  - ***de-jitter buffering***

Application mechanisms used to improve QoS



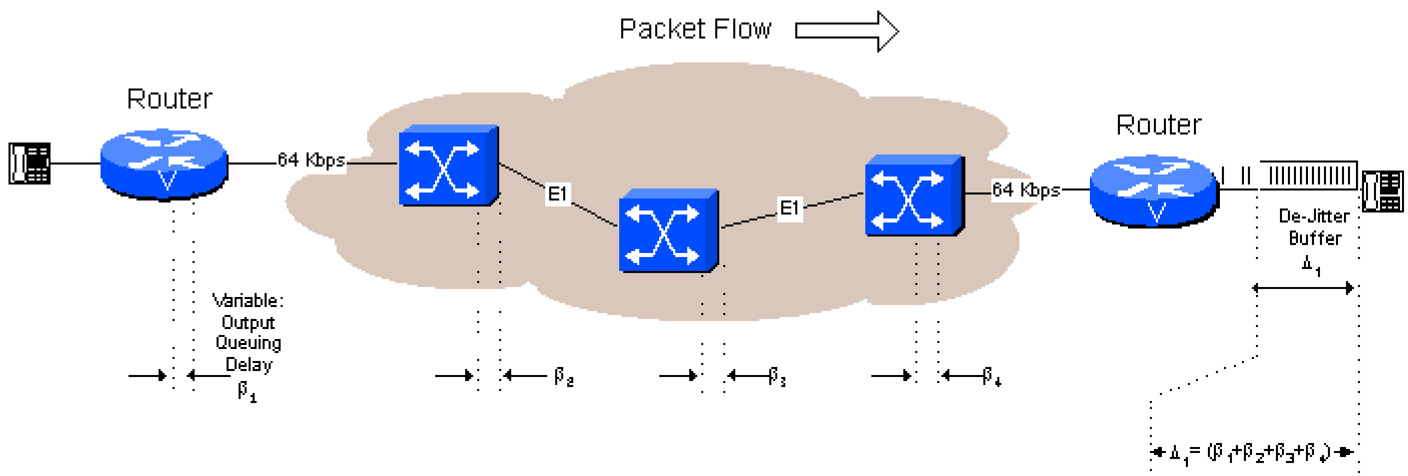
- **transmission impairments (delay, loss, jitter) affect conversational speech**
- **actions at the end-systems can greatly mitigate their effect:**
  - ***low bit-rate codecs***
  - ***packet loss concealment - PLC***
  - ***echo cancellation***
  - ***de-jitter buffering***

Variable delays and de-jitter buffer



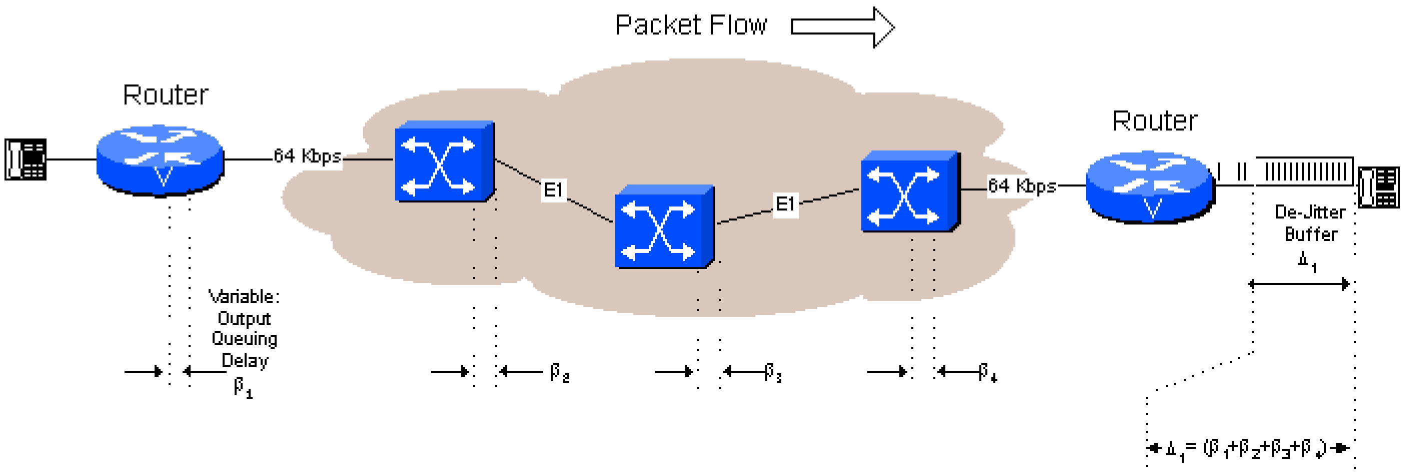
- Size of de-jitter buffer at VoIP receiver is critical!

Variable delays and de-jitter buffer



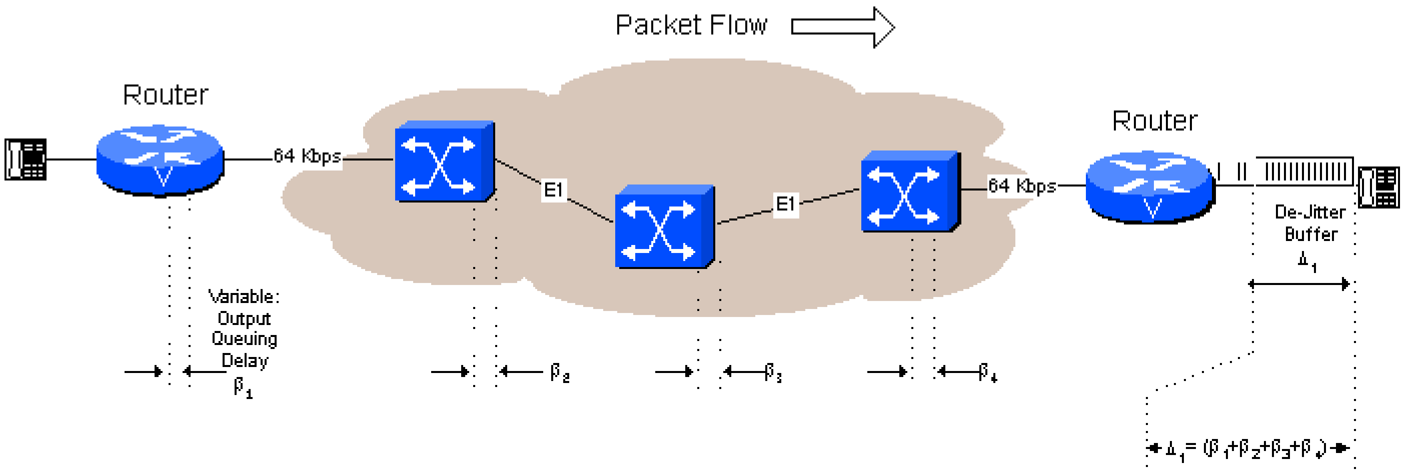
- Size of de-jitter buffer at VoIP receiver is critical!
  - **Too small** -> decreases e2e delay but increases late packet loss

Variable delays and de-jitter buffer



- Size of de-jitter buffer at VoIP receiver is critical!
  - **Too small** -> decreases e2e delay but increases late packet loss
  - **Too large** -> decreases late packet loss but increases e2e delay

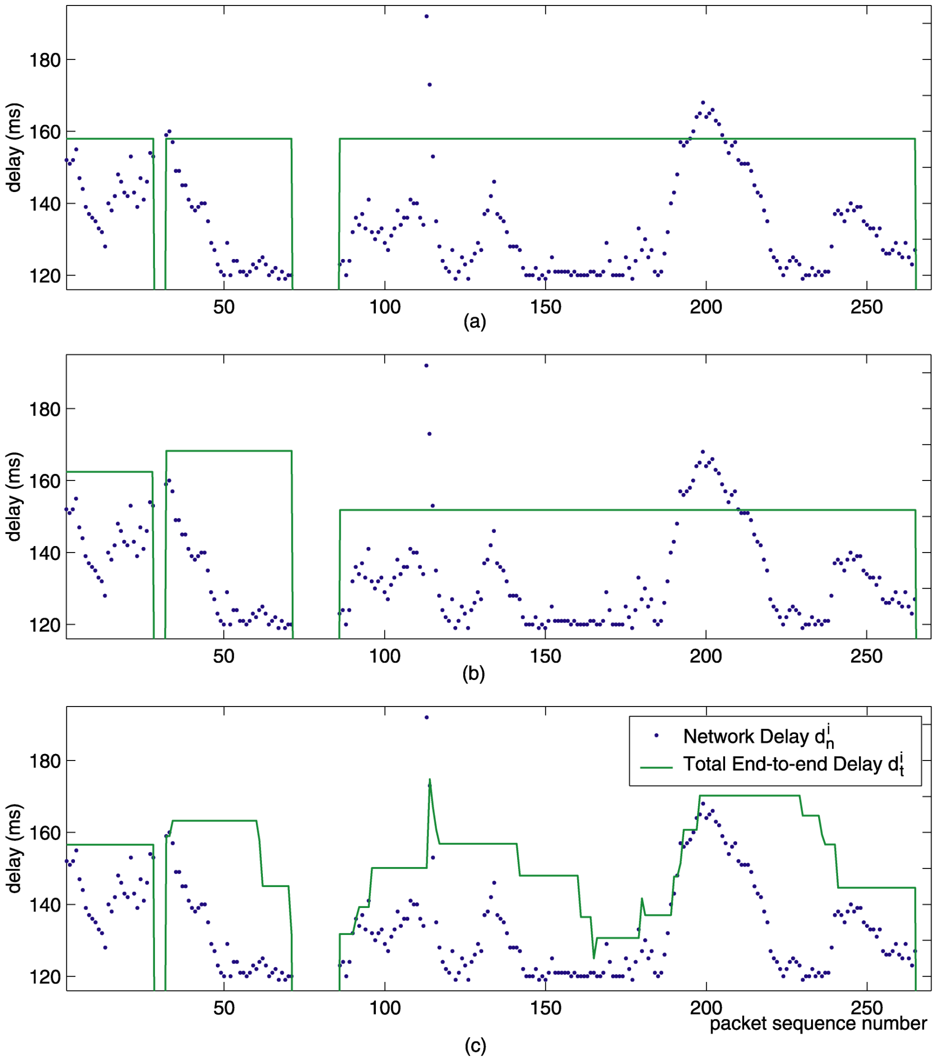
Variable delays and de-jitter buffer



- Size of de-jitter buffer at VoIP receiver is critical!
  - **Too small** -> decreases e2e delay but increases late packet loss
  - **Too large** -> decreases late packet loss but increases e2e delay

- Many adaptive playout algorithms!

Playout delay adjustment (when?)



1. fixed throughout the whole session;
2. adjusted during silence periods (per talkspurt) *[Ramjee '94, Moon '98, …]*;
3. adjusted within talkspurts (per packet) using packet scaling (TSM technique) *[Yi Liang '03]*.

Playout delay adjustment (how?)

- **Reactive algorithm *[Ramjee '94]***

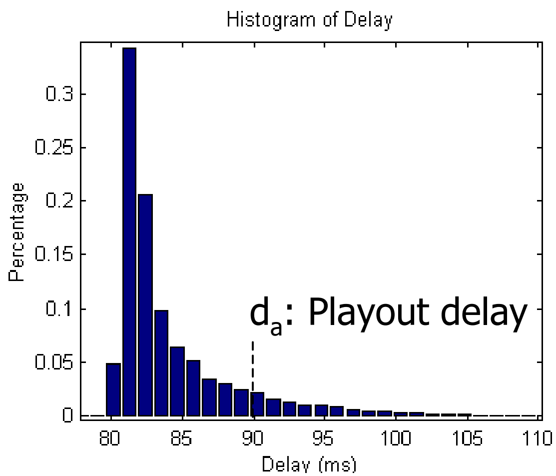
$$INPUT: delay_i: \text{ network delay of a current packet}$$

$$d\_est_i = \alpha \cdot d\_est_{i-1} + (1-\alpha) \cdot delay_i$$

$$v\_est_i = \alpha \cdot v\_est_{i-1} + (1-\alpha) \cdot \left| d\_est_i - delay_i \right|$$

$$OUTPUT: p_i = d\_est_i + \beta v\_est_i : \text{ playout delay of a current talkspurt}$$

- **Histogram-based algorithm *[Moon '98]***



- Past 10 – 1000 delays are stored in a circular array, both the array and the histogram get updated once a new packet is received.
- Loss rate specified by user.
- Playout time chosen from the histogram according to the desired loss rate.

Problem to solve

- management of the playout buffer is not specified by any standard and is vendor specific
- many fixed and adaptive playout schemes exist, each with a different parameter set
- information on the implementation of the playout buffer in commercial applications is practically nonexistent (it has a strategic value from vendor perspective)
- WE NEED A METHOD TO EVALUATE THEM FROM AN END USER PERSPECTIVE

How?

- “Listening-only” tests do not take into account delay impairments

- PESQ do not take into account end-to-end one-way delay in its rating

-> NOT RECOMMENDED to asses the effect of CONVERSATIONAL delay

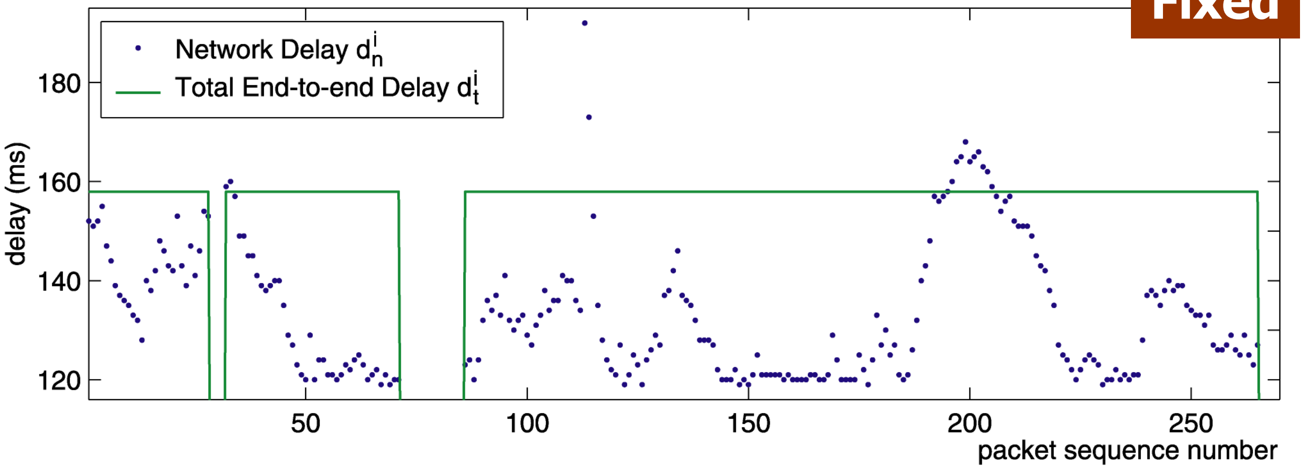
- E-model relies on static transmission parameters and do not take into account dynamics of changing transmission impairments (delay,loss)

- PESQ+E-model hybrid solution requires a reference speech signal and does not work in real time

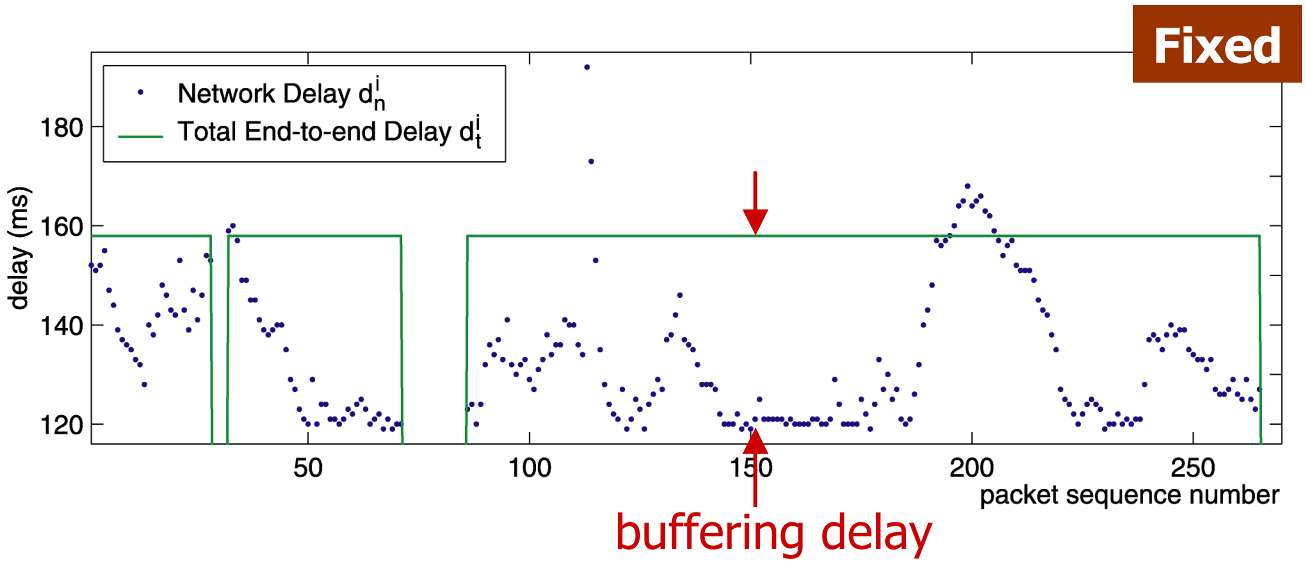
- old statistical methods focus on “late packet loss vs. buffering delay” trade-off only and do not provide a direct link to conversational speech quality

Trade-off: late packet loss/buffering delay

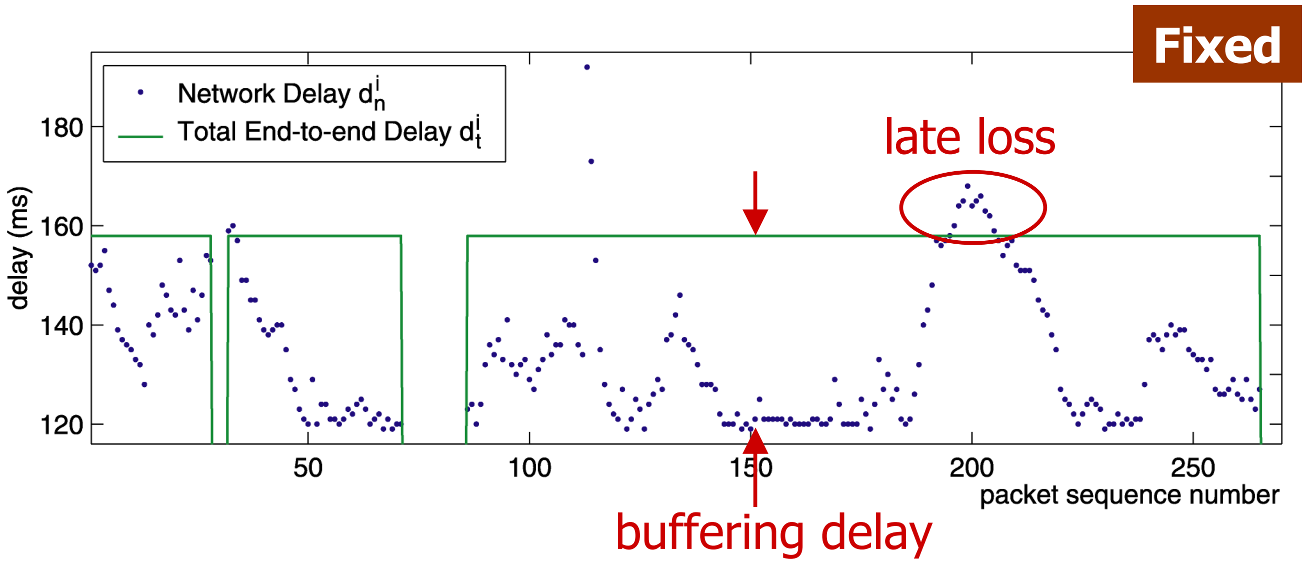
Fixed



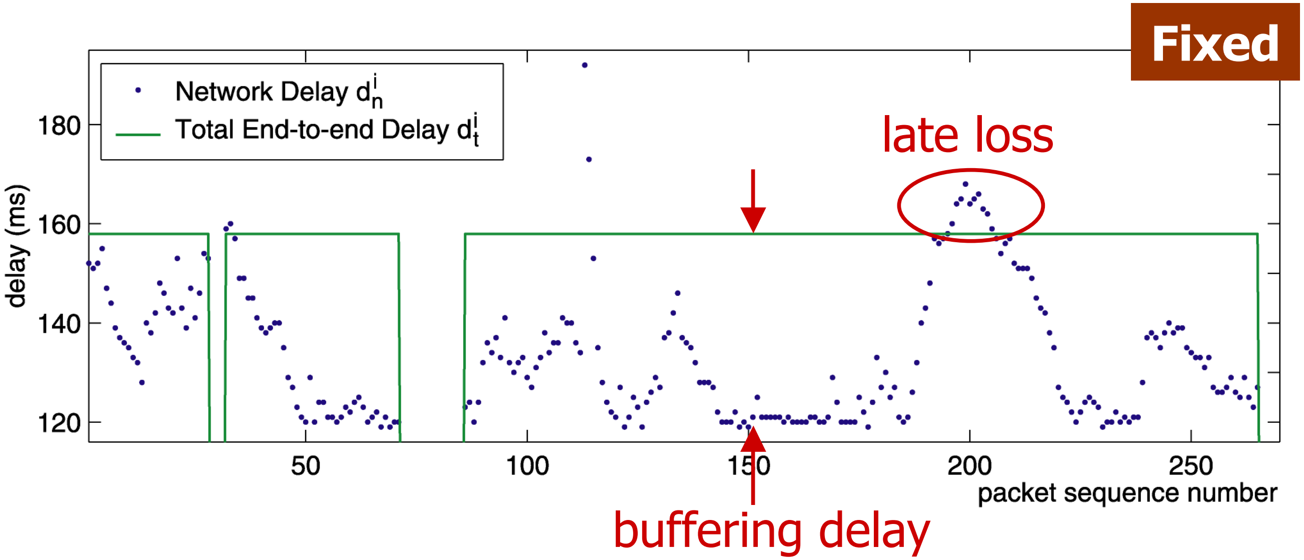
Trade-off: late packet loss/buffering delay



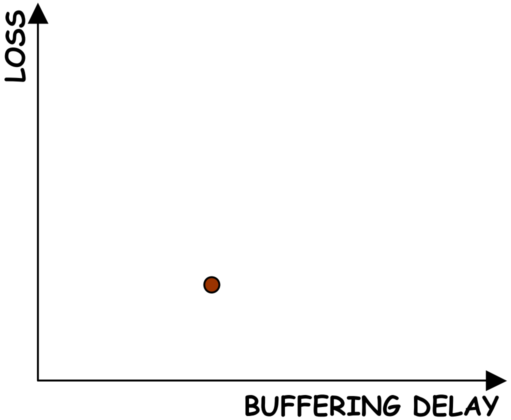
Trade-off: late packet loss/buffering delay



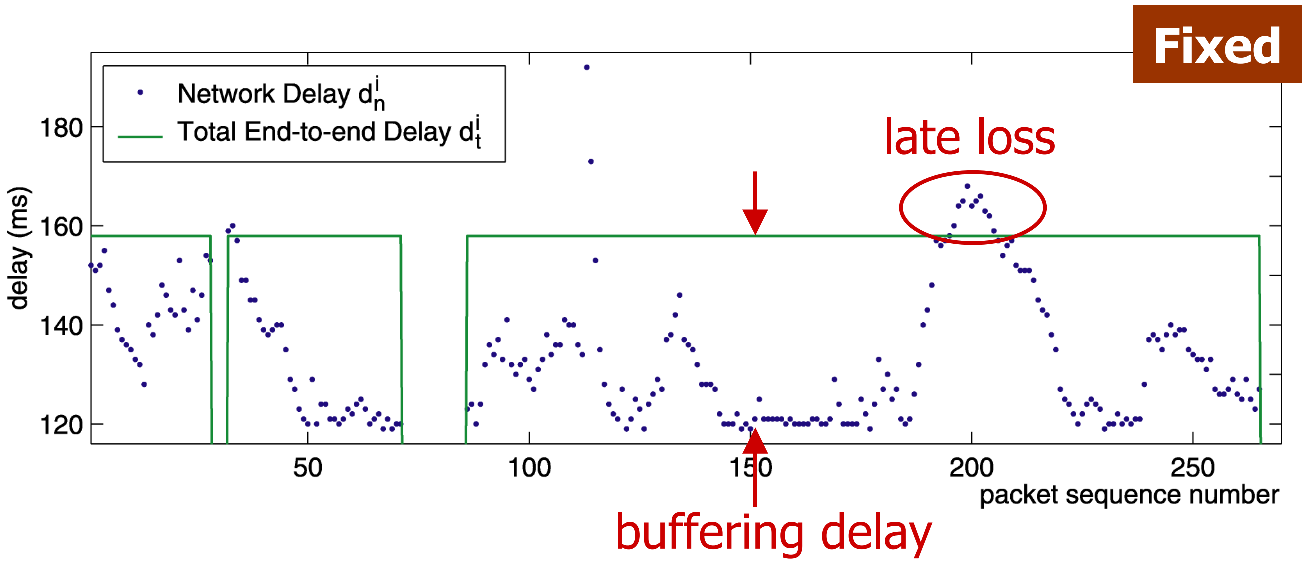
Trade-off: late packet loss/buffering delay

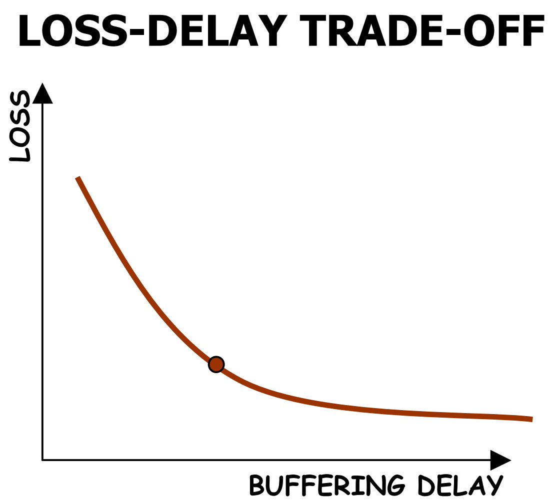


LOSS-DELAY TRADE-OFF

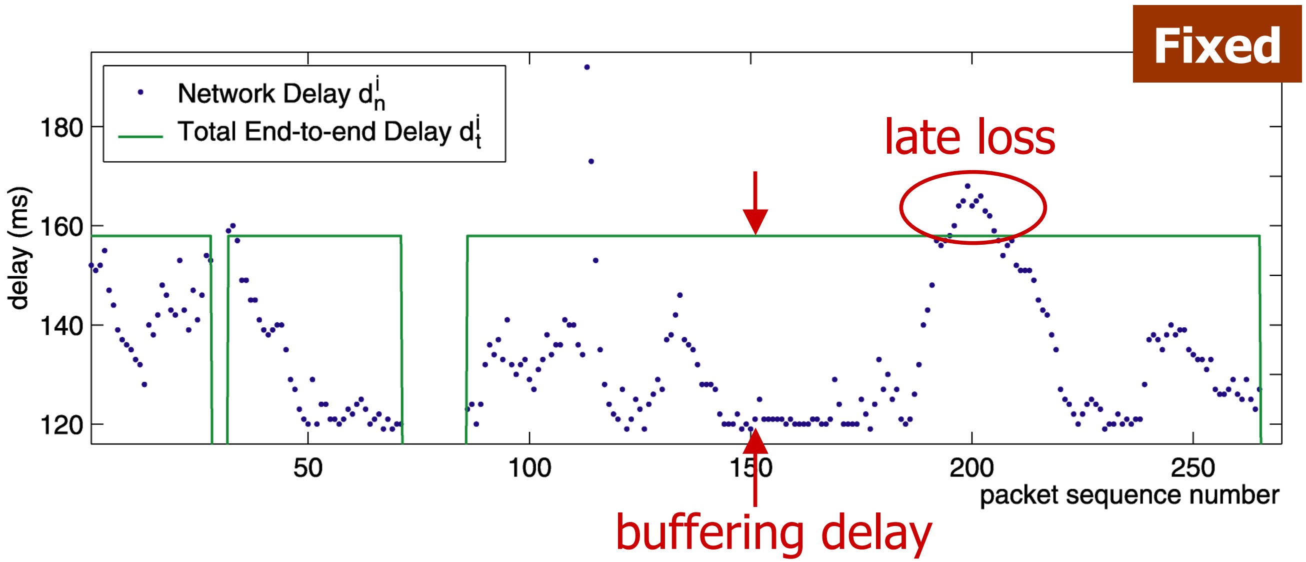


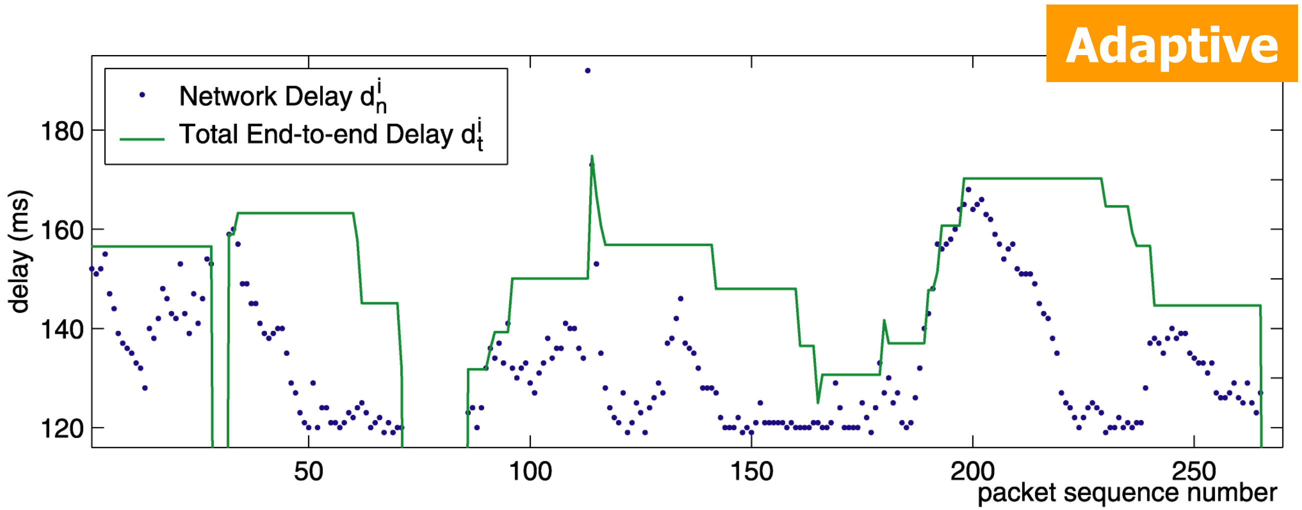
Trade-off: late packet loss/buffering delay

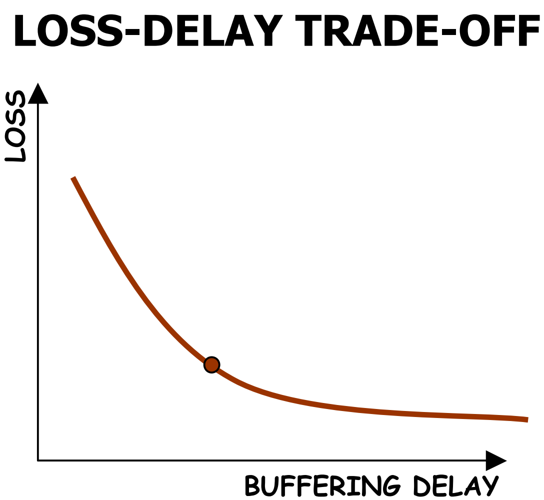




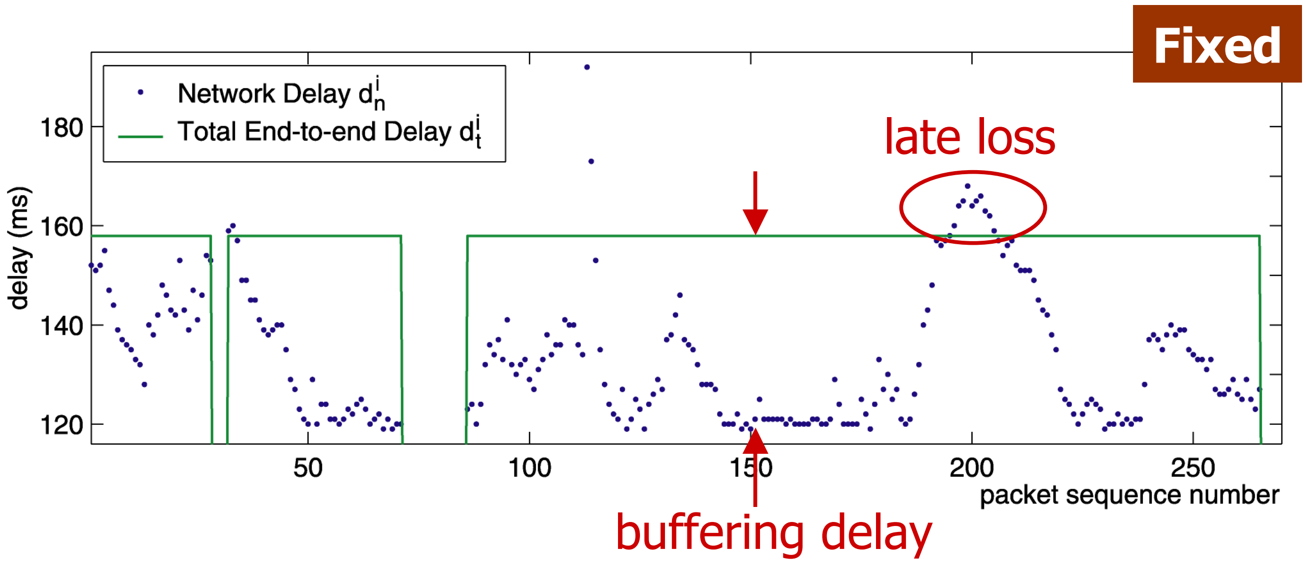
Trade-off: late packet loss/buffering delay

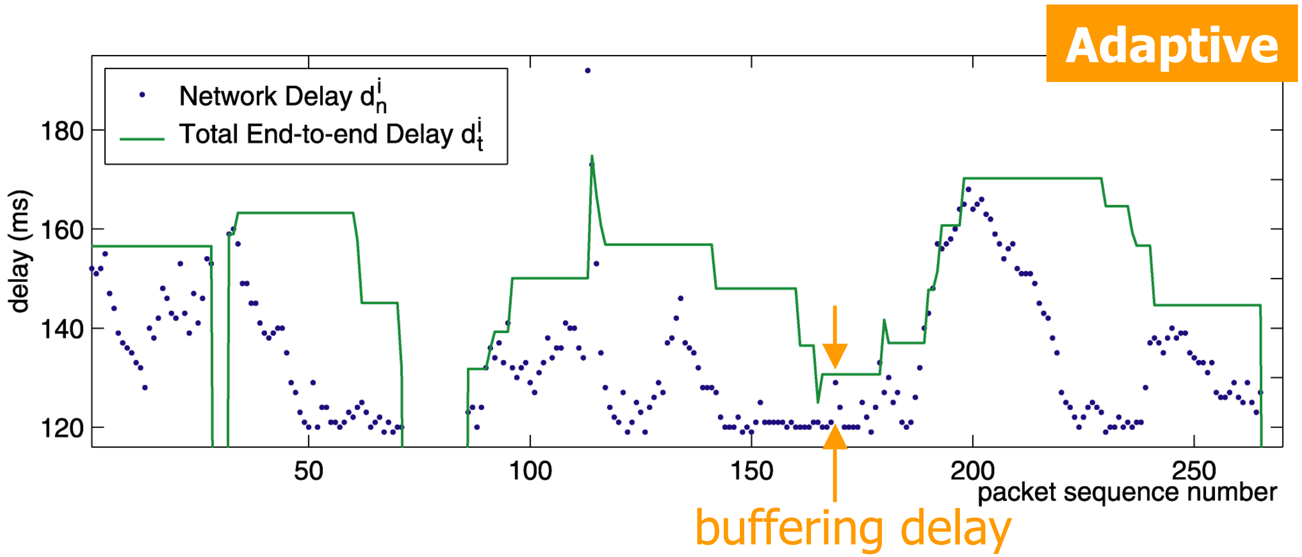


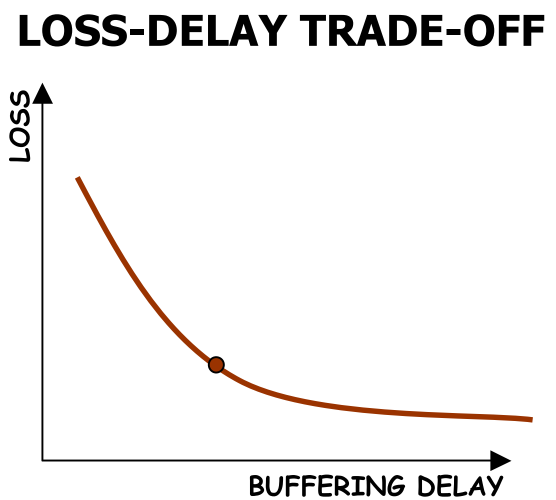




Trade-off: late packet loss/buffering delay

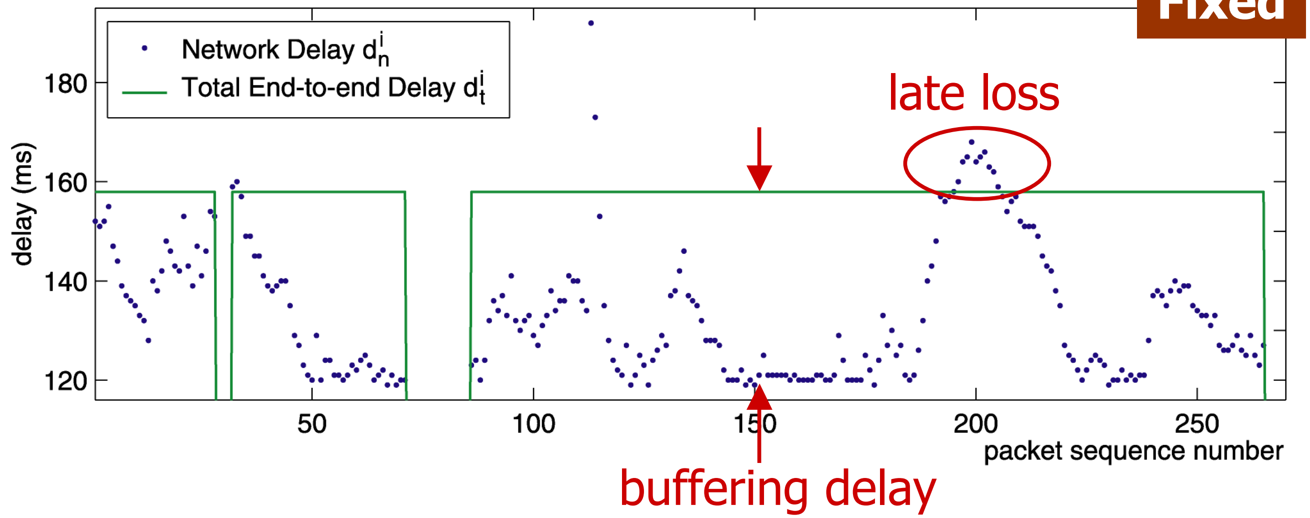




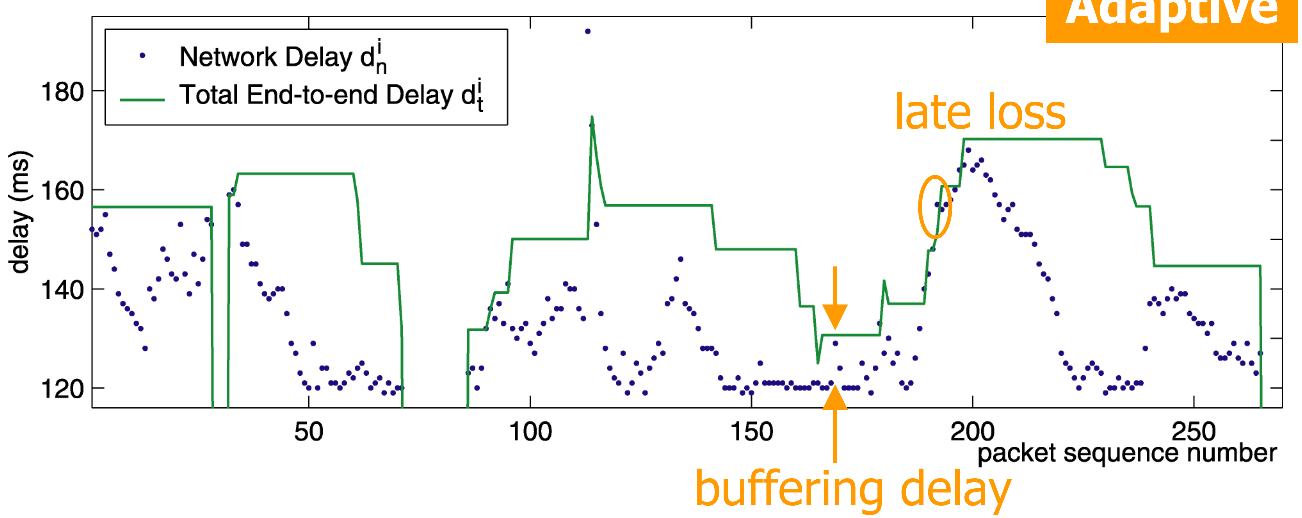


Trade-off: late packet loss/buffering delay

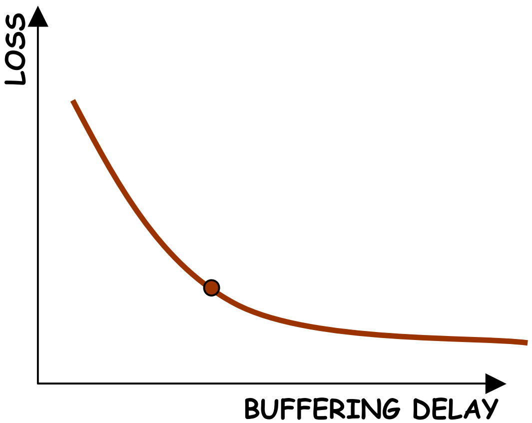
Fixed



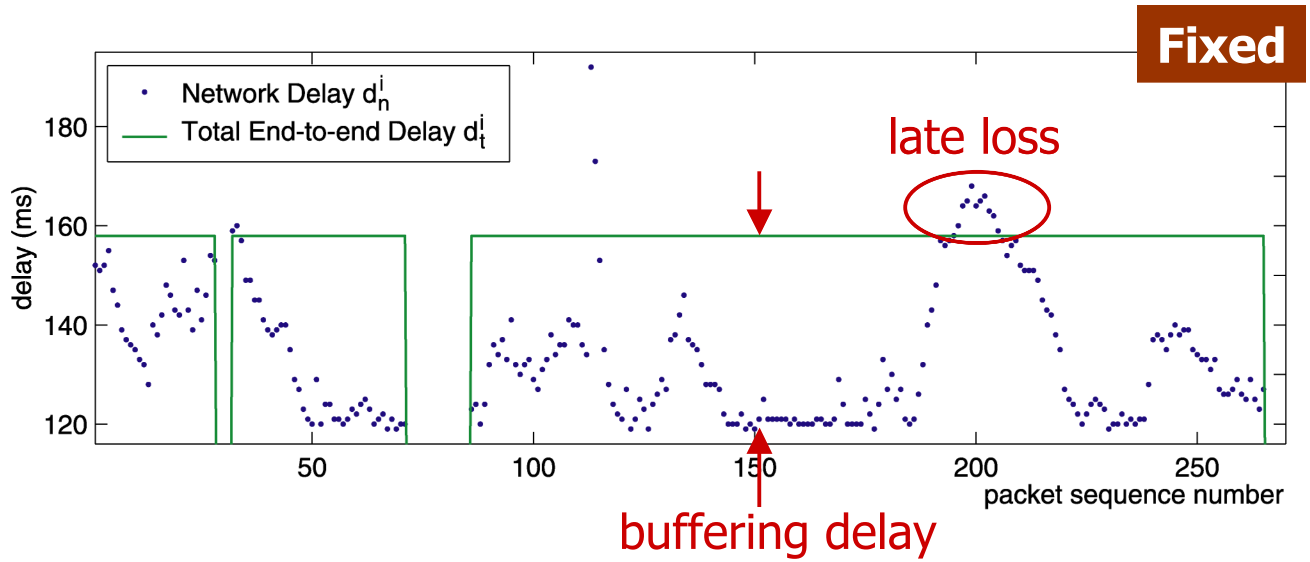
Adaptive

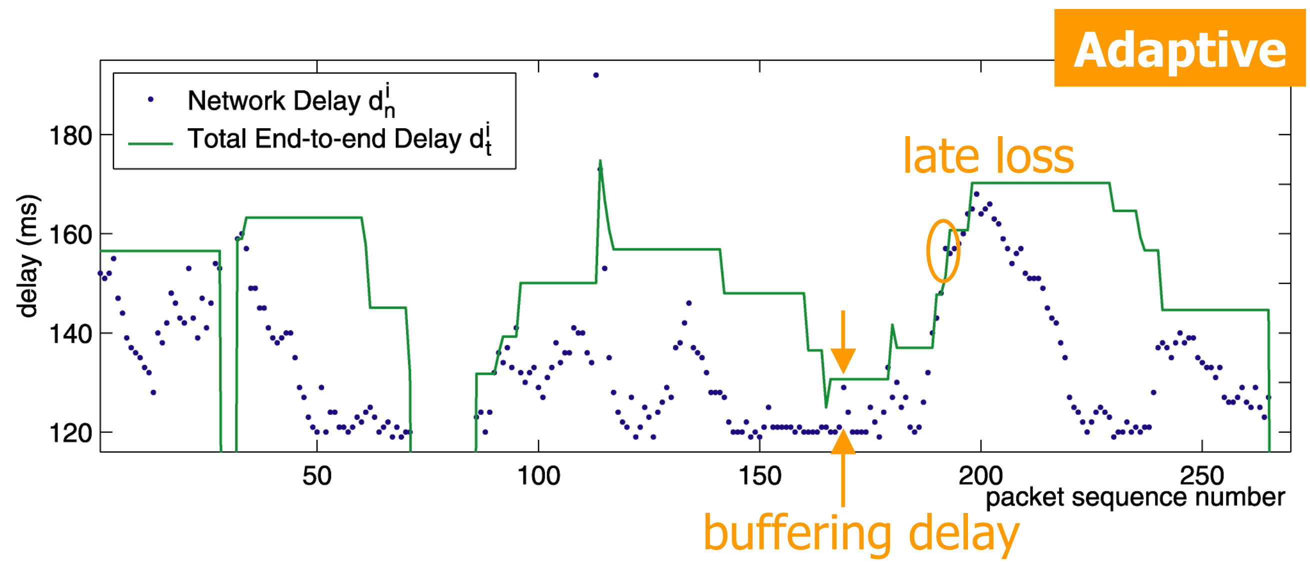


LOSS-DELAY TRADE-OFF

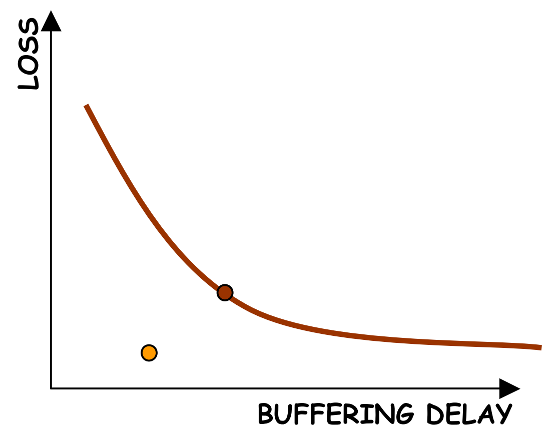


Trade-off: late packet loss/buffering delay



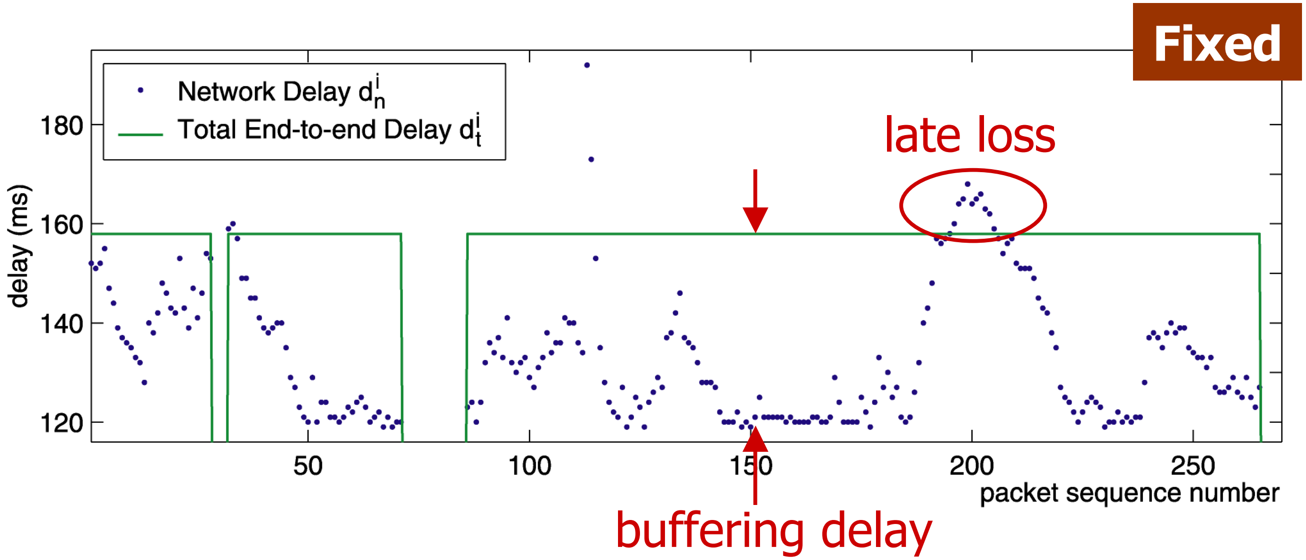


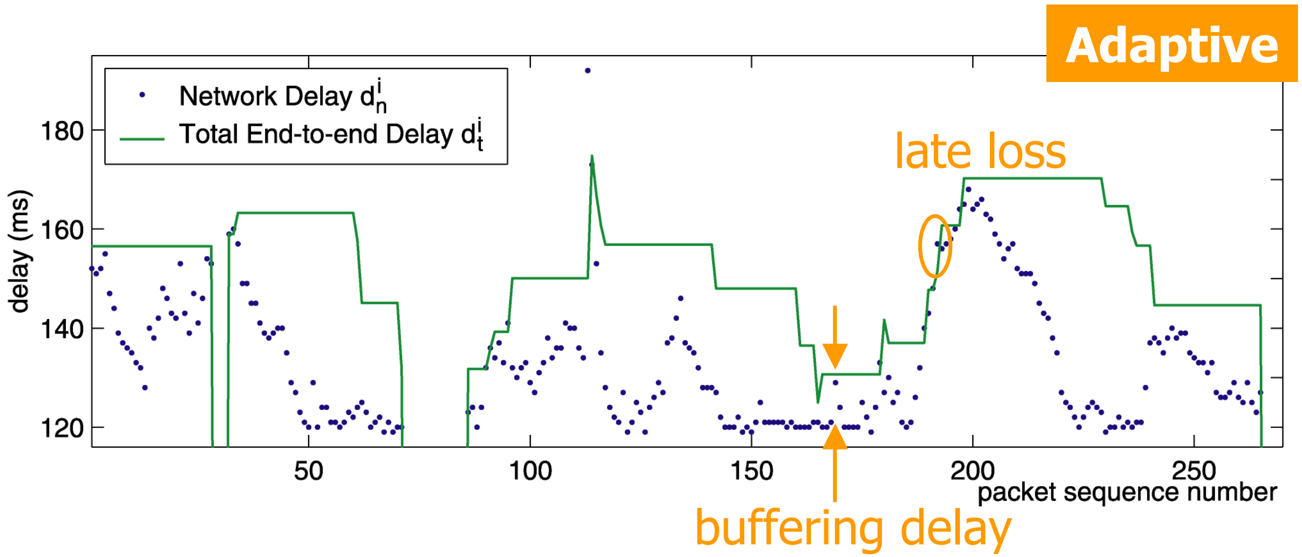
LOSS-DELAY TRADE-OFF

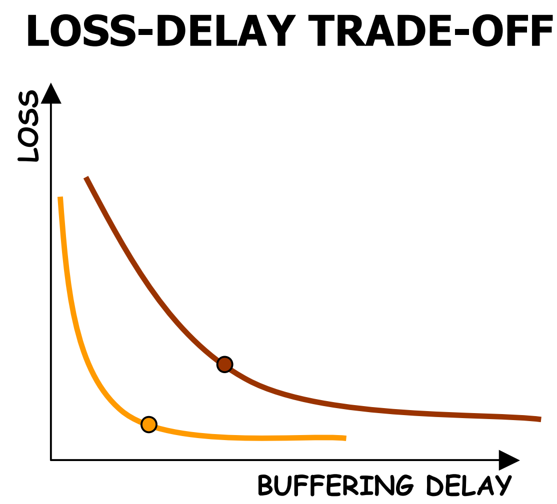


Assessing the quality of VoIP transmission affected by playout buffer scheme and encoding scheme

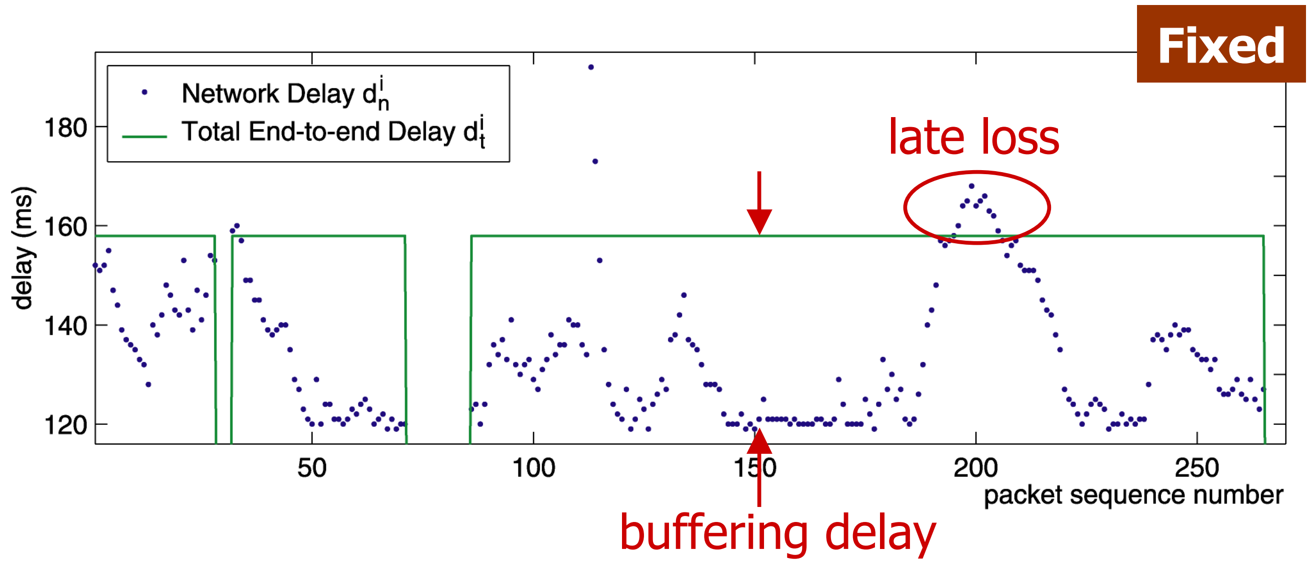
Trade-off: late packet loss/buffering delay

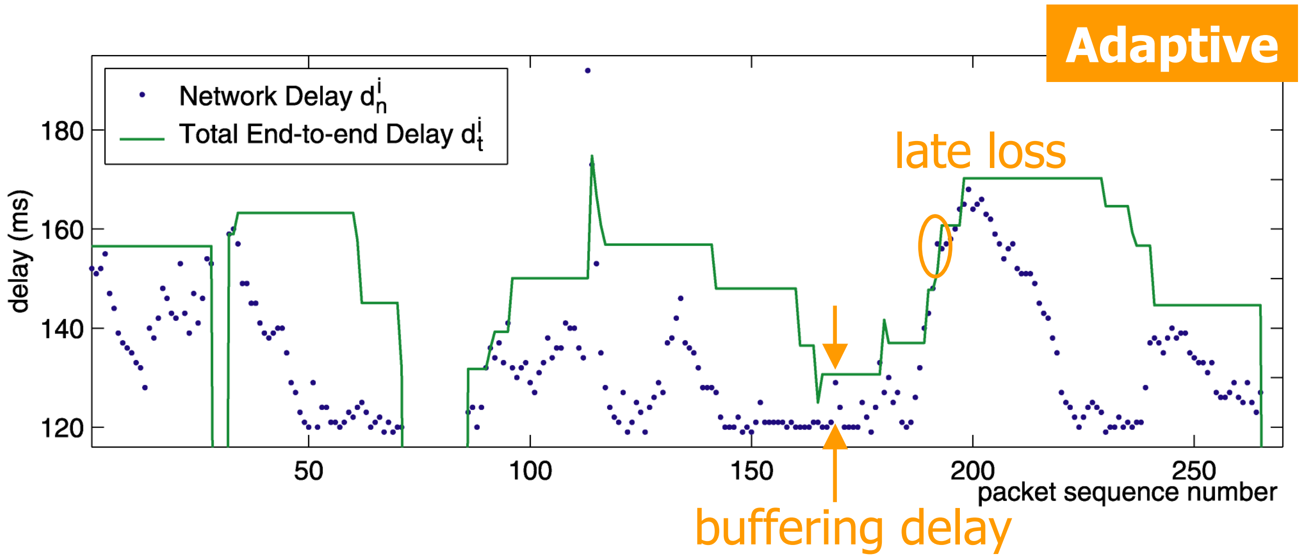


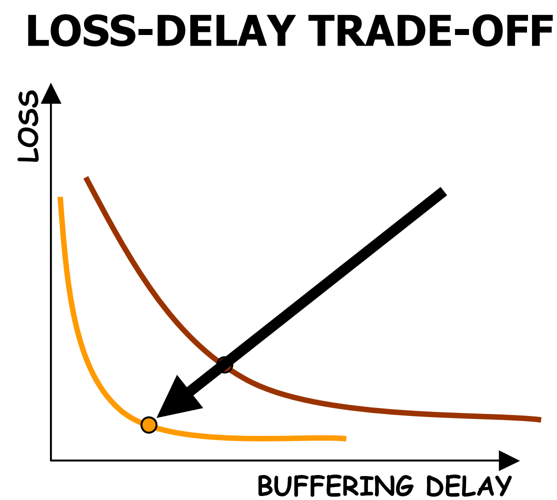




Trade-off: late packet loss/buffering delay





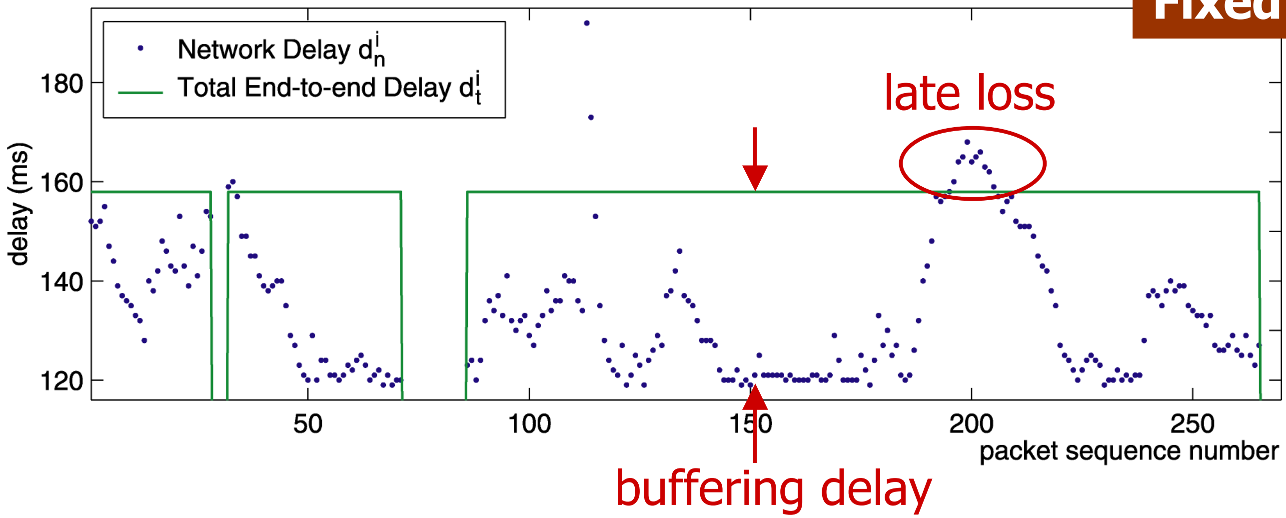


good algorithm SHOULD:

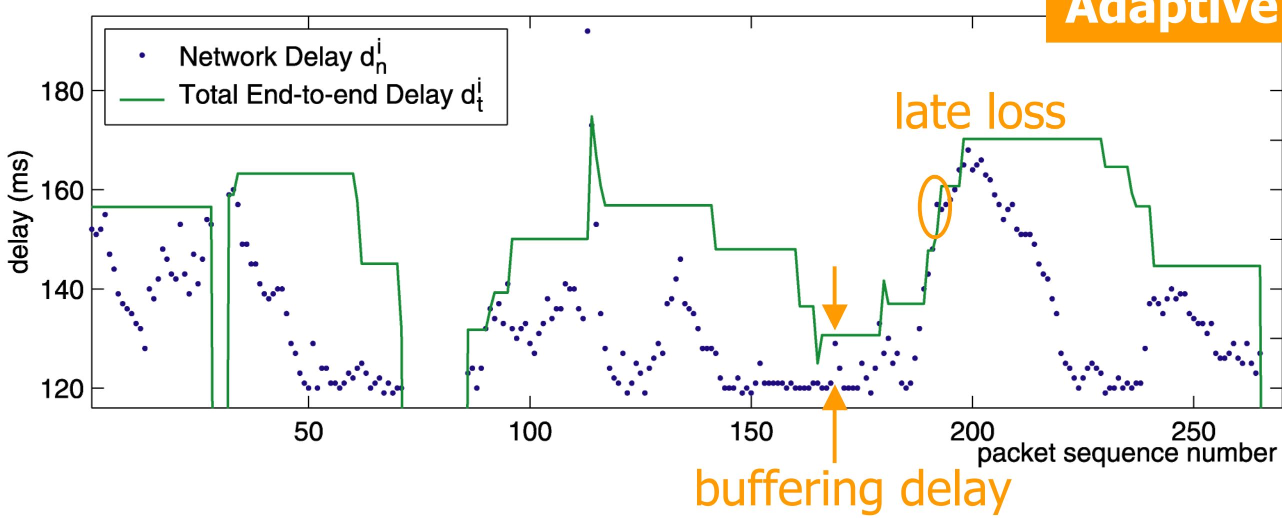
- achieve the best possible loss-delay trade-off

Trade-off: late packet loss/buffering delay

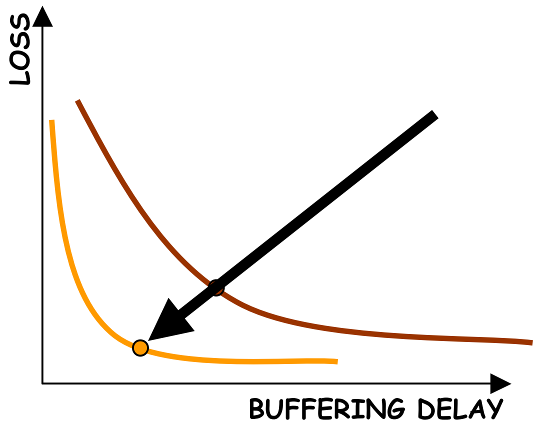
Fixed



Adaptive



LOSS-DELAY TRADE-OFF



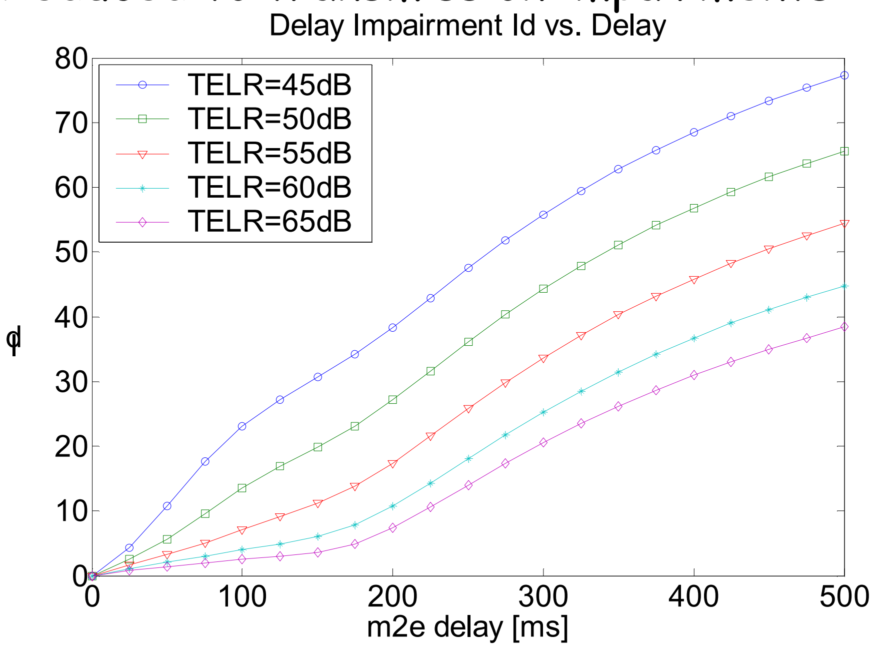
good algorithm SHOULD:

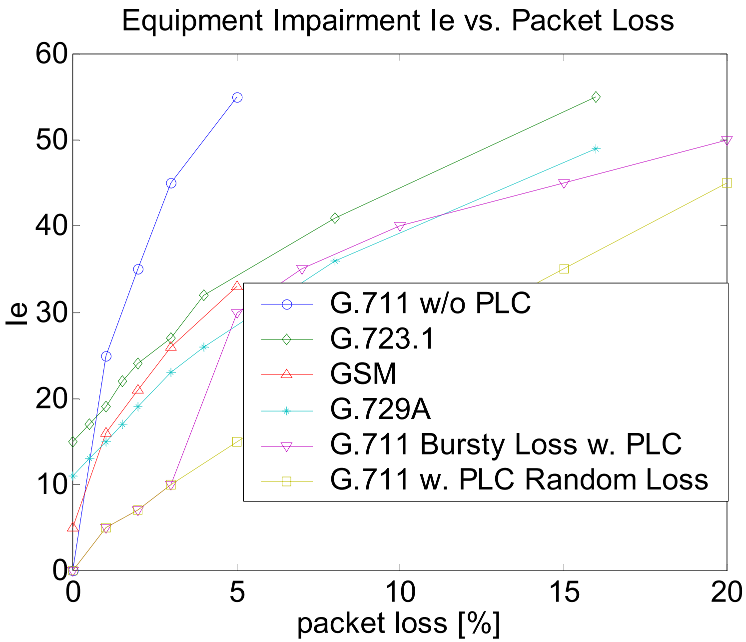
- achieve the best possible loss-delay trade-off
- be well tuned to **maximize user satisfaction**

E-model, speech transmission categories, user satisfaction

Transmission rating factor R = $(R_0 - I_s) - I_{d(echo,delay)} - I_{e(codec,\ loss)} + A$

Reduced to transmission impairments: R = 94.15 $- I_{d(echo,delay)} - I_{e(codec,\ loss)}$

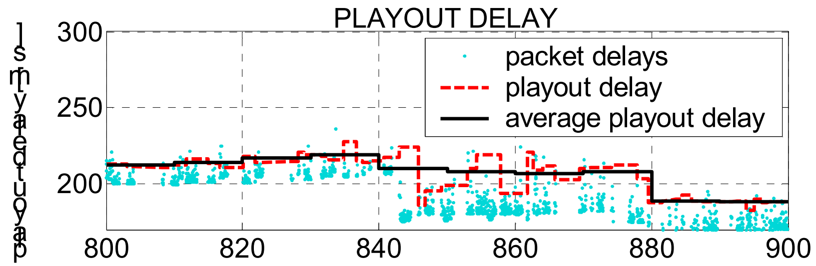




R-value	94.15 -90	90-80	80-70	70-60	60-50
Speech transmission quality	**Best**	**High**	**Medium**	**Low**	**Poor**
User satisfaction	**very satisfied**	**satisfied**	**some very dissatisfied**	**many very dissatisfied**	**almost all dissatisfied**

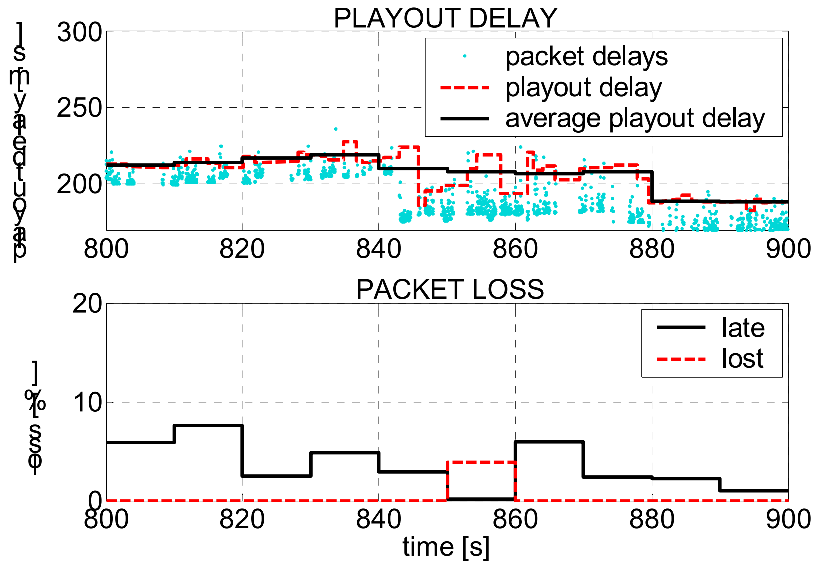
Predicting time varying speech transmission quality

Predicting time varying speech transmission quality



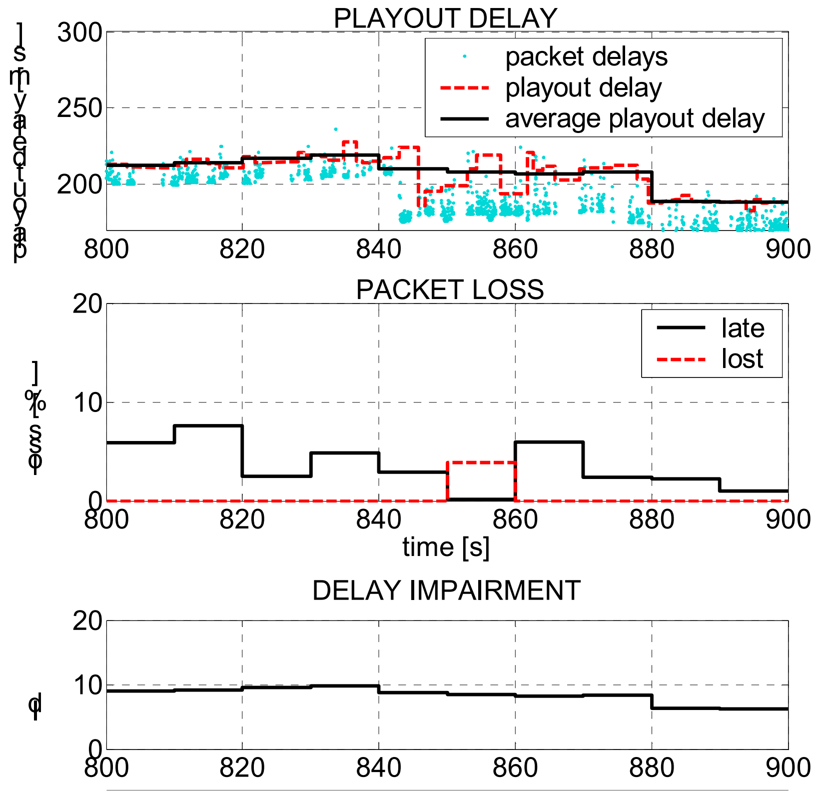
1. Playout buffer module calculates playout delays (m2e), average playout delays…

Predicting time varying speech transmission quality



1. Playout buffer module calculates playout delays (m2e), average playout delays…
2. … and resulting average packet loss values for fixed time windows

Predicting time varying speech transmission quality

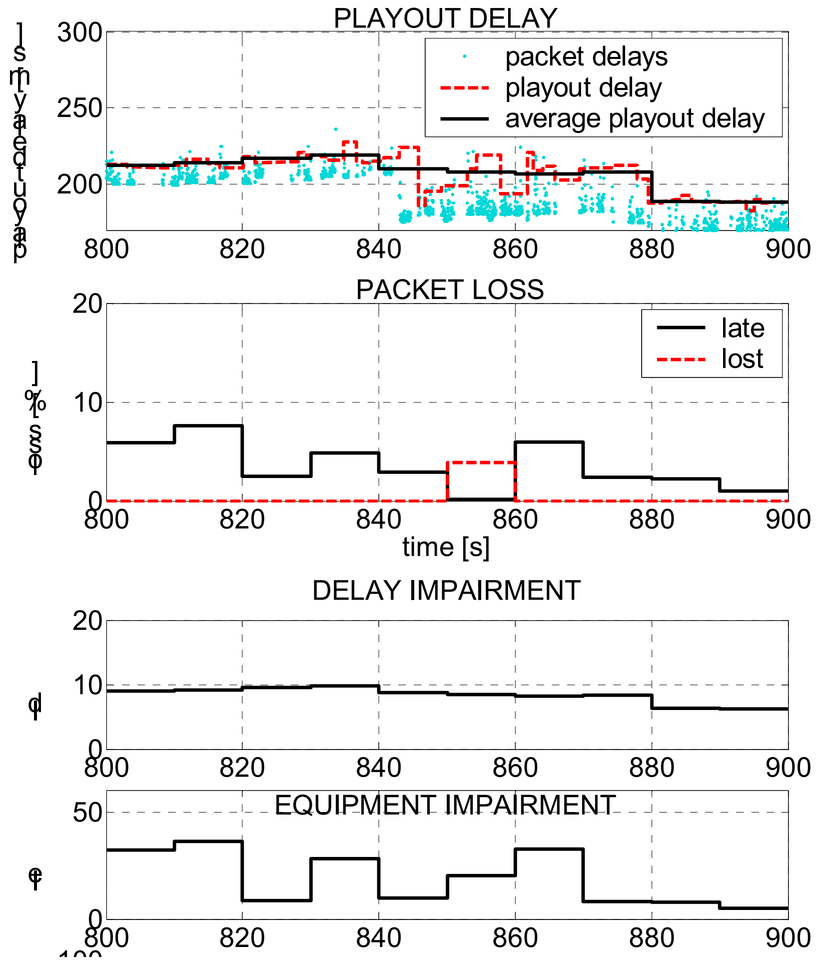


1. Playout buffer module calculates playout delays (m2e), average playout delays…

2. … and resulting average packet loss values for fixed time windows

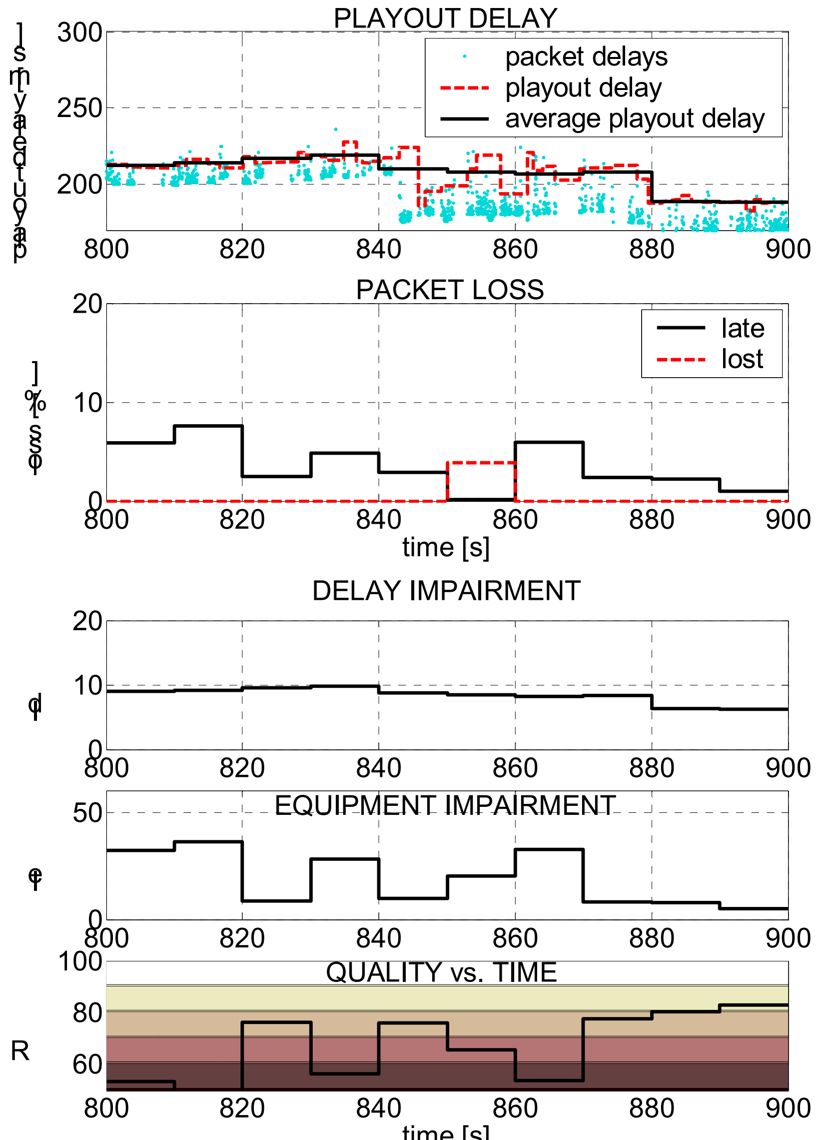
3. E-model simulator calculates corresponding delay impairments Id …

Predicting time varying speech transmission quality



1. Playout buffer module calculates playout delays (m2e), average playout delays…
2. … and resulting average packet loss values for fixed time windows
3. E-model simulator calculates corresponding delay impairments Id …
4. ... and equipment impairments Ie…

Predicting time varying speech transmission quality

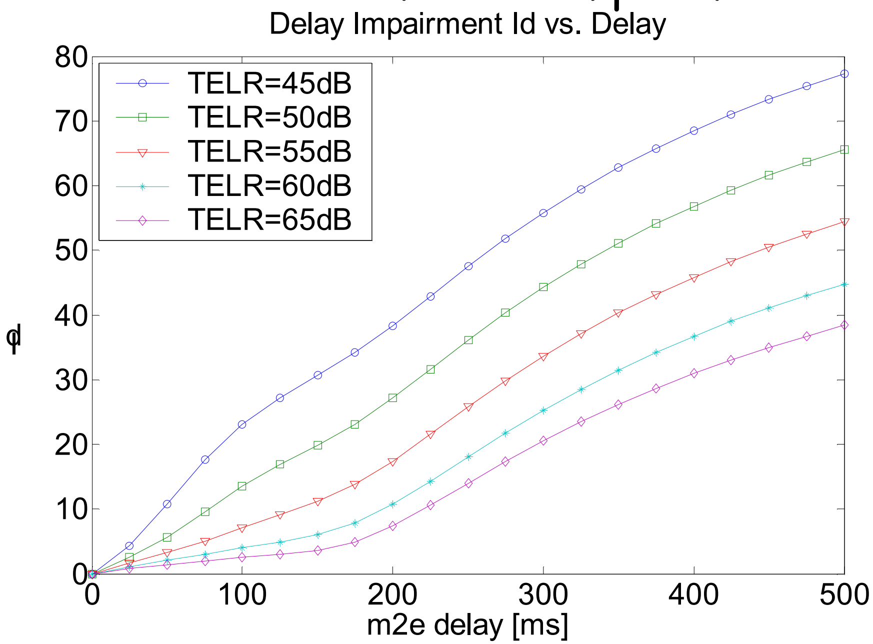


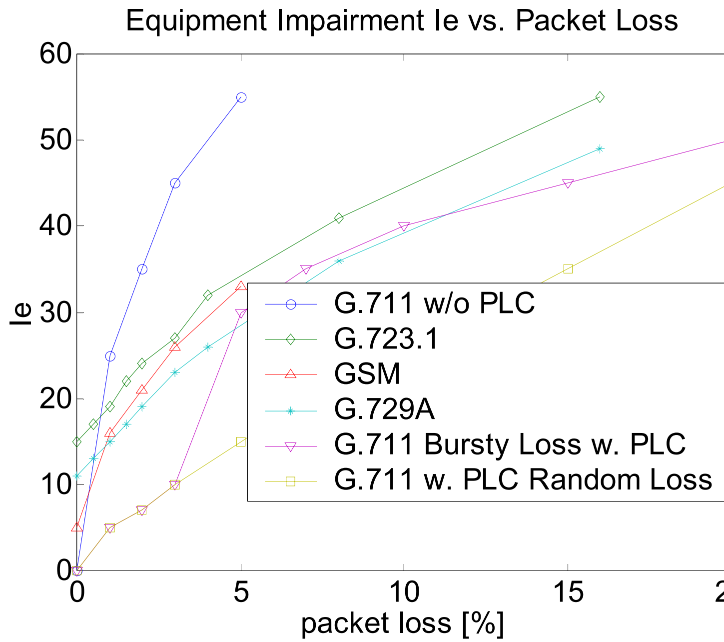
1. Playout buffer module calculates playout delays (m2e), average playout delays…
2. … and resulting average packet loss values for fixed time windows
3. E-model simulator calculates corresponding delay impairments Id …
4. ... and equipment impairments Ie…
5. …and resulting rating R

E-model, speech transmission categories, user satisfaction

Transmission rating factor R = $(R_0 - I_s) - I_{d(echo,delay)} - I_{e(codec,\ loss)} + A$

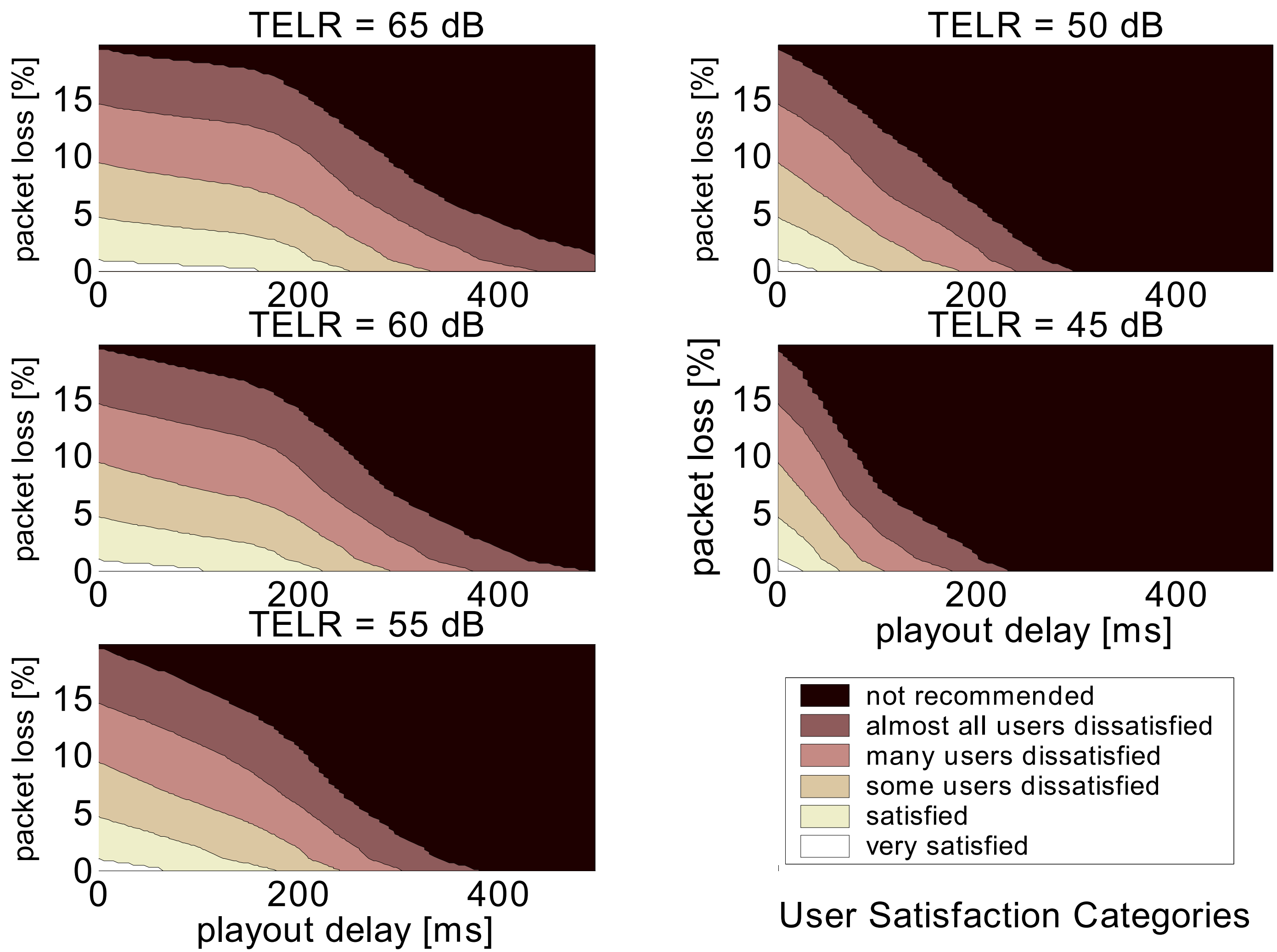
Reduced to transmission impairments: R = $94.15 - I_{d(echo,delay)} - I_{e(codec,\ loss)}$



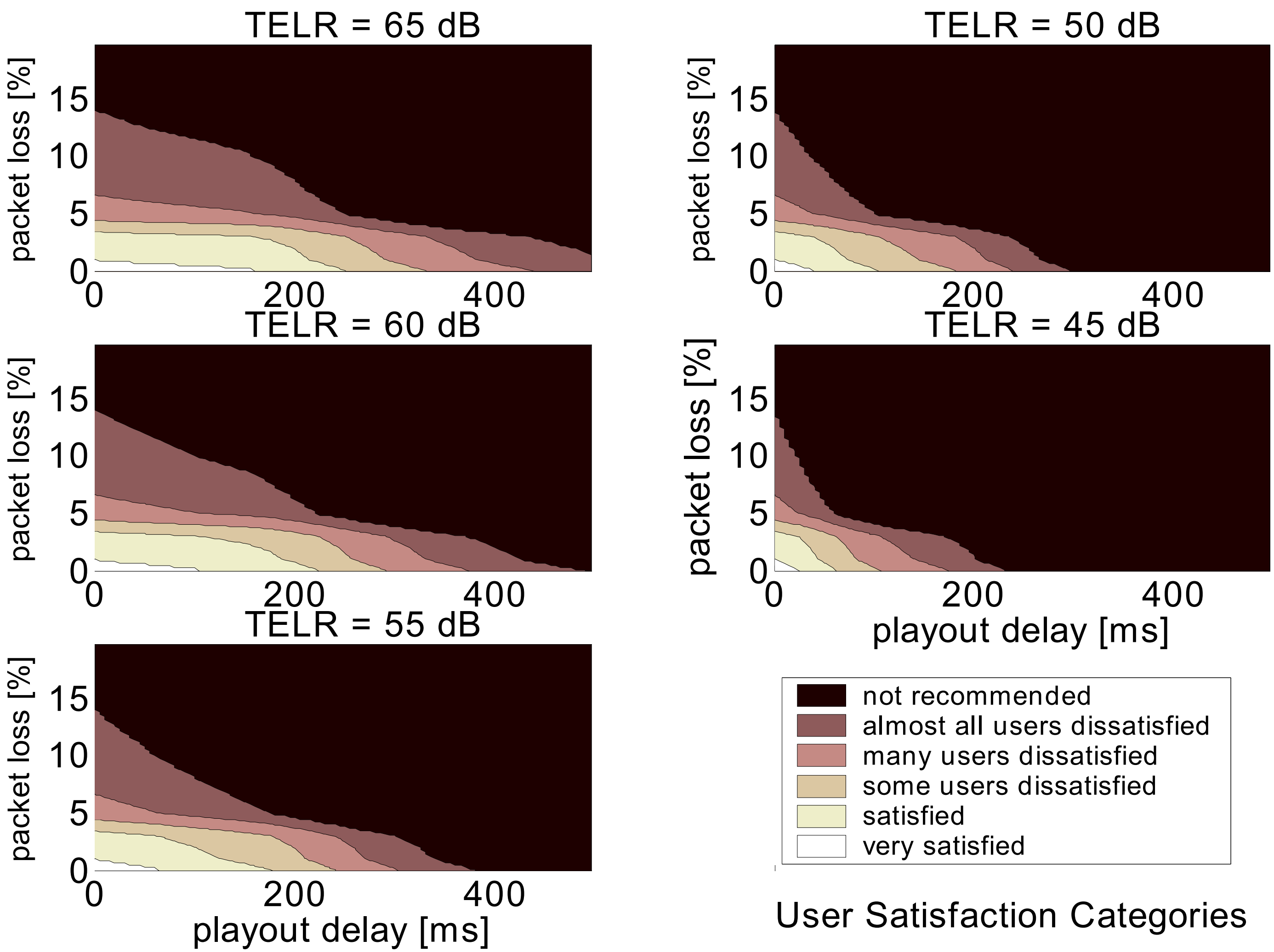


R-value	94.15 -90	90-80	80-70	70-60	60-50
Speech transmission quality	Best	High	Medium	Low	Poor
User satisfaction	very satisfied	satisfied	some very dissatisfied	many very dissatisfied	almost all dissatisfied

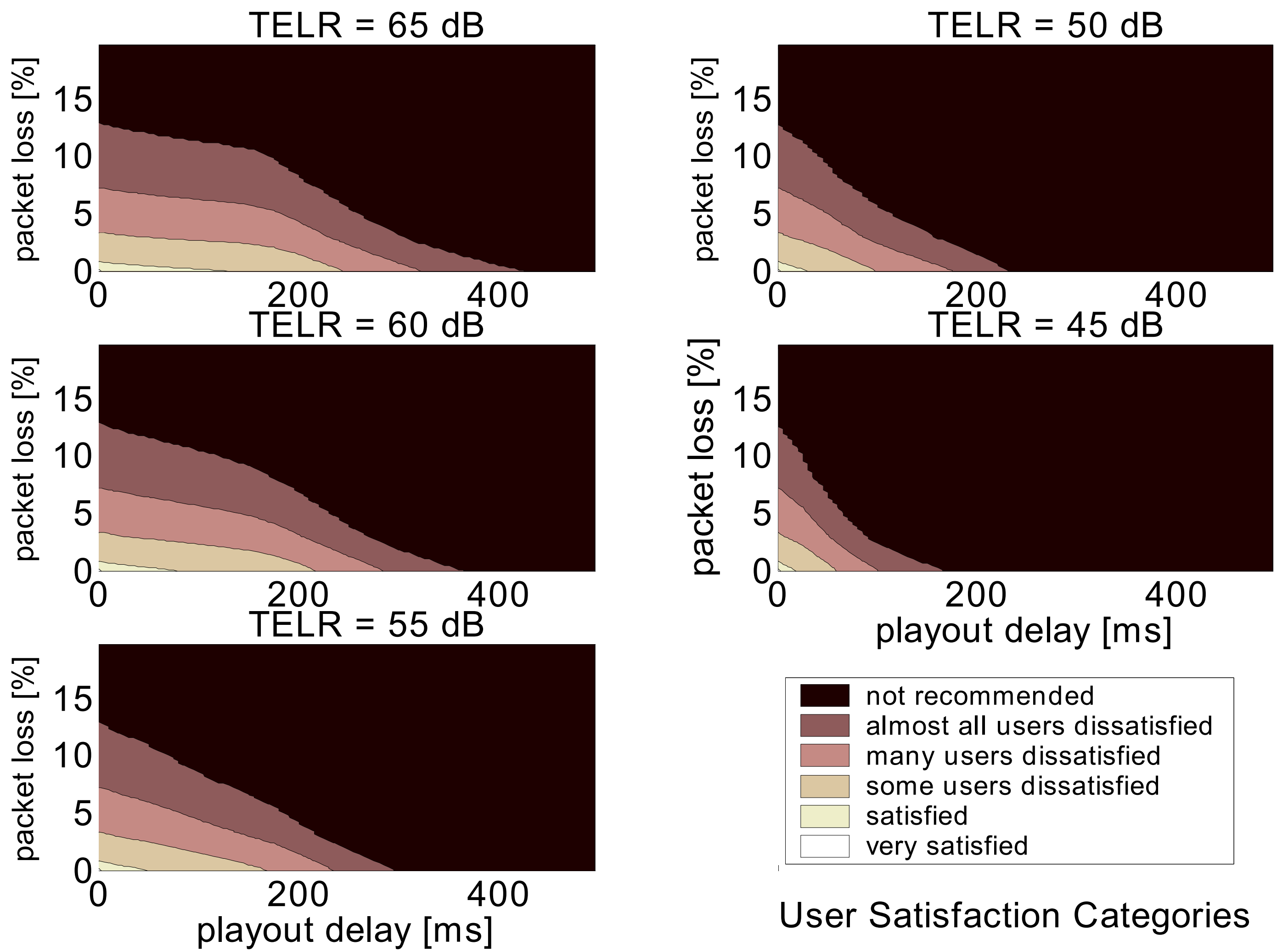
Quality contours for G.711 with PLC (random loss)



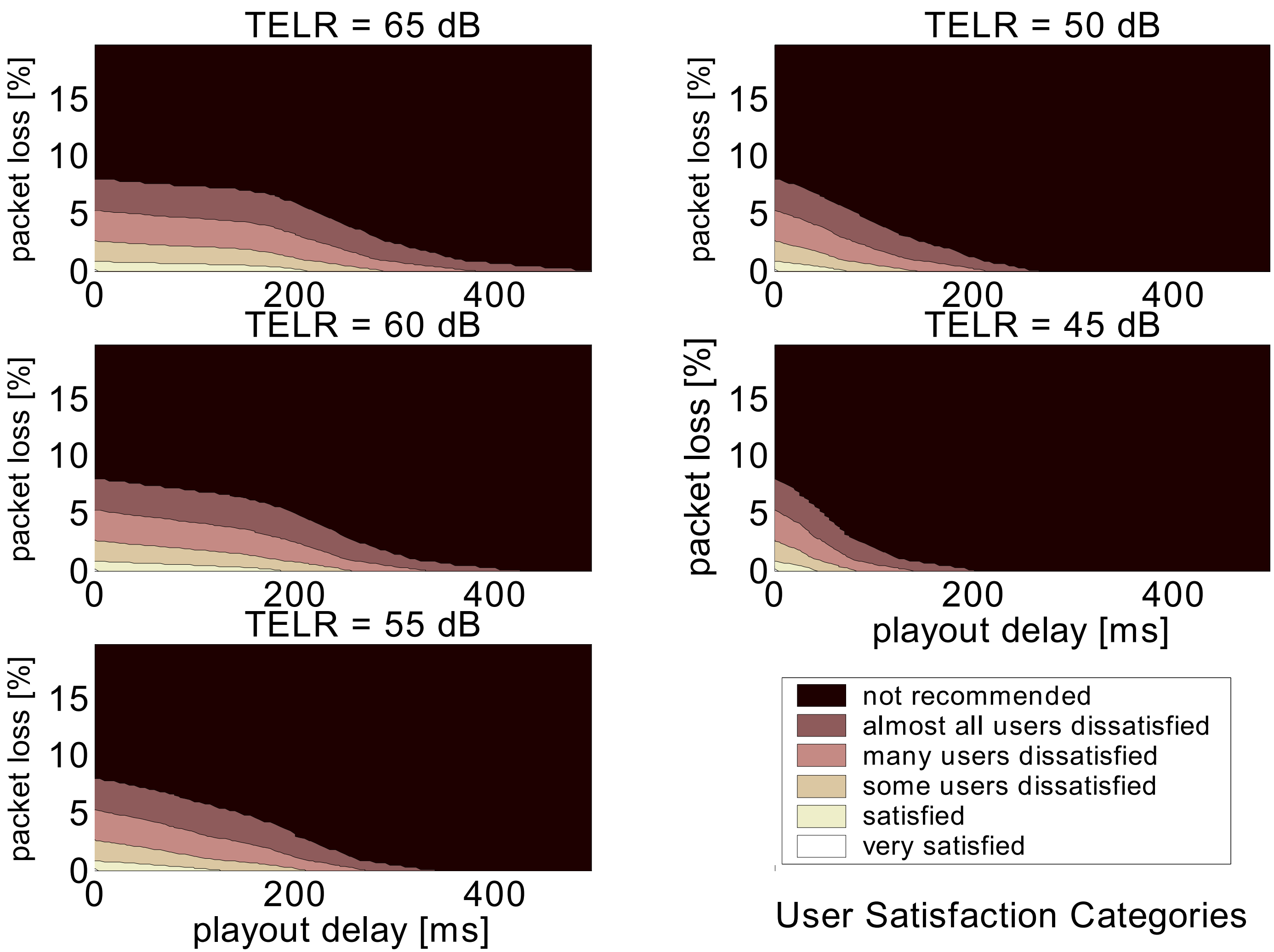
Quality contours for G.711 with PLC (bursty loss)



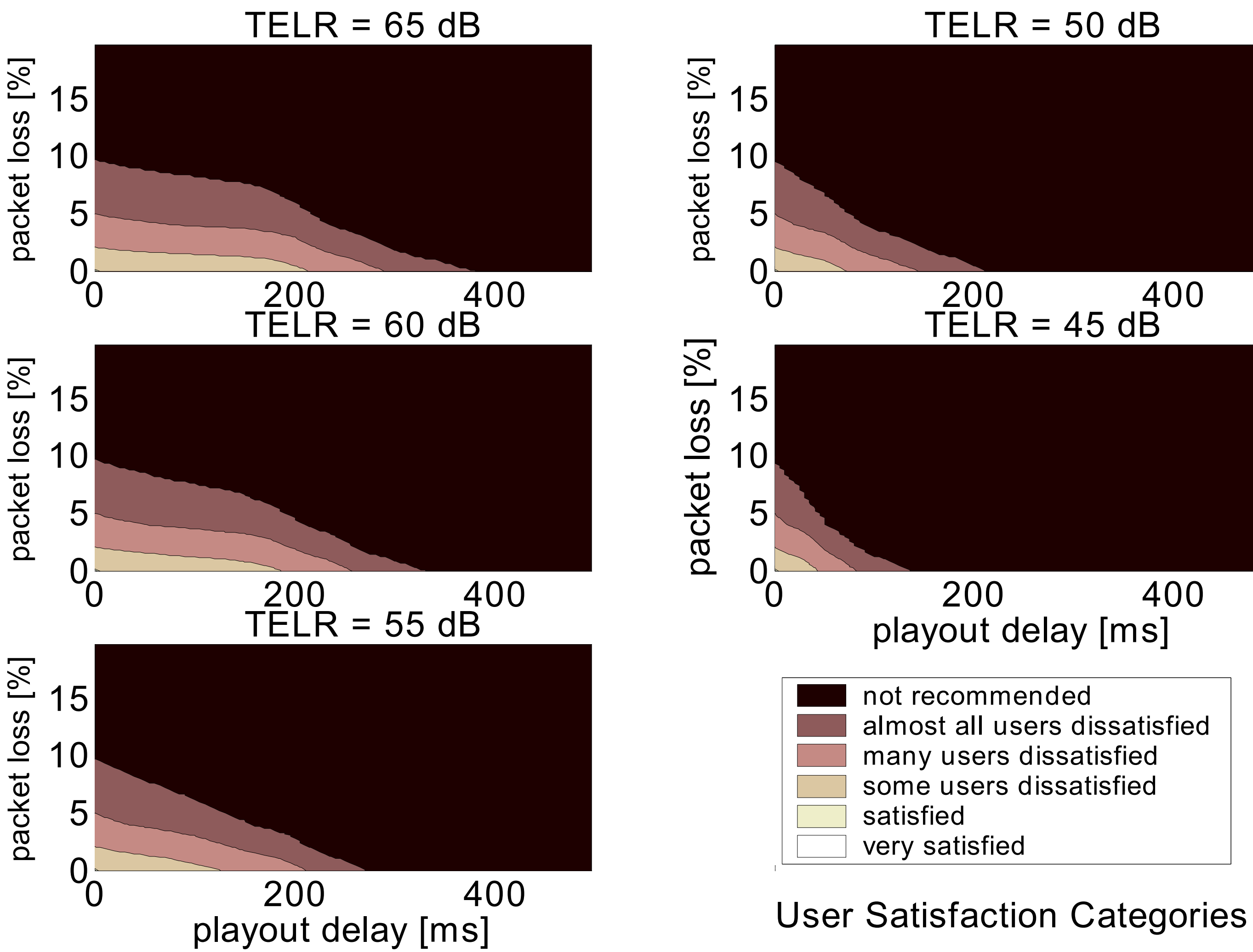
Quality contours for G.729A



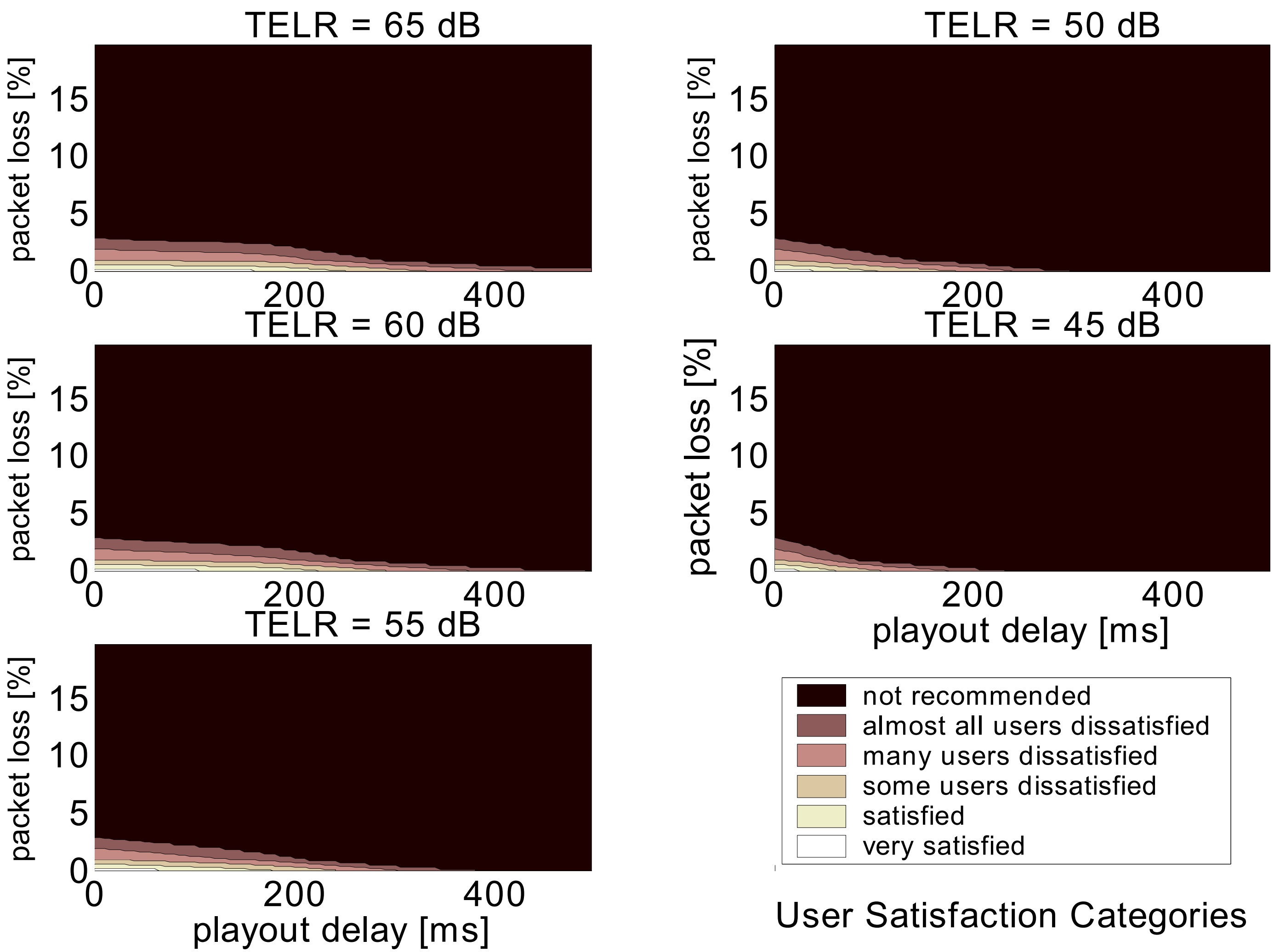
Quality contours for GSM



Quality contours for G.723.1 (6.4kbps)

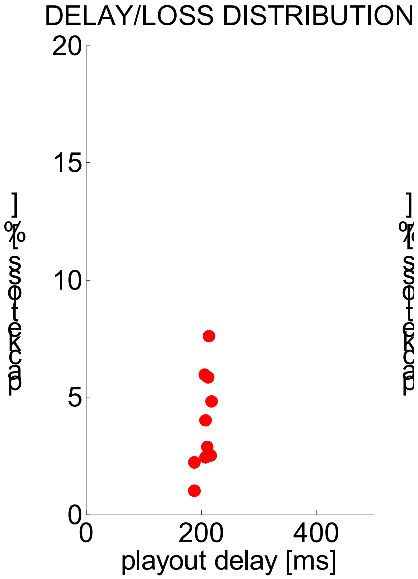


Quality contours for G.711 w/o PLC

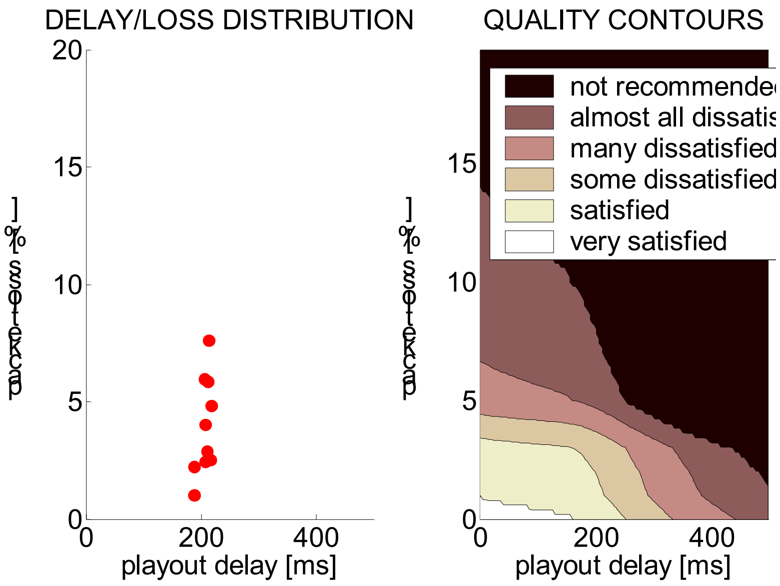


Predicting user satisfaction (proposed solution)

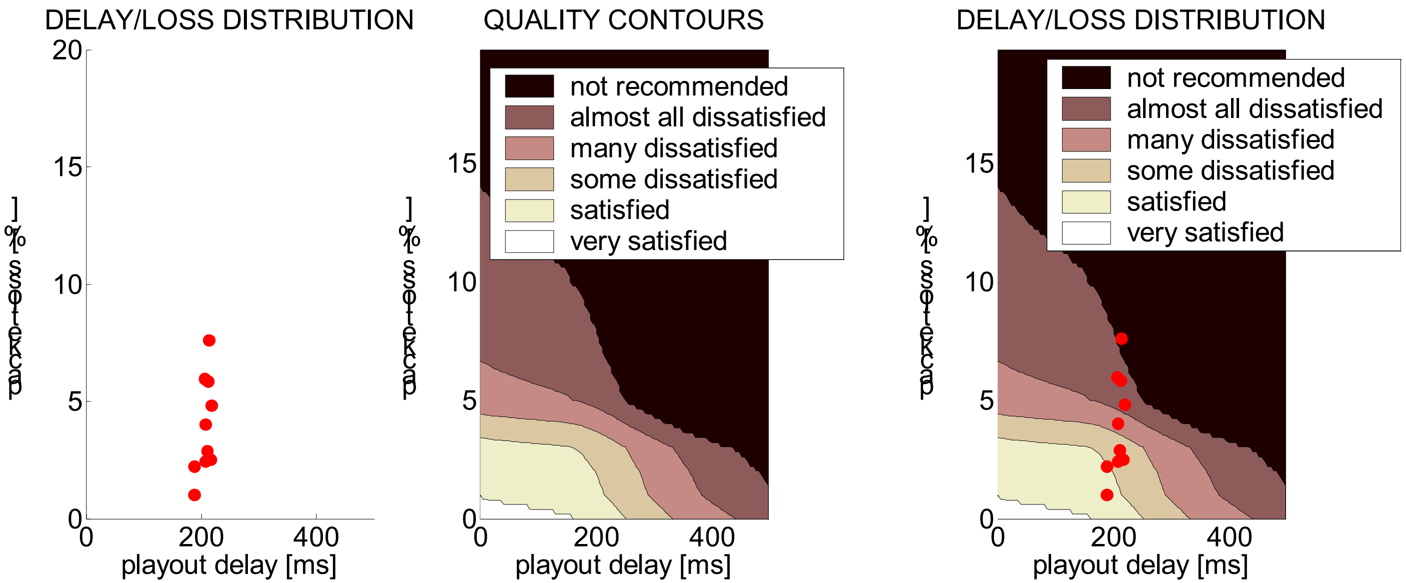
Predicting user satisfaction (our solution)



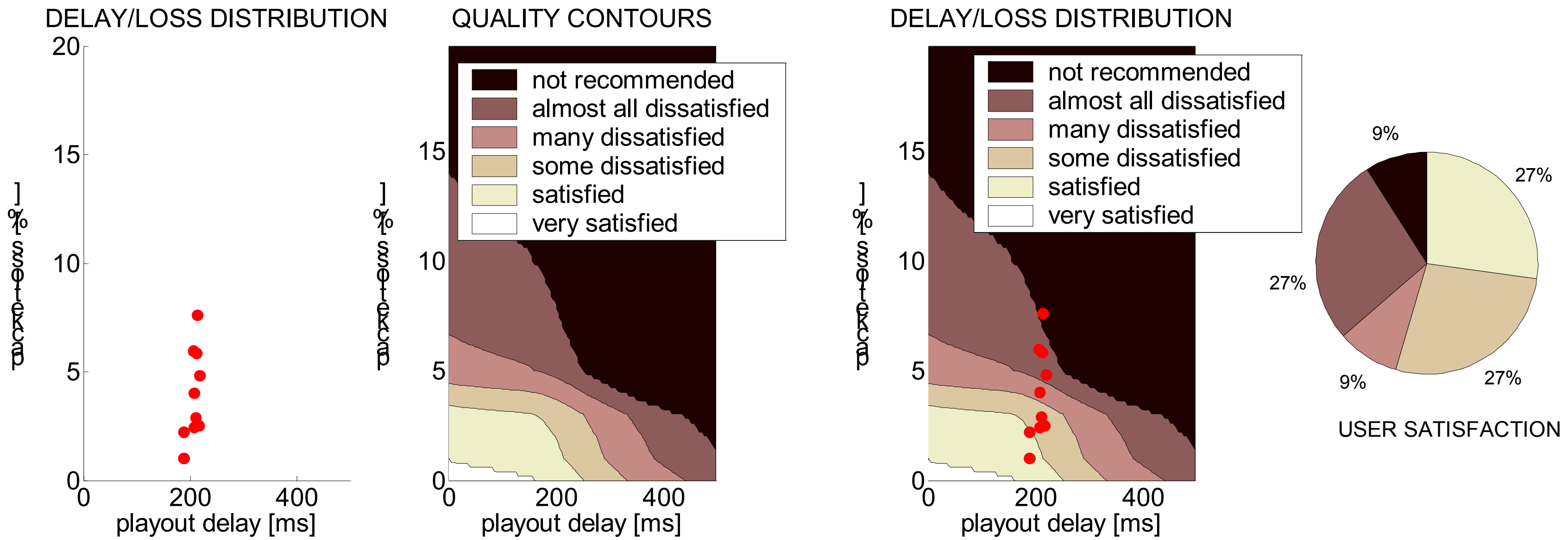
Predicting user satisfaction (our solution)



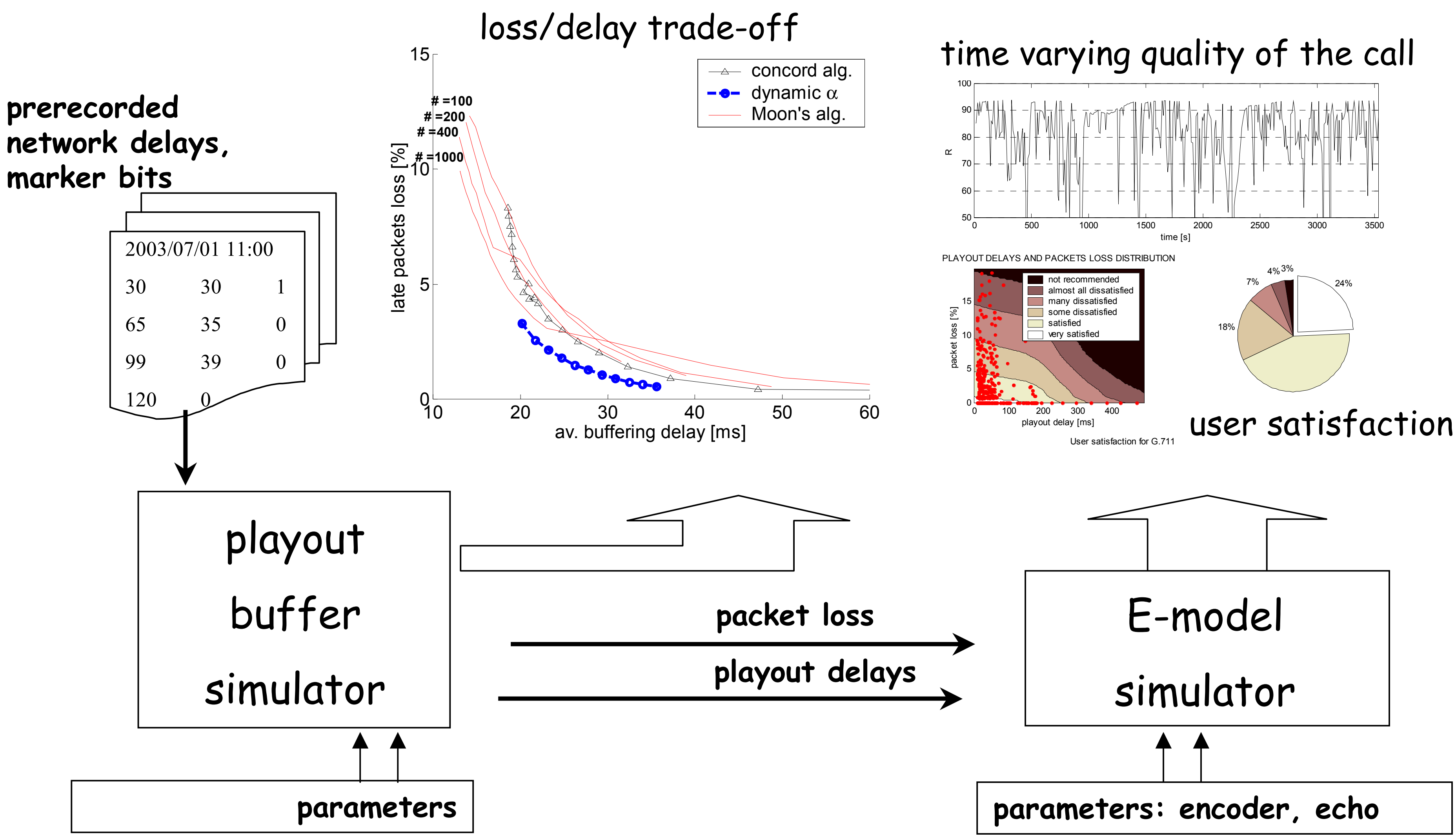
Predicting user satisfaction (our solution)



Predicting user satisfaction (our solution)

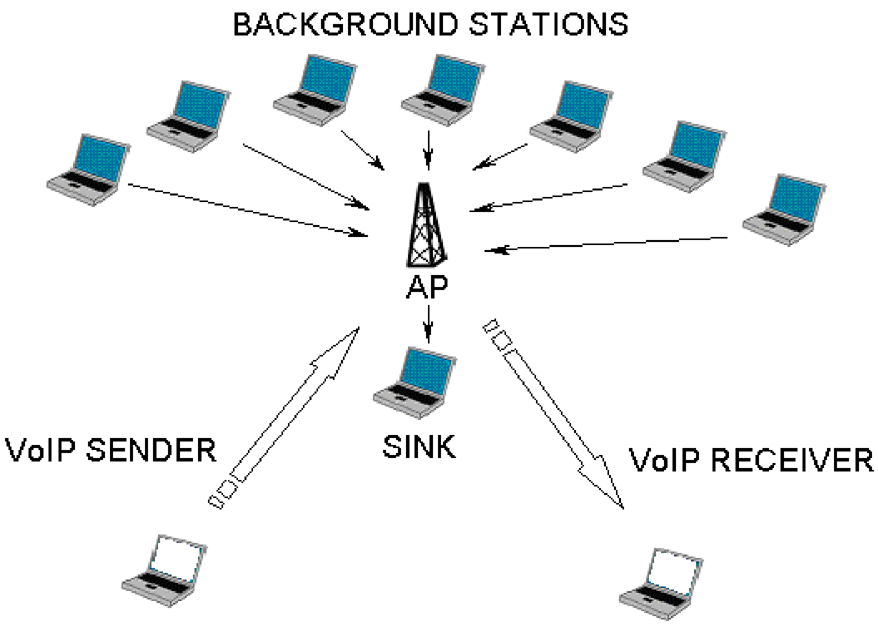


Algorithms evaluation – testing environment

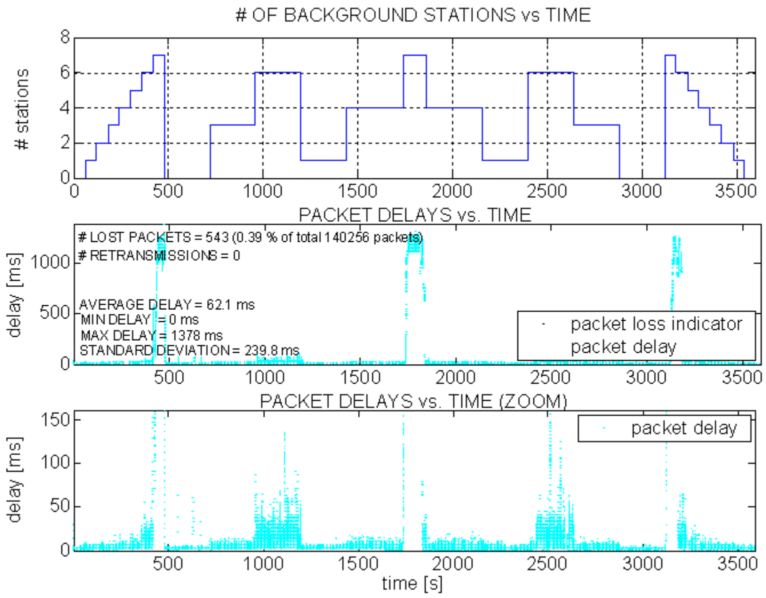


Assessing the quality of VoIP transmission affected by playout buffer scheme and encoding scheme

Experiment in the wireless LAN

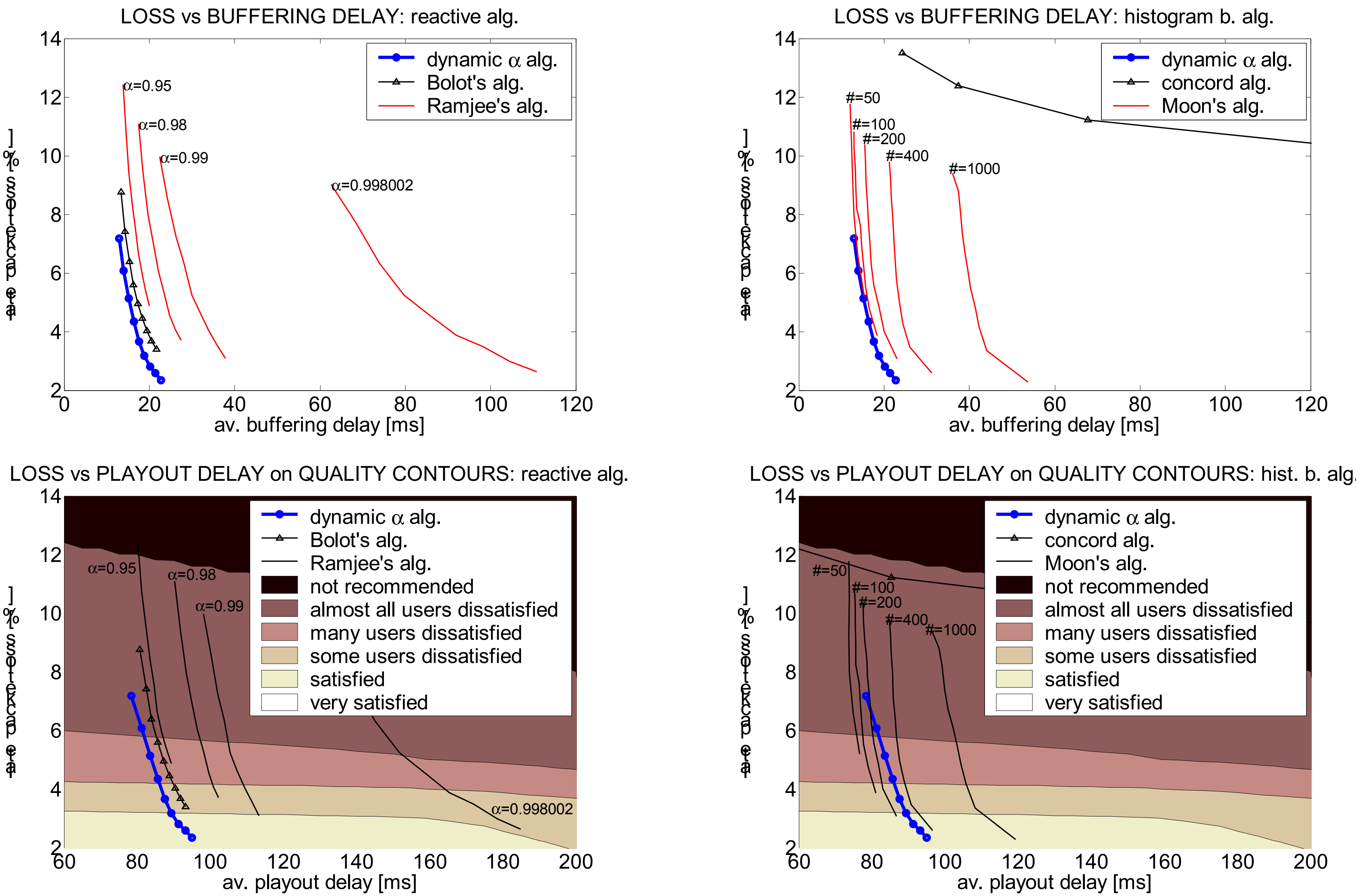


Measurement setup

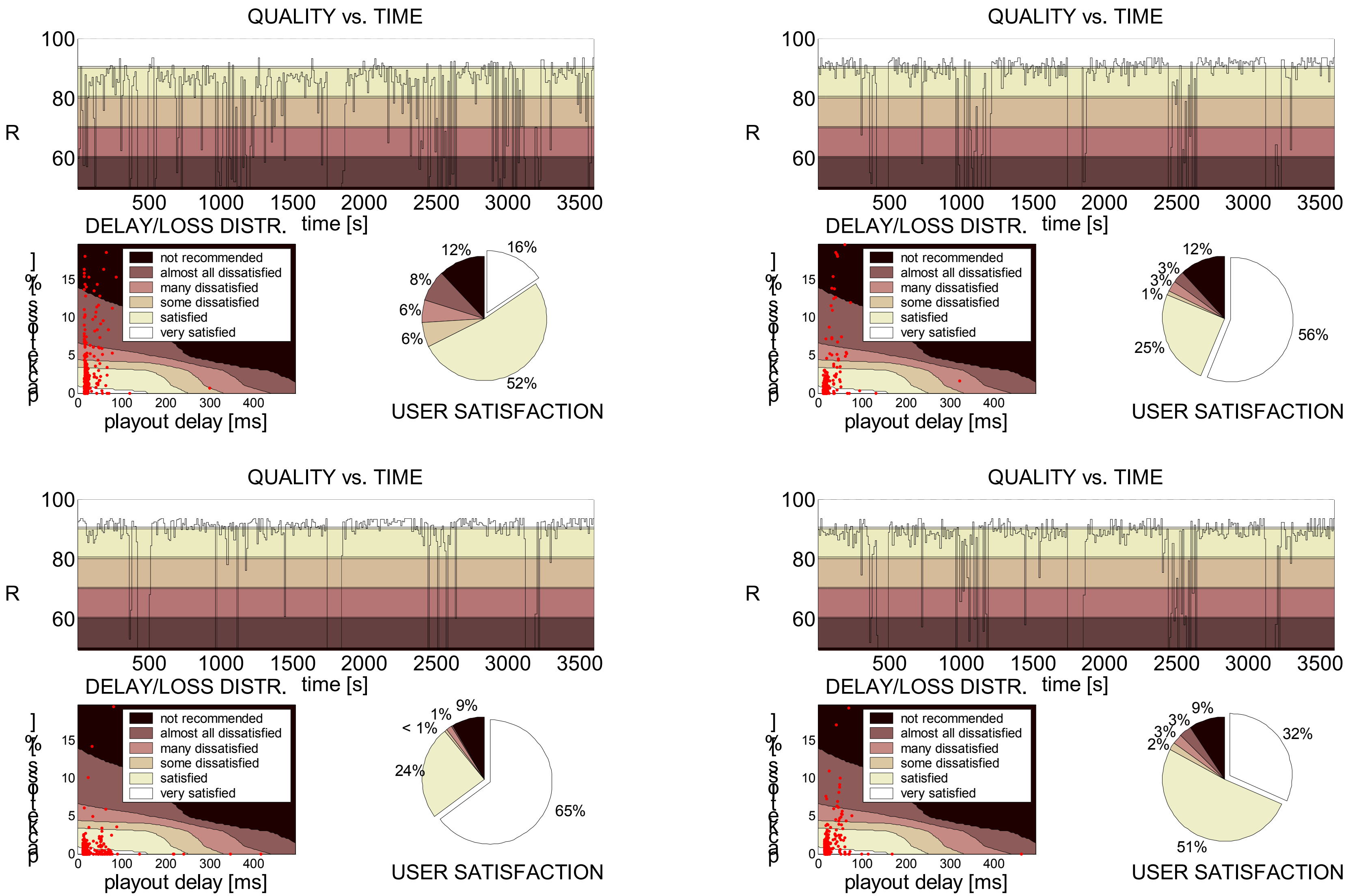


Influence of the background traffic on delay variation and jitter

Loss/delay trade-off (various algorithms)



Transmission quality vs. time and user satisfaction (Ramjee's: 0.95, Bolot's, "dynamic α", Moon's: #200/1%



Assessing the quality of VoIP transmission affected by playout buffer scheme and encoding scheme

		USER SATISFACTION CATEGORIES					
CODEC	**PLAYOUT MECHANISM**	not recommended [% time]	almost all dissatisfied [% time]	many dissatisfied [% time]	some dissatisfied [% time]	satisfied [% time]	very satisfied [% time]
G.711	Ramjee's alg. α=0.9980	11	1	2	3	42	40
	Ramjee's alg. α=0.9	10	3	6	11	60	11
	Concord alg.	89	0	1	3	6	1
	Moon's alg.	9	3	3	2	51	32
	Bolot's alg.	12	3	3	1	25	56
	dynamic α alg.	**9**	**0**	**0**	**2**	**24**	**65**
G.723.1	Ramjee's alg. α=0.9980	15	2	11	72	0	0
	Ramjee's alg. α=0.9	18	10	35	37	0	0
	Concord alg.	90	2	2	5	0	0
	Moon's alg.	10	2	8	80	0	0
	Bolot's alg.	14	4	8	80	0	0
	dynamic α	9	2	4	85	0	0
G.729A	Ramjee's alg. α=0.9980	12	3	4	45	36	0
	Ramjee's alg. α=0.9	15	6	19	50	10	0
	Concord alg.	89	1	3	6	1	0
	Moon's alg.	9	1	4	49	37	0
	Bolot's alg.	13	2	4	30	51	0
	dynamic α alg.	9	1	2	25	64	0

Summary

- We need a method to evaluate both playout buffer algorithms and encoding scheme in end-to-end VoIP system
- statistical loss/delay metrics (average delay/average loss) give little information on user satisfaction
- "listening-only" tests and PESQ don't take into account interactivity
- E-model relies on static transmission impairments that do not correspond to dynamics of de-jitter buffering
- We proposed extended version of the E-model (short-time version)
- It provides direct link to speech transmission quality by estimating user satisfaction. User satisfaction is estimated from varying transmission impairments (playout delay, resulting packet loss) and takes into account encoding scheme
- Pictorial representation of playout delays and packet loss gives more detailed view on the performance of a given playout algorithm
- The proposed method can work in real time and off-line on delay traces

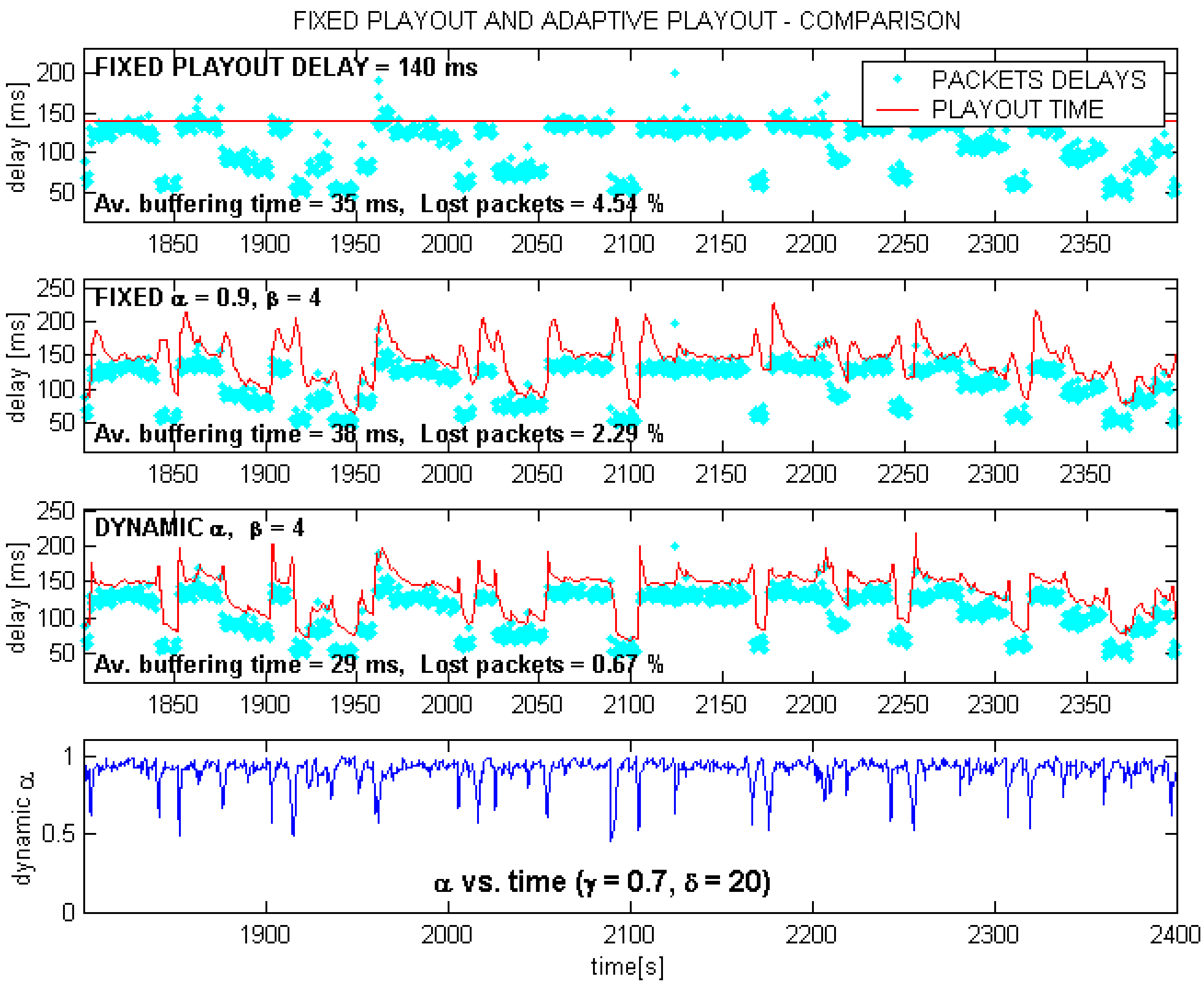
Conclusions

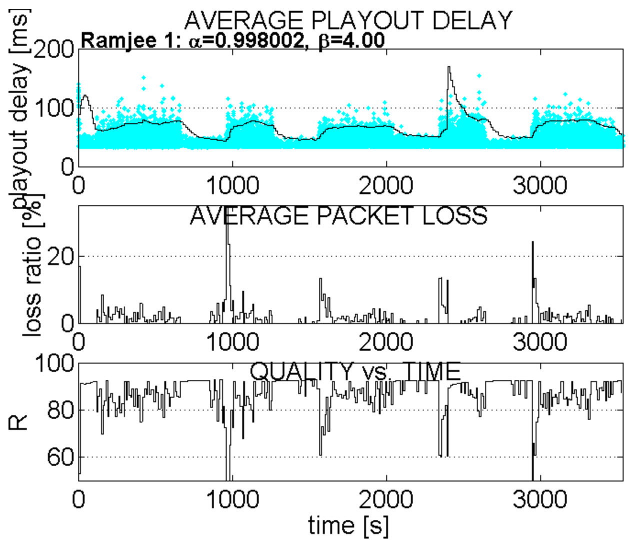
- We compared the performance of three audio codecs (ITU-T G.711, G.723.1, and G.729A) and various buffering schemes in a WLAN environment under varying load conditions using the extended version of the ITU-T E-model methodology.

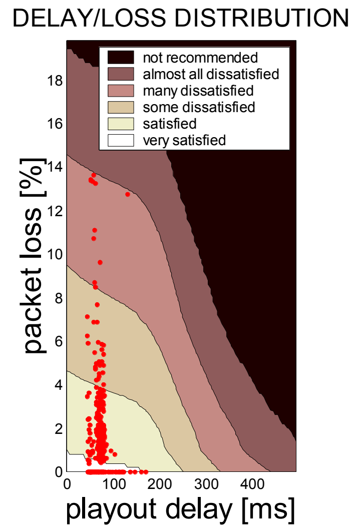
- Results show that the use of the G.711 audio codec in conjunction with "dynamic alpha" adaptive playout scheme gives the highest user satisfaction of the Voice over WLAN schemes considered.

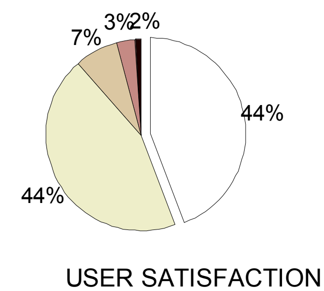
Thank you.

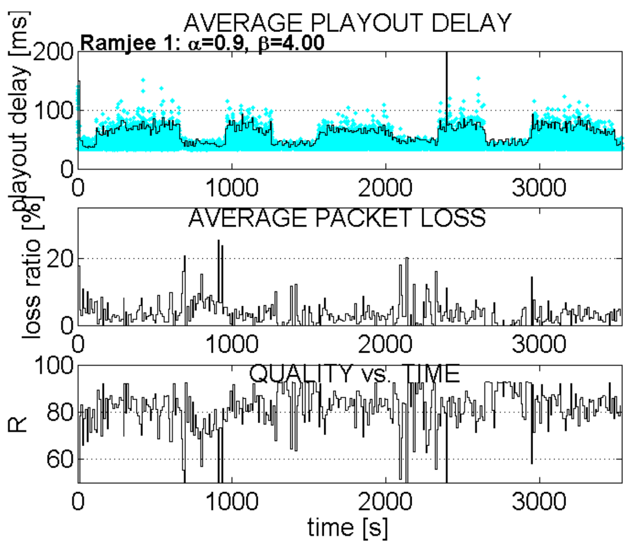
Performance comparison

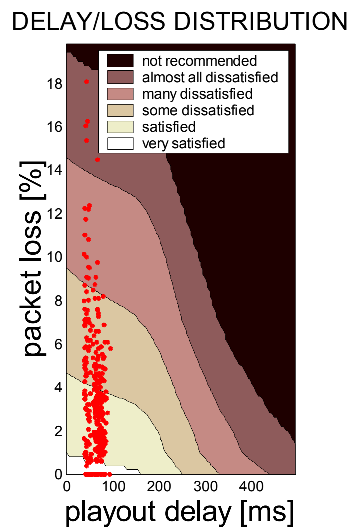


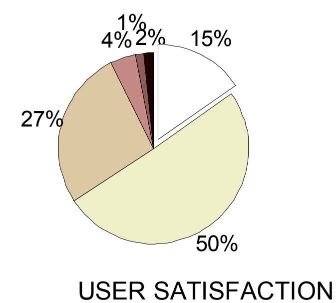












Assessing the quality of VoIP transmission affected by playout buffer scheme and encoding scheme

QoS Concerns and Challenges

Communication over best-effort networks ...

- *Delay* Impairs interactivity of conversational services
  Voice over IP: recommended one way delay < 150 ms
  [*ITU-T G.114*]

- *Packet loss* Impairs perceptual quality

- *Delay jitter* Obstructs sequential and continuous media output