Interactivity Assessment of Streamed Games over Heterogeneous Access Networks using Bayesian Networks

Henrique Souza Rossi*, Karan Mitra*, Christer Åhlund*, Niclas Ögren†, Per Johansson†
*Mobile and Pervasive Computing, Department of Computer Science, Electrical and Space Engineering,
Luleå University of Technology, Skellefteå, Sweden
†Infovista, Skellefteå, Sweden
Email: henrique.souza.rossi@ltu.se

***Abstract*—Interactivity is a metric that measures the level of control or manipulation users may exert over a system, software, or service. It is considered a key dimension in cloud-gaming services, measuring how effectively users can control and respond to game events in real-time. Although widely acknowledged by standards such as ITU-T Rec. G.1051, G.1072, its relationship to other quality dimensions such as video, audio, and overall Quality of Experience (QoE), remains not thoroughly examined. In this paper, we present a novel Bayesian Network-based framework to model and analyze interactivity alongside other quality factors under varied network conditions. Our method enables probabilistic inference and sensitivity analysis across perceptual variables, offering explainable insights into their mutual influence. Using data from two (*N1*=30, and *N2*=31 subjects) subjective studies — one for Virtual Reality Cloud Gaming (VRCG) and another for Mobile Cloud Gaming (MCG) — we show interactivity is most sensitive to round trip time (RTT) but resilient to jitter (RJ) effect. Further, an assessment of the seven interactivity-only variables shows that their distributions change uniformly under varying network conditions, suggesting that interactivity may be captured by a single metric. Finally, sensitivity analyses indicate that QoE is a more representative metric than interactivity for quality assessment in cloud gaming over heterogeneous access networks.**

***Index Terms*—Bayesian Network, Subjective tests, Interactivity, Quality of Experience, Quality of Service, Cloud Gaming, Heterogeneous Access Networks, Virtual Reality, Mobile Cloud Games**

I. Introduction

Interactivity is a metric that measures the quality of real-time manipulation of digital content between users and technology [1], [2]. Steuer, Biocca, and Levas define it as *"the extent to which users can participate in modifying the form and content of a mediated environment"* [3]. Currently, stakeholders such as network operators, application developers, and service providers are increasingly interested in understanding, measuring, and predicting this metric for emerging services, as these factors affect service adoption and revenue. Understanding these factors enables operators to optimize resource allocation and ensure efficient, user-centric network performance. In particular, cloud gaming (CG) services have emerged as a compelling platform for stakeholders to deliver gaming experiences without relying on high-end local hardware. With the market projected to reach USD 121.77 billion by 2032, CG will increase demand for low-latency, and immersive services [4]. What is more, CG is a key enabler pervasive computing for games: they can be played anywhere (e.g., location and device agnostic) and anytime. However, cloud gaming services are highly sensitive to quality of service (QoS) degradations since games demand real-time and fast responses to control and react to game events [5], [6].

For these services to achieve their full pervasiveness, the stakeholders must find ways to optimize their infrastructure to better support them, especially in QoS-sensitive heterogeneous access networks (HANs) such as 4G,5G, Wi-Fi, wired [7]. The ITU-T Rec. G.1051 Recommendation [8], recent works from [9], [10] and others [11]–[15], have attempted to measure how QoS factors such as RTT, random jitter (RJ), and/or packet loss (PL) affects interactivity metric, and also the standard quality of experience (QoE) metric, both for cloud-gaming services. These studies consider and report interactivity and QoE as separate perceptual metrics. Their results indicate that each metric is affected differently by QoS degradations, suggesting that users perceive these impairments in different ways [10], [11], [13], [15]. Furthermore, both QoE [6], [11], [16], [17] and interactivity [8], [9], [11] have been the focus of academic and industry efforts to develop predictive models based on underlying QoS parameters - models that are increasingly valuable to stakeholders seeking to optimize CG service performance and user satisfaction.

Although the literature treats interactivity and QoE as distinct constructs or metrics, we note a research gap in studying the relationship between them. Interactivity is typically assessed through specific constructs related to responsiveness [10], [11], [15], [18], [19], control [13], [20], and input–action mapping [21], [22], in contrast to QoE, which represents a broader, aggregated perception of overall service quality [10], [14], [17]. It remains unclear from previous work whether a relationship exists within interactivity constructs, between interactivity and QoE metrics, the magnitude and direction of these relationships, and their sensitivity towards QoS degradations in CG services. This leads to several important research questions: *Should stakeholders prioritize interactivity,*

978-3-903176-75-1

QoE, or both when provisioning cloud gaming services over HANs? Among the various interactivity dimensions, which should be prioritized in modeling and optimization efforts? Are interactivity and QoE influenced by QoS degradations in similar or fundamentally different ways? Should the research community invest in developing context-specific interactivity models tailored to cloud gaming use cases such as VR, PC, and mobile platforms?

This paper provides clarity on the aforementioned questions by proposing, for the first time, Bayesian Network (BN) probabilistic analyses that can model jointly the complex and hard-to-assess dependencies between interactivity (responsiveness, control, and feedback), audio, video, and QoE dimensions. BNs offer an explainable and interpretable framework that enables the inference of network conditions' effects on several quality dimensions [23]. Our BN models are evaluated through large-scale subjective tests considering both mobile and virtual reality (VR) cloud-gaming scenarios, revealing how network impairments affect not just isolated variables but their perceived interaction – an important step towards a holistic understanding of the interactivity construct in cloud gaming. In particular, our contributions are two-fold:

- We introduce a new Bayesian Network framework that jointly models interactivity, video, audio, and QoE, learning their complex inter-dependencies under stochastic network conditions such as RTT, jitter, and PL. Our approach considers two cloud-gaming datasets and provides a holistic, probabilistic interpretation and explainable structure to understand how impairments affect various quality dimensions.
- We provide a detailed investigation of interactivity by modeling and analyzing seven perceptual factors (e.g., control precision, input feedback, responsiveness) using Bayesian Networks that have never been considered before.

This paper is structured as follows: Section II reviews related work. Section III describes the datasets and experimental setup. Section IV outlines the Bayesian network modeling framework. Section V presents the main results and analysis. Finally, Section VI discusses the findings, and Section VII concludes with future work.

II. Related Work

The review of Weber *et al.* [18], studied and identified important game related interactivity dimensions: controller responsiveness, game features, customization, artificial intelligence, exploration, and perceptual persuasiveness. These were validated through questionnaires for users playing a custom first-person shooter game with various interactivity features. In parallel, Möller *et al.* [24] and ITU-T P.809 [25] studied and defined "interaction quality" as a central component of the QoE metric in games, drawing conceptual similarities with the notion of "playability" [26]. Interaction quality captures both the input (user-to-system) and output (system-to-user) experience, incorporating variables such as responsiveness, control precision, visual feedback, and multi-modal feedback fidelity. Their view of interaction is broader in its dependency on other quality factors such as video and audio.

TABLE I: Common influential factors on game interactivity.

Influencial Factors	Authors	Influencial Factors	Authors
Audio Quality	[17], [19], [24]	Network Jitter	[8], [9]
Video Quality	[17], [19], [24] [10], [15]	Network Delay	[8], [12], [14] [9], [10], [15]
Responsiveness	[11], [18] [10], [15], [19]	Packet loss	[8], [9] [14], [15]
Level of Content Control	[13], [20]	Bandwidth	[17], [19], [24] [9], [10]
Mapping	[21], [22]		

Multiple studies have evaluated interactivity based on network-layer QoS factors such as delay, jitter, and packet loss (PL), which affect responsiveness and perceived control. Zinno *et al.* [9] and ITU-T G.1051 [8] propose interactivity scoring models based solely on QoS metrics (RTT, PL, RJ) for cloud gaming. Ida *et al.* [12] show that delay impairs responsiveness and synchronization in collaborative VR and FPS games. In cloud gaming, Zadtootaghaj *et al.* [11] demonstrate that frame rate and bitrate influence control and immersion—both affecting perceived interactivity. Möller *et al.* [19] used a multidimensional QoE questionnaire to show that bandwidth impacts video and audio quality, while responsiveness drives control satisfaction. Wahab *et al.* [14] demonstrated that under delay and packet loss, responsiveness affects perceived QoE more than video quality, indicating that control factors dominate user perception during network issues. In VRCG context, Lee *et al.* [10] showed that interaction quality degrades differently from video and audio under network impairments, with QoE sometimes remaining high despite poor interactivity. Li *et al.* [15] found that minimizing interaction delay is essential to user experience in VRCG. In summary, although no studies directly explore the link between video, audio, and interactivity, findings show that interactivity is highly sensitive to network conditions.

Beyond network factors, controller action mapping plays a key role in interactivity. Shafer *et al.* [22] and Seibert and Shafer [27] show that "natural mapping" — where inputs intuitively match in-game actions — enhances perceived interactivity, spatial presence, and enjoyment. Juvrud *et al.* [20] find that real-time control increases arousal and engagement compared to passive viewing. Similarly, Schmidt *et al.* [13] show that passive tests can overestimate QoE by ignoring interactivity effects.

Game Interactivity Factors: Based on the state-of-the-art review, we have consolidated the factors commonly explored in interactivity game research and identified them as possible key influential factors, presented in Table I. We refer to the definitions of 'action mapping', 'level of content control', and 'responsiveness' as analogous to Steuer *et al.* [3] three fundamental interactivity requirements: 'mapping', 'range', and 'speed'. These factors have been studied separately [11], [14], [19]–[22], [27] and largely under QoS factors degradations [11], [14], [19] but always with a focus on assessing their influence independently. In addition, some studies suggest that perceived interactivity is influenced by control responsiveness as well as audio and video quality, with all three dimensions affecting the user's overall experience [11], [24], [25].

Interactivity Knowledge Gaps: While there is growing evidence that context factors from the network, device, and application layers can dynamically affect user-perceived video, audio, and interactivity quality, the interdependence among these modalities remains insufficiently understood. Clarifying these relationships is key to building a holistic view of interactivity measurement, leading to more robust, network-aware models that can guide future optimization and quality provisioning for CG services. *To the best of our knowledge: (1) no prior research has systematically examined the internal dependencies among core interactivity dimensions—such as action mapping, content control, and responsiveness; (2) nor has any study formally analyzed the mutual influence between interactivity, QoE, video, and audio quality in a probabilistic framework—particularly how changes in one dimension affect inferences about others.* Existing works typically evaluate these factors in isolation [11], [14], [19]–[22], [27], limiting our understanding of their combined effect on perceived interactivity and QoE. In this paper, we address these gaps by investigating which interactivity dimensions are most sensitive to QoS degradations, whether interactivity and QoE are affected in similar ways, and whether both should be prioritized by stakeholders. To do so, we present the first probabilistic joint assessment of seven interactivity components (responsiveness, immediate feedback, controllability) and their dynamic interdependence with QoE, video, and audio quality in both mobile cloud gaming (MCG) and virtual reality cloud gaming (VRCG) contexts.

III. Methods

To address the aforementioned research gaps, we investigate two of the most challenging contexts for cloud gaming services: VRCG and MCG. VRCG involves the cloud-based streaming of games to virtual reality devices, enabling highly immersive user experiences. However, this very immersion makes VRCG particularly susceptible to users' perception of quality degradation, as even minor streaming quality disruptions can break the sense of presence [10]. MCG, on the other hand, focuses on gaming using smartphones — the most widely used gaming platform — often accessed over mobile networks [7]. These networks are stochastic and can significantly impact the user's cloud gaming experience [5], making MCG and VRCG critical and dynamic use cases to study. To this end, we used two datasets from [28], [29], referred to as VRCG and MCG, respectively. Although those works focused solely on QoE – with setup and analyzes detailed therein – this paper extends the analysis by incorporating unpublished data on VideoQ, AudioQ, and Interactivity, *not previously reported in our studies*.

A. Datasets:

The *VRCG dataset* examines the impact of mobile network conditions— RTT, PL, RJ,and the combination (RTT, PL)—on user-perceived QoE, video, audio, and interactivity. A total of N=28 network conditions were tested, each rated individually by all participants (see Table II). N=30 participants, recruited from their university across various programs (age range: *18-35*, gaming experience: *21%-no, 60%-intermediate, 19%-advance*). The users played *Serious Sam VR*[1], using *Oculus Quest 2* and VR gamepads while standing in a lab environment. The game required players to shoot waves of incoming enemies from multiple directions, without physical locomotion. For **MCG**, we used a similar set of network conditions (RTT, PL, RJ), with additional levels (see Table II), totaling N=23 conditions, tested one at a time by each user. N=31 participants (age range: *18-35*, gaming experience: *16%-no, 65%-intermediate, 19%-advance*) they were recruited from the same campus, but different user pool from VRCG. They played *CS:GO version 1*[2], using an Xbox controller (supported by the game natively) and an Asus ROG Phone 5 Pro. This setup represents a typical CG on portable devices. In addition, CS:GO gameplay involved eliminating enemies on a fixed map. For both datasets, RJ (jitter) conditions follow a normal distribution, and PL conditions have a constant probability of loss. We selected CS:GO and Serious Sam VR because shooter games demand fast, precise responses [30]. Hence, users would be able to perceive network degradations and effect on quality clearly.

TABLE II: Emulated network conditions in both datasets, using NetEm, bi-directionally applied. Baseline: RTT=4ms.

QoS Factor (unit)	Both	VRCG	MCG	Total
RTT (ms)	4, 27, 52, 402	77, 177, 277, 352	102, 202, 302	**11**
PL (%) with RTT=4ms	6, 24	12	45	**4**
[RTT (ms); and PL (%)]	[27; 2], [27; 4], [52; 2], [52; 4]	[27; 6], [52; 6], [77; 2], [77; 4]	[27;0.2], [52;0.2], [102;0.2], [102;0.2] [102;4]	**13**
[RTT (ms); and RJ (std)]	[52; 3], [52; 6]	[27; 1], [27; 3], [27; 6], [52; 1], [77; 1], [77; 3], [77; 6]	[52; 9], [52; 12]	**11**
Total	**12**	**16**	**11**	**39**

TABLE III: Questions used to assess interactivity and other quality factors based on [31]. Both=VRCG,MCG

Code	Question	Gane
VideoQ	Overall video quality	Both
AudioQ	Overall audio quality	Both
QoE	Overall gaming experience	Both
CN1	Felt in control of game	Both
IF2	Immediate action notification	MCG

Code	Question	Game
IF1	Immediate feedback on actions	MCG
CN2	Control over interface/devices	MCG
CN3	Control over game actions	MCG
RE1	Input responsiveness met expectations	MCG
RE2	Inputs applied smoothly	MCG

B. VRCG and MCG Experiment Design:

In both user experiment groups, a within-subject design was followed, where each user experienced and rated all the network conditions. The conditions were randomized following

[1]https://store.steampowered.com/app/465240/ [Access date: June 2025]
[2]https://store.steampowered.com/app/730 [Access date: June 2025]

a balanced latin square design. For both experiments, before the start of the test, users were requested to sign a consent form, followed by a demographics and game warm up session of 5 minutes. The duration of each test was set at 90 seconds following ITU-T P.809 [25]. Regarding the quality question measurement, for both experiments, they used a questionnaire proposed in [31]. Both VRCG and MCG experiments included questions on video quality (VideoQ), audio quality (AudioQ), overall quality of experience (QoE), and controllability (CN1).

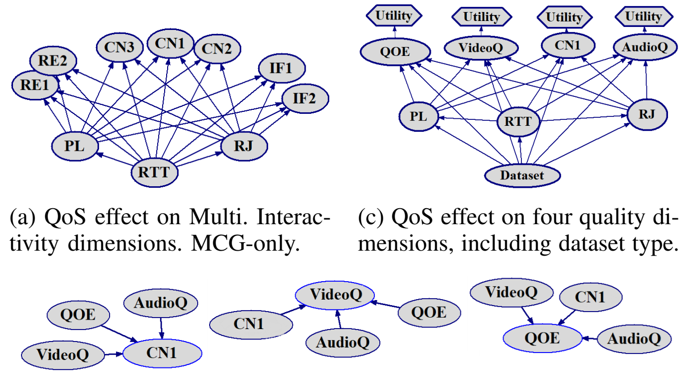
For the MCG experiment, seven interactivity-related questions were evaluated and mapped to three core constructs from previous studies as shown in Table III and Table I. The Responsiveness group (RE1, RE2), which captures input delay and motion smoothness, corresponds to the "Responsiveness" factor frequently cited in the literature [12], [19]. Immediate Feedback (IF1, IF2), reflecting the system's reactivity to user input, aligns with the "Mapping" construct—how well user actions are translated into outcomes [11]. The Controllability items (CN1–CN3), which assess the user's sense of control during gameplay, correspond to the "Control" dimension [20], [22]. In both VRCG and MCG the questions had a discrete 5-point scale ("very-poor", "poor", "Average", "Good", "Excellent") consistent with ITU P.809., and detailed in Table III.

IV. Probabilistic Modelling of QoE, Interactivity, and Media Quality in CG

A. Requirements:

Given the breadth of subjective responses collected across both experiments, we aim to understand how these perceptual variables relate to one another and how they are influenced by varying mobile network conditions. To support this analysis, the modeling framework must(1) capture the inherent probabilistic nature of user ratings, which typically appear as score distributions over discrete scale options (e.g., 1–5) [32], (2) handle both linear and nonlinear influences, since the relationship between QoS parameters and perceived quality may not be monotonic or threshold-dependent [6], [16]; (3) capture complex dependencies between multiple QoS and quality variables; (4) provide explainable results and; (5) enable bidirectional reasoning, both from QoS effect on quality factors and underlying QoS requirements given quality states.

We chose Bayesian Networks (BNs) for this study due to their ability to model uncertain, complex, and interdependent relationships among perceptual quality factors in a probabilistic yet interpretable way [23]. The most widely used regression methods offer prediction and interpretability but do not model causal or probabilistic dependencies between several factors [33]. Hidden Markov models are restricted to sequential data, and while neural networks are powerful, they often behave as black boxes with limited transparency. Mitra *et al.* [23] and Tasaka *et al.* [34] demonstrated that BNs provide a better combination of interpretability, cause-effect reasoning, and structural explanatory modeling than regression and black-box models in QoE contexts. In our case, BNs - being probabilistic graphical models, model how interactivity, QoE, video, and audio qualities are influenced by QoS factors (RTT, jitter, PL),



(a) QoS effect on Multi. Interactivity dimensions. MCG-only.

(c) QoS effect on four quality dimensions, including dataset type.

(b) Interdependence among four quality dimensions.

Fig. 1: BN models proposed to study QoS effect on quality dimensions on VRCG and MCG datasets. Analyses in Sec. V

supporting both top-down (e.g., fixing QoE to "Excellent" to infer required QoS) and bottom-up reasoning (e.g., assessing the impact of RTT)— a capability which is not supported by methods such as Belief Rule-Based systems, or fuzzy logic, which are limited to one-way rule evaluation.

B. Bayesian Network Design:

A BN can be defined as follows: **Definition 1.** *"A Bayesian network (BN) is a directed acyclic graph (DAG) where random variables form the nodes of a network. The directed links between nodes form a conditional dependency relationship. The direction of a link from X to Y means that X is the parent of Y. Any entry in the Bayesian network can be calculated using the joint probability distribution (JPD) denoted as:"*

$$P(x_1, \ldots, x_m) = \prod_{i=1}^{m} P(x_i \mid \text{Parents}(X_i)) \tag{1}$$

where, parents nodes X_i is the parent of node x_i [33]. To apply BNs to our datasets, we follow the design of *CaQoEM*, a context-aware method for BN-based QoE modeling, measurement, and prediction [23]. In their work, the authors combined BNs and utility theory to model and infer QoE under uncertain and dynamic mobile computing conditions. They structured the BNs with *nodes* representing both contextual parameters (e.g., network delay, jitter, packet loss) and QoE-related factors such as user satisfaction and technology acceptance. Each variable was defined as a *discrete node*, with context attributes folded into categorical states, and quality factors mapped to states $s_1, s_2, \ldots, s_n$ reflecting their scale values (e.g., Likert scores from "Very-Poor" to "Excellent"). We extend their design as follows:

- We incorporate four quality dimensions (VideoQ, AudioQ, QoE, and CN1 or "Interactivity" for both VRCG and MCG datasets, and seven interactivity-specific factors for MCG.
- We propose three BN models shown in Fig. 1 to study: Fig.1a how QoS factors affect 7 interactivity dimensions in MCG; Fig.1b the mutual relationship of the 4 quality

TABLE IV: QoS state discretizations grouping tested network conditions by effect and proximity. Tot.Count = number of tested conditions per state. See conditions in Table II

State	Has Conditions Where	Tot. Count
RTT=d4	RTT=4ms	5
RTT=d5_d27	RTT=27ms	8
RTT=d28_d52	RTT=52ms	10
RTT=d53_d102	RTT=77;102ms	10
RTT=d103_d202	RTT=177;202ms	2
RTT=d203_up	RTT=277;302; 352;402ms	4

State	Has Conditions Where	Tot. Count
PL=p0	PL=0%	22
PL=p02_p6	PL=0.2; 2; 4; 6%	14
PL=p7_p25	PL=12,24%	2
PL=p26_up	PL=45%	1
RJ=std0	RJ=0	28
RJ=std1_std3	RJ=1,3	7
RJ=std4_up	RJ=6,9,12	4

dimensions; and Fig.1c how QoS affects the 4 quality dimension across VRCG and MCG.

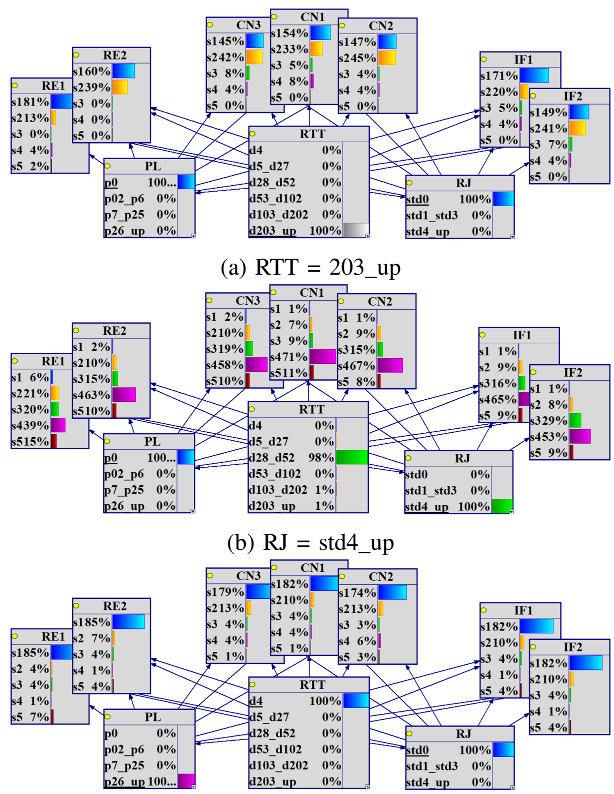
We constructed the Bayesian Network structures in Fig. 1 based on empirical dependencies observed in our datasets. Each subfigure targets a specific aspect of interactivity and quality. Arrow directions represent modeled and assumed conditional dependencies, where parent nodes influence the marginal probabilities of child nodes. In Fig. 1a, RTT, jitter (RJ), and packet loss (PL) are parent nodes, with RJ and PL conditioned on RTT. Each network condition connects to all interactivity components to infer changes in their posterior probabilities given network evidence. Fig. 1b captures interdependencies among interactivity, video, audio, and QoE using directional links. Fig. 1c combines network conditions and perceptual metrics, with arrows from RTT, RJ, and PL indicating direct effects. These models support the inference analyses in Section V.

In all models, the quality factor set $QF = \{VideoQ, AudioQ, QoE, IF1, IF2, CN1, CN2, CN3, RE1, RE2\}$ is represented as discrete nodes with five ordinal states (s_1 to s_5), corresponding to user ratings from "very-poor" to "excellent". Network conditions are also modeled as discrete nodes, where each state groups lab-emulated conditions based on effect magnitude and the presence other impairments. Details are provided in Table IV. For example, the RTT variable is discretized into six states. One such state, *RTT=d4*, includes all conditions with RTT=4ms, resulting in five conditions: the baseline (RTT=4ms) and four others where PL=6%, 12%, 24%, or 45% (each with RTT fixed at 4ms). Another state, *RTT=203_up*, includes only high-latency conditions where RTT $>$ 203ms (i.e., RTT = 277, 302, 352, and 402ms). Since these conditions do not involve jitter or PL and given their conditional dependence, reflecting out NetEm setup, setting evidence of *RTT=203_up* propagates this knowledge into all connected nodes. Then, BN by construction enforce the RJ and PL variables to their baseline states (*std0* and *p0*) with 100% probability. This discretization strategy preserves the fidelity to the experimental design while enabling meaningful probabilistic reasoning.

V. Results Analysis

This section presents probabilistic analyses for the three proposed models, fitted using VRCG (N=30) and MCG (N=31) user datasets. Models were trained in GeNIe[3] and CPTs estimated via the Expectation–Maximization (EM) algorithm [33]. While Bayesian networks support many evidence combinations, we focus on a representative subset. Specifically, we select high/low degradation or quality states to **1)** illustrate how probability distributions propagate under extreme network and quality conditions, and **2)** highlight the general direction of likely outcomes without exhaustively covering all combinations. *However, intermediate-state cases for each network variable (RTT, PL, RJ) (not depicted) were also analyzed to ensure the same conclusions of extreme cases would also apply to them.*

A. Effect of Network condition on Multiple Interactivity Factors for MCG:



(a) RTT = 203_up

(b) RJ = std4_up

(c) PL = p26_up

Fig. 2: BN inference results for the MCG dataset; interactivity components responsiveness (RE), Controllability (CN) and immediate feedback (IF) are conditioned on network factors. Figure labels use the state names: RTT=203_up $\in$ (277–402ms); RJ=std4_up $\in$ (6-12std) (at RTT=52ms); PL=p26_up $\in$ (45%). See details in Table III.

To address the question of which interactivity construct should be prioritize by stakeholders, our first analysis examines network conditions' impact on interactivity by focusing on interactivity-related factors set $QF' = \{\mathrm{IF1}, \mathrm{IF2}, \mathrm{CN1}, \mathrm{CN2}, \mathrm{CN3}, \mathrm{RE1}, \mathrm{RE2}\}$ where $QF' \subseteq QF$ reflects user perceptions of responsiveness, control, and input feedback. In this analysis, we *focus solely on the MCG dataset*. We model each interactivity-related factor as nodes

[3]https://www.bayesfusion.com/genie/ [Access date: Jul. 2025]

in the BN model, conditioned on the tested lab-emulated conditions RTT, RJ, and PL. The structure of the BN model is represented Fig.2, created based on the dataset (reefer to Table III for details on each interactivity question). Figure 2 presents BN inference results for interactivity-related factors under various network conditions. Each subfigure shows the inferred probability distributions for QF' factors across five ordinal levels (s1 to s5), given different evidence settings: high RTT (state=203_up corresponding on dataset to RTT= 277-402ms; Fig. 2a), high RJ (state=std4_up, corresponding on dataset to [RJ=4–6std, RTT=52ms]; Fig. 2b), and high packet loss (state=p26_up, corresponding on dataset to PL=45%; Fig. 2c). We note that high RTT leads to a probability redistribution across all interactivity questions to lower states (s1 and s2), indicating low quality of interactivity. For high RJ, it shows minimal impact, with distributions concentrated in s4 or "good", while high PL strongly skews interactivity metrics toward s1, indicating severe degradation. Despite these differences, the degradation pattern is consistent across all interactivity factors, i.e., none appears uniquely sensitive. This suggests that the effects of RTT, RJ, and PL on interactivity are uniform across the construct states of responsiveness, control, and feedback.

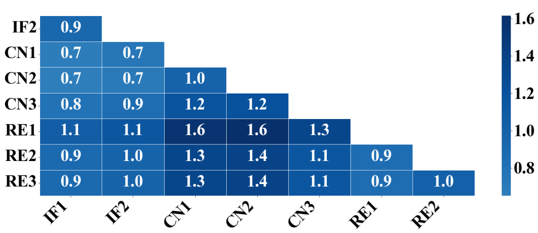


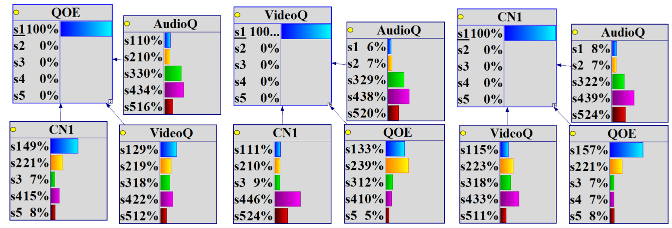
Fig. 3: Bayes Factor matrix to assess differences between interactivity factors.

We further investigate this uniformity, conducting pairwise Bayes factor tests [35] on all factors of QF' to assess evidence of differences in their corresponding marginal likelihoods (derived from each quality node CPT). Results are shown in Fig. 3, and we found BF_{Min}=0.7 and BF_{Max}=1.6; this ranges according to the literature [35] do not support a substantial differences across CPTs of QF'.

Summary: Although the questions focused on different interactivity constructs, our probabilistic analysis suggests that network conditions alone did not produce distinguishable effects across them. This supports the assumption of a shared underlying interactivity dimension and highlights the potential to reduce the number of interactivity questions in QoS-focused subjective tests. Therefore, we proceed to analyze inter-domain relationships using CN1 as the interactivity proxy for both VRCG and MCG datasets.

B. Quantifying Quality Factor Dependencies Through Sensitivity and Inference Analysis for MCG and VRCG:

A second research gap relevant to stakeholders with resource constraints —– and thus seeking to prioritize the



(a) QoE=s1|Factors (b) VideQ=s1|Factors (c) CN1=s1|Factors

Fig. 4: BN inference for QoE, VideoQ, AudioQ, and CN1 showing belief propagation and quality interdependence.

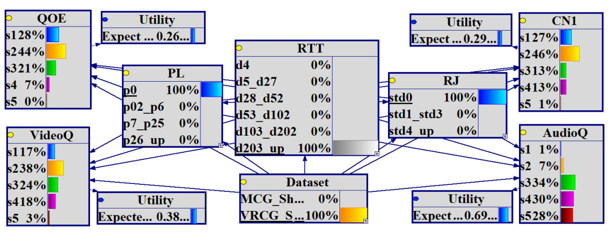
TABLE V: Conditional Sensitivity Measures for QoE, VideoQ, CN1, and AudioQ.

Model	Target	Conditional Variables Sensitivity			
		QoE	VideoQ	CN1	AudioQ
M1	QoE	—	Avg=0.112; Max=0.297;	Avg=0.142; Max=0.472;	Avg=0.047; Max=0.191;
M2	VideoQ	Avg=0.125; Max=0.308;	—	Avg=0.035; Max=0.09;	Avg=0.057; Max=0.154;
M3	CN1	Avg=0.158; Max=0.489;	Avg=0.031; Max=0.107;	—	Avg=0.051; Max=0.204;

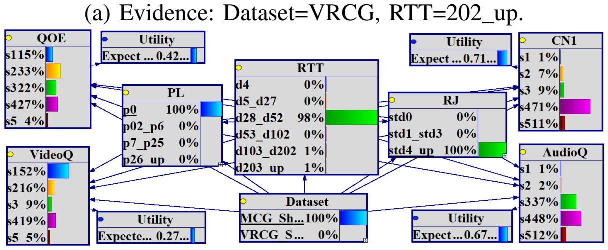
most informative quality metrics for CG provision —– relates to understanding how interactivity, QoE, and other media-specific quality factors (audio and video) influence each other, and whether they capture overlapping dimensions of user experience. To address this, we evaluate three Bayesian Network models (M1–M3) designed to capture potential interdependencies among interactivity (CN1), QoE, and audio/video quality (AudioQ, VideoQ), defined as: $M1 : P(\text{QoE} \mid \text{VideoQ}, \text{AudioQ}, \text{CN1})$, $M2 : P(\text{VideoQ} \mid \text{QoE}, \text{AudioQ}, \text{CN1})$, $M3 : P(\text{CN1} \mid \text{QoE}, \text{VideoQ}, \text{AudioQ})$., and detailed in Fig. 4. This figure presents the posterior distributions for each BN model (M1–M3), computed using belief propagation with target node set to the lowest quality state (s_1). In M1 (target: QoE), the strongest posterior belief under $QoE = s_1$ is associated with $CN1 = s_1$ and $VideoQ = s_1, s_2$, while $AudioQ$ remains more uniformly distributed across higher states. This confirms that poor QoE is most likely when interactivity and video quality are degraded, while audio quality exerts weaker influence. In M2 (target: VideoQ), conditioning on $VideoQ = s_1$ shifts the posterior for QoE to states s_1 and s_2, but CN1 and AudioQ retain high quality states s_5, s_4 distributions —– suggesting that poor video quality co-occurs with reduced overall QoE but not necessarily with poor interactivity or audio. In M3 (target: CN1), the posterior under $CN1 = s_1$ is dominated by $QoE = s_1$, while both $VideoQ$ and $AudioQ$ retain distributions between high quality states s_5, s_4.

Additionally, to complement and support the inference findings for M1, M2, and M3, we perform a sensitivity analysis using BNs to examine how changes in one quality factor influence the probabilities of others. We apply the Linear Fractional Transformation (LFT) method in GeNIe [36], which quantifies how changes in the states of input (parent) nodes affect the posterior distribution of a target node. Sensitivity is computed both for individual parent states and for all possible combinations of parent states, allowing us to capture isolated and joint effects, including non-linear dependencies. The results, shown in Table V, reveal that in M1 (target: QoE), CN1 (avg = 0.142) and VideoQ (avg = 0.112) exhibit the highest avg sensitivity, while AudioQ has limited influence (avg = 0.047). In M2 (target: VideoQ), QoE is the most influential factor (avg = 0.125), with CN1 and AudioQ showing lower sensitivity. In M3 (target: CN1), QoE again exerts the strongest effect (avg = 0.158), while VideoQ and AudioQ remain comparatively marginal.

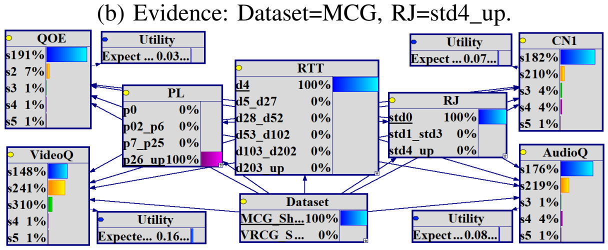
Summary: M1 which models QoE as dependent on CN1 (interactivity), VideoQ, and AudioQ; it presents the most balanced structure, capturing meaningful influence across quality dimensions. In contrast, M2 (video) and M3 (interactivity) show QoE as the most influential conditional factor. These findings support the view that interactivity alone may not sufficiently represent perceived quality in VRCG and MCG services under QoS degradation. Instead, QoE emerges as the more comprehensive and reliable metric.



(a) Evidence: Dataset=VRCG, RTT=202_up.



(b) Evidence: Dataset=MCG, RJ=std4_up.



(c) Evidence: Dataset=MCG, PL=p26_up.

Fig. 5: BN inference results for Interactivity (CN1), VideoQ, AudioQ and QoE conditioned on network factors.

C. Effect of Network Conditions on Interactivity and Quality Factors for MCG and VRCG:

Given that QoE emerged as the most influential factor in prior models, we further examine how QoE, VideoQ, AudioQ, and CN1 individually respond to network impairments. For each independent quality variable $X \in \{\text{VideoQ}, \text{AudioQ}, \text{CN1}, \text{QoE}\}$, we compute the expected utility $\mathbb{EU}[X]$ using a *linear utility function*, consistent with the decision-theoretic formulation in CaQoEM [23] and classical utility theory [37]. Each ordinal state s_i (e.g., "poor" to "excellent") is first mapped to a numeric value $V(s_i) \in \{1, 2, 3, 4, 5\}$, which is then normalized to the $[0, 1]$ interval using the formula from [23]. The resulting utility values are: $u(s_1) = 0$, $u(s_2) = 0.25$, $u(s_3) = 0.5$, $u(s_4) = 0.75$, and $u(s_5) = 1$. The expected utility is then computed as:

$$\mathbb{EU}[X] := \sum_{i=1}^{5} u(s_i) \cdot P(X = s_i) \tag{2}$$

This transformation ensures a consistent comparison and supportive interpretation of quality factors under varying network conditions. Next, we condition the model on each network variable independently, fixing them to representative high-impact states, to examine their marginal influence on the distribution of each quality factor.

RJ Effect: In Fig. 5b, we analyze the impact on RJ = std4_up and PL = $p0$, for the MCG dataset. The posterior distributions indicate that VideoQ is most likely in state $s1$ ("very poor"). QoE distributions also tend to the states $s2$ and $s3$, corresponding to mid to low quality levels. In contrast, AudioQ and CN1 remain largely unchanged, with the highest probability in state $s2$, suggesting a limited effect on RJ under these conditions. We conducted the same analysis on the VRCG dataset (not shown for brevity) and observed comparable patterns: VideoQ and QoE show increased probability mass in lower-quality $s2$, while AudioQ and CN1 remain stable and concentrated around state $s4$. **Summary**: The results suggest *that high levels of RJ negatively affect VideoQ and QoE, with VideoQ showing the strongest quality drop. However, the effect on AudioQ and CN1 appears small in both gaming scenarios and suggests weak dependence with RJ.*

RTT Effect: In Fig. 5a, the Bayesian network is conditioned on the VRCG dataset and network conditions RJ=std0 (no jitter), PL=p0 (no packet loss), and RTT=d202_up (high latency, 202 ms), focusing solely on the impact of high RTT. CN1 and QoE distributions are concentrated in states $s2$ and $s1$ (poor to very poor quality), with a low expected value of $\mathbb{EU}[CN1_{VRCG}] = 0.269$, confirming substantial degradation. VideoQ also performs poorly, leaning toward state $s2$ (poor) with $\mathbb{EU}[VideoQ_{VRCG}] = 0.383$. In contrast, AudioQ remains spread across state $s5 - s3$, indicating uncertainty of the users of the RTT effect on AudioQ. For MCG (not shown for brevity), the results show a similar distributions for CN1, with slightly lower expected value of $\mathbb{EU}[CN1_{MCG}] = 0.169$. VideoQ, however, favors state s_4 (good) with $\mathbb{EU}[VideoQ_{MCG}] = 0.643$, MCG is less sensitive to RTT degradation on VideoQ. AudioQ also spread mostly on $s5, s4$, aligned with the VRCG outcome. **Summary:** *These findings further confirm that RTT is the main cause of degradation in CN1, our reference interactivity variable, as observed in Section V-B.*

PL Effect: To evaluate the impact of packet loss (PL), the BN inference is set to RJ=$0std$ (no jitter), PL=$p26\_up$

or higher (high PL), and RTT=$d4$ (low RTT) to isolate PL effects. For MCG (Fig. 5c), CN1 is highly concentrated in state $s1$ (very-poor), with some probability in $s2$, resulting in an expected value of $\mathbb{EU}[CN1_{MCG}] = 0.07$. VideoQ, AudioQ and QoE distributions are skewed towards low quality states $s1, s2$, with $s1$ as the most probable. For VRCG (not shown due to brevity), CN1 is more concentrated in state $s2$ (poor), with a lower expected value of $\mathbb{EU}[CN1_{VRCG}] = 0.366$. Both VideoQ and AudioQ are more affected, skewed towards low-quality states, particularly $s1$. **Summary:** These findings highlight *the strong negative impact of PL on CN1, QoE, and VideoQ and AudioQ, in both VRCG and MCG contexts.*.

VI. Discussion

Our findings offer important implications for understanding how users perceive quality under varying network conditions in cloud gaming (CG) services. Although some of the analyses were conducted independently on VRCG and MCG datasets, results were consistent across both, highlighting patterns that are directly relevant for stakeholders and research community looking to improve and optimize CG services over HANs with quality.

A key question in this study is how specific network-level impairments—RTT, jitter, and packet loss (PL)—affect interactivity in cloud gaming (CG). Our results show that higher RTT and PL increase the probability of low interactivity-quality states. By contrast, jitter has minimal influence on interactivity quality but substantially lowers VideoQ. We interpret this pattern as consistent with video-layer degradations, based on the observed decreases in VideoQ. This was confirmed the QoE score distribution also concentrated to lower-quality states under jitter. These highlight the challenge of modelling interactivity quality considering QoS-only factors [8], which may underestimate interactivity quality loss in jitter-prone environments. In practice, mobile networks frequently experience jitter [38], [39], including on modern 5G. A recent review by Loh *et al.* [7] highlights persistently high jitter levels across European 4G and 5G networks. Since users worldwide increasingly rely on mobile connectivity, their access to MCG and VRCG services will require quality estimation from models that account for diverse and dynamic network impairments.

In addition, this paper results address how interactivity, QoE, and other perceptual quality metrics (audio and video) influence one another, and whether they capture overlapping dimensions of user experience. Our probabilistic analyses using Bayesian Network models (M1–M3) show that QoE consistently exhibits the strongest influence across all configurations. When QoE is the target (M1), both interactivity (CN1) and video quality contribute meaningfully, but when interactivity or video are modeled as dependent variables (M2, M3), QoE becomes the dominant factor. Audio quality has low influence across all cases. Although our findings did not show clear differentiation among the interactivity sub-components, this may be partly due to the nature of the experimental tasks, which focused on gameplay rather than configuration or environment manipulation. Control-related constructs—such as interface customization or object handling—may depend more on content-specific features than on network conditions alone, and thus remained less responsive to QoS variation in our design. This distinction merits further investigation.

Taken together, these results provide valuable guidance for both stakeholders and the research community in making informed decisions about where to focus efforts when designing, evaluating, or investigating CG services. Rather than treating all quality dimensions equally, our findings suggest that QoE offers more representative quality metric under network impairments, while interactivity can often be inferred from it. Further, the Bayesian Network approach used in this study enables these insights by modeling conditional dependencies and supporting probabilistic reasoning across perceptual and network variables. Our BN modeling framework and design may serve as a practical and inspirational tool for researchers aiming to isolate and investigate other quality metrics, and identify which matters most under specific conditions. Finally, as interest in CG services continues to expand across heterogeneous network environments, this modeling framework and the insights it reveals can support more efficient, user-centered and targeted approaches to both research and deployment.

VII. Conclusion and Future Work

This paper proposes Bayesian Network models to investigate how network impairments affect perceived interactivity in cloud-streamed gaming based on two distinct data sets regarding mobile cloud gaming and virtual reality cloud gaming. Through Bayesian inference, sensitivity, and utility-based analyses, we identified how quality factors such as interactivity, video, audio, and overall QoE are interrelated and affect each other. Our findings offer practical insights for refining QoE and interactivity subjective tests. They will directly impact future research and stakeholders to investigate the context of cloud-gaming quality measurement and prediction and mobile networks effect. In conclusion, we highlight:

- Results of seven interactivity-related variables show consistent degradation patterns across critical network conditions RTT, RJ, and PL; it suggests a single question can measure interactivity; it creates research opportunities to simplify and standardize interactivity assessment.
- Sensitivity and Bayesian inference analysis show that QoE is the most influenced perceptual metric, with its posterior probability strongly affected by interactivity, audio, and video quality in the tested models.
- Jitter was found to strongly influence video quality, while its influence on interactivity remained minimal. In contrast, RTT skewed towards low quality states both interactivity and QoE probability distribution in VRCG and MCG. These findings suggest that existing interactivity estimation models may underestimate the perceptual impact of jitter —- especially in CG services delivered over HANs – where jitter is both common and highly variable.

In the future, we plan to explore interactivity in various cloud gaming contexts (e.g., TV, PC) and study the control mapping effects.

REFERENCES

[1] M. Slater and S. Wilbur, "A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments," *Presence: Teleoperators and Virtual Environments*, vol. 6, no. 6, pp. 603–616, Dec. 1997.

[2] D. Z. Saltz, "The Art of Interaction: Interactivity, Performativity, and Computers," *The Journal of Aesthetics and Art Criticism*, vol. 55, no. 2, pp. 117–127, 1997.

[3] J. Steuer, F. Biocca, and M. R. Levy, "Defining Virtual Reality: Dimensions Determining Telepresence," *Communication in the age of virtual reality*, vol. 33, no. 37-39, p. 1, 1995.

[4] Fortune Business Inside, "Cloud Gaming Market Size, Value, Growth | Global Report [2025]," Tech. Rep. [Online]. Available: https://www.fortunebusinessinsights.com/cloud-gaming-market-102495 [accessed: Jun-2025]

[5] F. Metzger, S. Geißler, A. Grigorjew, F. Loh, C. Moldovan, M. Seufert, and T. Hoßfeld, "An Introduction to Online Video Game QoS and QoE Influencing Factors," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 3, pp. 1894–1925, 2022.

[6] H. S. Rossi, K. Mitra, C. Åhlund, and I. Cotanis, "QoE Models for Virtual Reality Cloud-based First Person Shooter Game over Mobile Networks," in *2024 20th International Conference on Network and Service Management (CNSM)*. Prague, Czech Republic: IEEE, Oct. 2024, pp. 1–5.

[7] F. Loh, F. Raithel, S. Raffeck, A. Seufert, R. Frohler, C. Heller, and T. Hossfeld, "QoS and QoE Study of the European 5G Mobile Networks for Next Generation of Applications," *IEEE Communications Magazine*, pp. 1–7, 2025.

[8] ITU-T Recommendation G.1051, "Latency measurement and interactivity scoring under real application data traffic patterns," *T G.*, Mar. 2023.

[9] S. Zinno, G. Caso, N. Pasquino, A. Botta, A. Brunstrom, and G. Ventre, "QoE for Interactive Services in 5G Networks: Data-driven Analysis and ML-based Prediction," in *2024 20th International Conference on Network and Service Management (CNSM)*, Oct. 2024, pp. 1–7.

[10] K.-Y. Lee, A. Singla, P. Cesar, and C.-H. Hsu, "Adaptive Cloud VR Gaming Optimized by Gamer QoE Models," *ACM Transactions on Multimedia Computing, Communications, and Applications*, Jul. 2024.

[11] S. Zadtootaghaj, S. Schmidt, and S. Möller, "Modeling Gaming QoE: Towards the Impact of Frame Rate and Bit Rate on Cloud Gaming," in *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*, May 2018, pp. 1–6, iSSN: 2472-7814.

[12] Y. Ida, Y. Ishibashi, N. Fukushima, and S. Sugawara, "QoE assessment of interactivity and fairness in First Person Shooting with group synchronization control," in *2010 9th Annual Workshop on Network and Systems Support for Games*, Nov. 2010, pp. 1–2, iSSN: 2156-8146.

[13] S. Schmidt, S. Möller, and S. Zadtootaghaj, "A Comparison of Interactive and Passive Quality Assessment for Gaming Research," in *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*, May 2018, pp. 1–6, iSSN: 2472-7814.

[14] A. Wahab, N. Ahmad, M. G. Martini, and J. Schormans, "Subjective Quality Assessment for Cloud Gaming," *J*, vol. 4, no. 3, pp. 404–419, Sep. 2021.

[15] Y. C. Li, C. H. Hsu, and C. H. Hsu, "Performance Measurements on a Cloud VR Gaming Platform," in *QoEVMA 2020 - Proceedings of the 1st Workshop on Quality of Experience (QoE) in Visual Multimedia Applications*, vol. 20. New York, NY, USA: Association for Computing Machinery, Inc, Oct. 2020, pp. 37–45.

[16] H. S. Rossi, K. Mitra, C. Åhlund, I. Cotanis, N. Ögren, and P. Johansson, "Objective QoE Models for Cloud-based First Person Shooter Game over Mobile Networks," in *2024 IEEE 21th Consumer Communications & Networking Conference (CCNC)*. Las Vegas, NV, USA: IEEE, Jan. 2024, pp. 550–553.

[17] ITU-T Recommendation G.1072, "Opinion model predicting gaming quality of experience for cloud gaming services," Tech. Rep., 2020.

[18] R. Weber, K.-M. Behr, and C. DeMartino, "Measuring Interactivity in Video Games," *Communication Methods and Measures*, vol. 8, no. 2, pp. 79–115, Apr. 2014.

[19] S. Möller, D. Pommer, J. Beyer, and J. Rake-Revelant, "Factors Influencing Gaming QoE: Lessons Learned from the Evaluation of Cloud Gaming Services," in *4th International Workshop on Perceptual Quality of Systems (PQS 2013)*, Sep. 2013, pp. 163–167.

[20] J. Juvrud, G. Ansgariusson, P. Selleby, and M. Johansson, "Game or Watch: The Effect of Interactivity on Arousal and Engagement in Video Game Media," *IEEE Transactions on Games*, vol. 14, no. 2, pp. 308–317, Jun. 2022.

[21] M. Sithu and Y. Ishibashi, "QoE assessment of joint haptic drum performance: Effect of local lag control," in *2013 IEEE 2nd Global Conference on Consumer Electronics (GCCE)*, Oct. 2013, pp. 461–465.

[22] D. Shafer, C. Carbonara, and L. Popova, "Controller Required? The Impact of Natural Mapping on Interactivity, Realism, Presence, and Enjoyment in Motion-Based Video Games," *Presence*, vol. 23, pp. 267–286, Oct. 2014.

[23] K. Mitra, A. Zaslavsky, and C. Åhlund, "Context-Aware QoE Modelling, Measurement, and Prediction in Mobile Computing Systems," *IEEE Transactions on Mobile Computing*, vol. 14, no. 5, pp. 920–936, May 2015.

[24] S. Moller, S. Schmidt, and J. Beyer, "Gaming taxonomy: An overview of concepts and evaluation methods for computer gaming QoE," in *2013 5th International Workshop on Quality of Multimedia Experience, QoMEX 2013 - Proceedings*, 2013, pp. 236–241.

[25] ITU-T Recommendation P.809, "Subjective evaluation methods for gaming quality," 2018.

[26] J. L. G. Sánchez, F. L. G. Vela, F. M. Simarro, and N. Padilla-Zea, "Playability: analysing user experience in video games," *Behaviour & Information Technology*, vol. 31, no. 10, pp. 1033–1054, Oct. 2012.

[27] J. Seibert and D. M. Shafer, "Control mapping in virtual reality: effects on spatial presence and controller naturalness," *Virtual Reality*, vol. 22, no. 1, pp. 79–88, Mar. 2018.

[28] H. S. Rossi, K. Mitra, S. Larsson, C. Åhlund, and I. Cotanis, "Subjective QoE Assessment for Virtual Reality Cloud-based First-Person Shooter Game," in *ICC 2024 - IEEE International Conference on Communications*, Denver, CO, USA, Jun. 2024, pp. 4698–4703, iSSN: 1938-1883.

[29] H. S. Rossi, N. Ogren, K. Mitra, I. Cotanis, C. Ahlund, and P. Johansson, "Subjective Quality of Experience Assessment in Mobile Cloud Games," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*. Rio de Janeiro, Brazil: IEEE, Dec. 2022, pp. 1918–1923.

[30] M. Claypool and K. Claypool, "Latency can kill: Precision and deadline in online games," in *MMSys '10*. ACM Press, 2010, pp. 215–222.

[31] ITU-T Contribution SG12.C293, "Data Assessment for an Opinion Model Predicting Gaming QoE (G.OMG)," Tech. Rep., 2018.

[32] T. Hossfeld, A. Seufert, F. Loh, S. Wunderer, and J. Davies, "Industrial User Experience Index vs. Quality of Experience Models," *IEEE Communications Magazine*, vol. 61, no. 1, pp. 98–104, Jan. 2023.

[33] S. J. Russell and P. Norvig, *Artificial intelligence: a modern approach*, third edition ed., ser. Prentice Hall series in artificial intelligence. Boston Columbus Indianapolis: Pearson, 2016.

[34] S. Tasaka, "Comparison of multimedia communications QoE models by Bayesian networks and Bayesian statistics: a case study," *SN Applied Sciences*, vol. 1, no. 9, p. 975, Sep. 2019.

[35] R. E. Kass and A. E. Raftery, "Bayes Factors," *Journal of the American Statistical Association*, vol. 90, no. 430, pp. 773–795, 1995.

[36] E. Castillo, J. Gutierrez, and A. Hadi, "Sensitivity analysis in discrete Bayesian networks," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 27, no. 4, pp. 412–423, Jul. 1997.

[37] R. A. Howard, "Uncertainty about Probability: A Decision Analysis Perspective," *Risk Analysis*, vol. 8, no. 1, pp. 91–98, 1988.

[38] E. Yu, J. Zhou, Z. Li, G. Tyson, W. Li, X. Zhang, Z. Xu, and G. Xie, "Mustang: Improving QoE for Real-Time Video in Cellular Networks by Masking Jitter," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 20, no. 9, pp. 1–23, Sep. 2024.

[39] S. Neumeier, E. A. Walelgne, V. Bajpai, J. Ott, and C. Facchi, "Measuring the Feasibility of Teleoperated Driving in Mobile Networks," in *2019 Network Traffic Measurement and Analysis Conference (TMA)*, Jun. 2019, pp. 113–120.