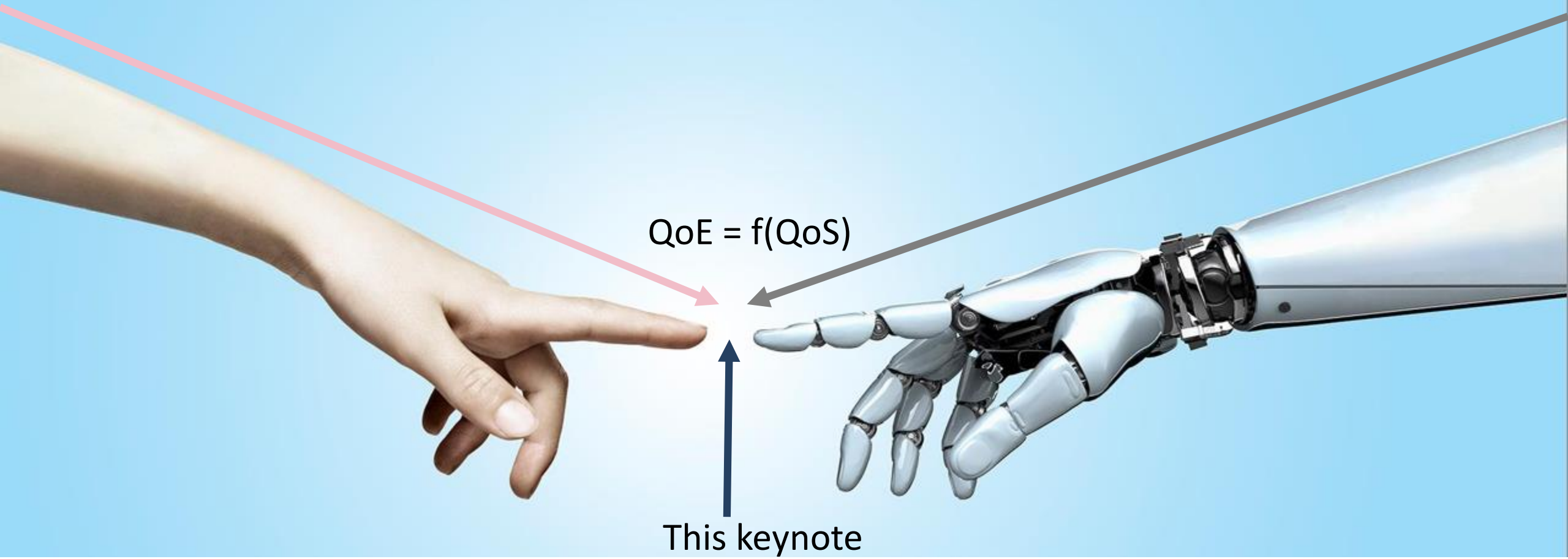
Human in the (Web) QoE loop

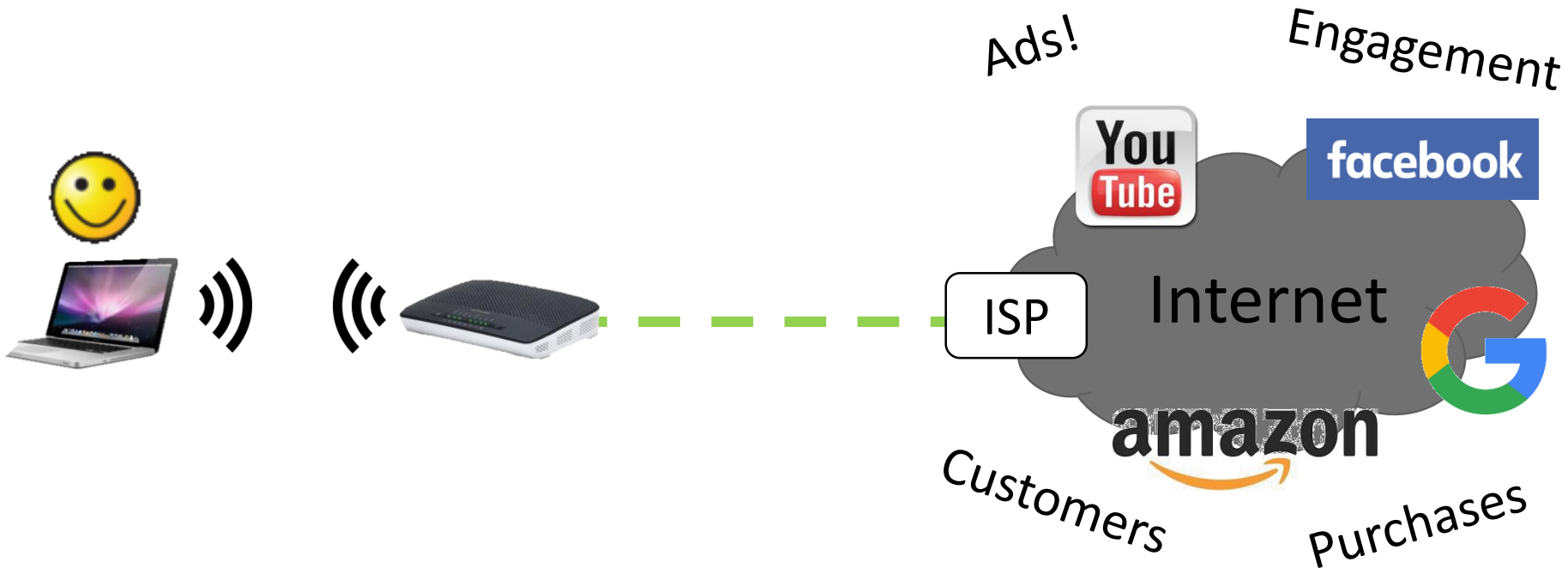


Dario Rossi
Chief Expert,
Huawei Technologies, Co. Ltd.
dario.rossi@huawei.com

Human in the (Web) QoE loop
aka the Wolf in Sheep's clothing

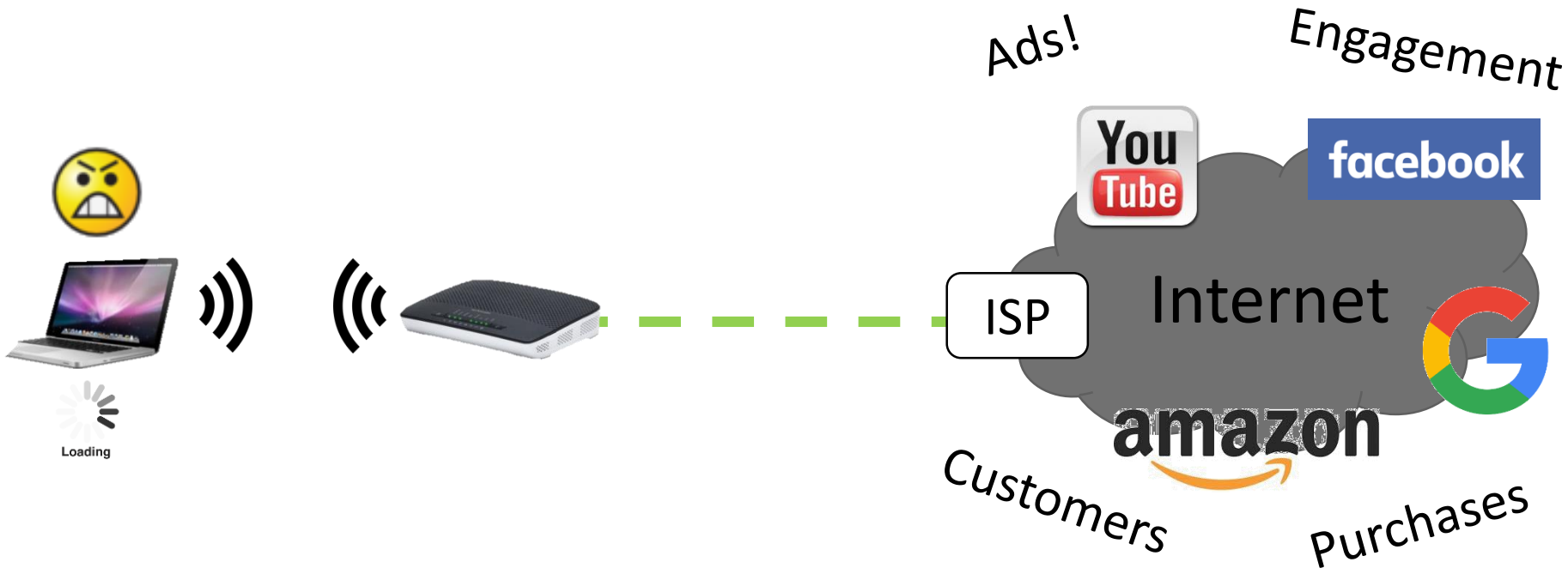
Dario Rossi
Chief Expert,
Huawei Technologies, Co. Ltd.

Quality of Experience (QoE) matters!



Good user QoE is a common goal for different stakeholders

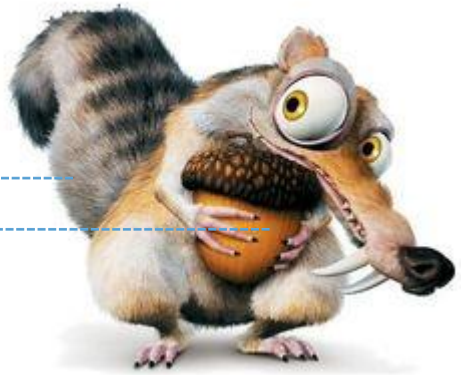
Quality of Experience (QoE) matters!



Detecting user QoE disruption is important.
Managing QoE to avoid disruption is even more!

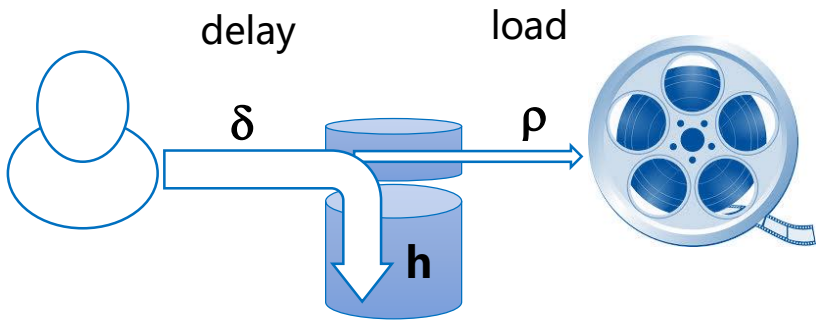
QoS vs QoE network management in a nutshell

max QoS
s.t. constraints



Engineer
Resource

Take video delivery as
a high-level example



Hit ratio

QoS vs QoE network management in a nutshell

Disclaimer:
I told you ☺

max QoS
s.t. constraints

Take video delivery as a high-level example

$$\max\ h = 1 - \sum_{\substack{l\in\mathcal{L}\\ o\in\mathcal{O}}} d_{l,o}^{out} \Big/ \sum_{o\in\mathcal{O}} d_o$$

$$\rho_o = F\left(\mathcal{Q}, P, \{x_{n,o} | n \in \mathcal{N}\}\right) \quad (1)$$

$$d_{core,o}^{out} = (1-\rho_o) d_o \quad (2)$$

$$d_{core,o}^{out} = \sum_{l\in\mathcal{L}} d_{l,o}^{in} \quad (3)$$

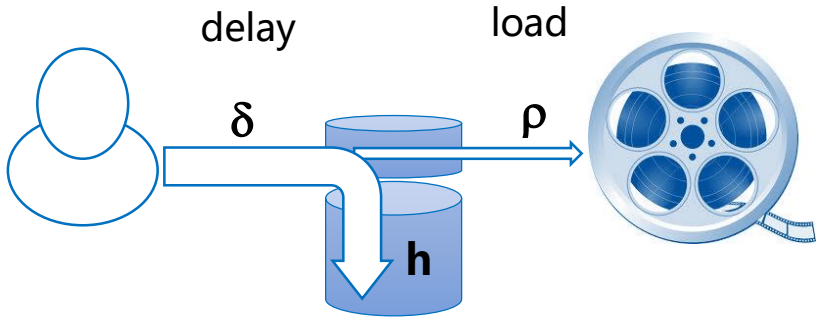
$$d_{l,o}^{in} = 0, \forall l \in \mathcal{L} \setminus \mathcal{L}_o \quad (4)$$

$$d_{l,o}^{out} = (1 - x_{l,o}) \cdot d_{l,o}^{in} \quad (5)$$

$$\sum_{o\in\mathcal{O}} x_{l,o} = cs_l, \forall l \in \mathcal{L} \quad (6)$$

$$\sum_{o\in\mathcal{O}} x_{n,o} = cs_n, \forall n \in \mathcal{N} \quad (7)$$

$$\sum_{l\in\mathcal{L}} cs_l + \sum_{n\in\mathcal{N}} cs_{n,o} \le C_{tot}. \quad (8)$$



Hit ratio

Network-centric QoS focus

[TPDPS-15] Araldo, A. and Rossi, D. and Martignon, F., Cost-aware caching: Caching more (costly items) for less (ISPs operational expenditures) In IEEE Transactions on Parallel and Distributed Systems, 2015

QoS vs QoE network management in a nutshell

Disclaimer:
I told you ☺

max QoS
s.t. constraints

max QoE = f(QoS)
s.t. constraints

$$\max \; h = 1 - \sum_{\substack{l \in \mathcal{L} \\ o \in \mathcal{O}}} d_{l,o}^{out} \Big/ \sum_{o \in \mathcal{O}} d_o$$

$$\rho_o = F\left(\mathcal{Q}, P, \{x_{n,o} | n \in \mathcal{N}\}\right) \quad (1)$$

$$d_{core,o}^{out} = (1 - \rho_o) d_o \quad (2)$$

$$d_{core,o}^{out} = \sum_{l \in \mathcal{L}} d_{l,o}^{in} \quad (3)$$

$$d_{l,o}^{in} = 0, \forall l \in \mathcal{L} \setminus \mathcal{L}_o \quad (4)$$

$$d_{l,o}^{out} = (1 - x_{l,o}) \cdot d_{l,o}^{in} \quad (5)$$

$$\sum_{o \in \mathcal{O}} x_{l,o} = cs_l, \forall l \in \mathcal{L} \quad (6)$$

$$\sum_{o \in \mathcal{O}} x_{n,o} = cs_n, \forall n \in \mathcal{N} \quad (7)$$

$$\sum_{l \in \mathcal{L}} cs_l + \sum_{n \in \mathcal{N}} cs_{n,o} \leq C_{tot}. \quad (8)$$

$$\max \sum_{o \in O} \sum_{q \in Q} \sum_{v \in V} n_v^{o,q} U^q$$

subject to:

$$\sum_{q \in Q} n_v^{o,q} = n_v^o \qquad \forall o \in O, v \in V \quad (2)$$

$$d^{o,q,v_d} = n_{v_d}^{o,q} \cdot r^q \qquad \forall o \in O, q \in Q, v_d \in V \quad (3)$$

$$d^{o,q,v_d} = z_{v_d}^{o,q,v_d} + w_{v_d}^{o,q,v_d} + \sum_{e \in BS(v_d)} y_e^{o,q,v_d} - \sum_{e \in FS(v_d)} y_e^{o,q,v_d} \qquad \forall o \in O, q \in Q, v_d \in V \quad (4)$$

$$z_{v_s}^{o,q,v_d} + w_{v_s}^{o,q,v_d} + \sum_{e \in BS(v_s)} y_e^{o,q,v_d} = \sum_{e \in FS(v_s)} y_e^{o,q,v_d} \qquad \forall o \in O, q \in Q, v_s \in V, v_d \in V, v_s \neq v_d \quad (5)$$

$$\sum_{o \in O} \sum_{q \in Q} \sum_{v_d \in V} y_e^{o,q,v_d} \leq b_e \qquad \forall e \in A \quad (6)$$

$$\sum_{v_d \in V} z_{v_s}^{o,q,v_d} \leq p_{v_s}^{o,q} \cdot bw_{v_s} \qquad \forall o \in O, q \in Q, v_s \in V \quad (7)$$

$$\sum_{v_d \in V} w_{v_s}^{o,q,v_d} \leq x_{v_s}^{o,q} \cdot bw_{v_s} \qquad \forall o \in O, q \in Q, v_s \in V \quad (8)$$

$$\sum_{o \in O} \sum_{q \in Q} x_{v_s}^{o,q} \cdot s^q \leq S_{v_s} \qquad \forall v_s \in V \quad (9)$$

$$\sum_{o \in O} \sum_{q \in Q} \sum_{v_s \in V} x_{v_s}^{o,q} \cdot s^q \leq S_{TOT} \quad (10)$$

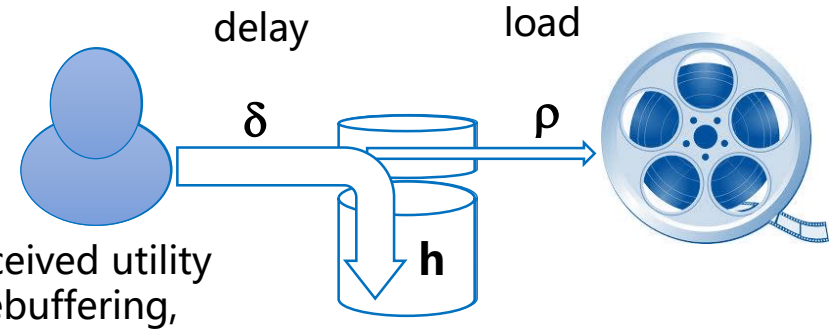
$$x_v^{o,q} \in \{0,1\} \qquad \forall o \in O, q \in Q, v \in V \quad (11)$$

$$n_v^{o,q} \in \mathbb{Z}^+ \qquad \forall o \in O, q \in Q, v \in V \quad (12)$$

$$y_e^{o,q,v_d} \in \mathbb{R}^+ \qquad \forall o \in O, q \in Q, v_d \in V, e \in A \quad (13)$$

$$d^{o,q,v_d} \in \mathbb{R}^+ \qquad \forall o \in O, q \in Q, v_d \in V \quad (14)$$

$$z_{v_s}^{o,q,v_d}, w_{v_s}^{o,q,v_d} \in \mathbb{R}^+ \qquad \forall o \in O, q \in Q, v_d \in V, v_s \in V. \quad (15)$$



User perceived utility
f(SSim, rebuffering, quality switch, …)

Hit ratio

User-centric QoE focus

[NETWORKING-16] Araldo, Andrea and Martignon, Fabio and Rossi, Dario, Representation Selection Problem: Optimizing Video Delivery through Caching IFIP Networking 2016,

[TPDPS-15] Araldo, A. and Rossi, D. and Martignon, F., Cost-aware caching: Caching more (costly items) for less (ISPs operational expenditures) In IEEE Transactions on Parallel and Distributed Systems, 2015

QoS vs QoE network management, reloaded

(Shamless placement)

max QoE = f(QoS)
s.t. constraints

QoE ≈ f(QoS)

delay δ

load ρ

h

User perceived utility
f(SSim, rebuffering, quality switch, …)

Hit ratio

Real users

Real loop

Higher levels of realism (eg DASH dynamics)

Deeper integration of application & network layer solutions

[TMM-17] J. Samain, G. Carofiglio, L. Muscariello, M. Papalini, M. Sardara, M. Tortelli, and D. Rossi, Dynamic Adaptive Video Streaming: Towards a systematic comparison of ICN and TCP/IP In IEEE Transactions on Multimedia, Vol. 19, pp.2166-2181, oct. 2017

[NOSSDAV-18] J. Samain, G. Carofiglio, M. Tortelli, D. Rossi, A simple yet effective network-assisted signal for enhanced DASH quality of experience 28th ACM SIGMM NOSSDAV, jun. 2018, *best paper award* ★

QoS vs QoE network management..... scary part coming!

max QoE = f(QoS)
s.t. ints

QoE management meaningful only as long as QoE ≈ f(QoS)

QoE ≈ f(QoS)

delay δ

load ρ

h

User perceived utility
f(SSim, rebuffering,
quality switch, ...)

Hit ratio

Real users

Real loop

Higher levels of realism
(eg DASH dynamics)

Deeper integration of
application & network
layer solutions

[TMM-17] J. Samain, G. Carofiglio, L. Muscariello, M. Papalini, M. Sardara, M. Tortelli, and D. Rossi, Dynamic Adaptive Video Streaming: Towards a systematic comparison of ICN and TCP/IP In IEEE Transactions on Multimedia, Vol. 19, pp.2166-2181, oct. 2017

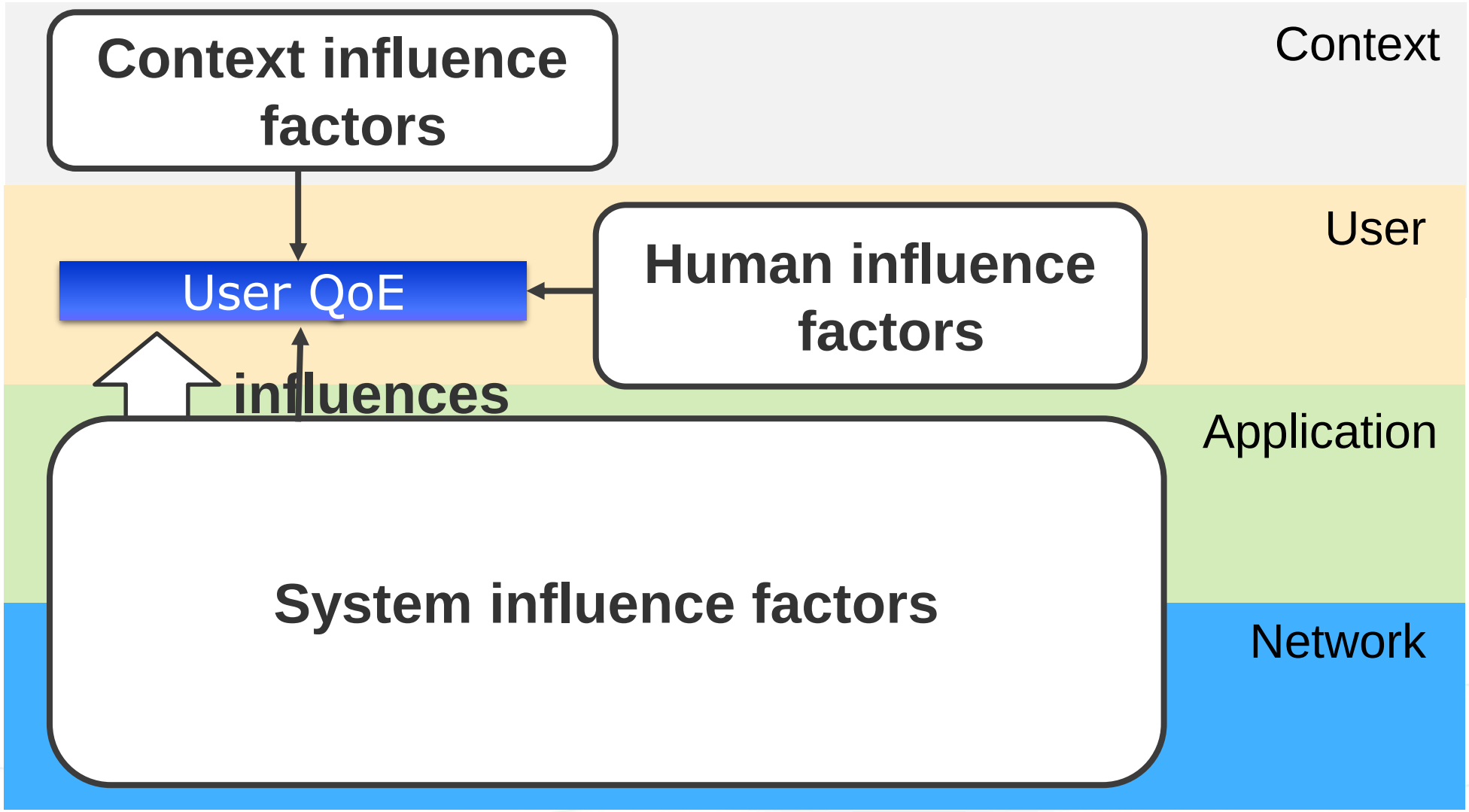
[NOSSDAV-18] J. Samain, G. Carofiglio, M. Tortelli, D. Rossi, A simple yet effective network-assisted signal for enhanced DASH quality of experience 28th ACM SIGMM NOSSDAV, jun. 2018, *best paper award* ★

QoE ≈ f(QoS) agenda: what can go wrong?

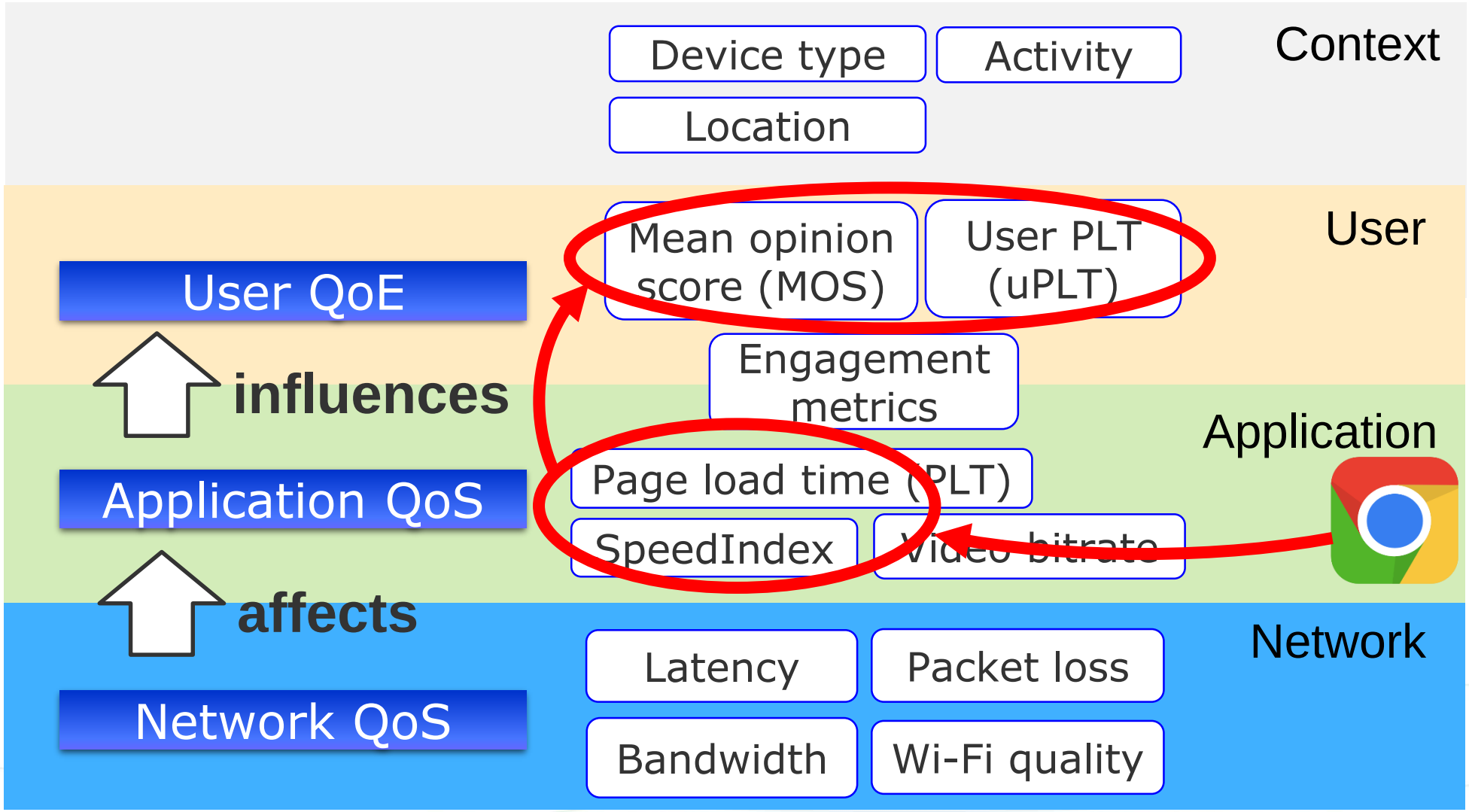
- Engineering/datascience viewpoint
  - Features engineering — Feature vector $\underline{x}$
  - MOS collection — Target labels y
  - QoE modeling — Regression $y = f(\underline{x})$
- Human perception viewpoint
  - Humans are not rational players
  - Crowdworkers vs Real users
- Painkiller prescriptions
  - Keep it sustained, stupid!
  - Removing features altogether
  - Accurate yet scalable QoE modeling

Using the Web as an example application

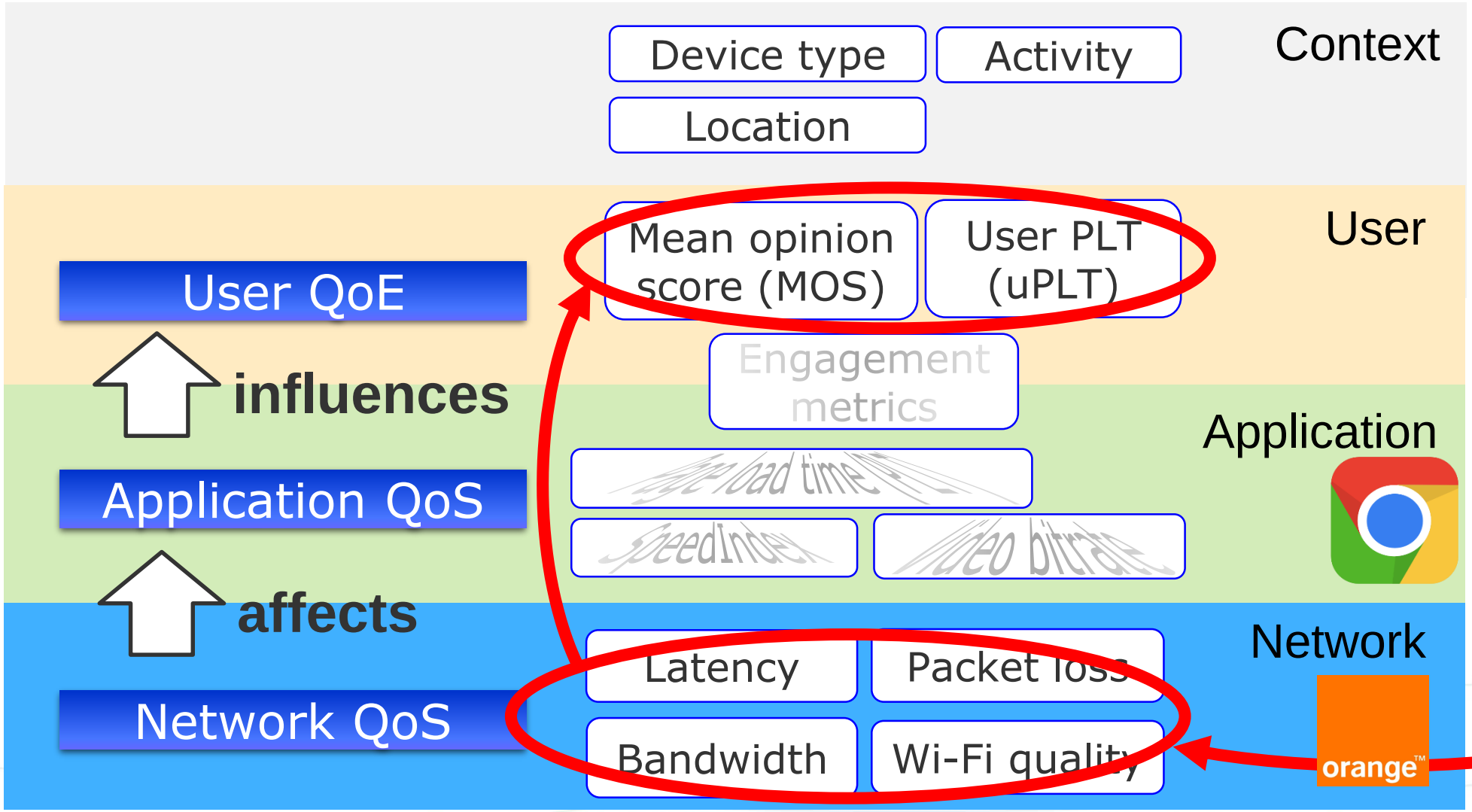
Quality at different layers



Quality at different layers



Quality at different layers



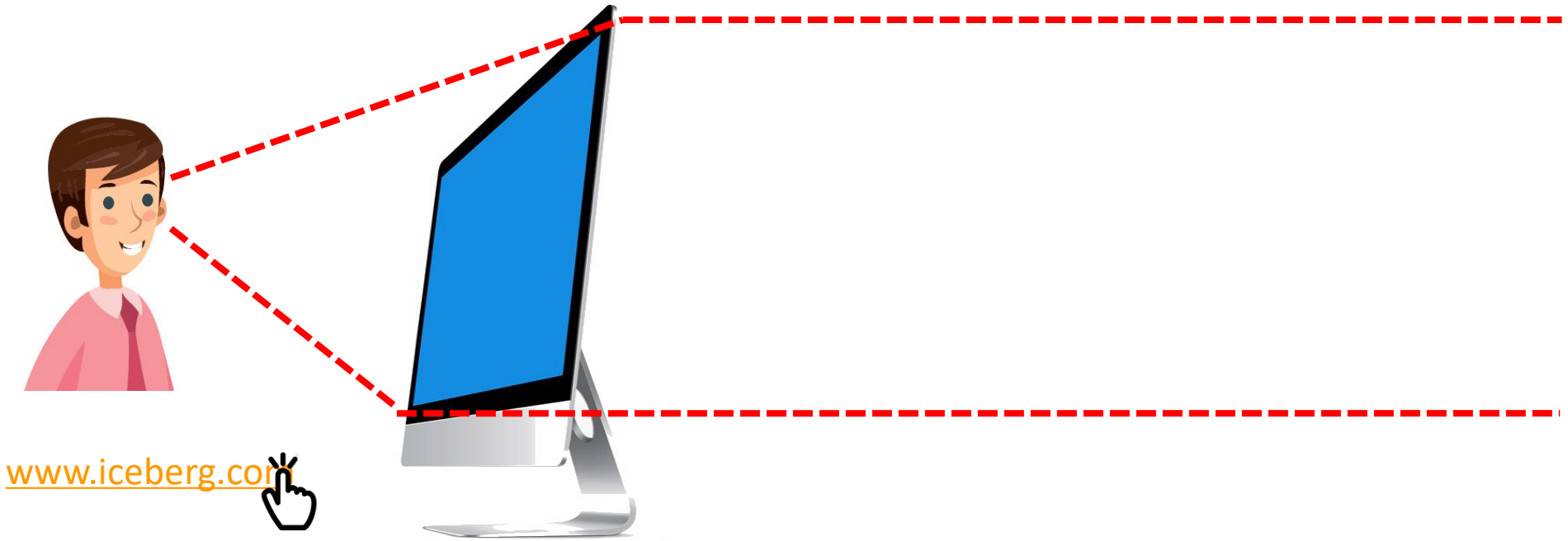
https://
HTTPbis, QUIC…

(true for any other application)

WIKIPEDIA

Feature engineering (1/2): Instant metrics

Stopping product placement, references at the end

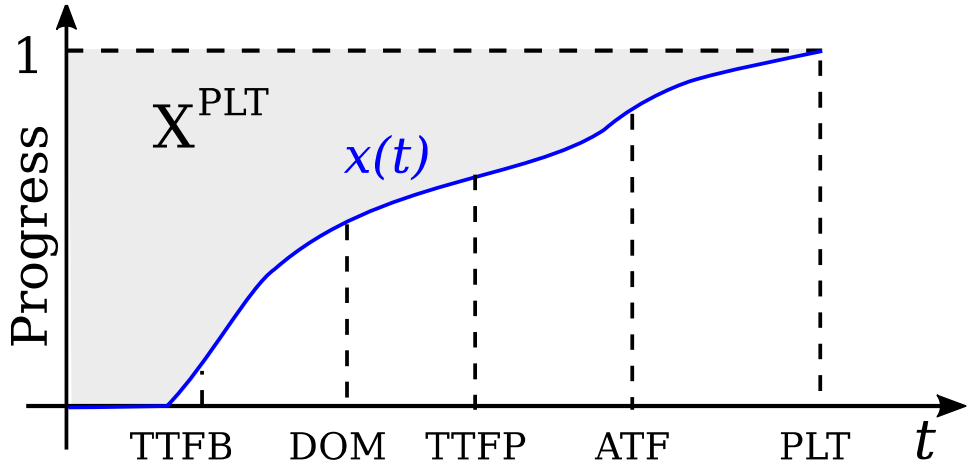


DOM ✓ ATF ✓ PLT ✓

* Images by vvstudio, vectorpocket, Ydlabs / Freepik

Feature engineering (2/2): Integral metrics

- SpeedIndex, RUMSI, PSSI
  - Processing intensive ✗
  - Only at L7 (in browser) ✗
  - Visual progress metric ✓
- ObjectIndex, ByteIndex and ImageIndex
  - Lightweight ✓
  - ByteIndex also at L3 (in network) ✓
  - Possibly far from user QoE ? ?



$$X = \int_0^{t_{end}} (1 - x(t))dt$$

SpeedIndex %of visual completeness (histogram, rectangles or SSim)

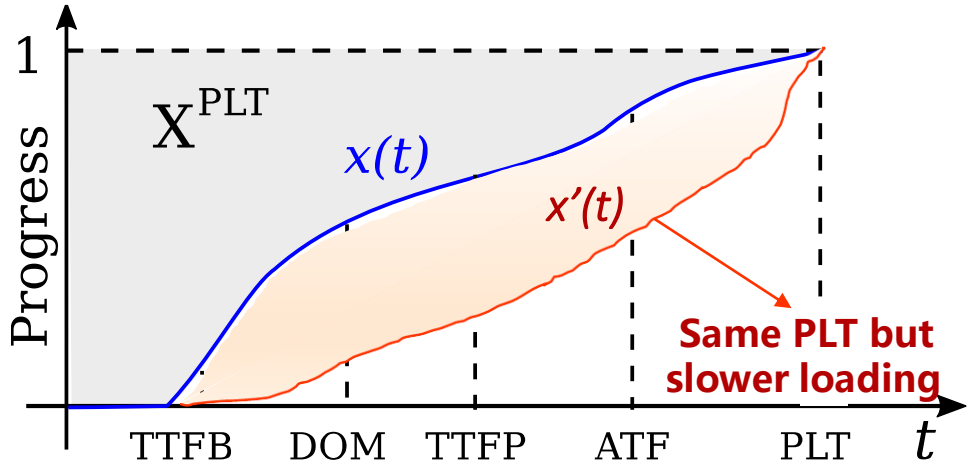
ObjectIndex % of objects downloaded

ByteIndex % of bytes downloaded

ImageIndex % of bytes of images

Feature engineering (2/2): Integral metrics

- SpeedIndex, RUMSI, PSSI
  - Processing intensive ✗
  - Only at L7 (in browser) ✗
  - Visual progress metric ✓
- ObjectIndex, ByteIndex and ImageIndex
  - Lightweight ✓
  - ByteIndex also at L3 (in network) ✓
  - Possibly far from user QoE ? ?



$$X = \int_0^{t_{end}} (1 - x(t))dt$$

SpeedIndex %of visual completeness (histogram, rectangles or SSim)

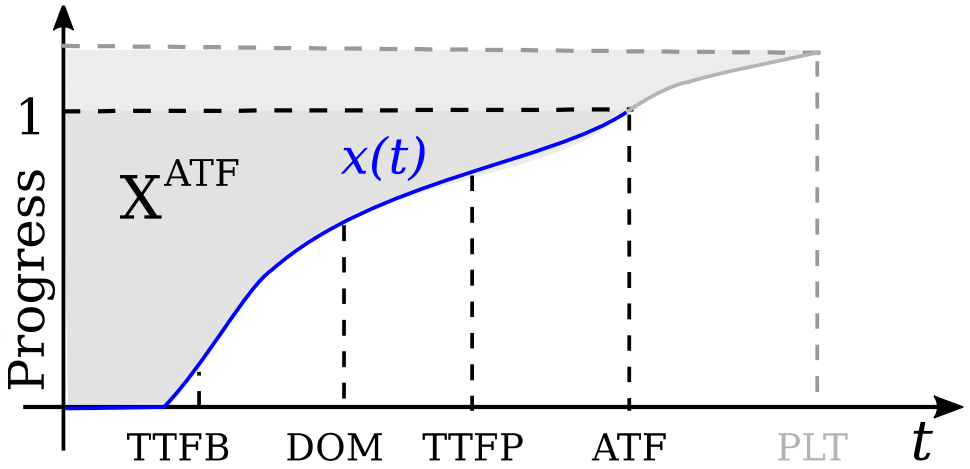
ObjectIndex % of objects downloaded

ByteIndex % of bytes downloaded

ImageIndex % of bytes of images

Feature engineering (2/2): Integral metrics

- SpeedIndex, RUMSI, PSSI
  - Processing intensive ✗
  - Only at L7 (in browser) ✗
  - Visual progress metric ✓
- ObjectIndex, ByteIndex and ImageIndex
  - Lightweight ✓
  - ByteIndex also at L3 (in network) ✓
  - Possibly far from user QoE ? ?



$$X = \int_0^{t_{end}} (1 - x(t))dt$$

Different cutoffs

SpeedIndex %of visual completeness (histogram, rectangles or SSim)

ObjectIndex % of objects downloaded

ByteIndex % of bytes downloaded

ImageIndex % of bytes of images

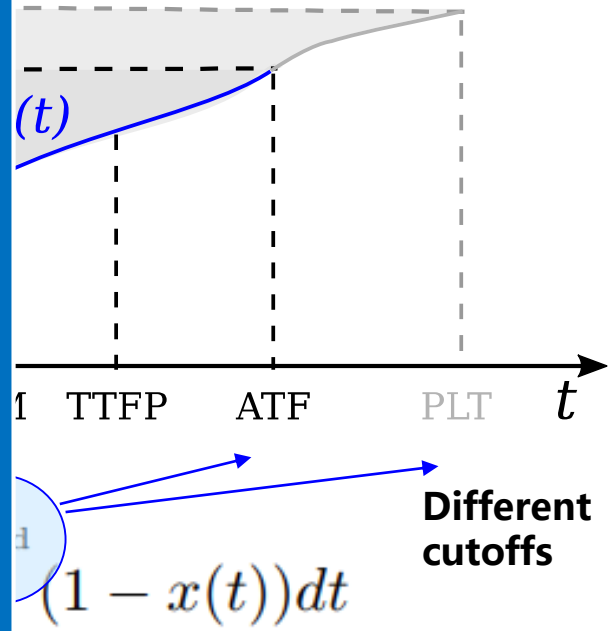
Feature engineering (2/2): Integral metrics

- SpeedIndex
  - Processing intens
  - Only at L7 (in bro
  - Visual progress m
- ObjectIndex, By
  - Lightweight ✓
  - Also at L3 (in net
  - Possibly far from user QoE ? ?

Takeaway message #1

In the age of ML & AI, still a fair amount of *feature engineering* !

(and this is at L7, w/o even considering encryption !!)



SpeedIndex %of visual completeness (histogram, rectangles or SSim)

ObjectIndex % of objects downloaded

ByteIndex % of bytes downloaded

ImageIndex % of bytes of images

Crowdsourcing user feedback

- Mean opinion score (MOS)

"Rate your experience from 1-poor to 5-excellent"

Lab experiments
- Small user diversity, volounteers
- Web browsing, but artificial websites
- Artificial controlled conditions

Award winning dataset released

- User perceived PLT (uPLT)

"Which of these two pages finished first ?"

Crowdsourcing (payed crowdworkers)
- Larger userbase, but higher noise
- Side-to-side videos ≠ Web browsing!
- Artificial controlled conditions

Collab with eyeorg

- User acceptance

"Did the page load fast enough ?" (Yes/No)

Experiments from operational website
- Actual service users
- Browsing in typical user conditions
- Huge heterogeneity (devices/browsers/nets)

Collab with WIKIPEDIA

Crowdsourcing  feedback (1/3)

(Dataset and paper available)

- Mean opinion score (MOS)

"Rate your experience from 1-poor to 5-excellent"

Lab experiments
- Small user diversity, volounteers
- Web browsing, but artificial websites
- Artificial controlled conditions

Humans are non rationale players

....

Some humans give *higher* scores to *slower* pages

Crowdsourcing feedback (1/3)

(Dataset and paper available)

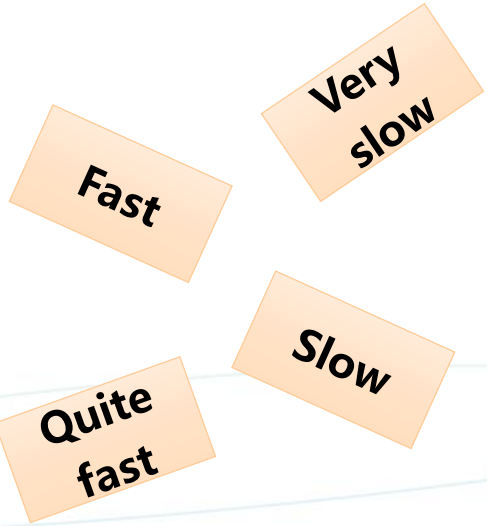
- Mean opinion score (MOS)

"Rate your experience from 1-poor to 5-excellent"

Lab experiments
- Small user diversity, volounteers
- Web browsing, but artificial websites
- Artificial controlled conditions

We also collected free-form comments for NLP processing but... ☺

I really liked how fast the images loaded. What a great user experiment. I feel like >I could do that all my life, you know clicking on webpages and watching them load... It provides a weird feeling. Weird but somehow appreciable, much like when you sit on a sand beach, alone, and you watch the waves coming and going. You can then think about your life, your goals and the meaning of life. Really, a way to retrieve your inner peace. Well, this comment is quite long. I think it is enough. Do you think this comment is long enough ? Hey ! I don't even know who you are ! Human ? Bot ? Cyborg ? Those things I just told you, do you REALLY understand them or will you just run an algorithm on it in order to find keywords such as "quick" "load" etc... I wonder, am I just whispering all my feelings, emotions, thougths (what makes me be me in fact) to a empty hole ? I don't like this idae, so I stop now. By the way, it was quite fun to have this little chat with you.



Crowdsourcing  feedback (2/3)

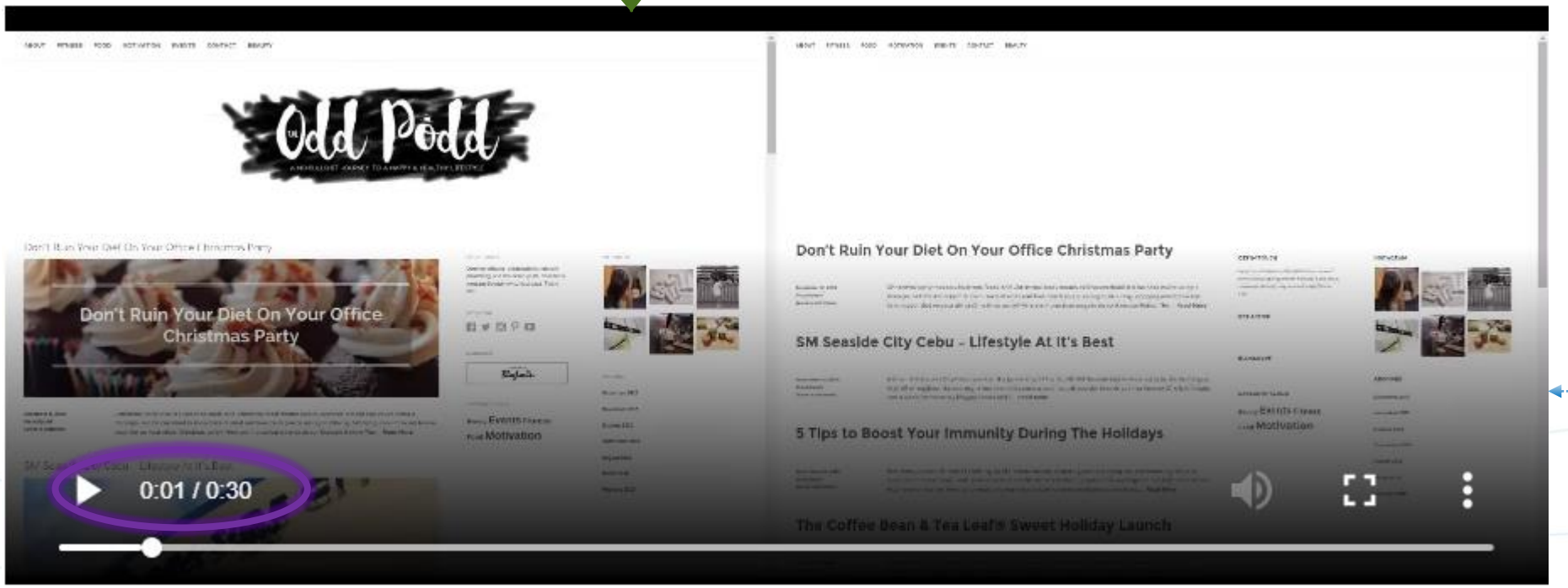
(Dataset and paper will come)

- User perceived PLT (uPLT)

"Which of these two pages finished first ?"

Crowdsourcing (payed crowdworkers)
- ➕ Larger userbase, but higher noise
- ➖ Side-to-side videos ≠ Web browsing!
- 🟰 Artificial controlled conditions

Collab with eyeorg



Crowdworkers
- Cheap but…
- Lot of work spent into label quality checking and filtering

Real instance
- Difference noticeable until just t= 1 sec…
- uPLT ≠ MOS !

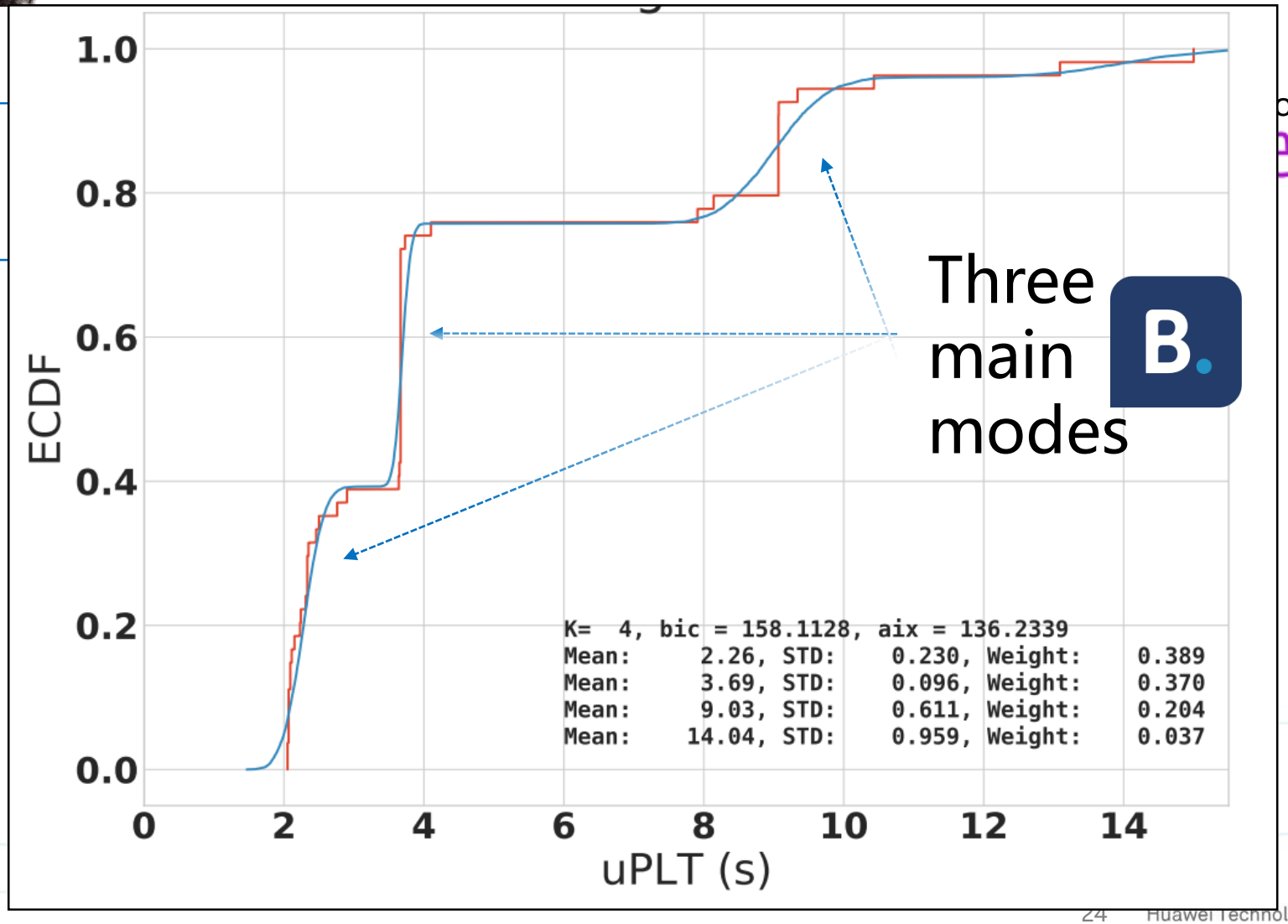
Crowdsourcing feedback (2/3)

(Dataset and paper will come)

collab with eyeorg

- User perceived PLT

"When does this page finishes loading?"



When do users think that a page is ready ?

Booking.com

Crowdsourcing feedback (2/3)

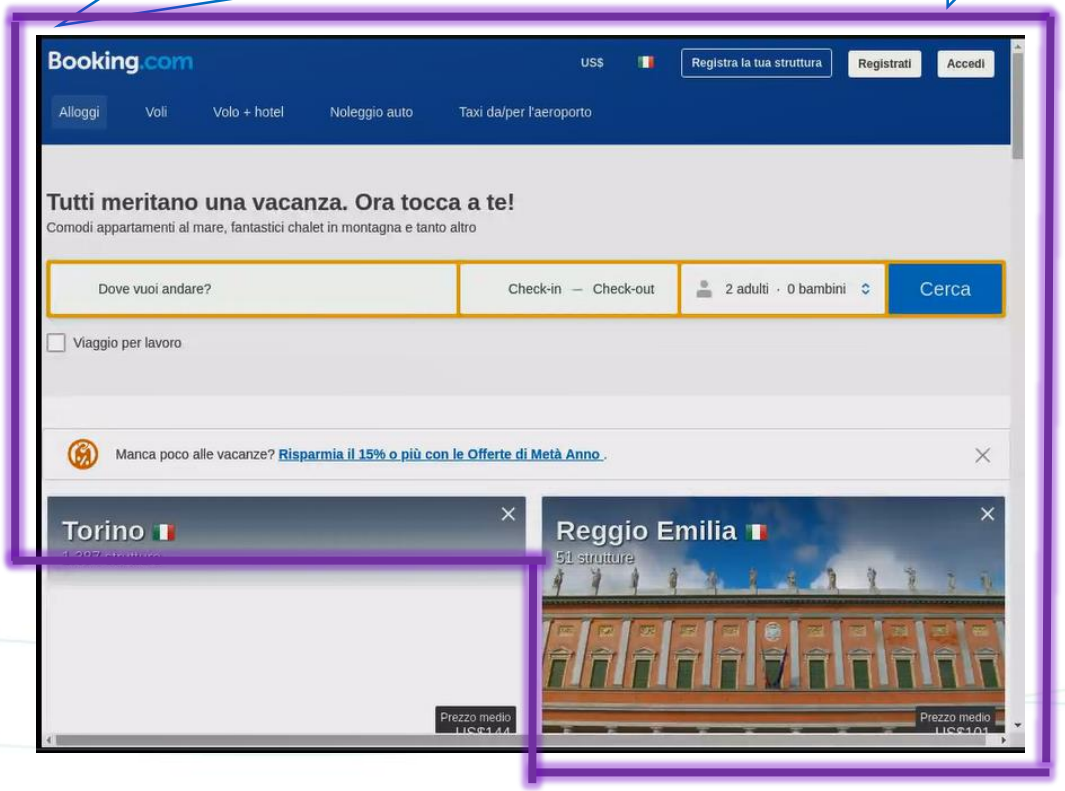
(Dataset and paper will come)

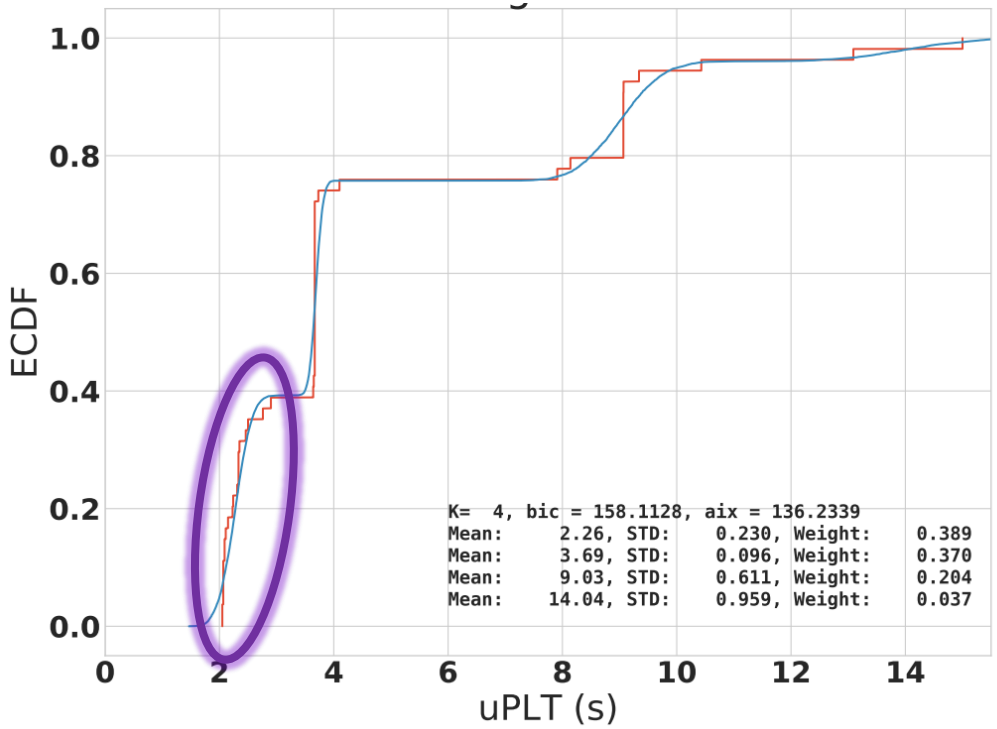
- User perceived PLT (uPLT)

"When does this page finishes loading?"

Crowdsourcing (payed crowdworkers)
- Larger userbase, but higher noise
- Side-to-side videos ≠ Web browsing!
- Artificial controlled conditions

Collab with eyeorg





Crowdsourcing feedback (2/3)

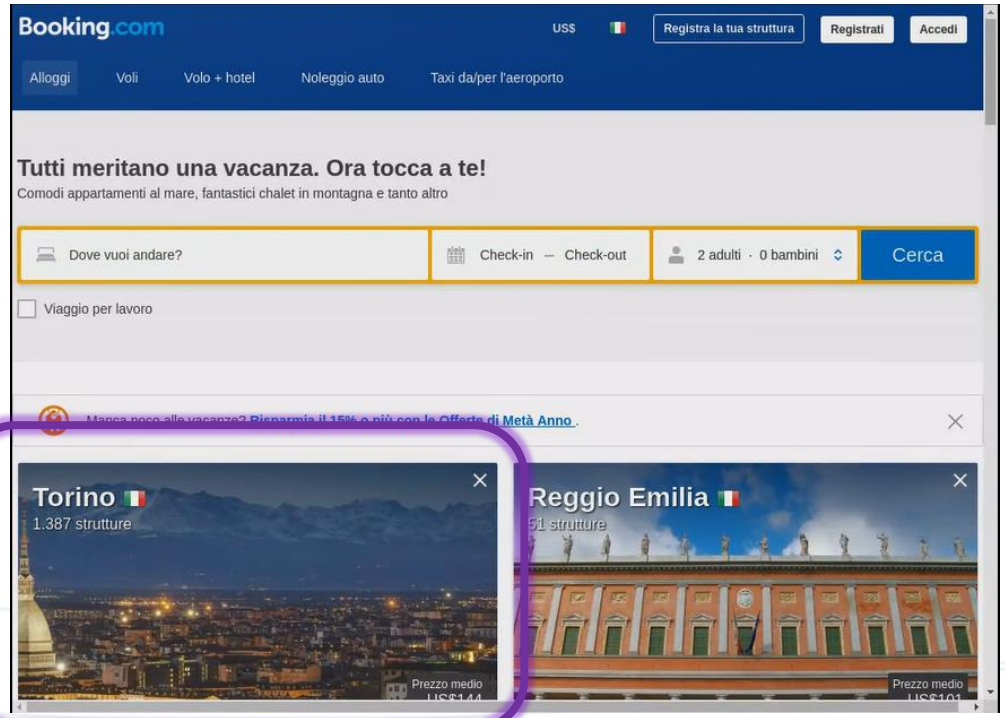
(Dataset and paper will come)

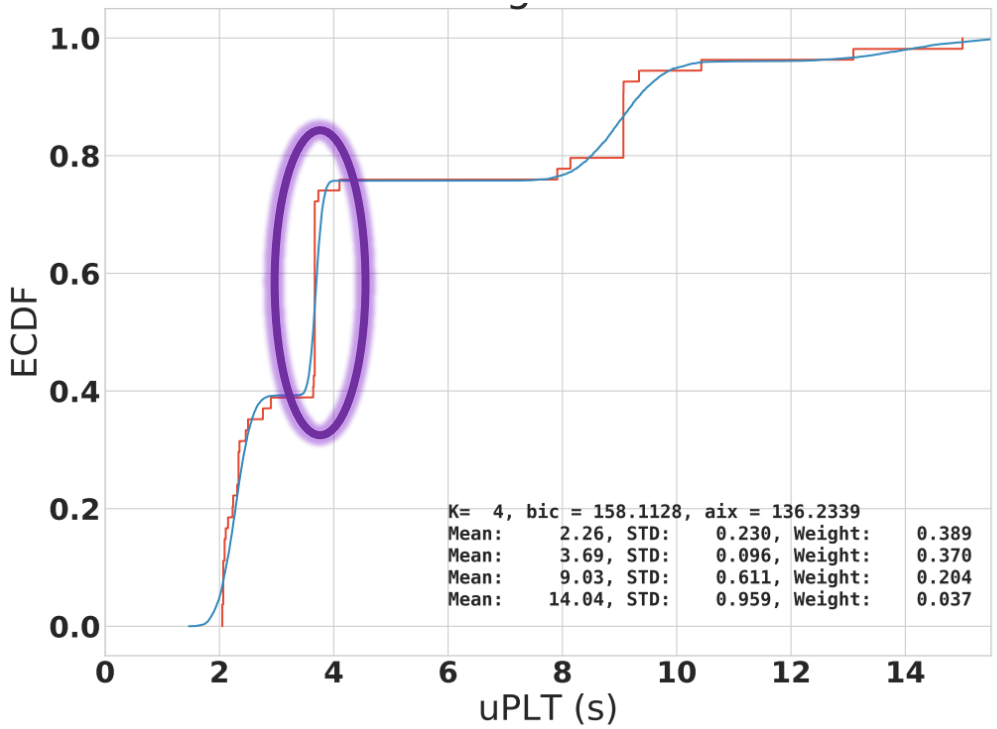
- User perceived PLT (uPLT)

"When does this page finishes loading?"

Crowdsourcing (payed crowdworkers)
- ➕ Larger userbase, but higher noise
- ➖ Side-to-side videos ≠ Web browsing!
- 🟰 Artificial controlled conditions

Collab with eyeorg





Crowdsourcing  feedback (2/3)

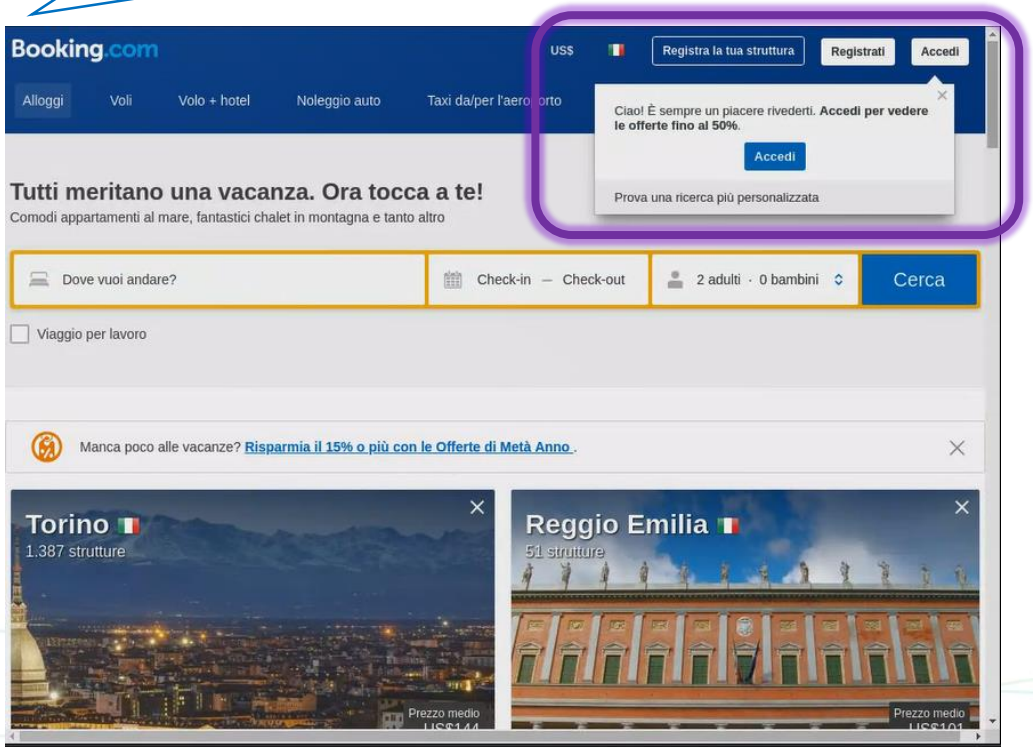
(Dataset and paper will come)

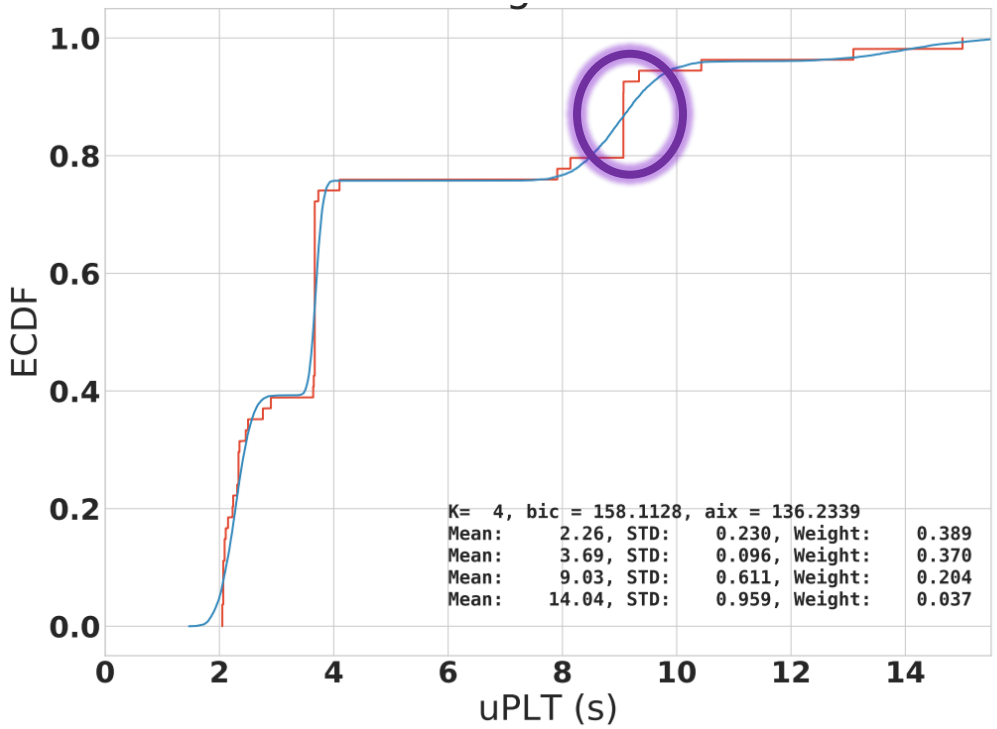
- User perceived PLT (uPLT)

"When does this page finishes loading?"

Crowdsourcing (payed crowdworkers)
- Larger userbase, but higher noise
- Side-to-side videos ≠ Web browsing!
- Artificial controlled conditions

Collab with eyeorg





Crowdsourcing  feedback (3/3)

(Paper out, dataset will come)

Collab with

WIKIPEDIA

User acceptance

"Did the page load fast enough ?" (Yes/No)

Experiments from operational website
- Actual Wikipedia users
- Browsing in typical user conditions
- Huge heterogeneity (devices/browsers/nets)

Year [ref]	Scale/heterogeneity						Experimental settings
	Lab + CW[1]	Pages	Network[2]	Sw[3]	Hw[4]	Samples	
2015 [17]	0 + 120	30	-	-	-	3.6k	ACR importance of Webpage elements
2016 [48]	100 + 1k	100	n.a.	1	1	6k	Side-by-side videos (of the same site)
2017 [13]	147 + 0	25	32	1	1	4k	Controlled browsing experiments
2017 [53]	28 + 323	28	3	1	1	2.5k	Side-by-side videos (unclear if different pages)
2017 [26]	0 + 5.4k	500	16	1	1	40k	Side-by-side videos (160 different website pairs)
2017 [30]	50 + 0	45	1	1	1	2.2k	Webcam, eye-tracking glasses
2018 [20]	241 + 0	12	n.a.	1	1	9k	Controlled browsing experiments
this study	*62k users*	*46k*	*3.8k ISPs*	*45*	*2.7k*	*62k*	User feedback from real browsing activity

[1]Crowdworkers, [2]Number of controlled network conditions, [3]Software browser, [4]Hardware device

Crowdsourcing  feedback (3/3)

(Paper out, dataset will come)

Collab with

- User acceptance

"Did the page load fast enough ?" (Yes/No)

Experiments from operational website
- Actual Wikipedia users
- Browsing in typical user conditions
- Huge heterogeneity (devices/browsers/nets)

Year [ref]	Scale/heterogeneity						Experimental settings
	Lab + CW[1]	Pages	Network[2]	Sw[3]	Hw[4]	Samples	
2015 [17]	0 + 120	30				3.6k	ACR importance of Webpage elements
2016 [48]	100 + 1k	100				6k	Side-by-side videos (of the same site)
2017 [13]	147 + 0	25	**Hard to even have have a MOS value per page ...**			4k	Controlled browsing experiments
2017 [53]	28 + 323	28				2.5k	Side-by-side videos (unclear if different pages)
2017 [26]	0 + 5.4k	500				40k	Side-by-side videos (160 different website pairs)
2017 [30]	50 + 0	45				2.2k	Webcam, eye-tracking glasses
2018 [20]	241 + 0	12				9k	Controlled browsing experiments
this study	62k *users*	46k	3.8k *ISPs*	45	2.7k	62k	User feedback from real browsing activity

[1]Crowdworkers, [2]Number of controlled network conditions, [3]Software browser, [4]Hardware device

Crowdsourcing  feedback (3/3)

(Paper out, dataset will come)

Collab with

WIKIPEDIA

User acceptance

"Did the page load fast enough ?" (Yes/No)

Experiments from operational website
- Actual Wikipedia users
- Browsing in typical user conditions
- Huge heterogeneity (devices/browsers/nets)

Year [ref]	Scale/heterogeneity						Experimental settings
	Lab + CW[1]	Pages	Network[2]	Sw[3]	Hw[4]	Samples	
2015 [17]	0 + 120	30	-			3.6k	ACR importance of Webpage elements
2016 [48]	100 + 1k	100	n.a.			6k	Side-by-side videos (of the same site)
2017 [13]	147 + 0	25	32			4k	Controlled browsing experiments
2017 [53]	28 + 323	28	3			2.5k	Side-by-side videos (unclear if different pages)
2017 [26]	0 + 5.4k	500	16			40k	Side-by-side videos (160 different website pairs)
2017 [30]	50 + 0	45	1			2.2k	Webcam, eye-tracking glasses
2018 [20]	241 + 0	12	n.a.			9k	Controlled browsing experiments
this study	*62k users*	*46k*	*3.8k ISPs*	45	*2.7k*	*62k*	User feedback from real browsing activity

Huge variability in ISP and conditions ...

[1]Crowdworkers, [2]Number of controlled network conditions, [3]Software browser, [4]Hardware device

Crowdsourcing  feedback (3/3)

(Paper out, dataset will come)

Collab with

WIKIPEDIA

User acceptance

"Did the page load fast enough ?" (Yes/No)

Experiments from operational website
- Actual Wikipedia users
- Browsing in typical user conditions
- Huge heterogeneity (devices/browsers/nets)

Year [ref]	Scale/heterogeneity						Experimental settings
	Lab + CW[1]	Pages	Network[2]	Sw[3]	Hw[4]	Samples	
2015 [17]	0 + 120	30	-	-	-		ACR importance of Webpage elements
2016 [48]	100 + 1k	100	n.a.	1	1		Side-by-side videos (of the same site)
2017 [13]	147 + 0	25	32	1	1		Controlled browsing experiments
2017 [53]	28 + 323	28	3	1	1		Side-by-side videos (unclear if different pages)
2017 [26]	0 + 5.4k	500	16	1	1		Side-by-side videos (160 different website pairs)
2017 [30]	50 + 0	45	1	1	1		Webcam, eye-tracking glasses
2018 [20]	241 + 0	12	n.a.	1	1		Controlled browsing experiments
this study	*62k users*	*46k*	*3.8k ISPs*	45	*2.7k*	*62k*	User feedback from real browsing activity

Huge variability in Browser software & devices ...

[1]Crowdworkers, [2]Number of controlled network conditions, [3]Software browser, [4]Hardware device

Crowdsourcing feedback (3/3)

(Paper out, dataset will come)

Collab with WIKIPEDIA

User acceptance

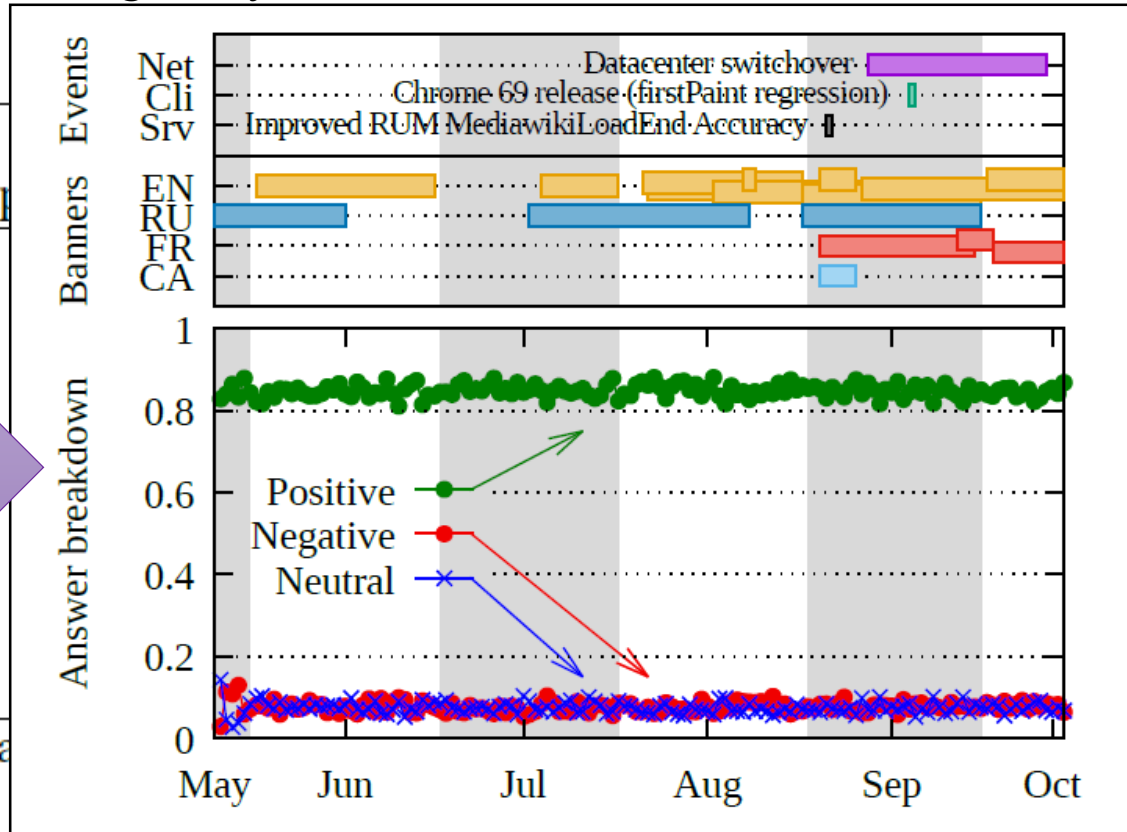
"Did the page load fast enough ?" (Yes/No)

Experiments from operational website
- Actual Wikipedia users
- Browsing in typical user conditions
- Huge heterogeneity (devices/browsers/nets)

Year [ref]	Scale/heterogeneity					
	Lab + CW[1]	Pages	Network[2]	Sw[3]	Hw[4]	Samp
2015 [17]	0 + 120	30	-	-	-	3.6k
2016 [48]	100 + 1k	100	n.a.	1	1	6k
2017 [13]	147 + 0	25	32	1	1	
2017 [53]						
2017 [26]						
2017 [30]	50 + 0	45	1	1	1	.2k
2018 [20]	241 + 0	12	n.a.	1	1	9k
this study	62k *users*	46k	3.8k *ISPs*	45	2.7k	62k

[1]Crowdworkers, [2]Number of controlled network conditions, [3]Softwa

Pretty unbalanced for modeling, but very good for accurate & factual operational insights



Crowdsourcing feedback (3/3)

(Paper out, dataset will come)

Collab with WIKIPEDIA

User acceptance

"Did the page loa ... *al website*

enough ?" (Yes/N ... *ditions*

... *s/browsers/nets)*

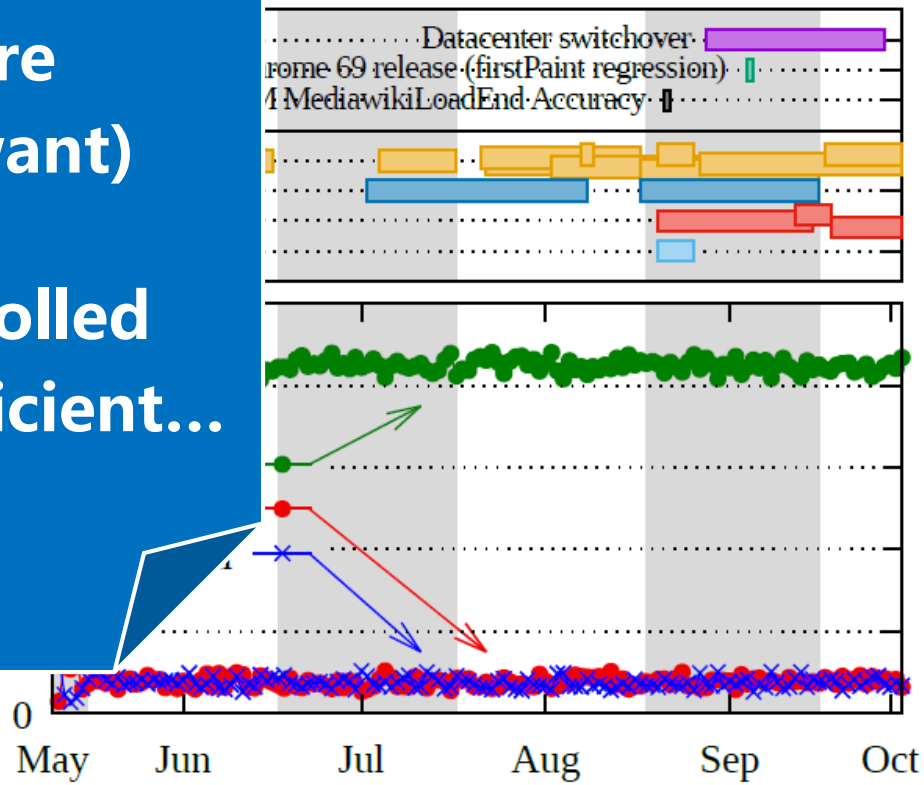
Year [ref]	Lab + CW[1]	P
2015 [17]	0 + 120	3
2016 [48]	100 + 1k	1
2017 [13]	147 + 0	2
2017 [53]	**Pretty unbala**	
2017 [26]	**for accurate**	
2017 [30]	50 + 0	4
2018 [20]	241 + 0	1
this study	*62k users*	4

[1]Crowdworkers, [2]Number of controlled network conditions, [3]Softwa

Takeaway message #2

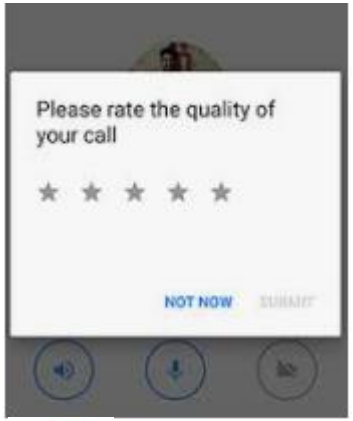
You need us(ers)more than you think (or want)

No finite amount of controlled user-panels will ever be sufficient…

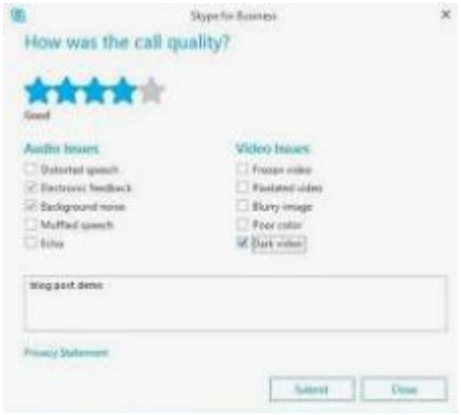


Painkiller prescription (1/3): Keep it sustained, stupid !

- Other applications/players are doing this already!

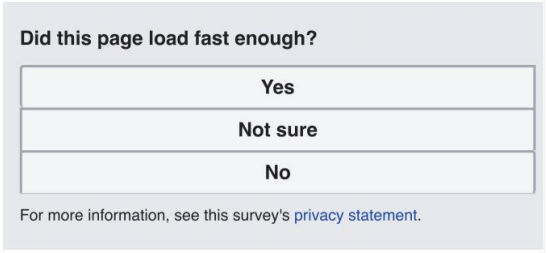


Facebook



Skype

Android



Wikipedia

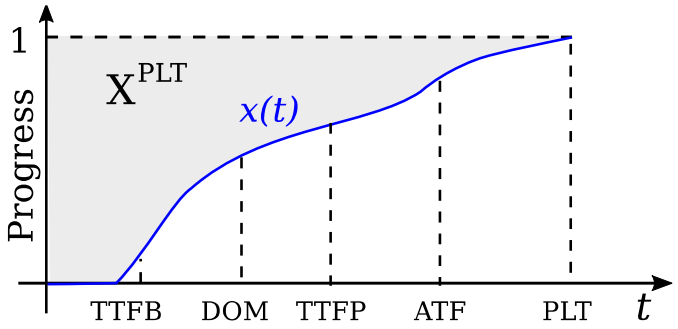
Physical world

- Sustained continuous user QoE indication benefits
  - Useful samples for QoE management assessment, troubleshooting, regression detection, etc.
  - Get continuous stream of samples for improving QoE = f(QoS) models on the long run
- Very limited downsides (risk of annoying users if leveraging small panels)

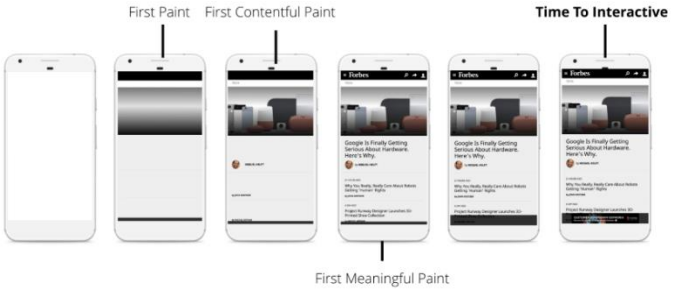
Painkiller prescription (2/3): Removing features altogether

- Expert-driven feature engineering
  - Explainable but inherently heuristic approach
  - Hard to keep in sync with application/network change
- Neural Networks
  - Less interpretable but more versatile
  - Downside: requires *lots* of samples....

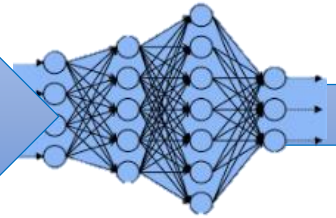
- Feed NN with *x(t) signal*
- Still lightweight



- Feed NN using a *filmstrip*
- More complex



Possible inputs



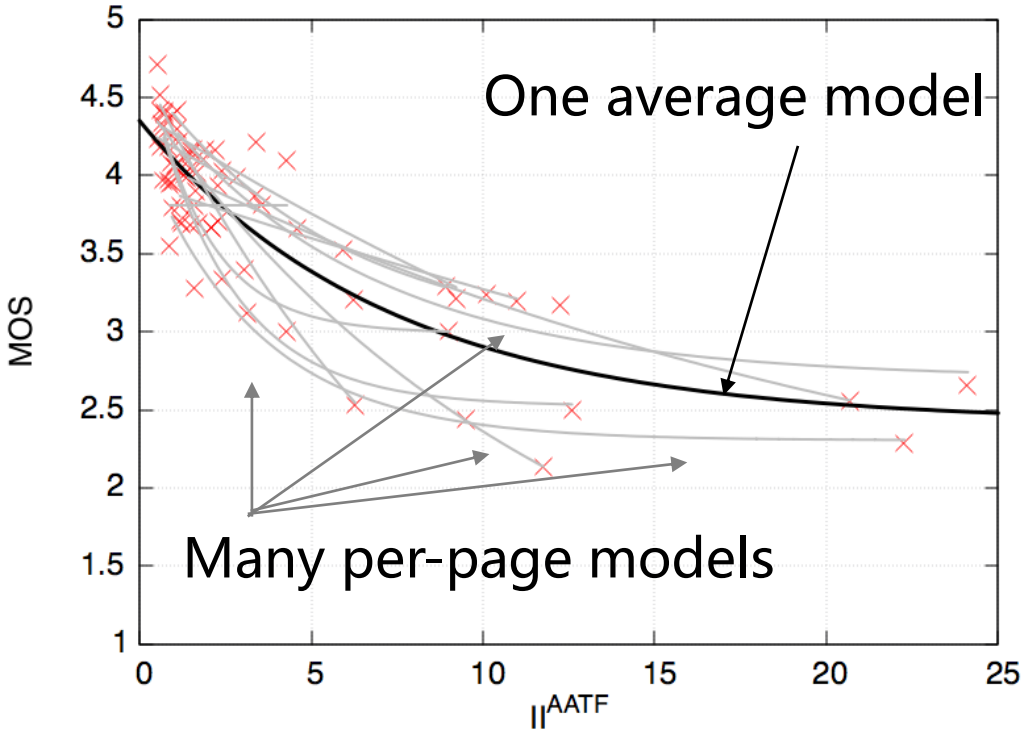
Possible outputs

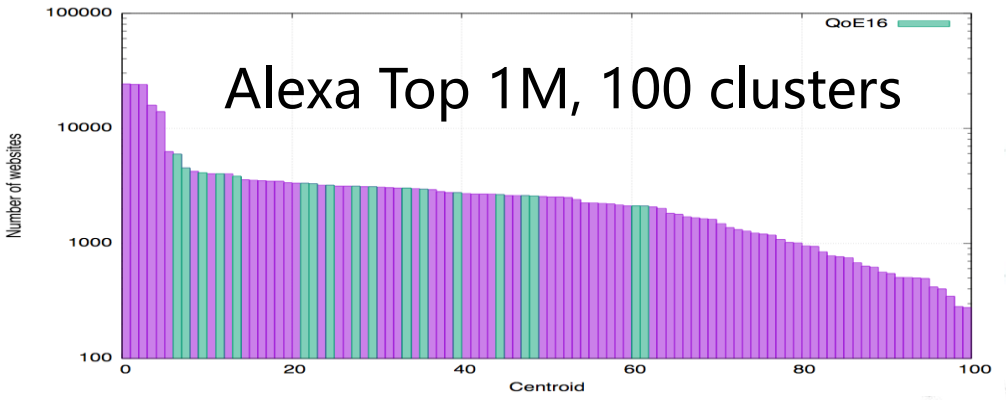
- User feedback (e.g. MOS, user PLT, etc.)
- Smartphone sensors (eg happiness estimation via facial recognition)

- Brain signals acquired with sensors
- Activity of brain areas correlated with user happiness

Painkiller prescription (3/3): Divide et impera

- World Wild Web
  - Huge diversity, not captured by single model
- Increase accuracy
  - Per-page QoE models
  - Inherently non scalable
- Increase accuracy & scalability
  - Per-page QoE models (eg Alexa top 100 pages)
  - Aggregate QoE models (eg 100 clusters top 1M)
  - Generic QoE model (for the tail up to 1B pages)





Wrap up

Move away from
feature engineering
(or be locked in forever,
for each application) !

Think hard about how
to continuously involve
users for QoE assessment,
and even more for quality
assessment of your QoE
management solution !!

A good dose of painkillers
(= continuous feedback +
automated learning +
divide et impera + …)

References

Final product placement. Dataset and code are free, though ☺

60,000+ real user grades

[WWW-19] F. Salutari, D. Da Hora, G. Dubuc and D. Rossi A large scale study of Wikipedia users' quality of Experience Proc. WWW, to appear

Chrome plugin implementation

[SIGCOMM-18] D. da Hora, D. Rossi, V. Christophides, R. Renata, A practical method for measuring Web above-the-fold time, ACM SIGCOMM Demo, aug. 2018,

[QOMEX-18] Hossfeld, Tobias and Metzger, Florian and Rossi, Dario, Speed Index: Relating the Industrial Standard for User Perceived Web Performance to Web QoE 10th International Conference on Quality of Multimedia Experience (QoMEX 2018) jun. 2018

9,000 real user grades

[PAM-18] D. da Hora, A. Asrese, V. Christophides, R. Teixeira and D. Rossi, Narrowing the gap between QoS metrics and Web QoE using Above-the-fold metrics Proc. PAM 2018, **Best dataset award ★**

[PAM-17] Bocchi, Enrico and De Cicco, Luca and Mellia, Marco and Rossi, Dario, The Web, the Users, and the MOS: Influence of HTTP/2 on User Experience Proc. PAM 2017

10,000 automated experiments

[SIGCOMM-QoE-16] E. Bocchi, L. De Cicco, D. Rossi, Measuring the Quality of Experience of Web users, ACM SIGCOMM Internet-QoE worshop 2016, **Best paper award ★**

?? || //

- Thanks for lis