Διαχείριση Edge και Cloud δικτύων βασισμένων στο λογισμικό

Ποιότητα Εμπειρίας

Δρ. Ειρήνη Λιώτου
eliotou@hua.gr

29/4/2025

INTRODUCTION TO QoS & QoE

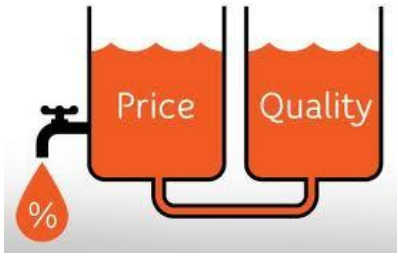
Some statistics

- Two “competing” entities:

Operator/Provider vs. Customer/User

↔

min(Cost) vs. max(Quality)



- Some facts:
  - **82%** of customer defections are due to frustration and the provider’s inability to deal with this effectively
  - For 1 person who calls with a problem, **29** never will
  - 1 frustrated customer will tell **13** others
  - **90%** abandons a service without even complaining

For which one would you rather pay?

Q: Which one represents QoS and which one QoE?

- "*Unlimited internet with speed up to 24Mbps*"

  → QoS (Quality of Service)

- "*Excellent user experience guaranteed*"

  → QoE (Quality of Experience)

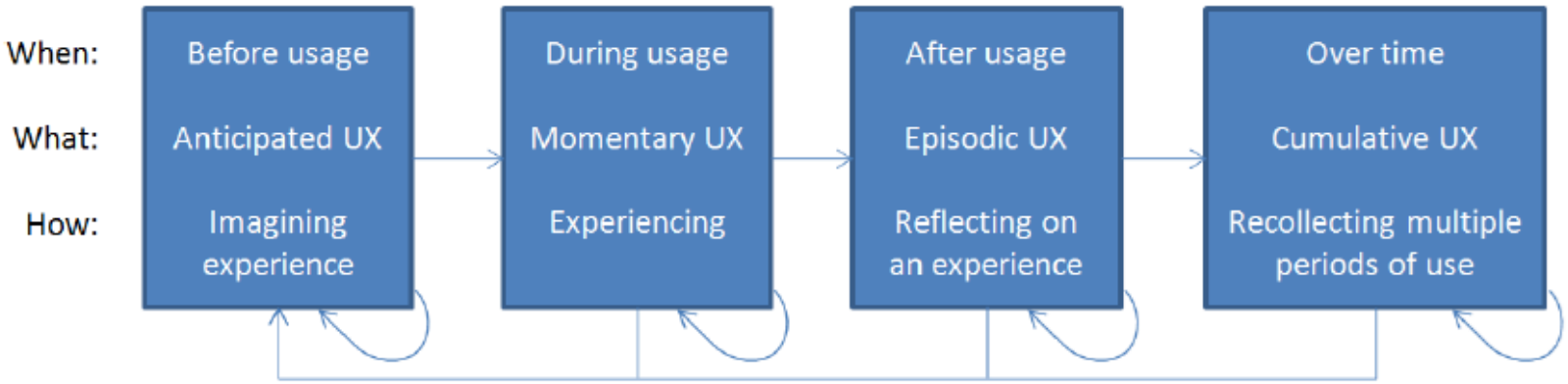
➢ **QoS** is "a set of technical quality requirements on the collective behaviours of one or more objects in order to define the required performance criteria". But:

- It handles pure technical aspects
- Same QoS values do not imply same customer experience
- QoS does not reflect the end-user satisfaction

QoE definition

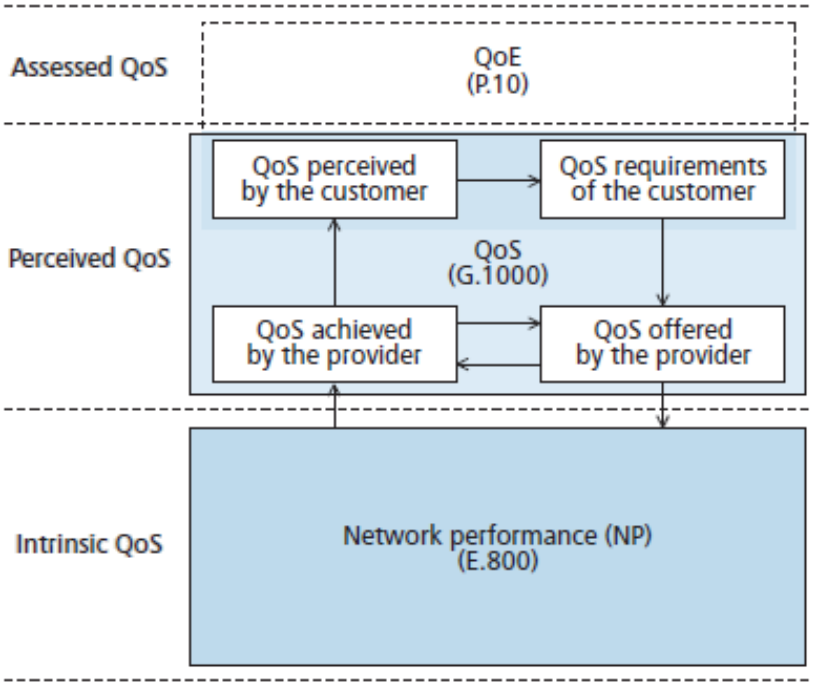
- ITU-T: “The overall acceptability of an application or service, as perceived subjectively by the end-user.”
- ETSI: “A measure of user performance based on both objective and subjective psychological measures of using an ICT service or product.”
- Practically: “The degree of your delight or annoyance over a product, application or service.” [Qualinet]

Time spans of UX



* “User Experience White Paper: Bringing clarity to the concept of user experience”, Dagstuhl Seminar

QoE: A multidisciplinary field



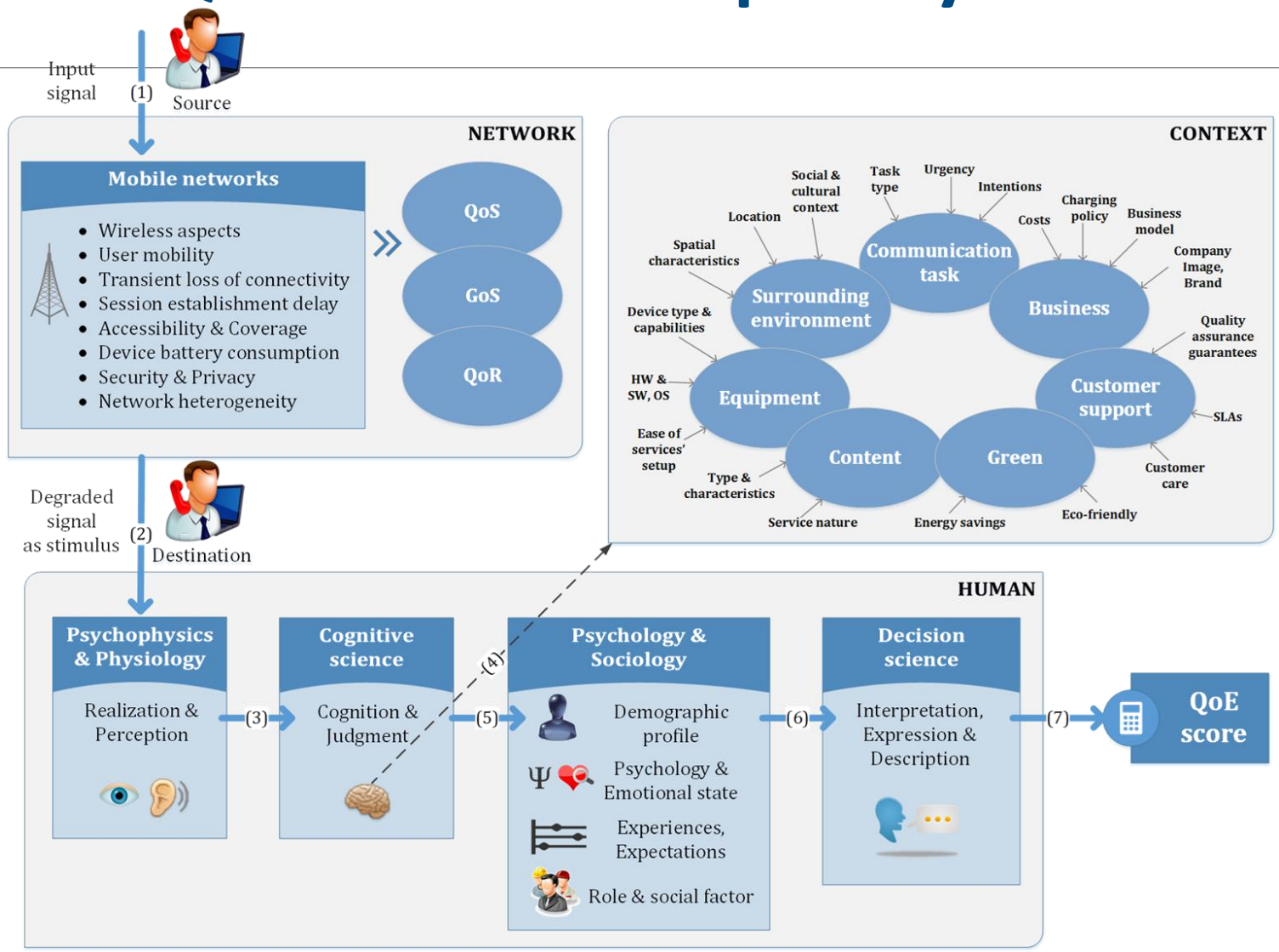
Main properties:

- User-dependent
- Application-dependent
- Terminal-dependent
- Time variant

- QoS: technology-centred
- QoE: user-centred

* R. Stankiewicz, P. Cholda, and A. Jajszczyk, "QoX: What is it really?," IEEE Communications Magazine, vol. 49, no. 4, pp. 148–158, Apr-2011.

QoE: A multidisciplinary field



1. Human

Age, gender, education level, cultural background, sociological and psychological factors, cognitive and perceptual abilities, user expectations, experiences, emotion, mood, perception, preferences

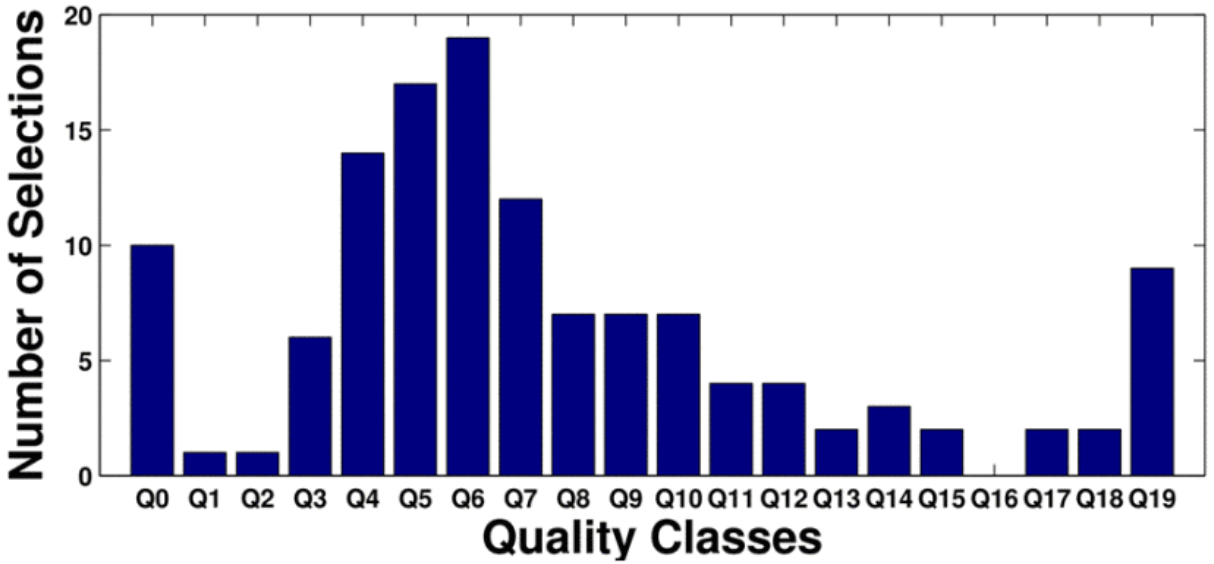
2. Network (Key Performance Indicators - KPIs)

Aspect	Quality Influence Factors	Aspect	Quality Influence Factors
Video specific	- Frame Rate - Video bit rate - Video content - Terminal type - Display size, type and resolution - Codec type and implementation - Video resolution and video format	**Transport / Network**	- Round trip / one-way delay - Jitter - Packet loss ratio - Delay burstiness distribution - Loss burstiness distribution - Bottleneck bandwidth - Congestion period
Video on Demand	- Number of stalling events - Duration of stalling events - Total video duration - Initial delay (start-up delay) - Time on highest layer (HTTP Adaptive Streaming - HAS) - Number of switches (HAS) - Altitude (HAS)	**Physical**	- SNR / SIR / SINR - Bit rate - BLER - Outage probability - Packet / Symbol / Bit Error probability - Outage capacity - Ergodic capacity / throughput - Diversity order / coding gain - Area spectral efficiency - Energy efficiency

3. Context

- Energy consumption
- Terminal type
- Human role
- Communication task, Urgency
- Customer support, ease of setup & use
- Charging policy & price
- Environment
- **Content**

Willingness to pay



Quality Class	Q0	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
VBR [kBit/s]	128	181	256	362	512	724	1024	1448	2048	2896
Priceplan A [€]	0	0.105	0.211	0.316	0.421	0.526	0.632	0.737	0.842	0.947

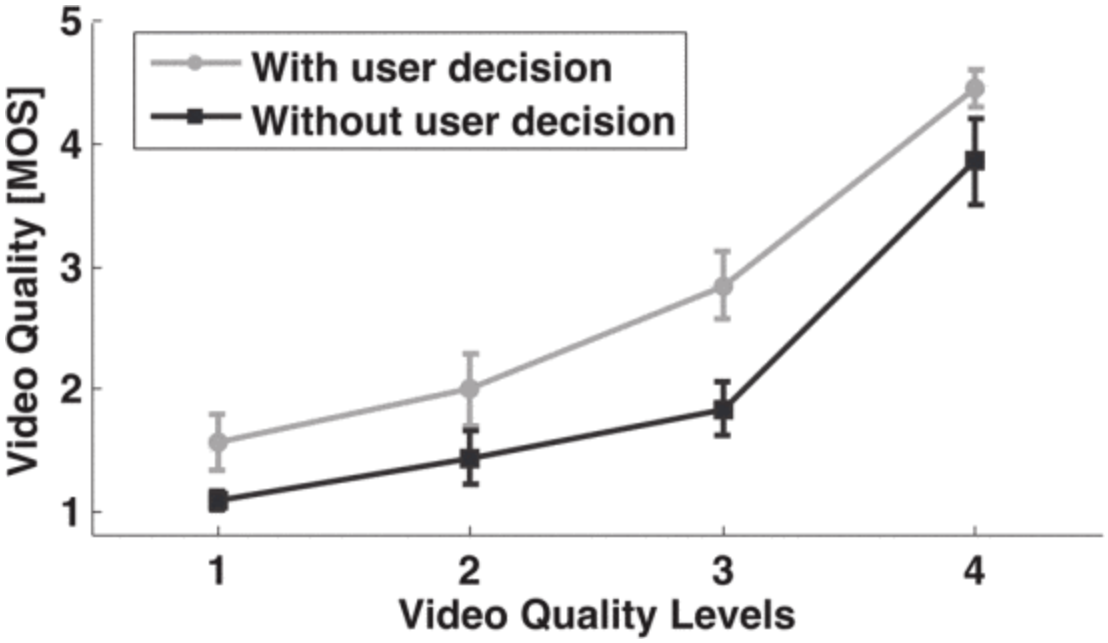
Quality Class	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19
VBR [kBit/s]	4096	5793	8192	11585	16384	23170	32768	32768	32768	32768
Priceplan A [€]	1.053	1.158	1.263	1.368	1.474	1.579	1.684	1.789	1.895	2

Q: What has influenced each user decision?

* A. Sackl, P. Zwickl , et al. "The trouble with choice: An empirical study to investigate the influence of charging strategies and content selection on QoE", IEEE CNSM, 2013.

Willingness to pay and QoE

- Users who decide to choose (and pay for) high quality multimedia services tend to evaluate this quality in a different way than if they are simply offered the same quality levels for consumption



Quality is evaluated more positively when preceded by a *monetary* decision

* A. Sackl, P. Zwickl , et al. "The role of cognitive dissonance for QoE evaluation of multimedia services", IEEE Globecom Workshops, 2012.

QoE intelligence potential

The absolute way to evaluate a service

Possibility to incorporate QoE intelligence in network decisions to better reflect user perception

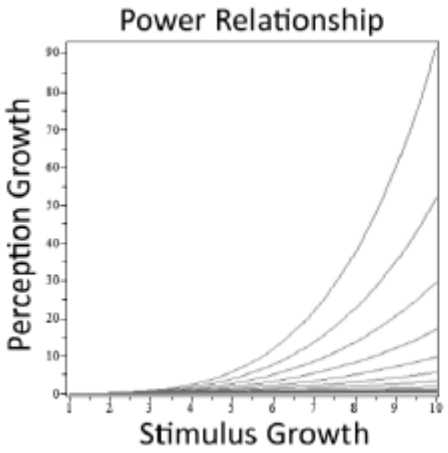
Propose innovative algorithms that focus on targeted QoE Key Performance Indicators (KPIs)

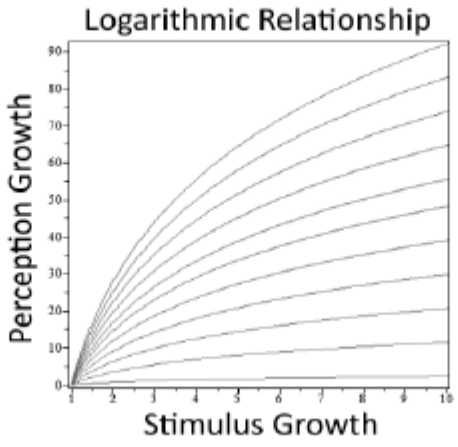
QoE-awareness may drive a more resource-efficient network operation

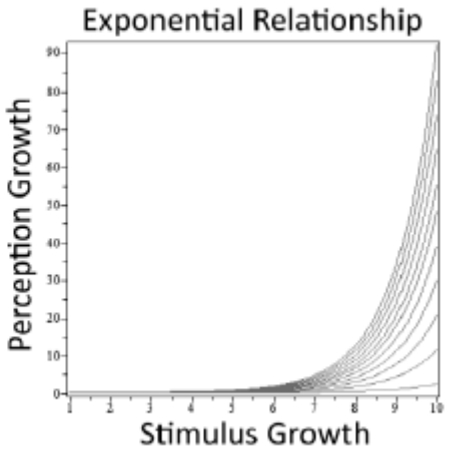
Improve Customer Experience Management, decrease churn, build meaningful Service/Experience Level Agreements

QoS - QoE relationship

	Name	Trend	Relation	Form
Adopted from Psychophysics	Stevens' Power Law	Stimulus-centric	$QoE = K.QoS^{b}$	Power
	Weber-Fechner Law	Stimulus-centric	$QoE = k.\ln(QoS)$	Logarithmic
Adopted from a Hypothesis	IQX	Perception-centric	$QoE = \alpha.e^{-\beta.QoS} + \gamma$	Exponential

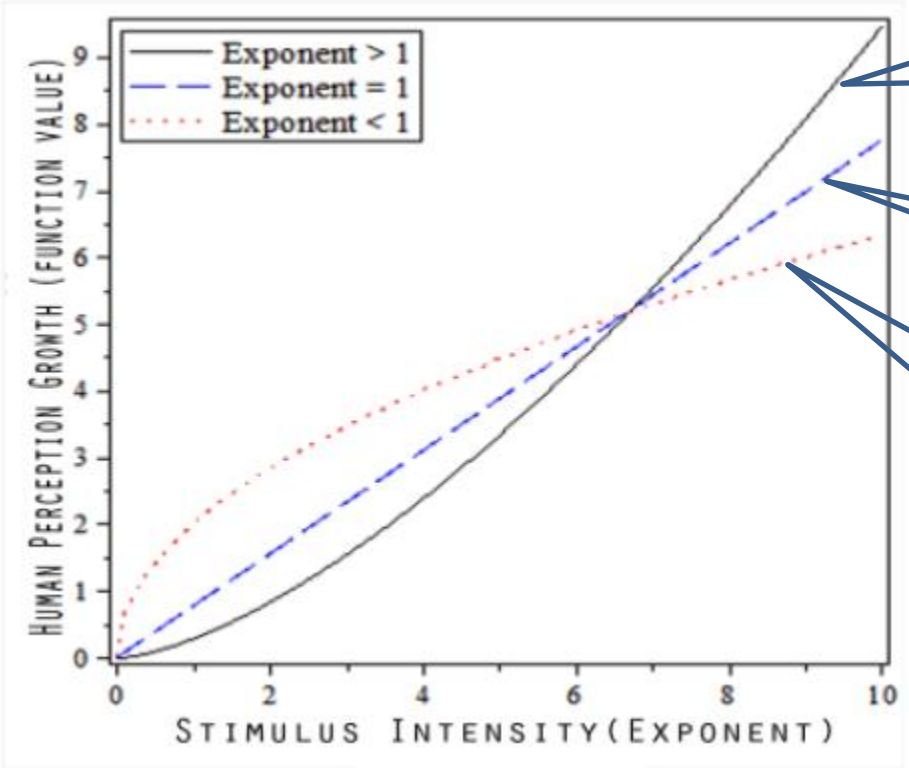






* S. Khorsandroo, et al, "A Generic Quantitative Relationship to Assess Interdependency of QoE and QoS", Ksii Transactions on Internet and Information Systems, 2013.

Steven's law



Human perception growth as a function of muscle force

How humans can perceive changes in visual length

Human perception as a function of smell

$$P(S) = K * S^b$$

Q: Which one represents muscle force, visual length and smell stimulus?

* S. Khorsandroo, et al, "A Generic Quantitative Relationship to Assess Interdependency of QoE and QoS", Ksii Transactions on Internet and Information Systems, 2013.

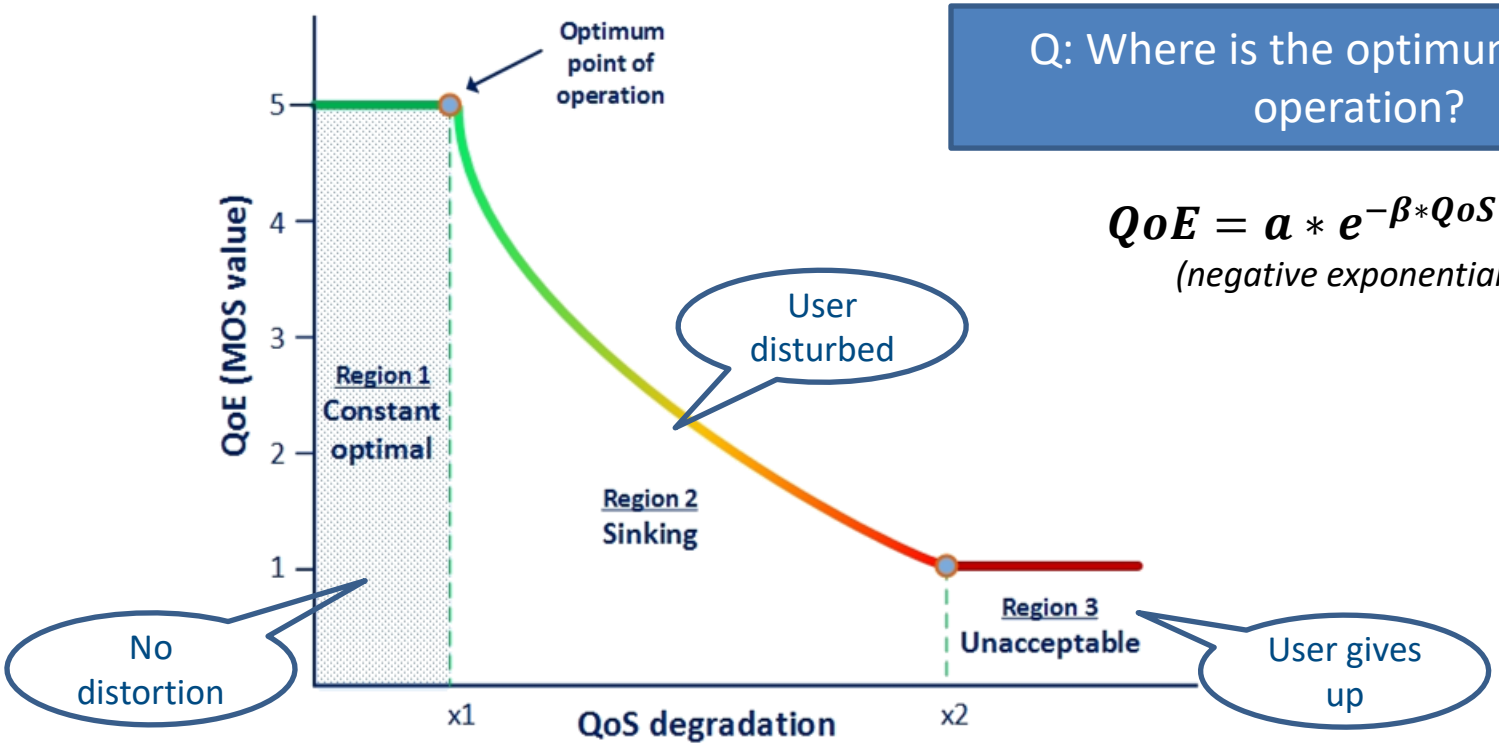
The IQX hypothesis

- The subjective sensitivity of QoE is more pronounced the higher the experienced quality is:
  - If QoE is very high, a small disturbance will strongly decrease it
  - If QoE is already low, a further disturbance will not be so perceived

Example: This relationship can be motivated considering a restaurant QoE: If we dined in a 5-star restaurant, a single spot on the clean white tablecloth would strongly disturb the atmosphere. The same incident would go unnoticed in a simple tavern…

The IQX hypothesis

- The change of QoE depends on its current level
- High QoE => small disturbances strong impact ≠ small QoE => unperceived

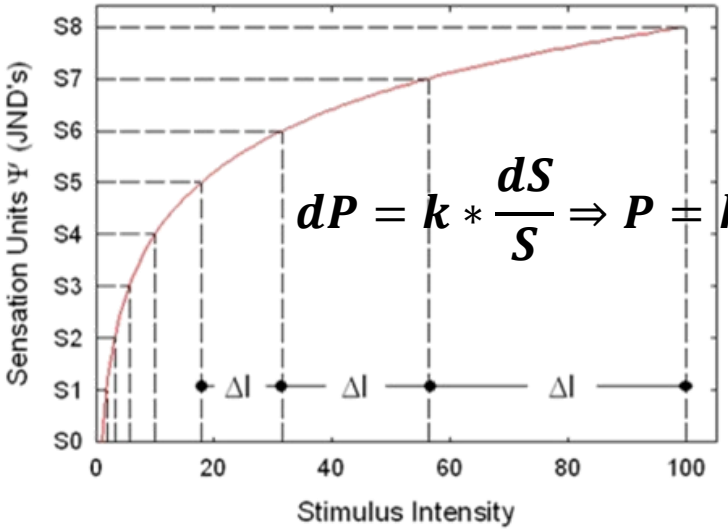


* M. Fiedler, T. Hossfeld, and P. Tran-Gia, "A generic quantitative relationship between quality of experience and quality of service," IEEE Network, vol. 24, no. 2, pp. 36–41, Mar-2010.

Weber Fechner Law

"Just noticeable differences" concept - jnd:

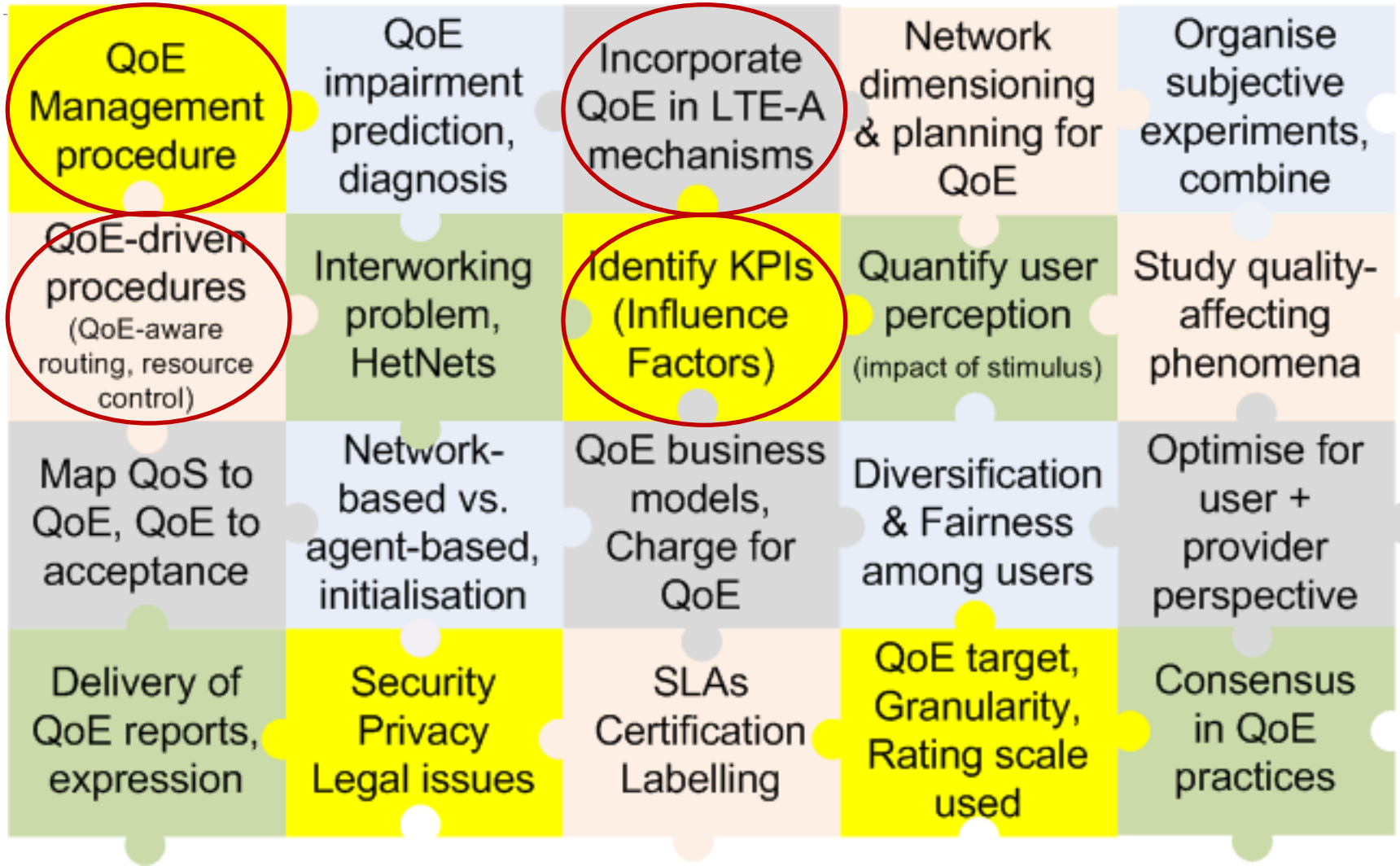
- Weight: 100gr distinguished from 105 gr, 200gr distinguished from 210gr => 5% is the "Weber fraction"
- Observed values need to change by at least some small but constant proportion of the current value to ensure humans will reliably detect it
- Brightness, loudness, numerical cognition, etc.



$$dP = k * \frac{dS}{S} \Rightarrow P = k * ln(\frac{S}{S_0})$$

- dP = differential change in perception
- dS = differential increase in the stimulus
- S = instantaneous stimulus
- S_0 = stimulus threshold
- k = constant, experimentally found

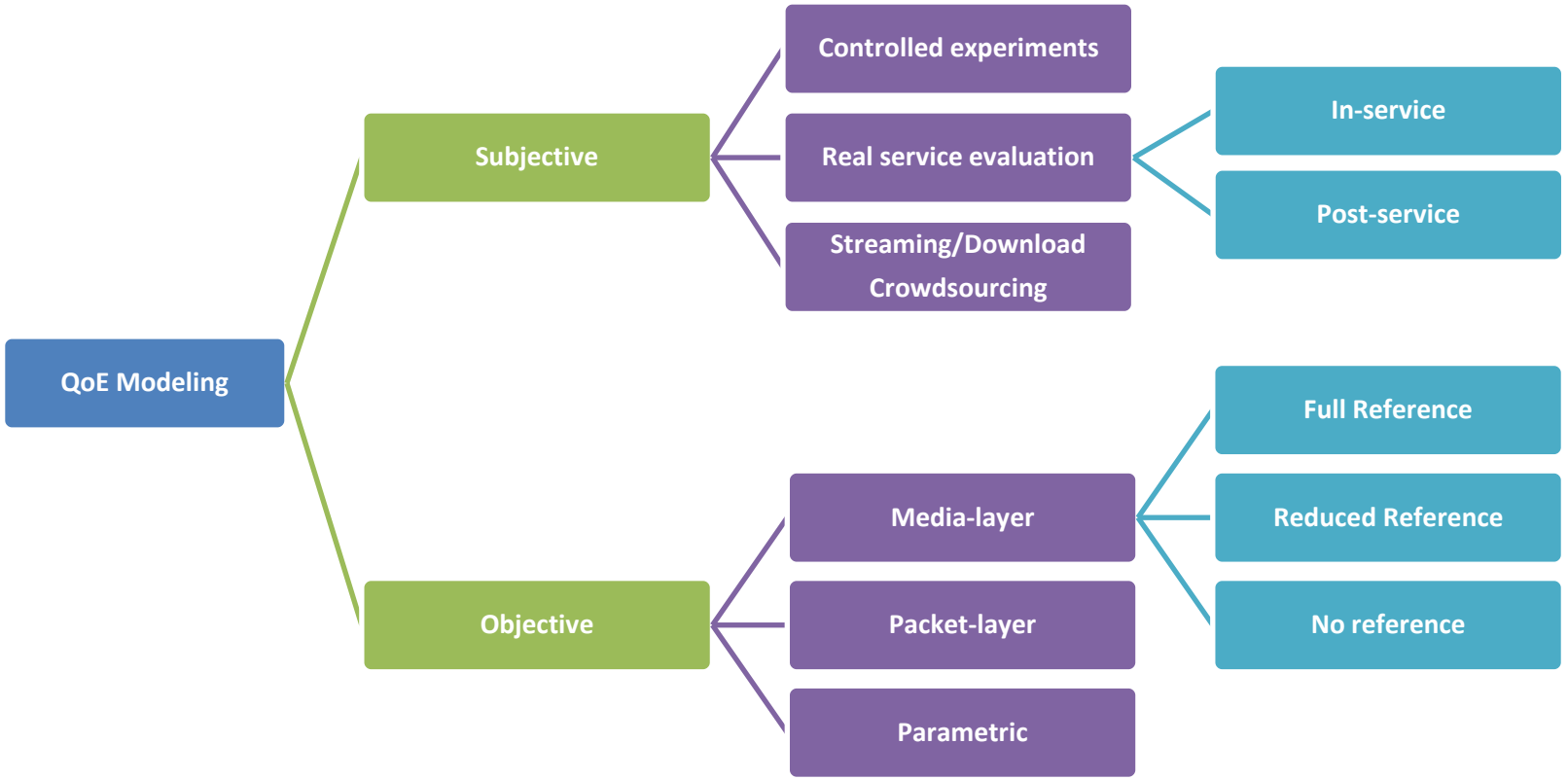
Research areas



QoE ESTIMATION

How can QoE be measured?

- The answer is via: QoE modeling!



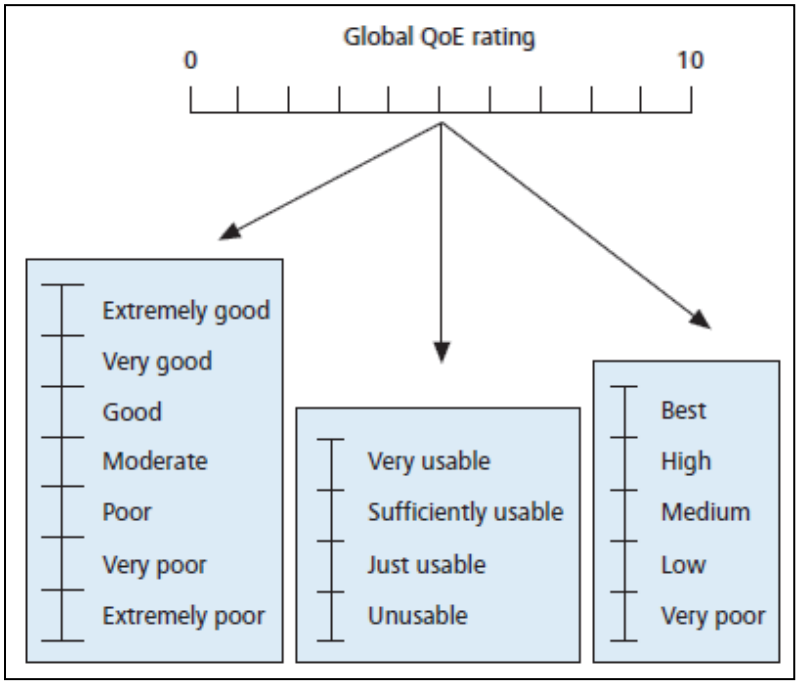
Comparison

Model	Advantages	Disadvantages	Restrictions
Subjective (controlled)	+ The most reliable QoE measurement model, highly accurate and valid + Ensures uniformity between subjective scores from different laboratories	- Not real-time (requires lab setting), not reproducible on demand - Time consuming and expensive - Needs thorough planning => complex - May be biased by user opinion, assumptions or unconscious psychological factors - Users may be greedy on their QoE demands and hence evaluations - Users' tiredness and lack/loss of concentration - Participants may just want to earn money and not be concise - Difficult for users to discriminate between e.g. "Bad" and "Poor" values in MOS scale	-> Experiments need to be conducted under strict requirements and controlled conditions: isolated sound room, dedicated equipment, suitably selected panel and number of participants, specific duration of signals, etc.
Objective (in general)	+ Automatically predict QoE + Same input always gives same output + Bypass the need for a human panel (the majority) + May be real-time, may be proactive	- Complexity - May not always highly correlate to reality - No universal generic quality model available, each one has a specific application scope - Need continuous validation against subjective data	-> Differ per application/service

Quality scales

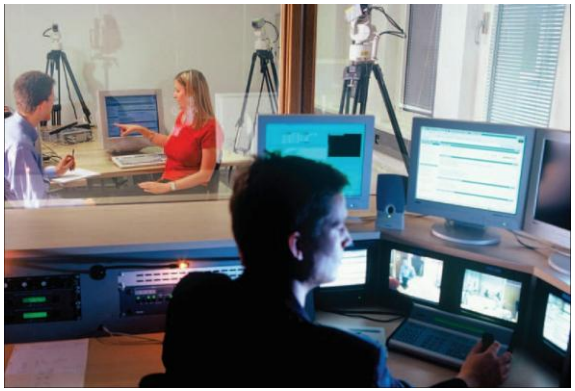
Absolute Mean Opinion Scores (**MOS**) / comparative

MOS	Quality	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

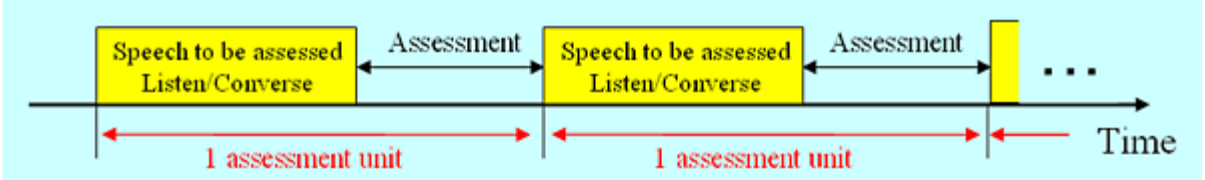


Score	Description
3	Much Better
2	Better
1	Slightly Better
0	About the Same
–1	Slightly Worse
–2	Worse
–3	Much Worse

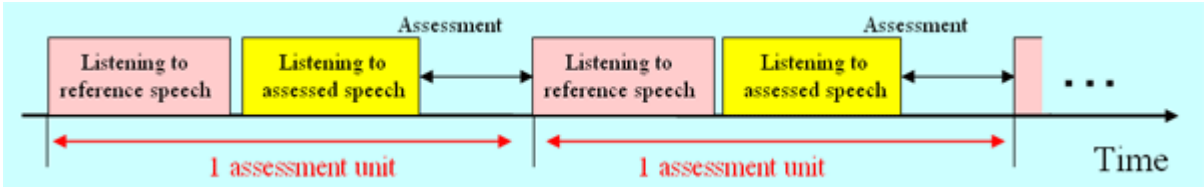
Subjective: controlled experiments



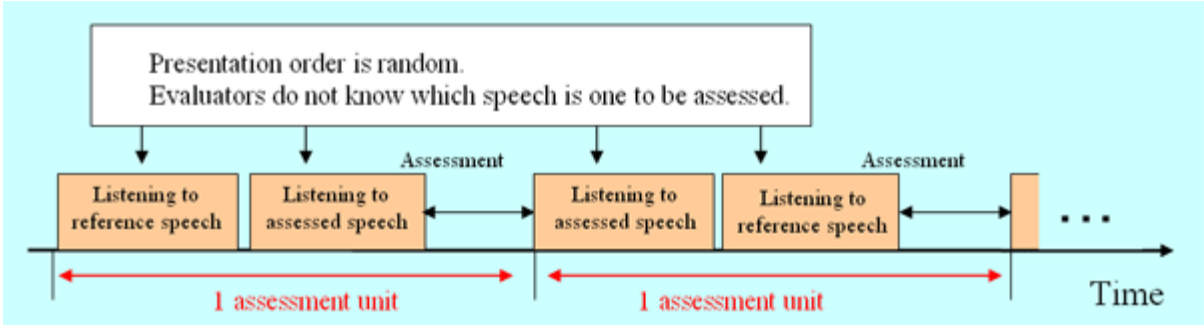
- MOS (Mean Opinion Score)



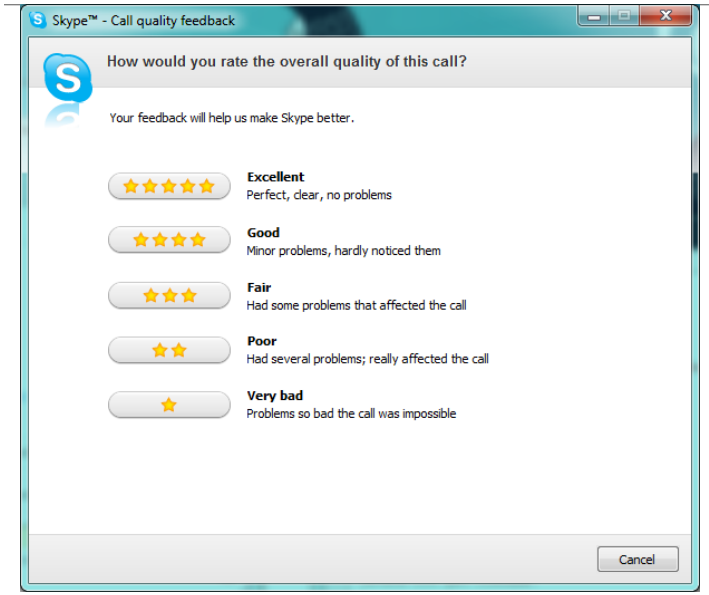
- DMOS (Degradation MOS)



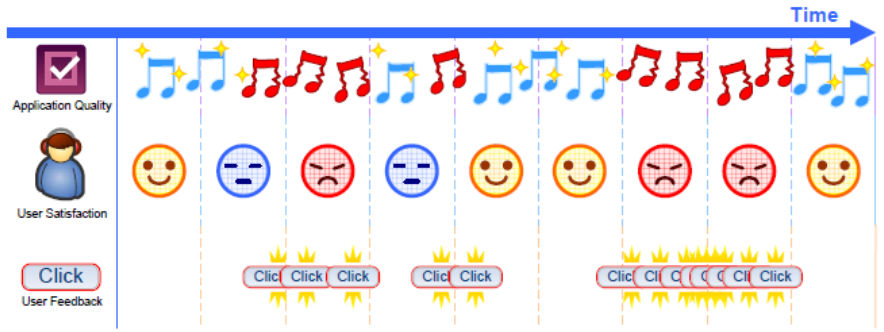
- CMOS (Comparison MOS)

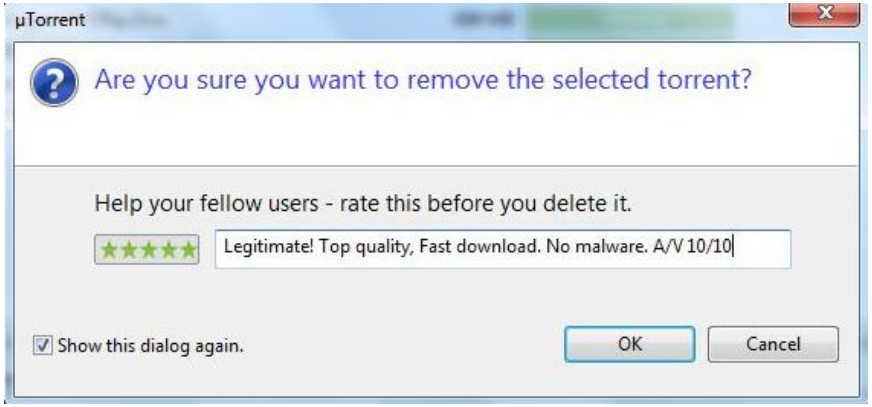


Subjective: real service evaluation



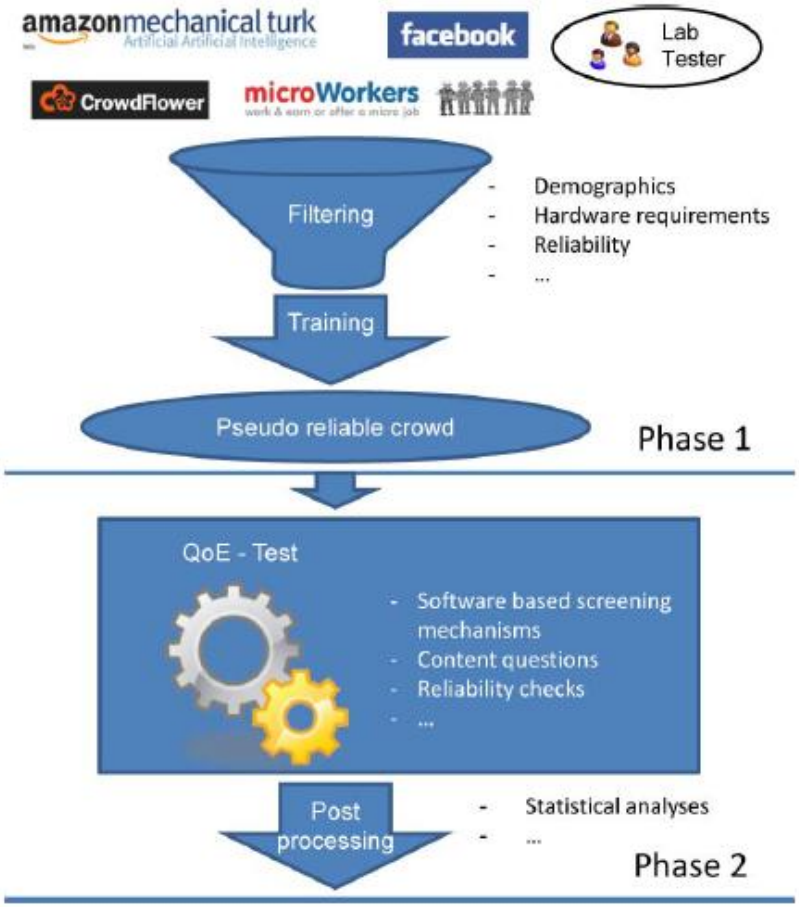
"OneClick"





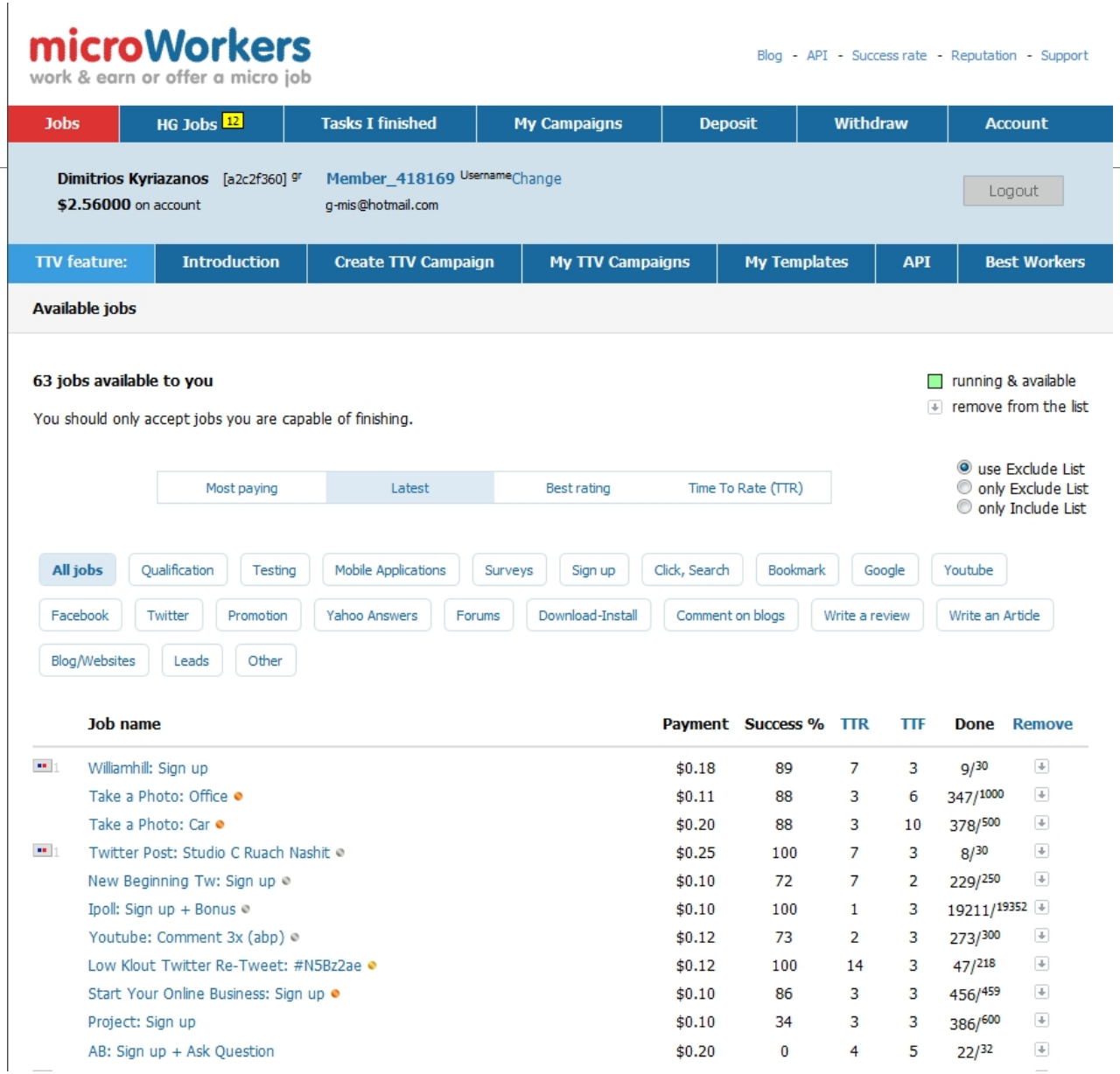
Subjective: Crowdsourcing

Q: What is "outsourcing"



Q: Have you ever used such a tool?

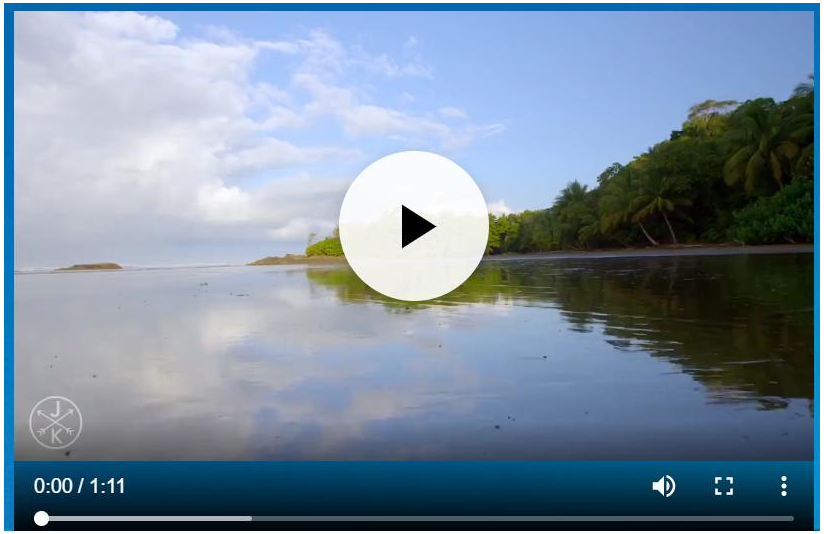
* T. Hossfeld, C. Keimel, M. Hirth, B. Gardlo, J. Habigt, K. Diepold, and P. Tran-Gia, "Best Practices for QoE Crowdtesting: QoE Assessment With Crowdsourcing," IEEE Trans. Multimed., vol. 16, no. 2, pp. 541–558, Feb. 2014.


microWorkers
work & earn or offer a micro job
Blog - API - Success rate - Reputation - Support
Jobs
HG Jobs 12
Tasks I finished
My Campaigns
Deposit
Withdraw
Account
Dimitrios Kyriazanos [a2c2f360] gr
$2.56000 on account
Member_418169 Username Change
g-mis@hotmail.com
Logout
TTV feature:
Introduction
Create TTV Campaign
My TTV Campaigns
My Templates
API
Best Workers
Available jobs
63 jobs available to you
You should only accept jobs you are capable of finishing.
running & available
remove from the list
Most paying
Latest
Best rating
Time To Rate (TTR)
use Exclude List
only Exclude List
only Include List
All jobs
Qualification
Testing
Mobile Applications
Surveys
Sign up
Click, Search
Bookmark
Google
Youtube
Facebook
Twitter
Promotion
Yahoo Answers
Forums
Download-Install
Comment on blogs
Write a review
Write an Article
Blog/Websites
Leads
Other
Job name
Payment
Success %
TTR
TTF
Done
Remove
Williamhill: Sign up $0.18 89 7 3 9/30
Take a Photo: Office $0.11 88 3 6 347/1000
Take a Photo: Car $0.20 88 3 10 378/500
Twitter Post: Studio C Ruach Nashit $0.25 100 7 3 8/30
New Beginning Tw: Sign up $0.10 72 7 2 229/250
Ipoll: Sign up + Bonus $0.10 100 1 3 19211/19352
Youtube: Comment 3x (abp) $0.12 73 2 3 273/300
Low Klout Twitter Re-Tweet: #N5Bz2ae $0.12 100 14 3 47/218
Start Your Online Business: Sign up $0.10 86 3 3 456/459
Project: Sign up $0.10 34 3 3 386/600
AB: Sign up + Ask Question $0.20 0 4 5 22/32

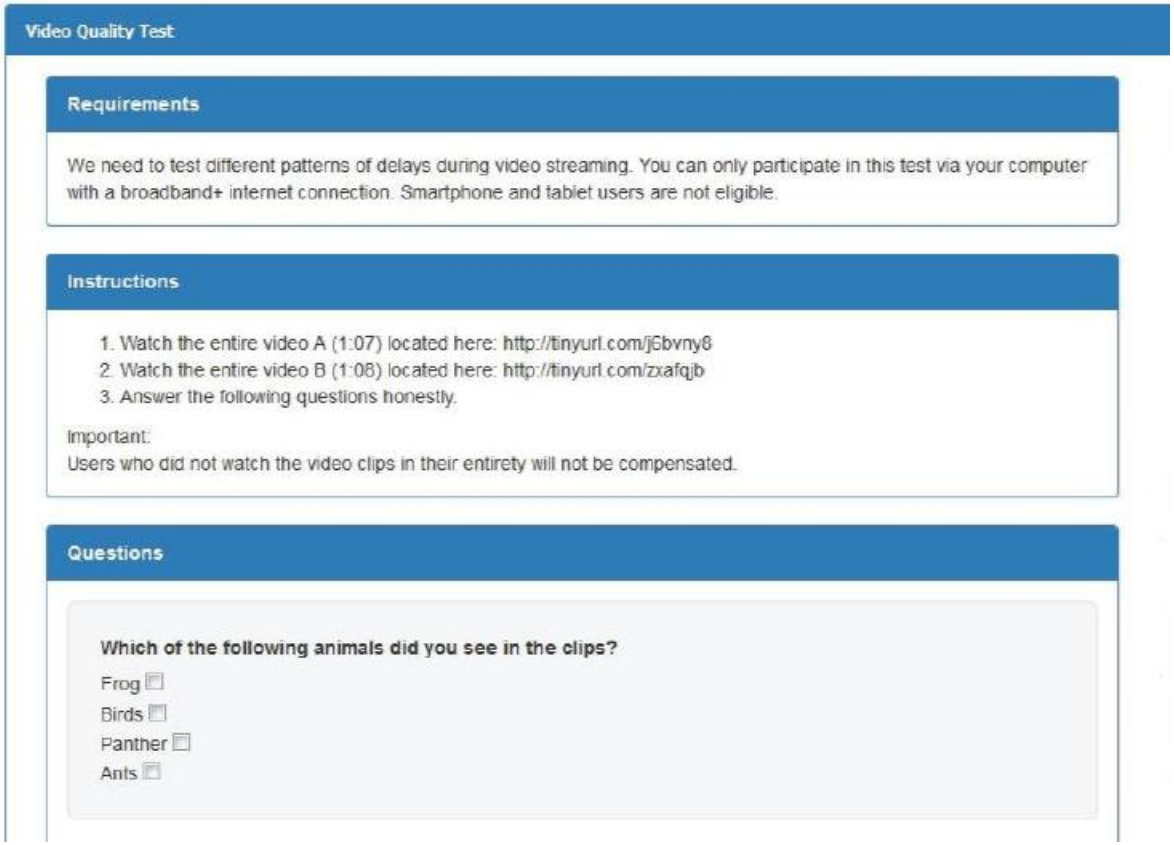
Crowdsourcing example 1 (BSc thesis)

- QoE videos\s2b\s2b.mp4
- QoE videos\s1c\s1c.mp4

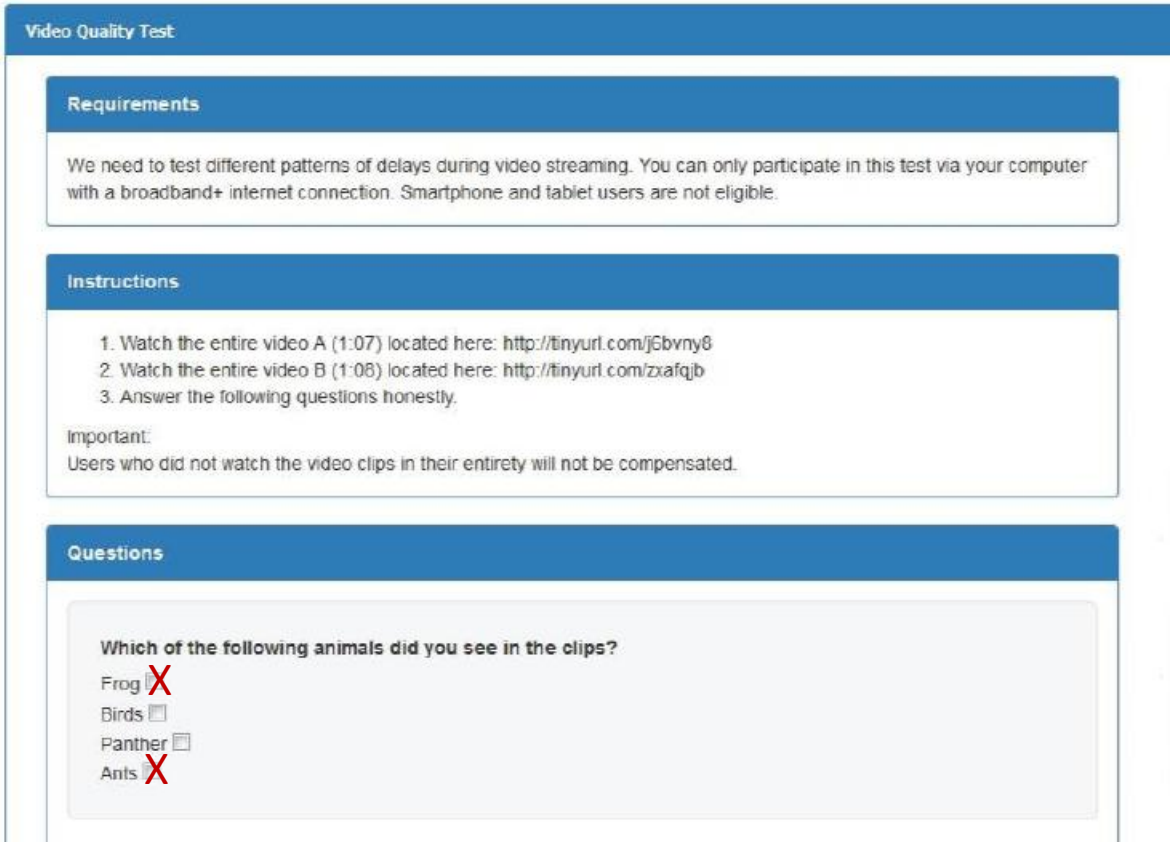
Hands on



Crowdsourcing example 1



Crowdsourcing example 1



Crowdsourcing example 1

How many stalling events (frozen video) did you notice in video A?

0
1
2
3
4
5

How many stalling events (frozen video) did you notice in video B?

0
1
2
3
4
5

Would you say that your viewing experience was better for A, B, or about the same?

A
B
About the same

How would you rate the quality of the video stream in A?

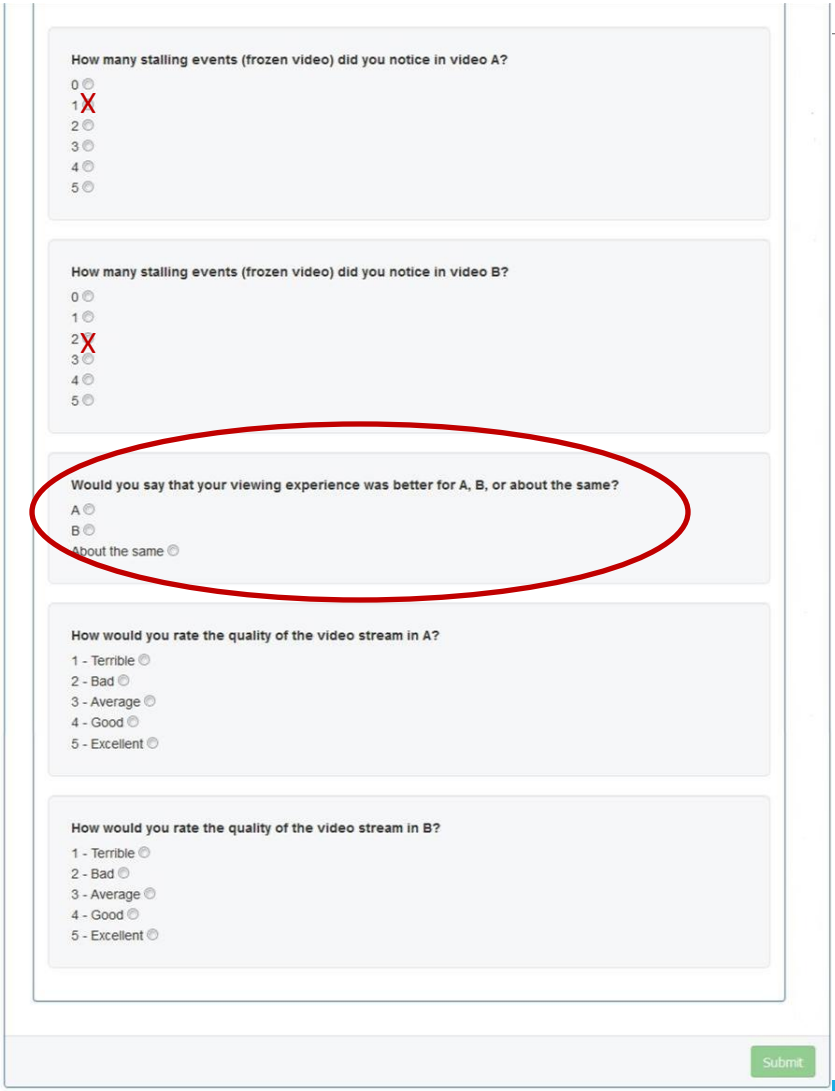
1 - Terrible
2 - Bad
3 - Average
4 - Good
5 - Excellent

How would you rate the quality of the video stream in B?

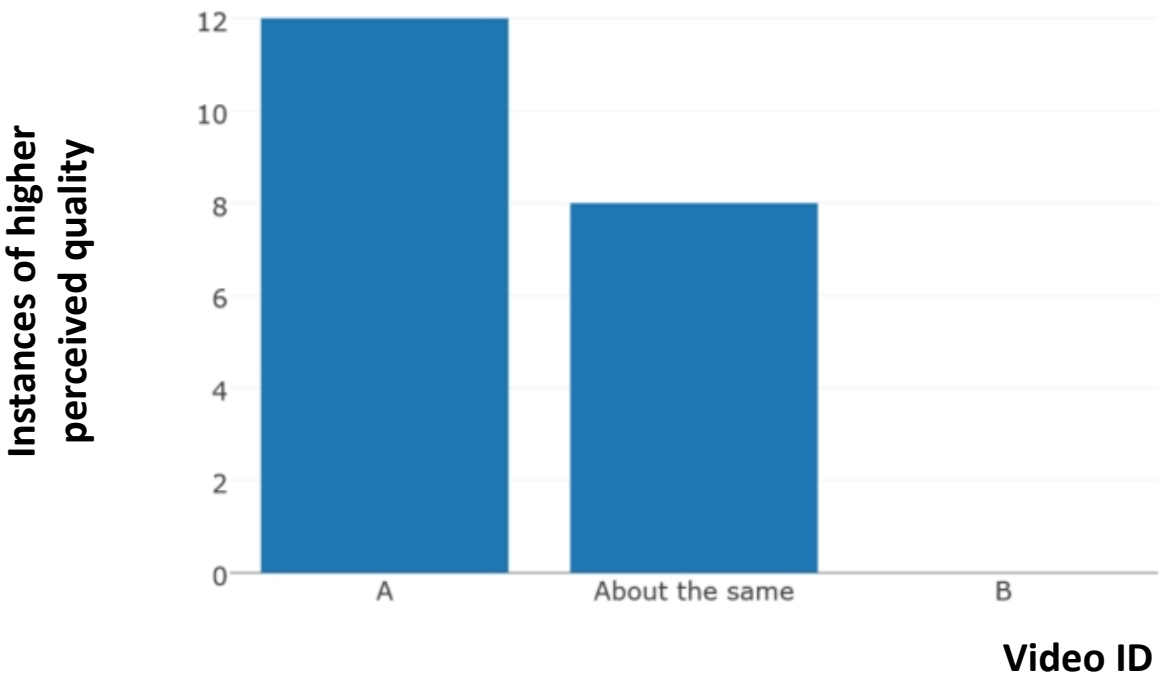
1 - Terrible
2 - Bad
3 - Average
4 - Good
5 - Excellent

Submit

Crowdsourcing example 1



Crowdsourcing example 1



1 stalling of 6s vs 2 stallings of 3s each

Crowdsourcing example 2

- QoE videos\357\357.mp4
- QoE videos\753\753.mp4

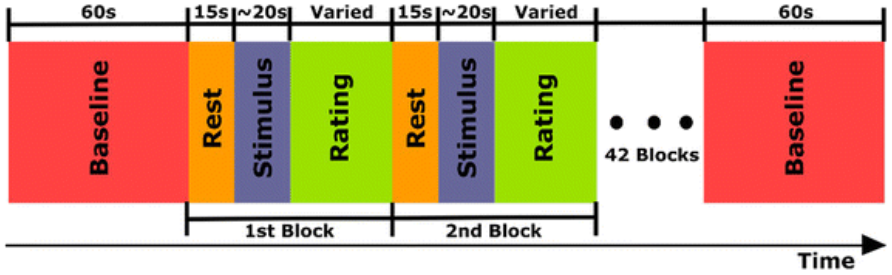
The purpose of this test is to assess whether users are more negatively influenced by their first impression of the clip or by the residual effect that the last part of the clip left them.

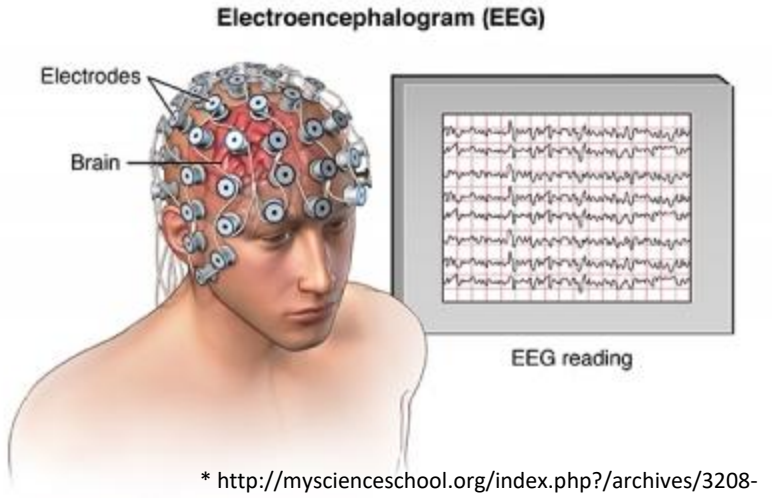
Crowdsourcing example 2

- http://gain.di.uoa.gr/kyr/357/qoe.html
- http://gain.di.uoa.gr/kyr/753/qoe.html

Electroencephalography (EEG)

- Human influence factors (HIFs) characterize the user's perception, emotional and mental state with respect to a service
  - Facial expressions, body posture, voice, eye measurement, electrocardiography (ECG), electrodermal activity (EDA)
  - EEG measures electrical activity in the brain





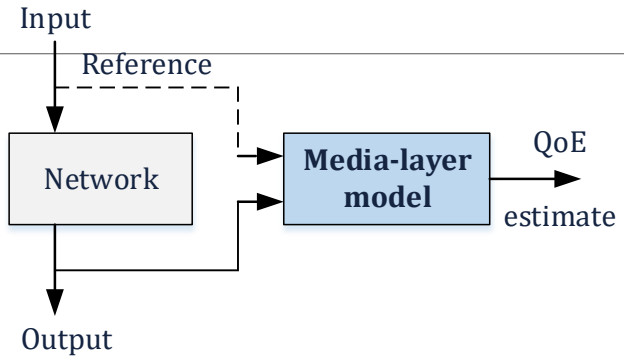
* http://myscienceschool.org/index.php?/archives/3208-What-is-Electroencephalography-EEG.html

* R. Gupta, K. Laghari, H. Banville, T. H. Falk, "Using affective brain-computer interfaces to characterize human influential factors for speech quality-of-experience perception modelling", Human-centric Computing and Information Sciences, 2016.

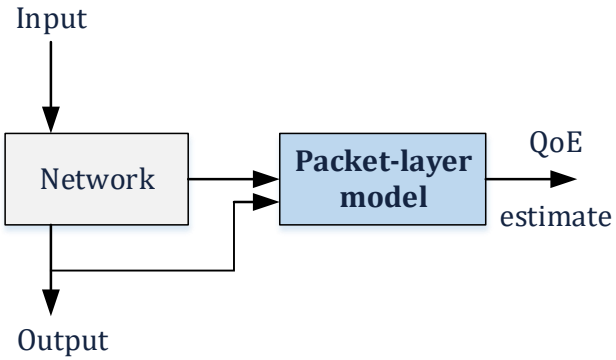
Objective: evaluation methods

- **Media-layer**

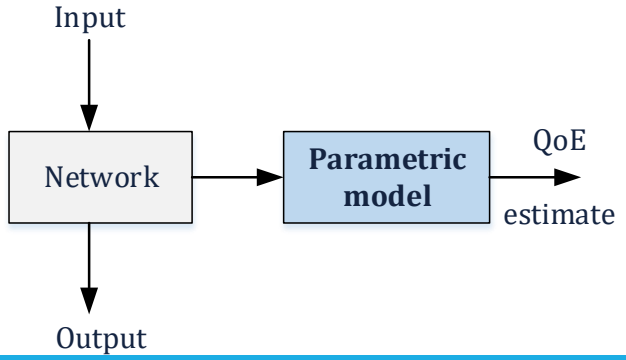
Q: Why is this difficult/impossible to implement in a real-time network?



- **Packet-layer / Bitstream**



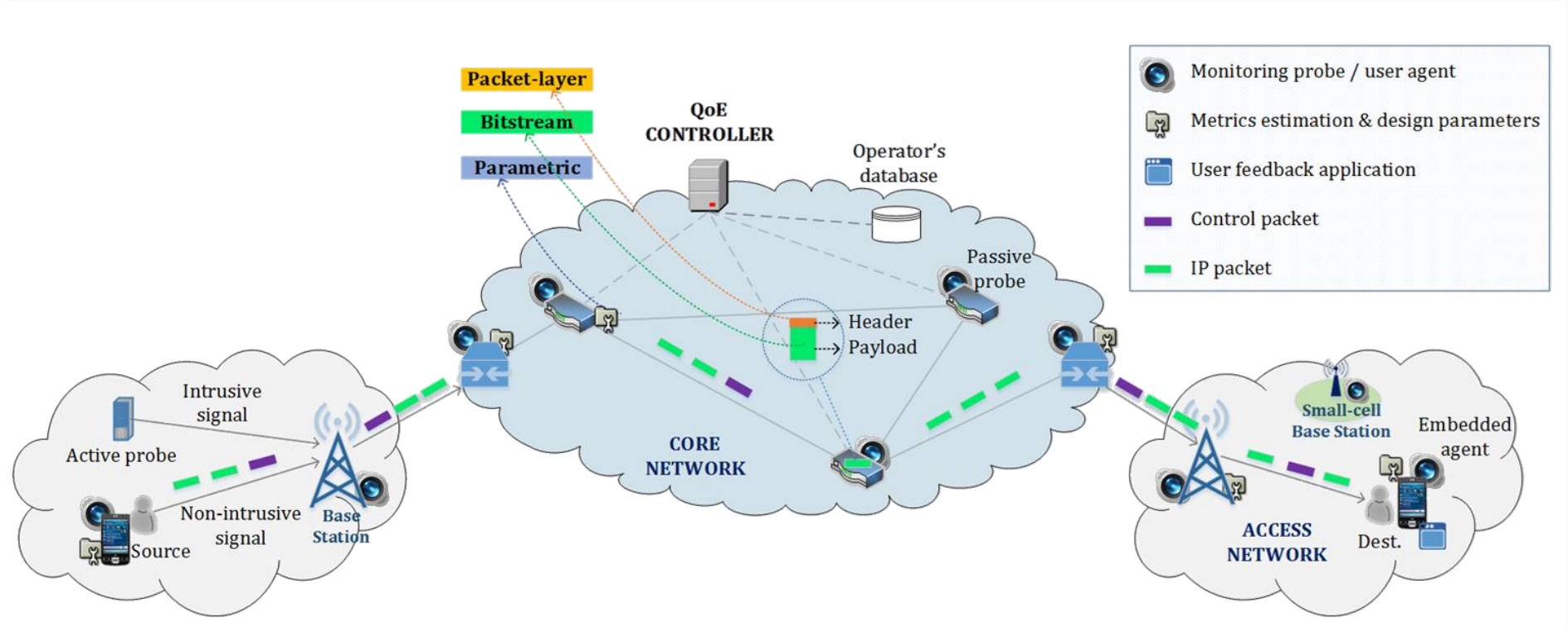
- **Parametric planning**



Objective: evaluation methods

Model	Advantages	Disadvantages
Media-layer: **Full Reference** **(e.g. PESQ)**	+ Do not require any a-priori knowledge or assumptions about the underlying network + Highly accurate and robust (based on psycho-acoustics)	- Require the reference signal (intrusive) - Very high computational effort - Practically impossible to implement at network midpoint - Do not enable insight into the internal system functionality & degradation causes (black-box) => diagnosis not possible - Neglect human dimensions, pure technical
Parametric planning: **E-model**	+ Ease of use and respect of privacy + The network is characterized by the technical specifications of its constituent elements, (non-intrusive approach) + Quantifies the human factor through the “Advantage factor”, & contextual factor + Mouth-to-ear complete transmission chain => conversational + No restrictions on the network with respect to size, configuration, hierarchy, technology used, nor on the components of the network	- Intended only for the planning phase of a system (extended format) - Good in theory, but difficult to include all the model parameters online - Accurate only under strict application scenarios: new subjective tests and regression analysis needed for different conditions - Speech independent - A-priori information requirement
Packet-layer: **ITU-T P.564**	+ Enables insight into the internal system functionality (glass-box) + Light in terms of computational effort + Multiple monitoring points help identify the root of a network impairment + Used not only for speech quality predictions but also for the production of diagnostic outputs + In-service, non-intrusive (privacy) + Quality followed and pooled over time	- Not standardized, models need to be created that comply with these recommendations - The model doesn’t know the characteristics of speech content to evaluate (speech level, echo, background noise etc.): assumes a generic voice payload - Only concerns impairments on the IP network (no end-to-end evaluation) - Large volume of QoE data - Models deployed require strict conformance testing

How can QoE KPIs be monitored?



* E. Liotou, D. Tsolkas, N. Passas and L. Merakos, "Quality of Experience in Mobile Cellular Networks: Modeling, Provisioning and Key Challenges", IEEE Communications Magazine , Network & Service Management Series, July 2015.

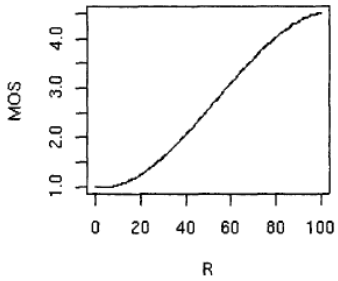
Examples of parametric models

- **VoIP:** $= 94.2 - [0.024d + 0.11(d - 177.3)H(d - 177.3)] - [11 + 40\,ln(1 + 10p)]$
  - d: delay
  - p: packet loss rate
- **YouTube (TCP):** $QoE = 3.5 * e^{-(0.15L + 0.19)*N} + 1.5$
  - L: duration of stalls
  - N: #of stalls
- **HTTP Adaptive Streaming (TCP):** $QoE = 0.003 * e^{0.064*t} + 2.498$
  - t: time on highest quality level
- **Real-time video (UDP):** $V_q = 1 + I_{coding} * I_{transmission}$
  - FR, BR, PLR
- **FTP:** $QoE = \alpha\, log_{10}(\beta R)$, $10kbps < R < 300kbps$
  - R: data rate

A. ITU-T G.107 “E-model” for voice

- Computes the transmission quality of VoIP by estimating the mouth-to-ear conversational quality as perceived by the receiver
- A parametric model that produces the so-called Rating factor R:

$$R = R_0 - I_s - I_d - I_{e-eff} + A$$



- R_o → basic signal-to-noise ratio, $R_o = 100$
- I_s → impairments due to the voice signal travelling in the network
- I_d → impairments caused by delay from end-to-end travelling signal
- $I_{e\text{-}eff}$ → equipment impairment factor & impairments due to packet loss
- **A** → advantage/expectation factor, in exchange for some user benefits or other factors difficult to quantify

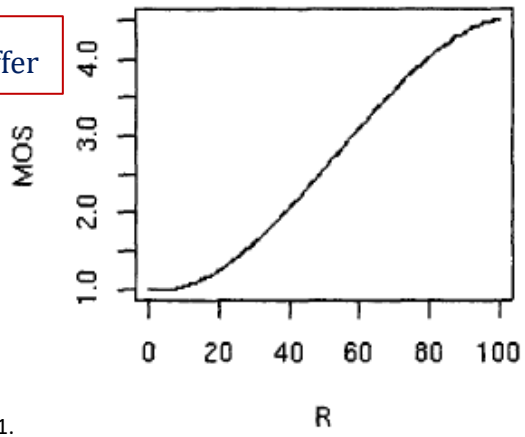
E-model: simplified version

$$R = R_0 - I_s - I_d - I_{e-eff} + A$$

- Under specific assumptions, the model may be simplified:
  - I_s → default values, A → neglected
    $\Rightarrow R = 94.2 - I_d - I_{e-eff}$
  - $I_d = 0.024\mathbf{d} + 0.11(\mathbf{d} - 177.3)H(\mathbf{d} - 177.3)$ → G.107
  - $I_{e-eff} = 11 + 40\ln(1 + 10\boldsymbol{e})$ → G.113
  - G.729a codec
  - more…

delay = $d_{network} + d_{codec} + d_{de\text{-}jitter\_buffer}$

packet loss = $e_{network} + e_{de\text{-}jitter\_buffer}$



- Then, R [0..100] is mapped to MOS [0..5]

➢ Purpose: monitoring the conversational voice quality

➢ Delay & Packet loss are isolated

* R. G. Cole, J. H. Rosenbluth, "Voice over IP performance monitoring," ACM SIGCOMM Comput. Commun. Rev., vol. 31, no. 2, p. 9, 2001.

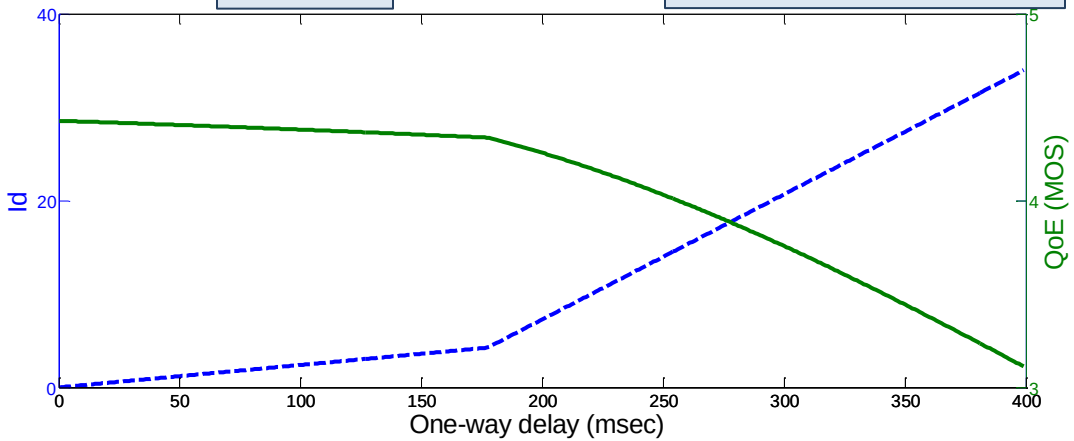
E-model: simplified version

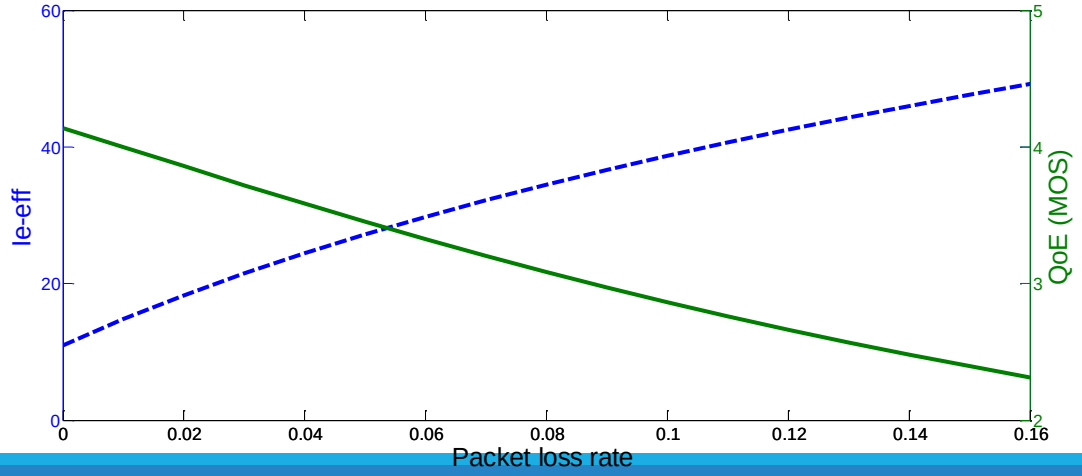
$$R = R_0 - I_s - I_d - I_{e-eff} + A$$

$$R = 94.2 - 0.024d - 0.11(d - 177.3)H(d - 177.3) - 11 - 40\,ln(1 + 10p)$$

delay

packet loss rate





B. ITU-T G.1070 “E-model” for video

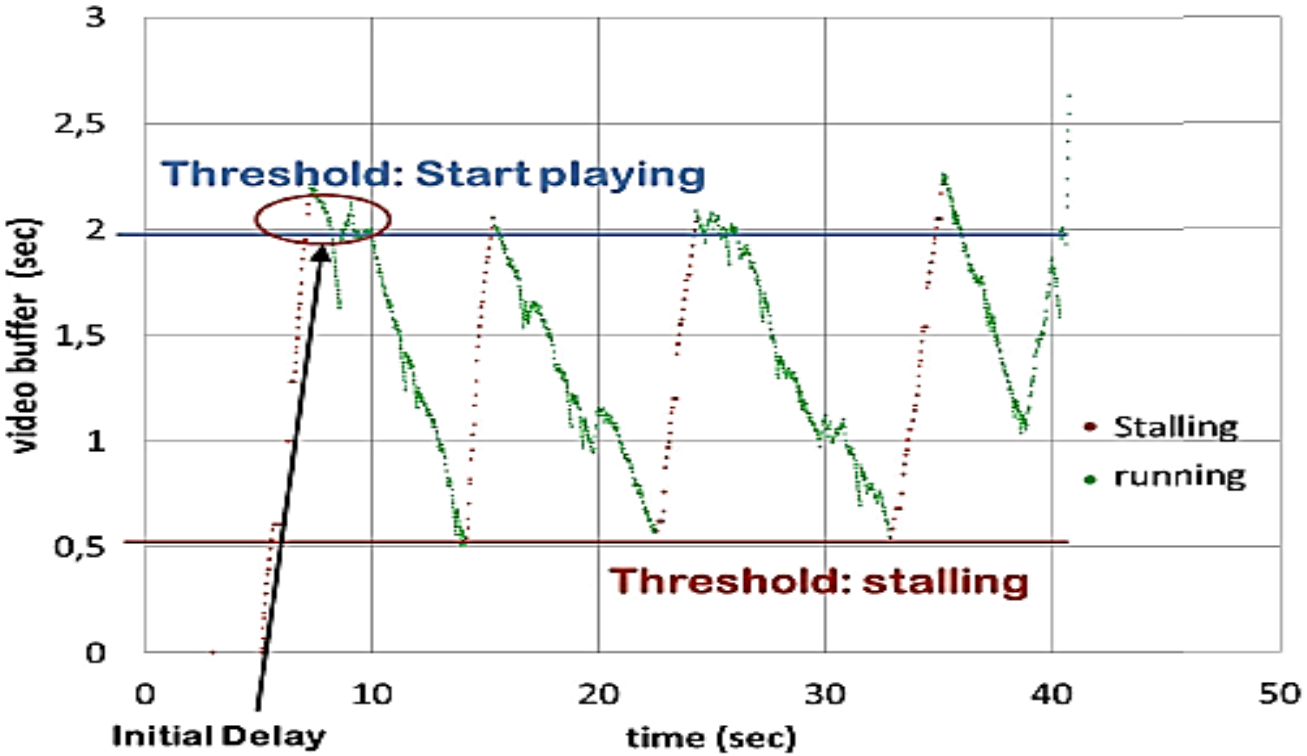
- A computational model for point-to-point interactive videophone applications over IP networks (UDP-based - lossy video)
- Network, Application & Terminal parameters incorporated
- Video quality =

$$V_q = 1 + I_{coding} * I_{transmission}$$

- I_{coding} = the video quality affected by the coding distortion
- $I_{transmission}$ = the video quality affected by the transmission process
- Ultimately everything is a function of:
  - the **video frame rate** (fps) - FR
  - the **video bit rate** (kbps) - BR
  - the **video packet loss rate** - PLR
  - 12 coefficients

C. QoE for YouTube

Buffering concept



Q: Why does a stalling happen?

QoE for YouTube

- Video on Demand (VoD), TCP-based connection (no losses)
- Quality influence factors (by crowdsourcing & lab tests):
  - Number of stalling events, ***N***
  - Duration of stalling events, ***L***
  - Total video duration, ***T*** (total stalling duration over video duration)
  - Initial delay (video start-up delay) → cache redirections' impact

QoE for YouTube

- Some conclusions:
  - The user demographics have no significant influence (!)
  - Initial delays have almost no influence on MOS for videos of duration 60s and 30s compared to the influence of stalling length

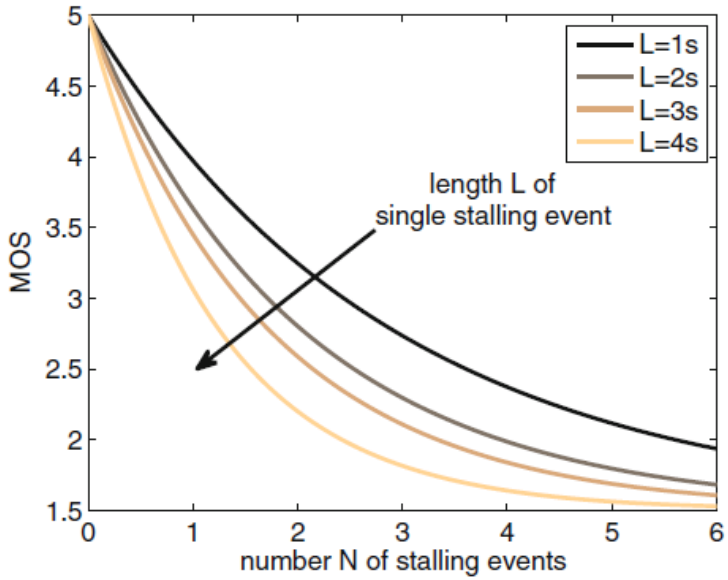
    Q: Is this your impression too?
  - The user ratings are statistically independent from video motion, type of content, the usage pattern of the user, access speed, etc.
  - The number of stalling events together with the stalling length are clearly dominating the user perceived quality
  - The video duration only plays a role if there are only a very few stalling events

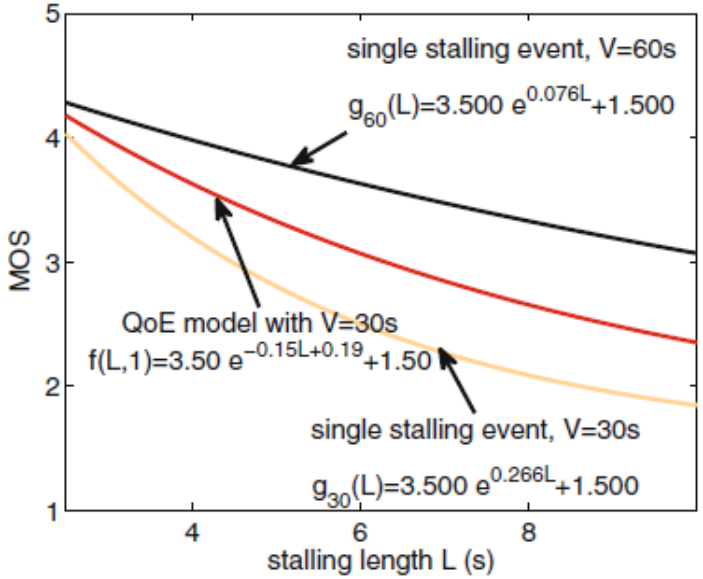
QoE for YouTube

IQX hypothesis validation:

$$QoE(L, N) = \alpha * e^{-\beta(L)*N} + \gamma,$$

$$\alpha = 3.5, \quad \beta(L) = 0.15L + 0.19, \quad \gamma = 1.5$$





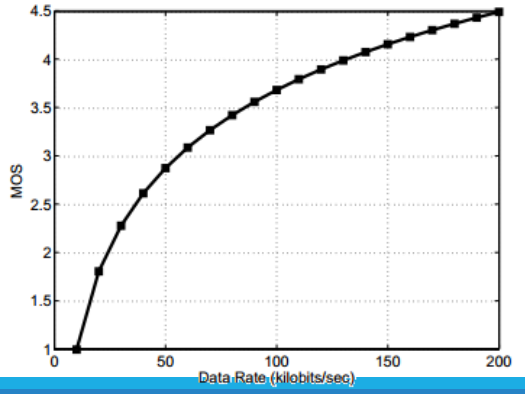
* T. Hossfeld, R. Schatz, E. W. Biersack, and L. Plissonneau, "Internet Video Delivery in YouTube: From Traffic Measurements to Quality of Experience," in Data Traffic Monitoring and Analysis, Eds. Springer Berlin Heidelberg, pp. 264–301, 2013.

D. QoE for file download services

- Elastic service, for which the utility function is an increasing, strictly concave, and continuously differentiable function of throughput
- The user satisfaction of a file transfer service is solely dependent on the provided data rate
- Logarithmic relationship between MOS and throughput:

$$MOS = \begin{cases} 1, & R < 10kbps \\ \alpha \log_{10}(\beta R), & 10kbps < R < 300kbps \\ 4.5, & 300kbps < R \end{cases}$$

- R is the data rate of the service
- α and β obtained from the upper and lower user perceived quality expectations

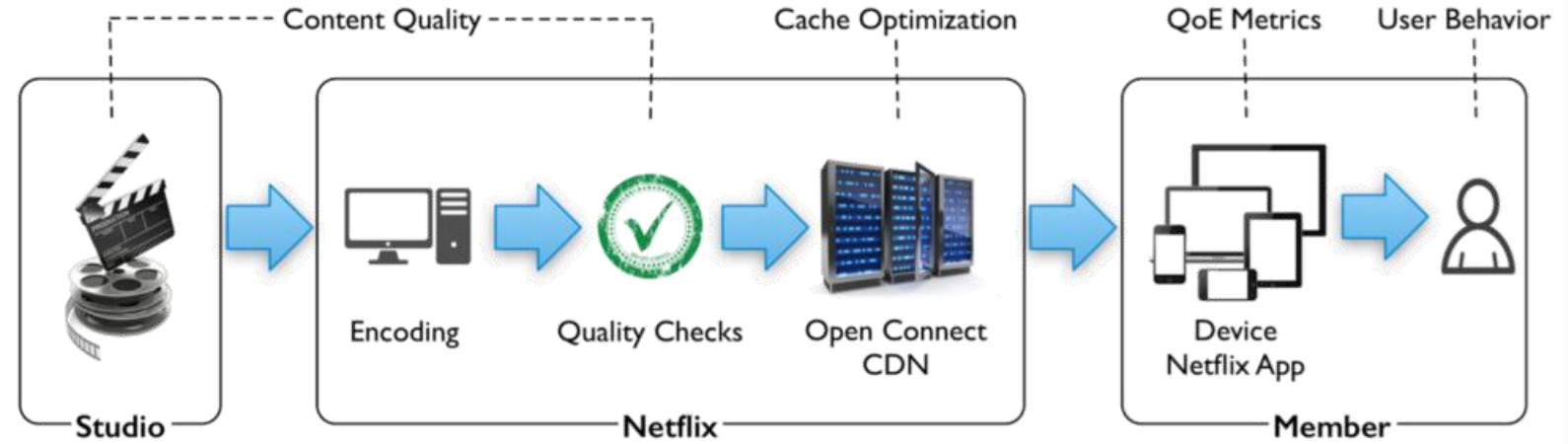


* S. Thakolsri, S. Khan, E. Steinbach, and W. Kellerer, "QoE-Driven Cross-Layer Optimization for High Speed Downlink Packet Access," J. Commun., vol. 4, no. 9, 2009

E. Netflix – challenges

- Understanding the impact of QoE on user behavior (compression artifacts, scaling artifacts, rebuffering rate, bitrate, etc.)
- Creating a personalized streaming experience
- Determining what movies and shows to cache on the edge servers based on member viewing behavior
- Improving the technical quality of the content using viewing data and member feedback

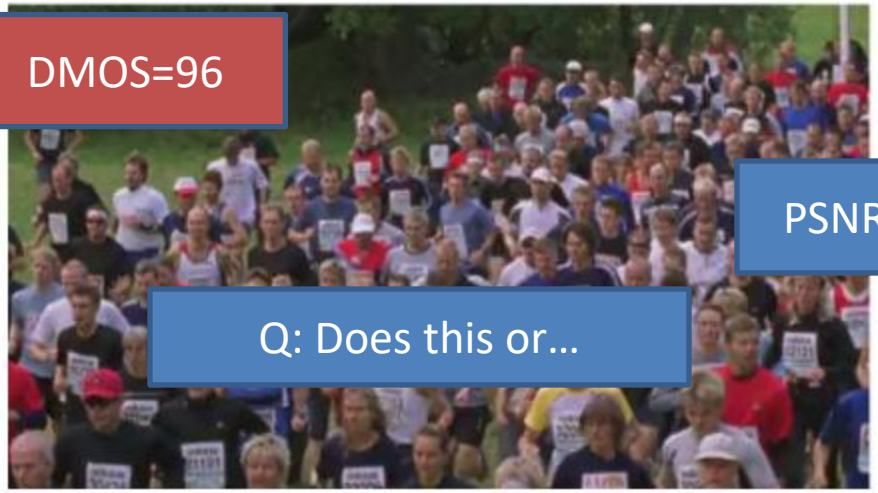
Q: How could Netflix infer that something is wrong?



* https://medium.com/netflix-techblog/optimizing-the-netflix-streaming-experience-with-data-science-725f04c3e834

Other QoE metrics

PSNR=31dB



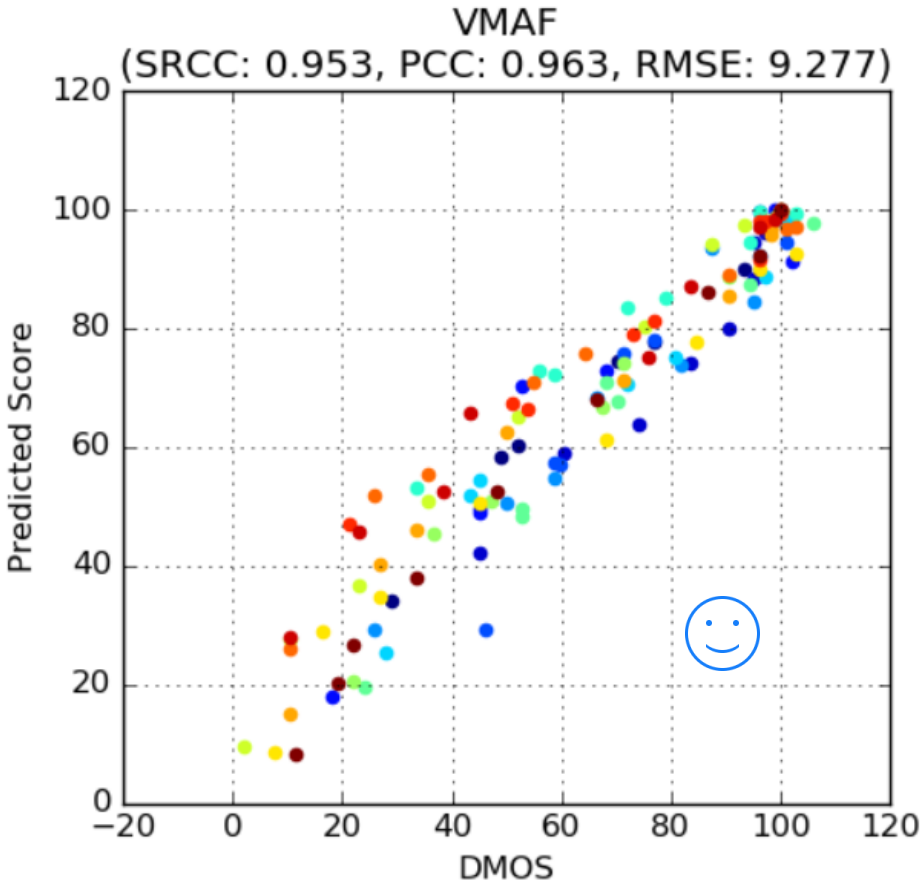
PSNR=34dB

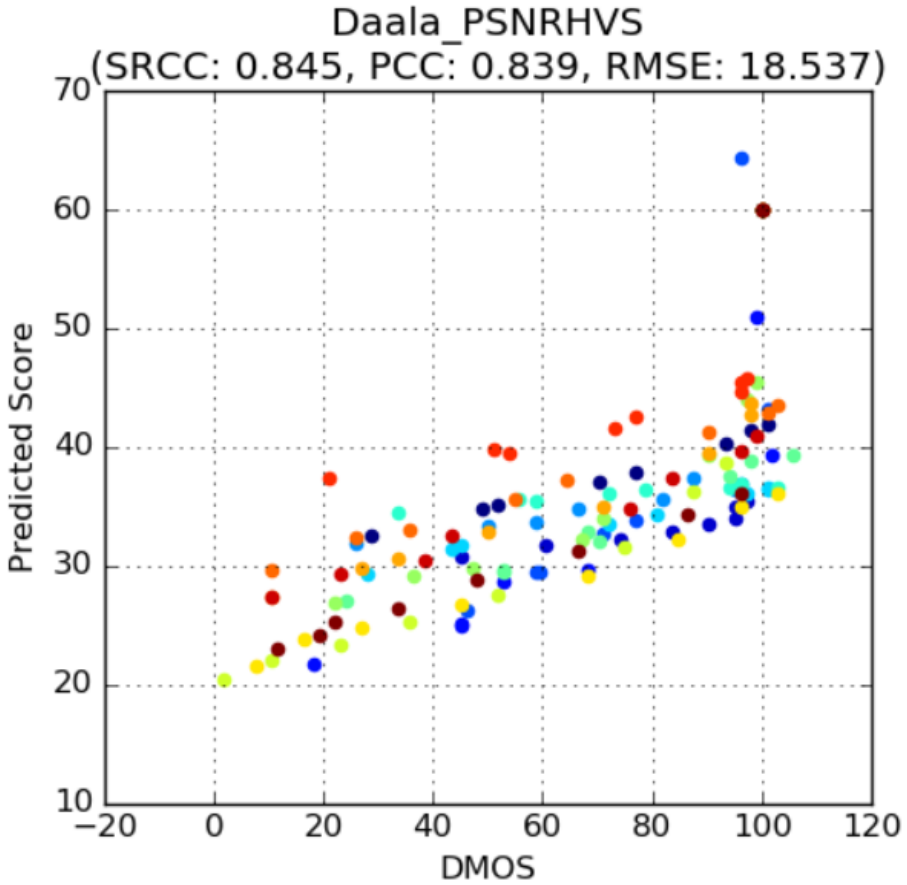
Q: Does this or…

Q: this look better?

DMOS is 100 for the reference video

Netflix Video Multimethod Assessment Fusion (VMAF) quality metric



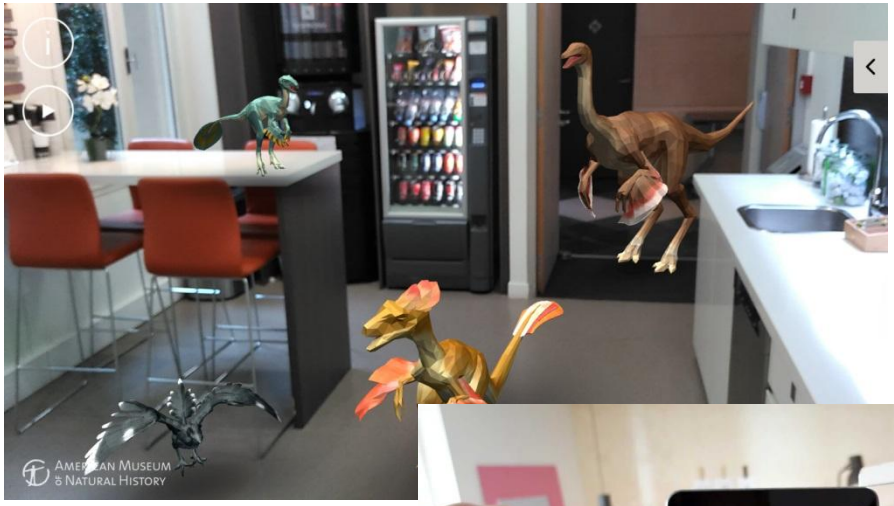


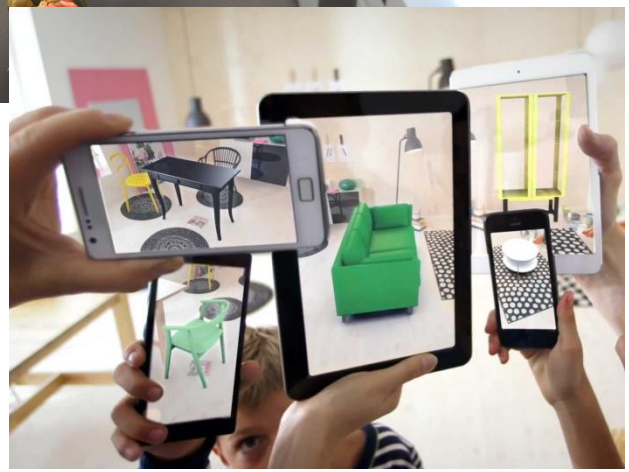
Q: Which one is better and why?

* https://medium.com/netflix-techblog/toward-a-practical-perceptual-video-quality-metric-653f208b9652

Augmented reality

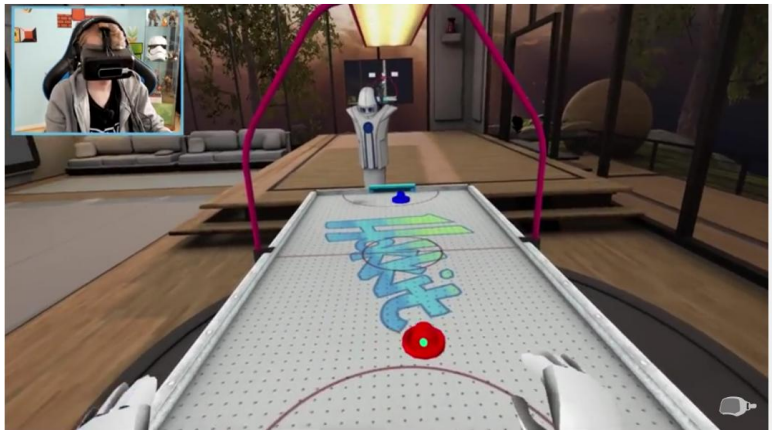
- It aims to "enrich" our world by inserting a layer of digital information between our eyes and the environment





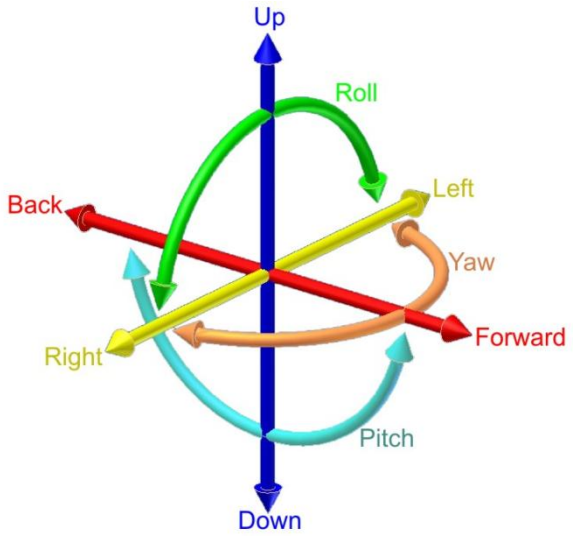
Virtual reality

- It tries to “transfer” the user to a world different from the real one



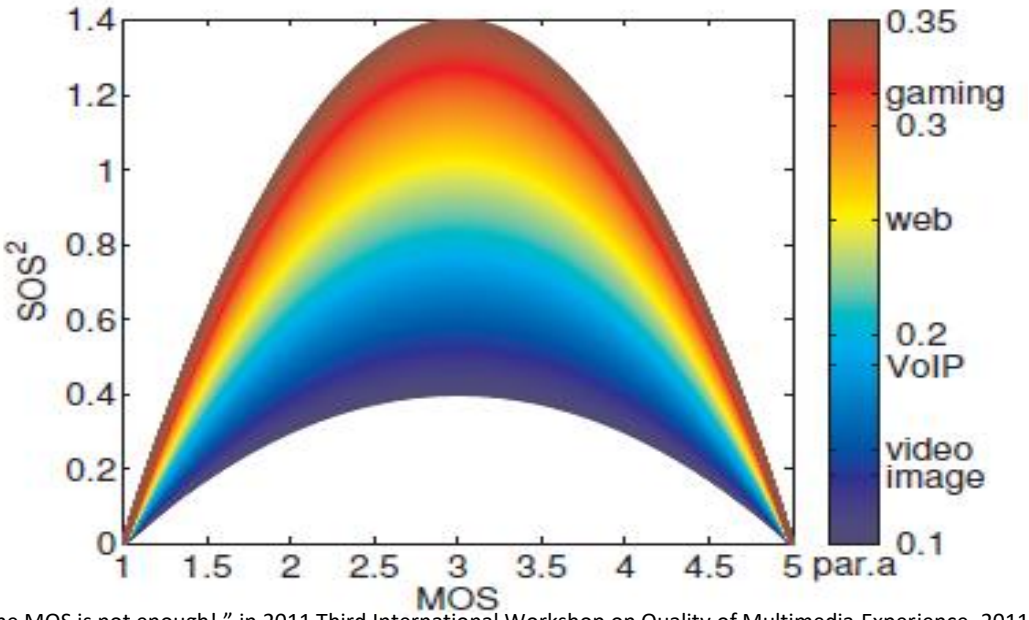
QoE in immersive services

- Feeling of “presence”
- Motion sickness
- 6 degrees of freedom
- Stereoscopic impression
- Delay / Non-synchronisation
- Image distortion
- Image freezing



SOS – The MOS is not enough

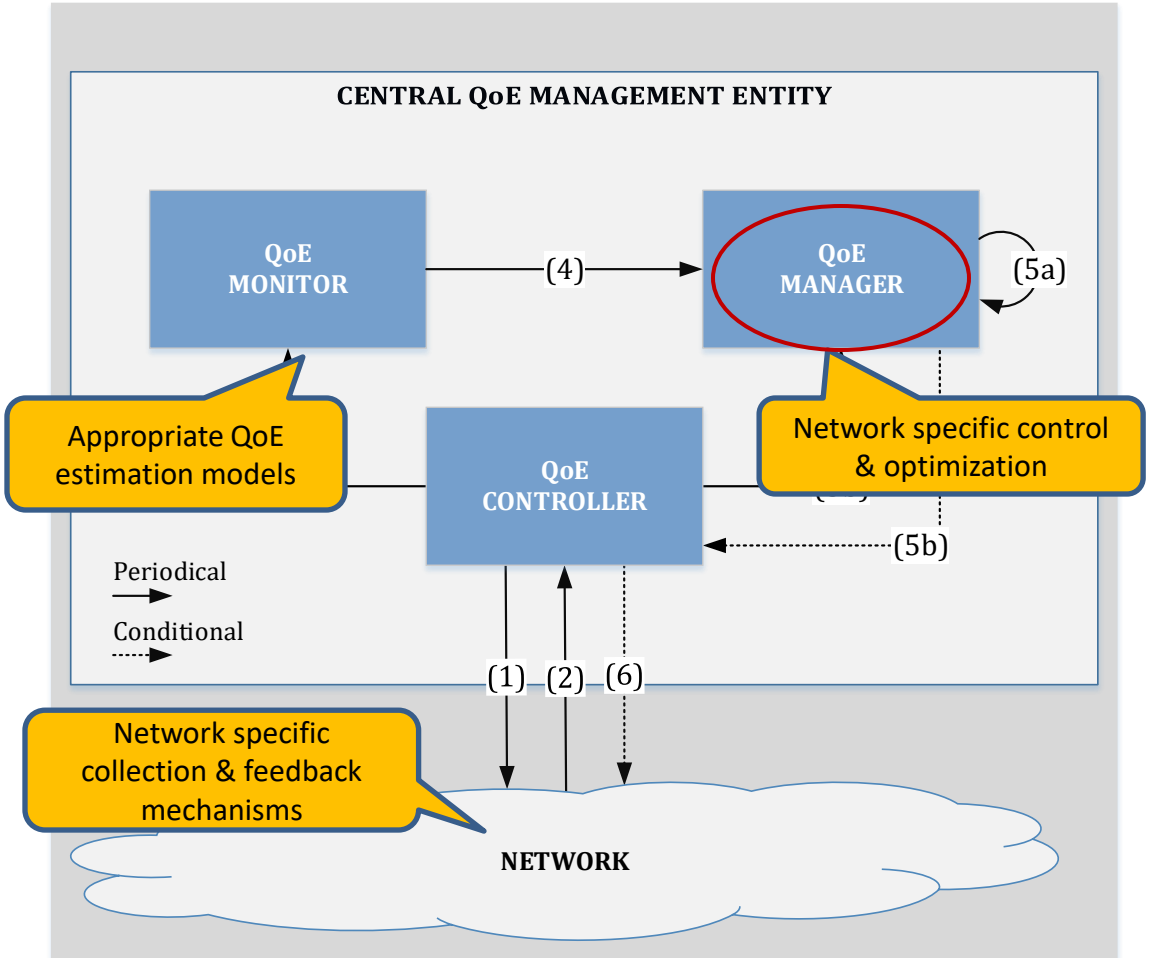
- Standard deviation of Opinion Scores (SOS)
- Reflects the level of rating diversity
- A square function of MOS → SOS hypothesis
- No diversity at the edges and maximal diversity at MOS = 3



* T. Hossfeld, R. Schatz, and S. Egger, "SOS: The MOS is not enough!," in 2011 Third International Workshop on Quality of Multimedia Experience, 2011, pp. 131–136.

QoE MANAGEMENT

QoE management framework



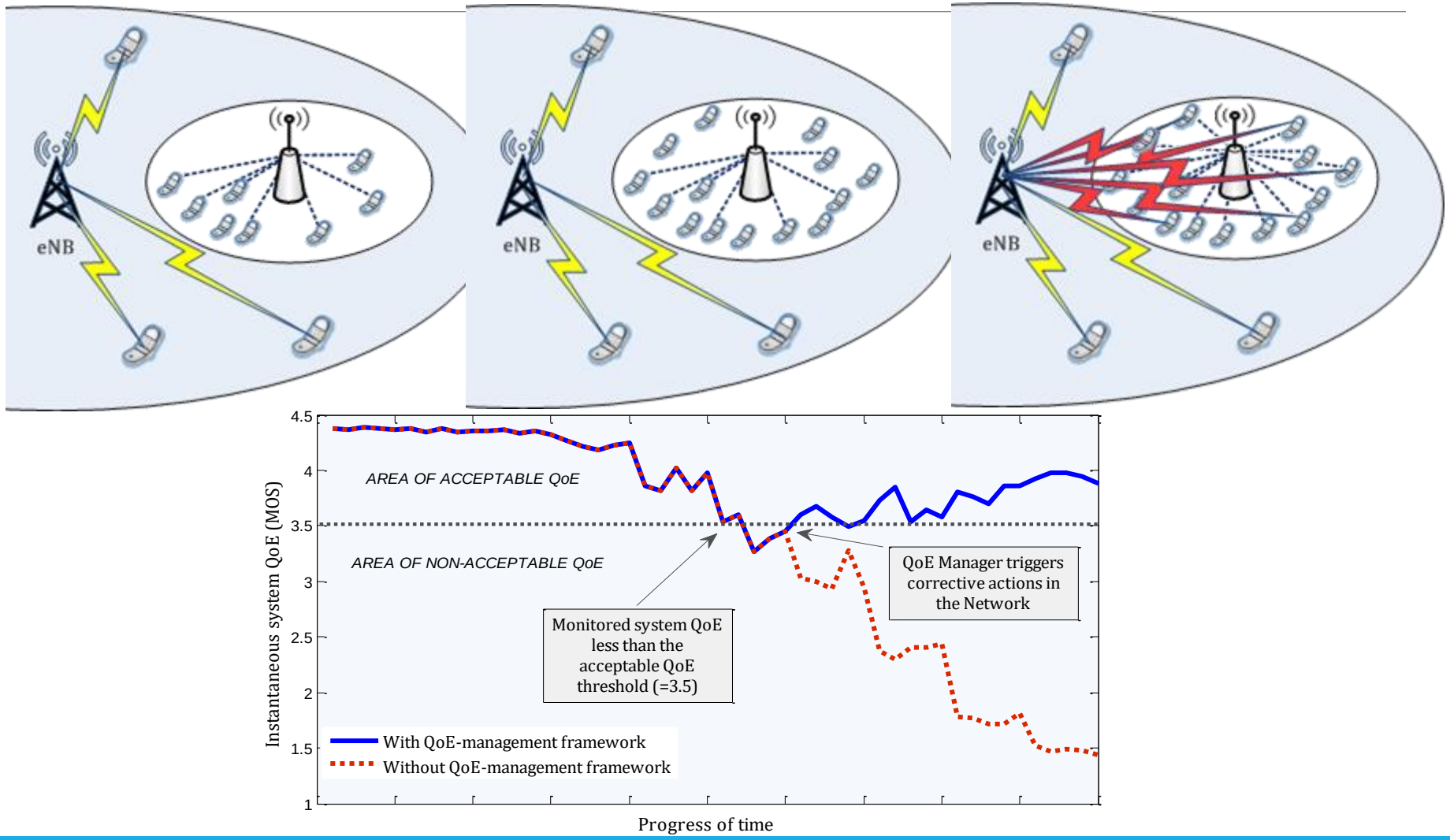
OBJECTIVE:

Enable a QoE-centric network management framework to:

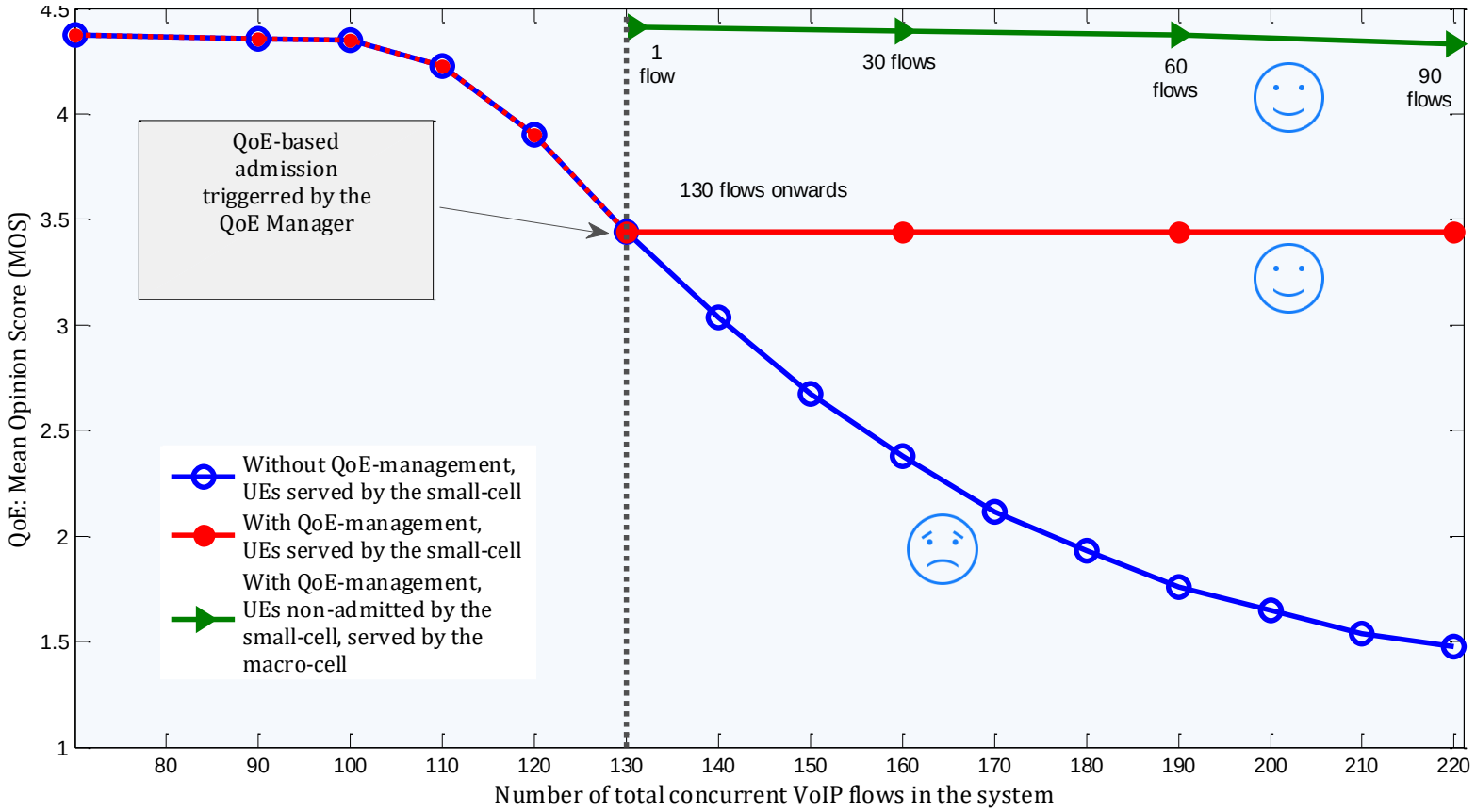
1. Monitor the end-users' QoE
2. Enhance their experience
3. Improve the network's efficiency (spectrum, energy)

* E. Liotou, D. Tsolkas, N. Passas and L. Merakos, "Quality of Experience management in mobile cellular networks: Key issues and design challenges," IEEE Communications Magazine, Network & Service Management Series, July 2015.

Improving QoE – Admission control



Improving QoE – Admission control



* E. Liotou, D. Tsolkas, N. Passas and L. Merakos, "Quality of Experience management in mobile cellular networks: Key issues and design challenges," IEEE Communications Magazine, Network & Service Management Series, July 2015.

The impact of fluctuations on QoE

Novel downlink-throughput related KPIs have to be developed for proper QoE-based evaluation