STQ Workshop

QUANTIFYING NETWORK PERFORMANCE FOR EMERGING SERVICES AND NON-HUMAN USERS

Dr. Jens Berger, Rohde & Schwarz

22/11/2022

4G AND 5G NSA NETWORK POLICIES TODAY

High attention on peak data rates in todays 4G/5G networks

► Today's mobile networks are designed and **optimized for human users**

► What human users are doing today?

  ► Retrieving of web content

  ► Streaming video

  ► Posting to social media

  ► Voice and video calls over IP

  ► …

► **Operators today mostly prioritize high data rates in DL direction**
► **'Performance' is often defined by 'maximum or average bitrate'. Is this sufficient?**

4G AND 5G NSA NETWORK POLICIES TODAY

High attention on peak data rates in todays 4G/5G networks

► Today's mobile networks are designed and optimized for human users

► What will happen with next steps of 5G and even 6G?

    → Many buzz words

Public networks

Smartphone users

Verticals / Private Networks

Mission critical communication

Open-RAN

VR/AR/XR

IoT / mMTC

Vehicles

Network Slicing

3D radio network evaluation

Critical Infrastructure

Campus networks

EXPANSION OF MOBILE NETWORKS AND USE CASES

Cellular ('mobile') networks

Public networks

Private networks

Slice
(dedicated resources)

Independent ('dedicated infrastructure')

Consumer
"Traditional"

Emergeny,
eCall,…

Non-critical
applications

Critical applications
in Critical Infrastructure

Verticals

Non-critical

Smartphone

IoT Devices

"Modem"

Smartphone, IoT Devices

…as today
(Telephony, streaming, retrieving web-content, file upload, video conferencing,…)

Dominated by non-human use cases
- **Non-smartphone devices**
- **Different traffic and load patterns**
- Very wide range of applications
- Very wide range of service requirements
- Potential use of low-layer protocols

…as today
(Telephony, streaming, retrieving web-content, file upload, video conferencing,…)

Real-time interactive services
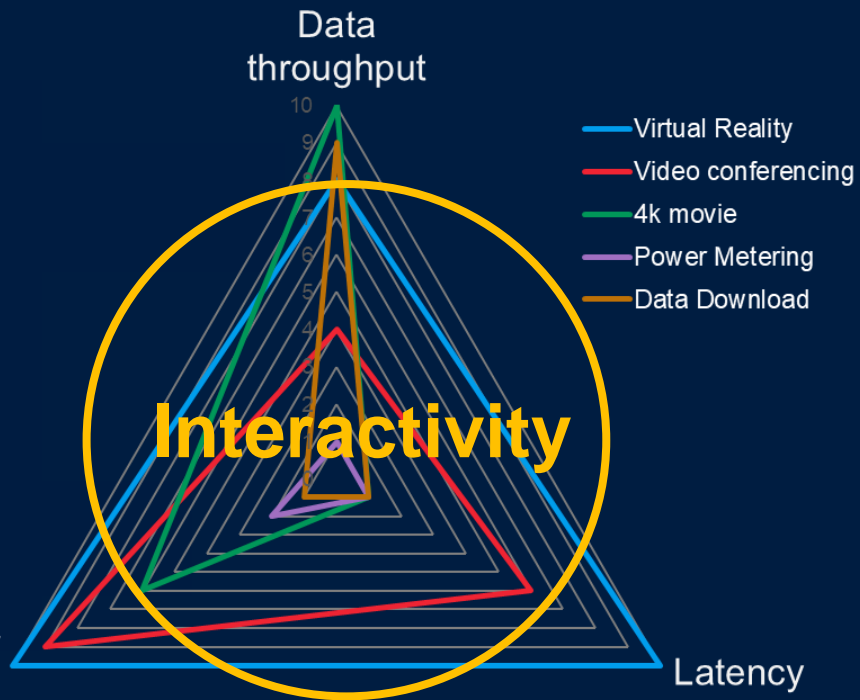(Gaming, VR/XR applications, Smart home, remote control,…

Real-time interactive services
(Gaming, VR/XR applications, Smart home, remote control,…

EXPANSION OF MOBILE USE CASES

► Real-time access and feedback

► Tons of new interactive applications

► Non-human users

► Non-smartphone access devices

► Critical applications

► What do we need?
  – More than **Bitrate**
  – **Latency** and **Reliability** are key

► Measured under **realistic load conditions**

► Demands of individual use cases are different

Data throughput

Virtual Reality
Video conferencing
4k movie
Power Metering
Data Download

Interactivity

Reliability/Continuity

Latency

CONSEQUENCES FOR TESTING NETWORK PERFORMANCE

► Technical KPIs to test?

  ► Real-time, interactive services need KPIs based on short-term evaluation

► How to define and to model QoE for new use cases and applications?

  ► Subjective Evaluation of QoE

  ► Objective QoE models for automated field testing

  ► What is about 'QoE' for non-human use cases?

► Increasing amount of testing? Many new use cases and heterogenous network topologies.

  ► Testing individual slices and network configurations (o-RAN)

  ► Many more use cases and applications (in different load situations)

► **Target is efficient network optimization, trouble shooting and proof of performance**
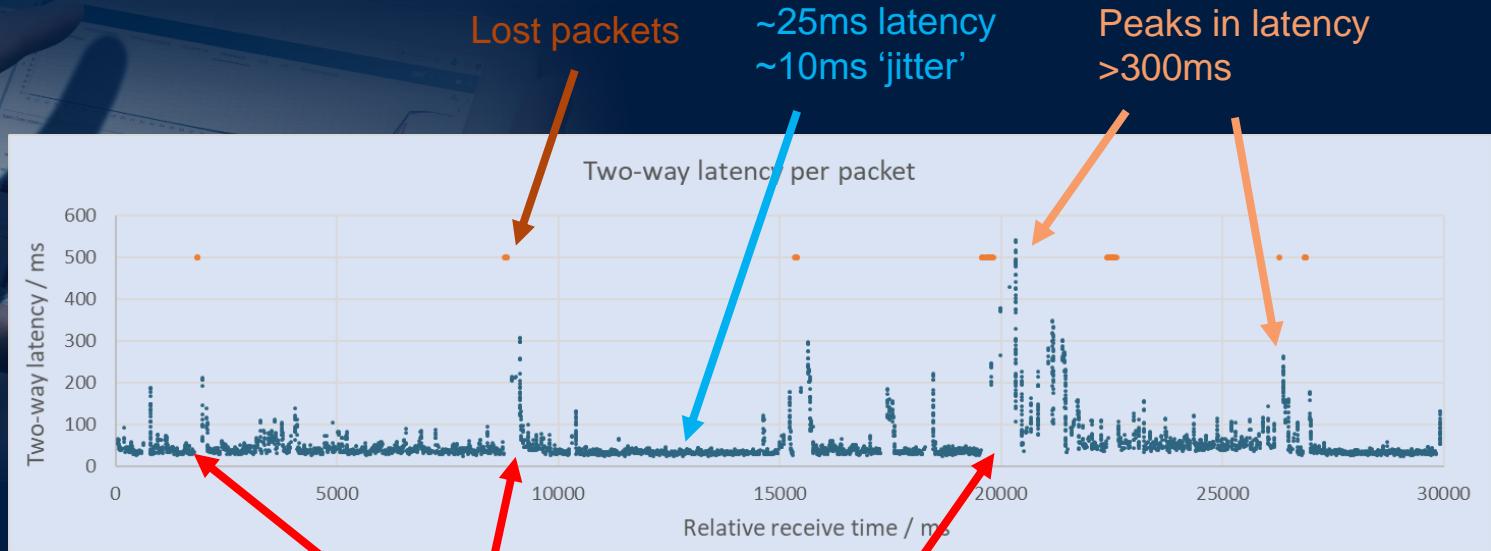
TECHNICAL PERFORMANCE – MORE THAN 'DATA RATE'

► Today's KPIs based on (relatively) long observation periods (e.g. Call setup time, MOS, Upload time, Data rate,…)

► Short term effects are masked

► No problem for file down-/upload based services (browsing, uploading photos, watching videos,…)

Data rate A = B = C

e.g. 8s

► Real-time, interactive applications are sensitive to short term degradations

► Using shorter observation periods? OK.

   ► Average, median, variation does not solve the problem

   ► KPIs describing reliability as 'Interruptiveness', 'Continuity', 'Burstiness' should be studied

Rohde & Schwarz
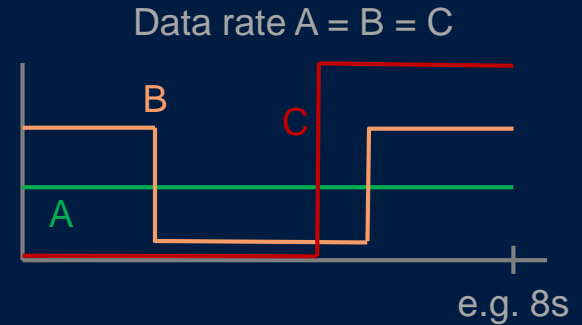
TECHNICAL PERFORMANCE – MORE THAN 'DATA RATE'
EXAMPLE: SMART LATENCY MEASUREMENT

► Measuring packet latency under realistic mobile conditions → real field UDP data stream

- Emulating real traffic by thousands of UDP packets (acc. IETF RFC 5357)
- Latency is not a constant, each packet has an individual packet transport time

Lost packets

~25ms latency
~10ms 'jitter'

Peaks in latency
>300ms

Two-way latency per packet

Interruptions / discontinuities in data flow

TECHNICAL PERFORMANCE – MORE THAN 'DATA RATE'
EXAMPLE: SMART LATENCY MEASUREMENT

► Simple statistics (average, median, standard deviation) do not describe the situation well

► Visual inspection by experts will provide much more information

► How to automatize? → Machine Learning models / Artificial Intelligence.
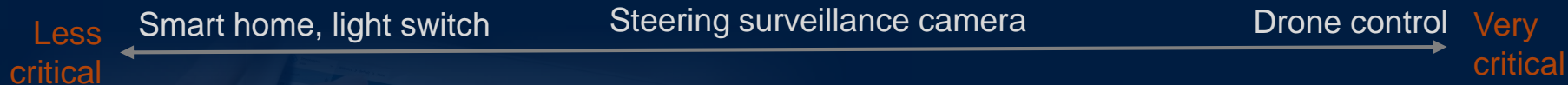
All happens in 30s while driving or walking
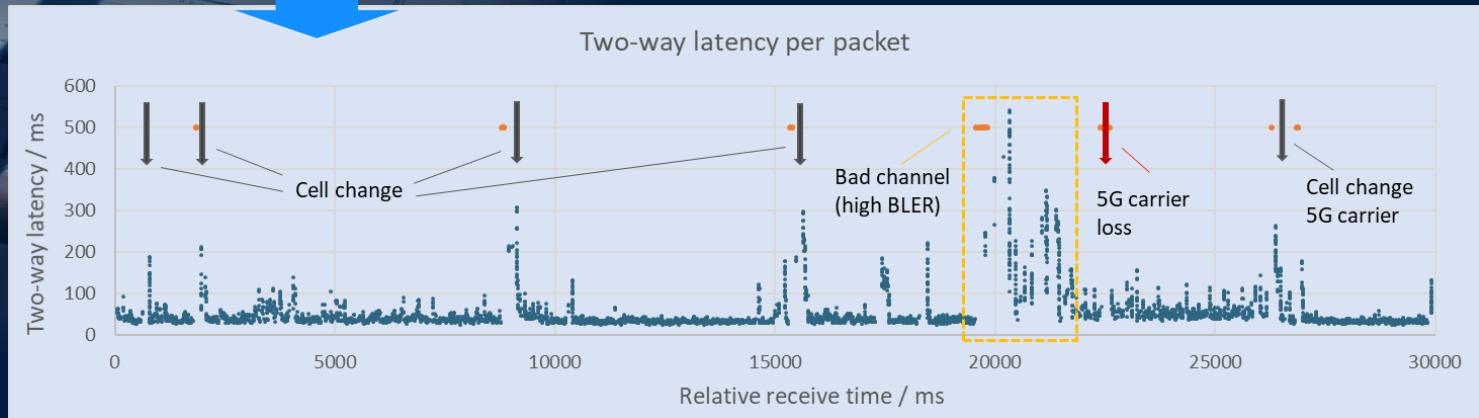
QoE FOR NEW USE CASES AND APPLICATIONS?

► QoE as an integrative measure based on (human) experience

    ► Target: Objective QoE models for automated field testing based on technical measurements

    ► Start: Subjective Evaluation of QoE

► 'Today' there is a narrow scope of services and a similar expectation by humans

    ► Example Voice telephony:

        ► Well defined and narrow situation (listening to a remote person) by handset or headphone

        ► Clear expectation about human voice or waiting time for setup

        ► Relatively small dependency on scoring person and content/context

    ► Similar for watching (2D-) videos or browsing / retrieving web-content

► Is it the same as narrow for a VR application, cloud gaming or remote control of something?

QoE FOR NEW USE CASES AND APPLICATIONS

► QoE depends strongly on use case / content (e.g. game and scene) and expectation

Remote control a device

Less critical   Smart home, light switch   Steering surveillance camera   Drone control   Very critical

Cloud gaming

Less critical   Online chess   Settlers   Ego shooter   Very critical

► QoE can be evaluated and modelled for a given setup (app, game, scene,…)

► Can we generalize QoE for an entire use case as e.g. Cloud gaming or Virtual Reality?

► Would a balanced mixture of scenes / games help?

    → *No. Average results does not match to full range of use cases*

QoE FOR NEW USE CASES AND APPLICATIONS

► Clustering of use cases and applications

→ Based on technical demands on the network

► QoE evaluation (<u>and modelling</u>) per cluster

Reliability

Data rate

Less critical

Very critical

Latency
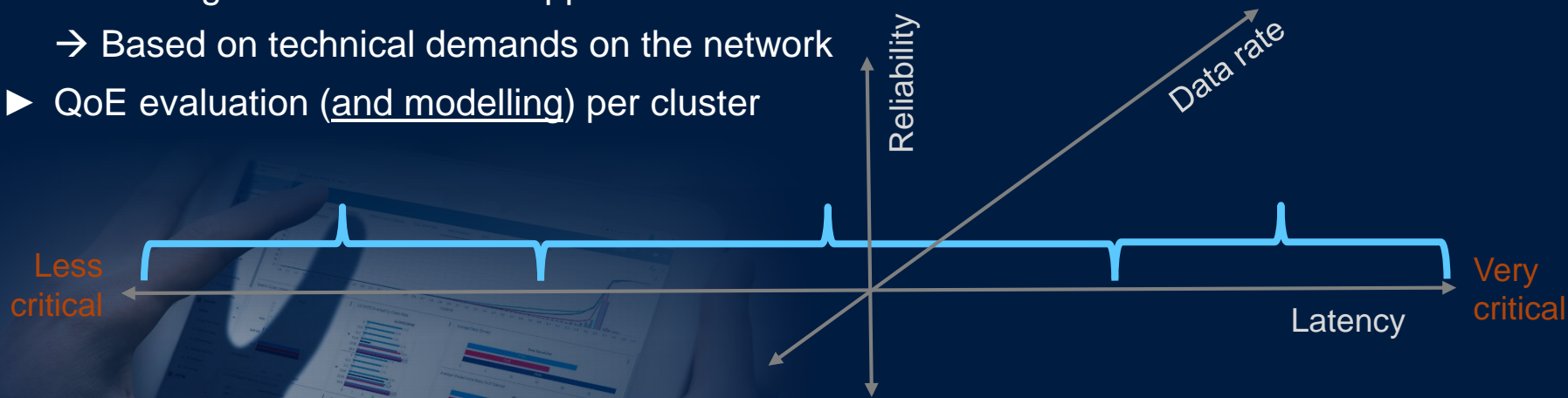
► Per cluster modelling and evaluation

► typical, popular application ('archetypes')

► generic application

► proxy tests

► **Traffic (load) pattern represents 'all' apps of the cluster**

► **Required KPIs for the QoE model must be obtainable**

QoE FOR NEW USE CASES AND APPLICATIONS

► Clustering of use cases and applications

  → Based on technical demands on the network

► QoE evaluation (<u>and modelling</u>) per cluster

Reliability

Data rate

Less critical

Very critical

► **How to quantify integrative performance for non-human use cases?**

  ► We need more than 'pass/fail' for individual technical KPIs

  ► Can we develop a 'machine-QoE' ?

  ► By expert knowledge or
   e.g. a fabrication tolerance in an automated production process ?

► Per cluster modelling and evaluation

  ► typical, popular application ('archetypes')

  ► generic application

  ► proxy tests

► **Traffic (load) pattern represents 'all' apps of the cluster**

► **Required KPIs for the QoE model must be obtainable**
  (real apps are often black-box)

CONSEQUENCES FOR TESTING NETWORK PERFORMANCE

▶ Increasing amount of testing? Many new use cases and network topologies.

▶ More test effort in the field? More test equipment?

▶ **More efficient data collection and analysis!**

   ▶ Generalize test cases

      ▶ Use of 'archetype' or generic tests instead of dozens different apps doing the same

      ▶ Definition of proxy tests emulating real applications' traffic and behavior

      ▶ Providing more details by grey- or glass-box implementations

      ▶ Obtaining more results and KPIs per test

   ▶ Use machine learning and artificial intelligence for new type 'smart-KPIs'

      ▶ Smart-KPIs provide more and better insights than legacy KPIs

      ▶ Smart-KPIs can be tailored for trouble-shooting and optimization
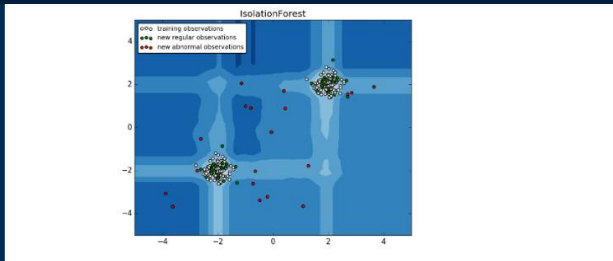
MACHINE-LEARNING BASED MOBILE NETWORK ANALYSIS
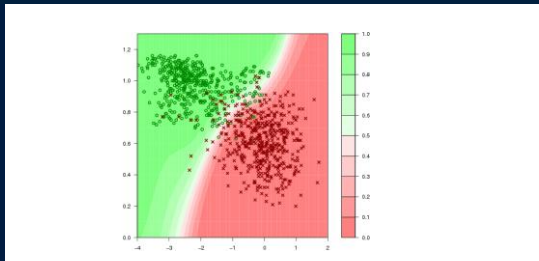
CALL STABILITY SCORE (CSS)

Scoring network stability
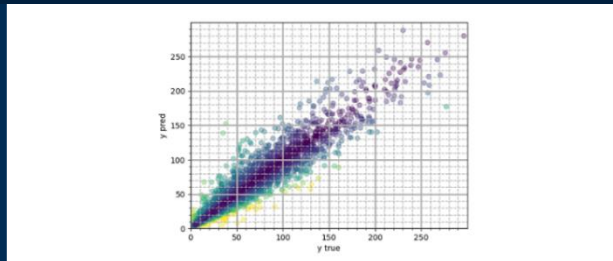**Probability of a call drop or fail**

► Pro-active recognition of problematic areas before drop or failed calls happen

ANOMALY DETECTION (AD)

Guided network optimization
Automatically detected anomalies in the network

► Self-detection unusual situations in the network

► Combined w/ root cause nalysis

NETWORK UTILIZATION RATING (NUR)

Rating local performance relative to 'state of the art' under comparable radio conditions and allocated resources

► Rating exhausted network resources

► Improvement potential without additional resources

► CSS, NUR and AD are a **revolutionary a new type of innovative 'meta KPI'**

► Prediction of a comprehensive value that does **not exist today** based on multiple input features

► If developed, understood and applied correctly → Much more efficient network optimization and trouble-shooting

EXPECTED MAJOR TRENDS TO BE CONSIDERED FOR 5G

► Telephony / conferencing
  – Stays a relevant service, transition to **VoNR** requires field testing
  – **OTT VoIP** telephony services become equivalent in use and quality
  – Telephony extends towards **interactive and audio-visual** experience

► Media / Web content delivery
  – Streaming stays very relevant, transition to **interactive use cases**
  – Online and **cloud gaming** require very short interaction times
  – Social Media become more **interactive**
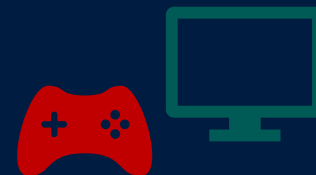
► 5G beyond smartphones
  – Mobile networks as access point for home networks
  – Office applications, **connected cars**, **non-smartphone devices**
  – **Private 5G networks**

THANK YOU!