MULTIMEDIA COMMUNICATIONS TECHNICAL COMMITTEE
http://www.comsoc.org/~mmc

MMTC Communications - Frontiers

Vol. 18, No. 4, Jul 2023

CONTENTS

SPECIAL ISSUE ON Sustainable multimedia communications and services-II

Guest Editors: Gülnaziye Bingöl[1] and Frank Loh[2]
[1]University of Cagliari, Italy, gulnaziye.bingol@unica.it
[2]University of Würzburg, Germany, frank.loh@informatik.uni-wuerzburg.de

The escalating environmental repercussions arising from the energy consumption and greenhouse gas emissions of information and communication technologies (ICT) have spurred a growing emphasis on researching and adopting Green communications and computing practices. The primary objective is to develop sustainable ICT-based services that meet users' needs and minimize their environmental impact. Despite the rapid progress in sustainable solutions, managing energy usage remains a pressing concern. Sustainable multimedia communications and services entail designing and implementing multimedia-based communication systems with a strong focus on reducing their environmental footprint while optimizing efficiency and user experience. Research efforts often centre around devising algorithms to enhance the utilization of renewable energy sources, decrease computational complexity, optimize resource allocation, minimize data storage, and enable automatic device switch-off when not in use. In addition to adhering to device-imposed energy and resource constraints, sustainable multimedia communications and services must also meet essential Quality of Service (QoS) and Quality of Experience (QoE) standards to meet user expectations. Thus, the Special Issue (SI) on Sustainable multimedia communications and services aims to tackle emerging concepts and challenges concerning energy efficiency and sustainability for multimedia-based applications and services. The SI endeavours to pave the way for a greener and more environmentally responsible ICT landscape by investigating these crucial research areas.

The third paper introduces a new hybrid P2P-CDN framework that leverages Network Function Virtualization (NFV) and edge computing for live video streaming. The framework aims to reduce server energy consumption, support low latency, and enhance video quality by efficiently distributing tasks across a hybrid P2P-CDN network.

The first paper addresses the challenge of deriving operational points for configuring Internet services, balancing QoE and energy consumption. It discusses the potential of creating a joint QoE-energy model for Internet services, considering both QoE and energy consumption in video streaming and content processing. The author also highlights that conducting joint video quality (VQ) and energy consumption (EC) tests is necessary to rate the joint VQ-EC conditions and to derive comprehensive QoE models.

The second paper introduces a novel architecture for Always-Listening Personal Digital assistants (ALPDAs) that address challenges related to user choice between prioritizing Quality of Experience (QoE) for precise context organization and sustainability for privacy and environmental concerns. The approach utilizes advanced natural language processing techniques, including Large Language Models (LLMs), to provide accurate responses while balancing data collection and processing with privacy and sustainability considerations.

The third paper outlines a potential solution, focusing on reducing energy consumption while maintaining acceptable Quality of Experience (QoE) levels for end-users. The proposed approach involves an end-device application that tracks energy consumption patterns and motivates users towards more environmentally friendly behaviors by reducing unnecessary high-quality video streams.

In the fourth paper, an examination was conducted by the authors into how video quality (specifically spatial video resolution) and latency (referring to video buffer size) impact the user experience and energy consumption. A balance exists between user comfort and energy consumption within the lab-based Tele-operation system. Energy consumption was gauged by tracking the voltage drop over a fixed time frame, keeping the current constant under the conditions of maximum spatial resolution and the highest video latency. Notably, there were no substantial variations observed in user comfort levels.

Gülnaziye Bingöl earned her B.Sc. degree in Electrical and Electronic Engineering from Omer Halisdemir University in Turkey in June 2019. In 2018, she also had the opportunity to study as a visiting student at the University of Oradea and worked as a summer intern at the University of Cagliari. She further advanced her education by completing her Master's Degree in Internet Engineering at the University of Cagliari in 2021, achieving a full score of 110/110. Currently, Gulnaziye is dedicated to her Ph.D. studies in the Department of Electrical and Electronic Engineering (DIEE/UdR CNIT) at the University of Cagliari. Her research interests revolve around Quality of Experience (QoE), Green Multimedia Services, WebRTC, Data Science, Machine Learning, and Human Behaviour.

Frank Loh is a PhD student of the chair of communication networks at the University of Würzburg, Germany with Prof. Tobias Hoßfeld. His research includes Quality of Service monitoring and management of different applications, especially on-demand and live streaming platforms. The focus is on improving Quality of Experience of different applications. Furthermore, he works with Low Power Wide Area, cellular, and mobile networks. With his main research on LoRaWAN, he develops novel channel access approaches and studies energy consumption towards a greener and a more energy efficient communication.

The Need for a Comprehensive Video Quality and Energy Consumption QoE Model

Tobias Hoßfeld
Institute of Computer Science, University of Würzburg, Germany
tobias.hossfeld@uni-wuerzburg.de

Abstract

A key problem in QoE management is the derivation of operational points concerning the configuration and parameterization of Internet services. However, this requires appropriate models to resolve tradeoffs, e.g., QoE vs. energy consumption. While the QoE research community focused on QoE models, recent literature tackles also the energy consumption, especially of video streaming services, at the end user device or for processing and storing contents, e.g., in a content distribution network (CDN). The key question is how to utilize existing QoE models, but enrich them with energy consumption, towards a joint QoE-energy model for Internet services. This article discusses the possibilities to deriving such a model and its implications in practice.

Keywords: QoE management, Energy consumption, Video streaming, QoE models.

1. The Need for Extended QoE Models

The derivation of operational points concerning the configuration and parameterization of Internet services requires appropriate models to resolve tradeoffs like the QoE and the energy consumption. With higher bitrates, the perceived quality of the user will increase. The QoE research community focuses on the derivation of such QoE models, which are especially considering the perceptual dimensions, like the video quality, the frame rate, the impact of the video codec, etc.

Perceptual Video Quality (VQ) Model: VQ QoE. A simple relationship $Q_v(x)$ between the video bitrate x and the perceptual video quality, determined in subjective studies and quantified through Mean Opinion Scores (MOS), is depicted in Figure 2. In [1], MOS scores are provided depending on the video bitrate x following a logarithmic relationship, as suggested by the Provisioning-Delivery Hysteresis [2] and the WQL hypothesis [3] due to the logarithmic nature of QoE and the role of the Weber-Fechner law in QoE assessment [4]. A generic logarithmic QoE model for video streaming is then provided by [5] which allows parameterizing the maximum video bitrate resulting in a MOS score of 5 (excellent quality). Figure 2 shows an example of such a video quality (VQ) related QoE model (or VQ QoE in short) $Q_v(x)$.

Tradeoff between Video Quality and Energy Consumption. Recently, the QoE research community also focused the energy consumption, e.g. [5]–[10], especially of video-streaming services, at the end user device or for processing and storing contents, e.g., in a content distribution network (CDN). In one of the first user studies, [6] found that the experience of mobile users was consistently affected by battery efficiency, as it restricted their phone usage, particularly towards the end of the day when the phone was fully discharged. However, the higher the video bitrate, the higher the energy consumption. [7], [11] showed that the choice of the video bitrate and the video codec have a significant influence on the power consumption at the end user device. They provide a linear model with corresponding coefficients for the trained variables, e.g., the video bitrate. The total energy consumption can be described by a linear model involving the service usage time and data traffic volume. The electricity intensity parameters for the various components are based on suggestions from [12] and [13]. Hence, we have an obvious trade-off between the VQ QoE $Q_v(x)$ and the energy consumption $E(x)$.

Perceptual Energy Consumption (EC) Model: EC QoE. However, such a QoE model $Q_v(x)$ is only considering the perceptual video quality dimensions, but ignoring the results from literature, e.g. [6], that the energy consumption and battery drainage is a relevant contextual influence factor, especially for smartphone users. Among mobile users, depleting the battery is regarded as one of the most undesirable experiences [14]. Consequently, energy consumption stands out as one of the most significant influencing factors on QoE and should be considered in mobile application design [14], [15]. Therefore, comprehensive QoE models should

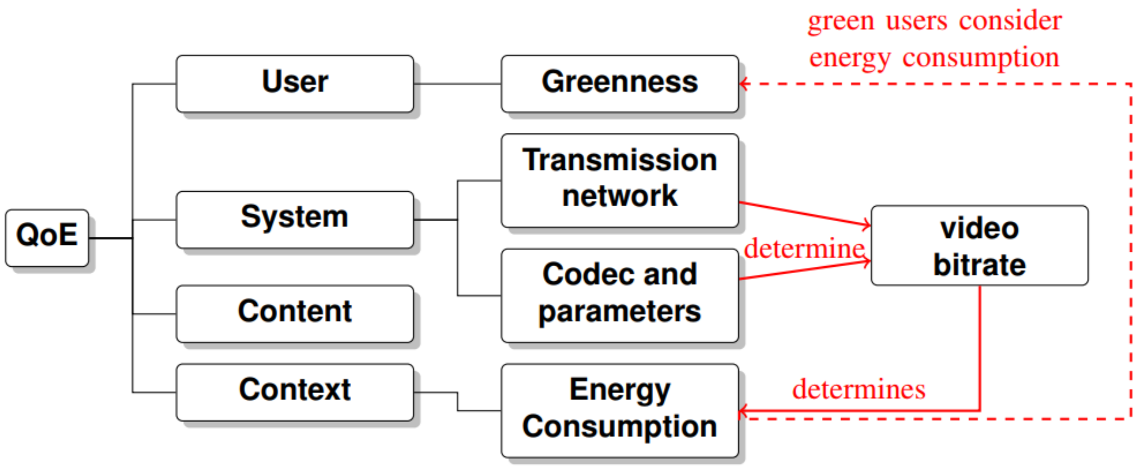


Figure 1: Factors influencing QoE. A comprehensive QoE model needs to consider video quality related features as well as context factors like energy consumption.

address the influences of energy consumption as well as mechanisms like reduced video bitrates which affect the QoE as well as the battery time [16].

Research Question. The key question is how to utilize existing video quality related QoE models $Q_v(x)$, but enrich them with energy consumption (i.e., energy consumption related perception of users), towards a joint video quality and energy consumption QoE model $Q(x)$ (or VQ-EC QoE in short) for Internet services. Such a combined VQ-EC QoE model and the influence factors considered in the following is visualized in Figure 1. Do we need to consider such comprehensive QoE models $Q(x)$ when deriving optimal trade-offs between QoE and the energy consumption?

2. Comprehensive VQ-EC QoE Models

Additive Models. As discussed above, the energy consumption may be seen as a QoE context factor and therefore directly influencing QoE [6], [16]. Then, additive models [17] may be appropriate. From the perspective of quality perception, impairment additivity implies distinguishable perceptual features, also referred to as orthogonal quality dimensions in a multidimensional perceptual feature space. There are several options for additive models. For example, the overall QoE may be (1) a weighted sum of video quality related QoE etc. and energy consumption related QoE. Or the overall QoE is (2) the VQ QoE as upper QoE bound, which may be degraded due to the amount of energy consumption, expressed as EC QoE.

Formally, the additive QoE model may be expressed as

Additive QoE Model: $$Q^+(x) = w_v^+ Q_v(x) + w_e^+ Q_e^+(x) \quad (1)$$

with corresponding weights w_v^+, w_e^+ for the VQ QoE and the EC QoE, respectively. However, subjective studies are required to determine the weights. This may be a simple questionnaire asking the users about the importance of those two factors w_v^+, w_e^+. Nevertheless, a dedicated subjective study needs to be conducted to derive a perceptual model $Q_e^+(x)$ of the energy consumption. The quantification of $Q_e^+(x)$, requires some appropriate subjective assessments and questions that users can rate the perceived QoE for a particular amount of energy consumption. It has to be noted that the energy consumption or battery drain must be visualized accordingly during the subjective test. It is an interesting research question how to properly set up subjective user studies to determine $Q_e^+(x)$ as well as the weights w_v^+, w_e^+.

If there are any interaction effects between the VQ QoE and the EC QoE, then the additive model must be extended

TABLE I: Notation of variables for the illustrative modelling

Variable	Description
x	Average video bitrate as a result of the application's implementation and the network transmission
$Q_v(x)$	Video quality (VQ) perception model maps the video bitrate to MOS scores; only VQ related perceptual factors are influencing $Q_v(x)$
$E(x)$	Energy consumption of user's device when consuming the video service with an average video bitrate x
$Q_e(x)$	Energy consumption (EC) related perception of a user is modelled as a function of the video bitrate to MOS scores: $Q_e(x) = Q_e^*(E(x))$
$Q(x)$	Comprehensive QoE model taking into account video quality and energy consumption: $Q(x) = f(Q_v(x), Q_e(x))$
$Q^+(x)$	Additive QoE model $Q^+(x) = w_v^+ Q_v(x) + w_e^+ Q_e^+(x)$ with corresponding weights w_v^+, w_e^+ in Eq. (1)
$Q^-(x)$	Degradation QoE model $Q^-(x) = Q_v(x) - (5 - Q_e^-(x))$ with QoE degradation due to energy consumption in Eq. (3)
$Q^\gamma(x)$	QoE model for different user classes with greenness factor γ

Additive QoE Model with Interaction: $Q^+(x) = w_v^+ Q_v(x) + w_e^+ Q_e^+(x) + w_{v,e}^+ Q_v(x) Q_e^+(x)$ (2)

by integrating the interaction term $Q_v(x) \cdot Q_e^+(x)$ and an appropriate weight $w_{v,e}^+$. However, this means that subjective user studies need to be conducted in which the video quality and the energy consumption are modified jointly. The subjects then need to rate the overall QoE, such that the interaction term and its weight can be determined. As a consequence, the existing VQ related QoE model $Q_v(x)$ is however not required anymore, since the subjective study needs to investigate the relevant combinations of VQ and EC. The resulting subjective scores can then be utilized to derive the comprehensive QoE model.

Next, we consider the (2) degradation QoE model $Q^-(x)$. The video quality $Q_v(x)$ determines the upper bound of the overall QoE, which may be degraded, however, due to the amount of energy consumption. In its simplest form, i.e., without any weights, the degradation QoE model is as follows.

Degradation QoE Model: $Q^-(x) = Q_v(x) - Q_e^-(x)$ (3)

This degradation $Q_e^-(x)$ needs to be investigated employing subjective user studies. While the perception of the energy consumption $Q_e^+(x)$ may be obtained with an absolute category rating scale (ACR), a degradation may be obtained with Degradation Category Rating (DCR). Then, the quality scales used for the ACR can be used for the DCR method by replacing the quality adjectives (5: excellent, 4: good, 3: fair, 2: poor, 1: bad) with the corresponding impairment adjectives (5: imperceptible; 4: perceptible, but not annoying; 3: slightly annoying; 2: annoying; 1: very annoying).

From a mathematical perspective, the degradation term may be simply the difference of excellent quality and the perception-based EC QoE model $Q_e^+(x)$ from the additive model in Eq. (1).

QoE Degradation Term: $Q_e^-(x) = 5 - Q_e^+(x)$ (4)

However, this postulates that the ACR rating of a test condition and the DCR rating of the same test condition always result in $Q_e^-(x) + Q_e^+(x) = 5$. The subjective testing methodology should fit to the QoE model to be developed and tested.

For the degradation QoE model, a Double Stimulus Impairment Scale (DSIS) method may be appropriate. Then, test conditions (i.e., different amounts of energy consumption while not changing the video quality) may be presented in pairs. Then, the first stimulus presented in each pair serves as reference with the minimal energy consumption for the provided video quality. The users are asked to rate the level of annoyance of the increased energy consumption that she/he observes for the second test condition.

As for the additive QoE model, potential interactions need to be quantified through subjective user studies.

Degradation QoE Model with Interaction: $Q^{-}(x) = Q_v(x) - w_e^{-} Q_e^{-}(x) - w_{v,e}^{-} Q_v(x) Q_e^{+}(x)$ (5)

Again, this means that combinations of video quality and energy consumption need to be tested and rated by the subjects. Then, the comprehensive QoE model can be derived from the user ratings without reusing the existing VQ QoE model $Q_v(x)$. In general, for testing video quality and energy consumption at the same time, an appropriate subjective methodology and tool for visualizing the energy consumption or battery drain need to be developed. Regarding the subjective testing methodology, a pairwise comparison may be easier for subjects to deal with complementary perceptual dimensions (VQ and EC) in a multidimensional perceptual features space. In [18], a probabilistic model is proposed that can bring the results of pairwise comparison and rating experiments into a unified quality scale. This allows to map the results of pairwise comparisons to an ACR scale, which may be used for resolving the tradeoff between QoE and energy consumption. Paired comparison has the problem that the number of test conditions may explode, since all tuples / pairs of VQ and EC need to be investigated. To this end, active learning for crowdsourced QoE modeling is promising to explore the complicated interaction between QoS factors more efficiently. [19] developed active learning algorithms for multidimensional QoE model where the subjective user studies were conducted by means of crowdsourcing, while following the best practices for QoE crowdtesting [20], [21].

In summary, the different comprehensive QoE models require appropriate subjective methodologies. Later, we will consider, if those comprehensive QoE models also will lead to different operation points in terms of optimal video bitrate.

Green User Classes and Parameterizable VQ QoE Model. Another approach of a comprehensive QoE model which considers VQ and EC jointly is a parameterizable VQ QoE model. Thereby, the parameters of the VQ model depend on the greenness of users. This “green user model” was suggested in [5].

For a *green user*, we adjust the logarithmic QoE model and introduce a *greenness factor* γ as an advantage factor, assuming users rate lower video quality with a higher score if they know it saves energy. In Figure 2, the provided video bitrate to MOS mapping function considers the minimum bitrate $x_1 = 200\,kbps$ (yielding MOS 1) and maximum bitrate $x_5 = 6\,Mbps$ (yielding MOS 5). The greenness factor γ considers that a green user is satisfied with maximum bitrate $x_5' = x_5/\gamma$ yielding MOS 5. We extend the logarithmic QoE model in [1] for the green user QoE model:

Green User QoE Model: $Q^{\gamma}(x) = \frac{4}{\log x_5' - \log x_1} \log x + \frac{\log x_5' - 5 \log x_1}{\log x_5' - \log x_1}$ (6)

In Figure 2, we assume $\gamma = 2$ for a green user. Thus, the green user is satisfied with maximum bitrate 3 Mbps ($\gamma = 2$), while the high quality (HQ) user desires 6 Mbps for excellent video quality ($\gamma = 1$). For the numerical results, the video quality related perception $Q_v(x)$ corresponds to the MOS mapping function of the HQ user with $\gamma = 1$:

HQ QoE Model: $Q^1(x) = Q_v(x)$ with $\gamma = 1$. (7)

In subjective experiments, for different types of users the corresponding factors γ need to be derived. This may be done using a simple questionnaire to cluster the users. However, then subjective experiments need to be carried out for those user groups to derive γ. And still, it is unclear how well the clustering based on a questionnaire works.

A better methodology is that the test conditions are rated by subjects, in which the VQ and energy consumption are varied. With appropriate clustering algorithms, the user classes may be determined, and the coefficient γ is derived. However, then the subjective user scores may be directly used to derive a comprehensive QoE model. Again, the existing VQ related QoE models are then not utilized.

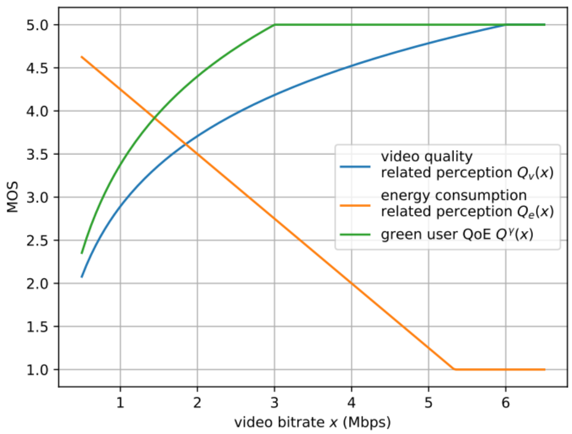


Figure 2: The Perceptual video quality model $Q_v(x)$ and the perceptual energy consumption model $Q_e(x)$ are illustrated here as logarithm and linear mapping function for a given video bitrate x, respectively. Both perceptual features (VC, EC) are ingredients of a comprehensive QoE model $Q(x)$.

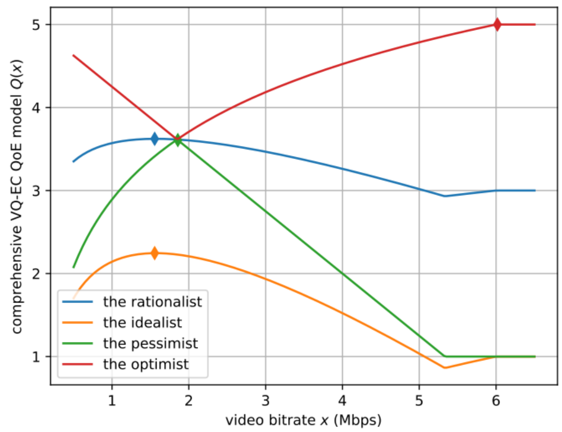


Figure 3: Examples of comprehensive QoE models taking into account the video quality and the energy consumption as QoE context factor. Using Kleinrock's power metric, operational points are identified in terms of optimal bitrate and are depicted with a marker.

3. Resolving the Tradeoff between QoE and Energy Consumption

For identifying the optimal tradeoff between QoE and energy consumption, we consider the following comprehensive QoE models. The VQ QoE model $Q_v(x)$ is provided in Eq. (7) and Eq. (6). The EC QoE model $Q_e(x)$ is depicted in Figure 2 and follows a linear relationship: $Q_e(x) = 5 - 0.75 \cdot x$.

Comprehensive QoE Models for Numerical Results. There are many possibilities how such a comprehensive QoE model could look like. The ignorant is only considering the VQ and ignores the EC: $Q(x) = Q_v(x)$. The rationalist represents an additive QoE model with equal weights for VQ and EC. The idealist is a simple degradation QoE model without any interactions. The pessimist takes the minimum of VQ and EC QoE, while the optimist takes the maximum.

The rationalist: $Q(x) = Q_v(x)/2 + Q_e(x)/2$ **The idealist:** $Q(x) = Q_v(x) - (5 - Q_e(x))$

The pessimist: $Q(x) = \min\,(Q_v(x), Q_e(x))$ **The optimist:** $Q(x) = \max(Q_v(x), Q_e(x))$

Figure 3 provides the QoE scores for the different models depending on the video bitrate.

Determination of the Optimal Video Bitrate. A simple approach is Kleinrock's power metric to solve the problem of identifying operational points, as discussed in [5]. Kleinrock's power metric is an auxiliary utility function, which combines the comprehensive QoE model $Q(x)$ and the energy consumption $E(x)$ into a single metric $U(x)$. The ratio of goodness (QoE) to badness (energy consumption) is the power metric which needs to be maximized. For the energy consumption, we approximate $E(x) = 6.97 \times 10^{-6}$ W/kbps $\cdot x + 5.31$W based on the numbers of [7].

Kleinrock's Power Metric: $U(x) = \frac{Q(x)}{E(x)}$ **Optimal Video Bitrate:** $\frac{d}{dx}U(x) = 0 \Rightarrow \mathrm{x}$ (8)

Just considering the VQ QoE model leads to an optimal video bitrate of 6 Mbps, i.e., the maximum bitrate. The same result is observed for the optimist. For the rationalist and the idealist, Kleinrock's approach leads to 1.56 Mbps. The optimal video bitrate for the pessimist is 1.86 Mbps. Those optimal video bitrates are depicted as markers in Figure 3.

The first observation is that the comprehensive QoE models result in different optimal video bitrates. Hence, it is important to fully understand the impact of energy consumption as a context factor on the overall QoE. The second observation is that the different models also result in different operational points. Hence, a careful modelling and joint VQ-EC test methodology is required to derive the multidimensional QoE model properly.

Please note that Kleinrock's power metric is quite sensitive regarding, e.g., changes of the energy consumption concerning the minimum energy consumption in idle mode. Such changes in $E(x)$ in the denominator may have a significant impact on the identified video bitrate.

Multi-objective optimization is a wide field that may provide rankings of configurations. However, first studies indicated that all those approaches lead to different results regarding the operational points. Thus, it appears that joint VQ-EC QoE models need to be developed and tested systematically and that joint VQ-EC studies are unavoidable.

Video Quality Sufficiency. Another approach would be to consider the acceptance of a minimum video quality, which directly provides the operational points in terms of video bitrate. In addition, the acceptance of a maximum energy consumption may be considered. This requires, however, acceptance studies, while literature typically considers Mean Opinion Scores (MOS). A direct relationship between acceptance and MOS is to be developed, if existing at all [22].

However, again, this requires subjective user studies to determine which video quality is sufficient for users. Furthermore, interactions between energy consumption and acceptance need to be investigated. Similar issues as raised above for the joint VQ-EC QoE model need to be addressed for a joint VQ-EC sufficiency or VQ-EC acceptance model.

4. Conclusions and Discussions

Although it is tempting to utilize existing perceptual video quality models to obtain comprehensive VQ-EC QoE models, the suggested models may have significant limitations (e.g., no interactions between VQ and EC perception models) or the models require joint VQ-EC test conditions (e.g., to quantify interactions or green user QoE models). To validate the limitations, joint VQ-EC test conditions need to be rated in subjective user studies. Thus, in practice, it appears that there is a need for conducting joint video quality and energy consumption tests and to derive comprehensive QoE models. Such comprehensive QoE models need to incorporate the influences of energy consumption as well as mechanisms like reduced video bitrates which affect the QoE as well as the battery time. Still, it is an open research question how to properly set up subjective experiments and derive an appropriate testing methodology for such comprehensive VQ and EC QoE models. The need for such comprehensive QoE models was demonstrated by quantifying the optimal video bitrate to tradeoff video quality and energy consumption with different comprehensive QoE models.

References

[1] J. P. Lopez, D. Martin, D. Jimenez, and J. M. Menendez, “Prediction and Modeling for No-Reference Video Quality Assessment Based on Machine Learning,” in *14th Int. Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*. IEEE, 2018, pp. 56–63.

[2] M. Fiedler and T. Hoßfeld, “Quality of experience-related differential equations and provisioning-delivery hysteresis,” in *21st ITC Specialist Seminar on Multimedia Applications-Traffic, Performance and QoE*. IEICE, 2010.

[3] S. Egger, P. Reichl, T. Hoßfeld, and R. Schatz, ““time is bandwidth”? narrowing the gap between subjective time perception and quality of experience,” in *2012 IEEE international conference on communications (ICC)*. IEEE, 2012, pp. 1325–1330.

[4] P. Reichl, S. Egger, R. Schatz, and A. D'Alconzo, “The logarithmic nature of QoE and the role of the Weber-Fechner law in QoE assessment,” in *2010 IEEE Int. Conference on Communications*. IEEE, 2010, pp. 1–5.

[5] T. Hoßfeld, M. Varela, L. Skorin-Kapov, and P. E. Heegaard, “A greener experience: Trade-offs between qoe and co2 emissions in today's and 6g networks,” *IEEE Communications Magazine*, 2023.

[6] S. Ickin, K. Wac, M. Fiedler, L. Janowski, J.-H. Hong, and A. K. Dey, “Factors influencing quality of experience of commonly used mobile applications,” *IEEE Communications Magazine*, vol. 50, no. 4, pp. 48–56, 2012.
[7] C. Herglotz, W. Robitza, M. Kranzler, A. Kaup, and A. Raake, “Modeling of energy consumption and streaming video qoe using a crowdsourcing¨ dataset,” in *2022 14th International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2022, pp. 1–6.
[8] V. Gnanasekaran, H. T. Fridtun, H. Hatlen, M. M. Langøy, A. Syrstad, S. Subramanian, and K. De Moor, “Digital carbon footprint awareness among digital natives: an exploratory study,” in *Norsk IKT-konferanse for forskning og utdanning*, no. 1, 2021, pp. 99–112.
[9] S. Afzal, R. Prodan, and C. Timmerer, “Green video streaming: Challenges and opportunities,” *ACM SigMM Records*, 2023.
[10] G. Bingol, “Sustainability vs. quality of experience: Striking the right balance for video streaming,” *ACM SigMM Records*, 2023.
[11] C. Herglotz, W. Robitza, A. Raake, T. Hossfeld, and A. Kaup, “Power reduction opportunities on end-user devices in quality-steady video streaming,” in *2023 15th International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2023.
[12] The Shift Project, “Lean ICT: Towards Digital Sobriety,” directed by Hugues Ferreboeuf, Tech. Rep., 2019, last accessed: August 2022. [Online]. Available: https://theshiftproject.org/en/article/lean-ict-our-new-report/
[13] A. S. G. Andrae and T. Edler, “On global electricity usage of communication technology: Trends to 2030,” *Challenges*, vol. 6, no. 1, pp. 117–157, 2015.
[14] J. Hadden, A. Tiwari, R. Roy, and D. Ruta, “Computer assisted customer churn management: State-of-the-art and future trends,” *Computers & Operations Research*, vol. 34, no. 10, pp. 2902–2917, 2007.
[15] Q. A. Chen, H. Luo, S. Rosen, Z. M. Mao, K. Iyer, J. Hui, K. Sontineni, and K. Lau, “Qoe doctor: Diagnosing mobile app qoe with automated ui control and cross-layer analysis,” in *Proceedings of the 2014 conference on internet measurement conference*, 2014, pp. 151–164.
[16] P. Reichl, S. Egger, S. Moller, K. Kilkki, M. Fiedler, T. Hoßfeld, C. Tsiaras, and A. Asrese, “Towards a comprehensive framework for qoe and user¨ behavior modelling,” in *2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*. IEEE, 2015, pp. 1–6.
[17] T. Hoßfeld, L. Skorin-Kapov, P. E. Heegaard, M. Varela, and K.-T. Chen, “On additive and multiplicative qos-qoe models for multiple qos parameters,” in *Proceedings of the 5th ISCA/DEGA Workshop on Perceptual Quality of Systems PQS 2016*. ISCA, 2016.
[18] M. Perez-Ortiz, A. Mikhailiuk, E. Zerman, V. Hulusic, G. Valenzise, and R. K. Mantiuk, “From pairwise comparisons and rating to a unified quality scale,” *IEEE Transactions on Image Processing*, vol. 29, pp. 1139–1151, 2019.
[19] H.-S. Chang, C.-F. Hsu, T. Hoßfeld, and K.-T. Chen, “Active learning for crowdsourced qoe modeling,” *IEEE Transactions on Multimedia*, vol. 20, no. 12, pp. 3337–3352, 2018.
[20] T. Hoßfeld, C. Keimel, M. Hirth, B. Gardlo, J. Habigt, K. Diepold, and P. Tran-Gia, “Best practices for qoe crowdtesting: Qoe assessment with crowdsourcing,” *IEEE transactions on multimedia*, vol. 16, no. 2, pp. 541–558, 2013.
[21] T. Hoßfeld, M. Hirth, J. Redi, F. Mazza, P. Korshunov, B. Naderi, M. Seufert, B. Gardlo, S. Egger, and C. Keimel, “Best practices and recommendations for crowdsourced qoe-lessons learned from the qualinet task force” crowdsourcing”,” 2014.
[22] T. Hoßfeld, P. E. Heegaard, M. Varela, and S. Moller, “Qoe beyond the mos: an in-depth look at qoe via better metrics and their relation to mos,” ¨ *Quality and User Experience*, vol. 1, pp. 1–23, 2016.

Authors Bio-data

Tobias Hoßfeld is professor at the Chair of Communication Networks at the University of Würzburg, Germany, since 2018. He finished his PhD in 2009 and his professorial thesis (habilitation) "Modeling and Analysis of Internet Applications and Services" in 2013 at the University of Würzburg, where he was also heading the "Future Internet Applications & Overlays" research group. From 2014 to 2018, he was head of the Chair "Modeling of Adaptive Systems" at the University of Duisburg-Essen, Germany. He received several best conference paper awards, among others the Best Paper Award at WMNC 2021 and QoMEX 2022, three awards for his PhD thesis, and the Fred W. Ellersick Prize 2013 (IEEE Communications Society) for one of his articles on QoE. He received the QoMEX 2017 Reviewer Recognition Award and the NetSoft 2023 Best Reviewer Award. Among others, he was general Co-Chair of ITC 30, TPC-Co-Chair of the IEEE NoF 2020, QoMEX 2019, NetSys 2019, ITC 28, and sustainability chair of QoMEX 2023. He is a member of the editorial board of IEEE Communications Surveys & Tutorials, ACM SIGMM Records, Springer Quality and User Experience, and a senior member of the IEEE.

Always-Listening Personal Digital Assistants: The Trade-Off between QoE and Sustainability

Luigi Serreli, Michele Nitti
DIEE, University of Cagliari, Italy
{luigi.serreli, michele.nitti}@unica.it

Abstract

The trade-off between Quality of Experience (QoE) and sustainability is a critical consideration for always-listening Personal Digital Assistants (ALPDAs). These devices, which constantly listen to user conversations to provide accurate and relevant responses, raise concerns about privacy and environmental impact. In this contribution, we propose a novel architecture that addresses these challenges by allowing users to choose between prioritising QoE for a more precise system to organise their daily context or sustainability to safeguard privacy and minimise environmental impact. Our approach employs advanced natural language processing techniques, including Large Language Models (LLMs), to provide accurate and relevant responses while balancing the need for data collection and processing with concerns for privacy and sustainability. We discuss the potential of our proposed architecture and its implications for developing more intelligent and sustainable PDAs.

Keywords: Sustainability, QoE, Personal Digital Assistants, Always Listening, LLM.

1. Introduction

The recent advances in natural language processing (NLP) have made it possible to improve Human to Machine (H2M) conversations, enabling more sophisticated interactions between users and Intelligent Systems [1]. Contemporary machine learning enhancements have also facilitated the progress of Personal Digital Assistants (PDAs), equipping them with the capability to engage in meaningful dialogues with humans and provide more refined responses. This progress has not only enhanced PDAs but also laid the foundation for a new era of multimedia communication and services.

The recent emergence of Large Language Models (LLMs) represents a paradigm shift in NLP research, enabling the processing of vast amounts of textual data and the generation of highly coherent and context-aware responses. Unlike traditional rule-based systems, LLMs learn from massive datasets, capturing intricate patterns and semantic nuances that empower them to provide more natural and human-like interactions. Thanks to this progress, machines can increasingly understand the context of a conversation leading to improvements in PDAs, which could listen to entire conversations to assist users better and predict their needs. This favours the development of devices known in the literature as always-listening or always-on devices [2].

The concept of “always listening”, applied to personal assistants, has been a subject of discussion in the realm of artificial intelligence and natural language processing [3]. However, until recently, developing a model capable of accurately understanding context and predicting user actions and decisions remained elusive. Previous iterations of PDA demonstrated proficiency in handling individual user requests but lacked the crucial element of memory. As a result, users could not engage in natural and fluid conversations with PDAs, as they lacked the ability to recall previous interactions and build upon them. Consequently, a significant challenge lies in enhancing the intelligence of PDAs to enable them to make automatic decisions based on the user’s needs. This gives rise to several challenges: understanding the domain of speech from other conversations is still considered a challenging topic by literature and requires computational resources to connect the various topics of interest to the user. In addition, there is the challenge of handling the conversation of each user, which impacts environmental sustainability, as all conversations give a significant workload requiring more and more computational resources. Furthermore, there are many privacy concerns since the PDA must store all user speeches. In particular, crossing the delicate balance between harnessing the power of PDAs to enhance user experiences while safeguarding sensitive information poses a critical challenge.

In this contribution, we present a novel approach to addressing these challenges using advanced natural language processing techniques and a proposed architecture. Our approach considers the trade-off between sustainability and Quality of Experience (QoE), allowing users to choose between prioritising a more precise system to organise their daily context or safeguarding privacy and minimising environmental impact. We begin by exploring the background and motivation driving our research, encompassing the challenges and concerns associated with always-listening PDAs (ALPDAs). Then we present a reference scenario and a proposed architecture strategically designed to address privacy concerns while allowing end-users to choose between prioritising QoE or sustainability. Then we delve into a comprehensive discussion pertaining to the delicate trade-off between sustainability and QoE. We conclude by discussing the implications of our work for the development of future always-listening digital assistants, as well as directions for future research.

2. Background and Privacy Concerns

The concept of always-listening functionality in PDA has been previously considered in the literature, primarily for commercial purposes rather than to enhance the user quality of experience and service [4].

PDAs are portable and non-portable devices that enable users to access and process information. They have been in use for years and have recently gained widespread popularity. However, several issues have been reported in the literature on PDAs, including technical challenges and privacy concerns as always-listening devices.

Several studies shed light on the topic of always-listening interfaces and commercially available devices that implement this feature. One study explored the potential of always-listening interfaces to create an assistant named Articulate+, capable of contextualising user queries for generating graphs, visualising data, and facilitating data exploration [5]. In addition, the market presents numerous always-listening devices, with prominent examples being Alexa and Google Home, as highlighted in the literature by the works [6],[7]. Despite claims that these devices should only record and transmit data when triggered by a Wake-up Word (WuW), a study [8] revealed concerning findings regarding privacy, in which the authors discovered instances where private conversations were recorded without using the WuW. Luger *et al.* [9] discovered a significant disconnect between user expectations and the capabilities of present-day voice assistants. Users anticipate conversational agents to grasp the context of ongoing tasks based on previous interactions while simultaneously expressing concerns about privacy issues. A prevailing lack of awareness exists among many users regarding the instances when intelligent devices are actively listening and the destination of transmitted data.

However, lack of transparency is not the only privacy issue for future listening devices. Depending on their location within the home, they can listen in on different conversations with varying degrees of sensitivity. In addition, some works related to the experience of users in employing these devices reveal that users may also want to change their sentence structure when addressing their devices in order to make the conversation natural, such as saying, "*Turn on the lights, Siri*" instead of "*Siri, turn on the lights*" [10]. Although the literature focuses heavily on privacy-related aspects and user expectations, there are also technical challenges to consider. For example, Cao *et al.* [11], mention that it is particularly challenging to produce a classification from a speech where there are no boundaries between sentences, i.e., punctuation, and it is not known when a sentence ends. Moreover, processing longer texts can be computationally intensive and may require more resources and energy [12]. Furthermore, the challenge in topic classification is particularly high when addressing a problem where the number of classes to be covered is high due to the lack of labelled datasets [13]. However, an unsupervised approach, i.e., without employing annotated datasets, often does not produce results that are comparable to when labelled datasets are employed to train a machine learning model.

While users desire enhanced conversations and prioritise QoE, privacy remains a paramount concern. Thus, this scenario prompts the examination of a potential compromise between optimising QoE and safeguarding privacy.

3. Reference Scenario

The proposed concept offers a comprehensive view of the interaction between an ALPDA and its user. With recent advancements in NLP, engaging in conversations with an assistant that closely resembles human-to-human (H2H) interactions is possible. For instance, by remembering previous conversations, the assistant can better support the user's needs and may even anticipate them over time without requiring any input.

In the reference scenario, the always-listening PDA is integral to the user's daily interactions. For the sake of simplicity, we consider a single user as the exclusive individual accessing the device, even though advancements in speaker recognition and identification allow for multi-user interactions [14]. Therefore, the system is designed to cater specifically to the user's needs, providing a personalised experience tailored to their preferences and requirements.

As users go about their day, they can issue voice commands or ask questions to the ALPDA, which constantly listens for its WuW. Upon detection, the system processes the user's request and generates a response based on their preferences and context. The user can choose to prioritise either QoE, for a more precise method of organising their daily context, or sustainability to safeguard privacy and minimise environmental impact.

Let us consider Bob, who interacts with his ALPDA named Alice at two different times. One day, Bob is undecided about his next travel destination and says, "*I can't decide between Paris or London*". At a later time, which could be hours or days later, Bob comments, "*The croissant this morning wasn't very good*", and asks Alice to recommend a good croissant nearby. Alice processes Bob's request and generates a response. It recommends a nearby location for finding a good croissant. Also, it connects the context perceived the first time by replying something like, "*If you're interested in travelling, choosing Paris would definitely lead you to delicious croissants*".

4. Proposed Architecture

The proposed architecture, depicted in Figure 1, begins with the system constantly listening to the user through an ASR block, translating speech into text. Subsequently, the textual content is checked for the presence of a WuW, which allows the system to determine if the user expects a response from the assistant. Depending on the user's choice between QoE by empathising with the system's performance or safeguarding privacy, an embedding layer is applied to have a limited environmental impact. This layer is responsible for translating the text into vector format. However, this layer is not applied in the case of sustainability to avoid storing entire user conversations but rather to have an overview of detected speech preferences.

The Topic Ontology System allows for the creation of an ontology instance composed of nodes and edges. Nodes represent topics and keywords retrieved from the user conversation. The edges represent connections between various topics and consider the user's interest in a specific topic.

When the WuW is not detected, this block only updates the ontology; however, in the case the WuW is expressed by the user, the ontology instance is first updated and then sent to the cloud to produce feedback on the last requested topic. If the user emphasises QoE, context construction is more precise as the request will be attached to the entire conversation describing the context it refers to. In this case, various embeddings representing entire detected conversations are stored in the cloud, specifically in a database designed to store embeddings. The ontology still plays a fundamental role here, as it not only weighs the occurrence of specific topics against others but also operates as a filter key for context extraction.

Finally, the last block is represented by the LLM, which takes as input user requests with the WuW and context, if available, in order to generate a response to send back to the user.

5. Trade-Off between Sustainability and QoE

The trade-off between sustainability and QoE is an important consideration for ALPDAs. First, the environmental impact of continuous data transmission and the need to reduce energy consumption must be considered. Second, users expect accurate and relevant information and assistance from their digital assistants,

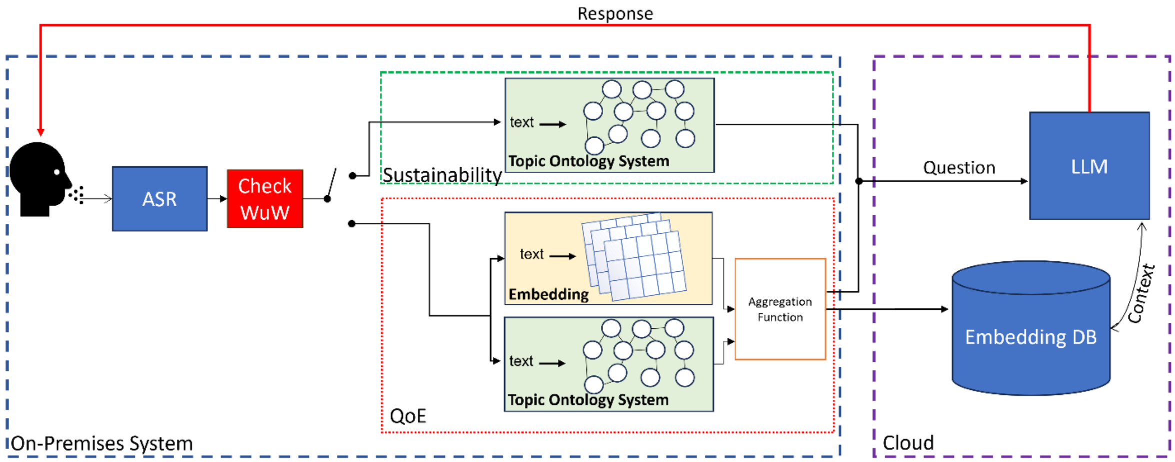


Figure 1: Proposed architecture.

which requires the collection and processing of data.

One potential solution to this trade-off is to allow users to choose between prioritising QoE for a more precise system that can organise their daily context or sustainability to safeguard privacy and minimise environmental impact. The first proposed solution could be achieved by employing an ontology instance to approximate the context, send it without storing the entire conversation, and generate a suitable response while sacrificing some QoE. To this, LLMs can be well-suited for this purpose as they are capable of generating human-like responses, even if the speech context is not well-defined. Therefore, users who choose sustainability may have access to limited functionality due to the reduced amount of data available for processing.

On the other hand, users who prioritise QoE may need to agree to specific terms under the General Data Protection Regulation (GDPR) to allow for more extensive data collection and processing. As discussed in the previous section, prioritising QoE means increasing the data collected. Therefore, the ontology system is supported by the embedding layer to translate the text into a vector form. This procedure raises interest in the literature, matching text-embedding with security and privacy concerns [15]. In fact, embedding can prevent clear text from being transmitted and make it difficult for attackers who do not have the same model to understand the data.

However, the constant transmission of data consumes a significant amount of energy. For instance, cloud computing uses a large number of data centres and servers to serve many clients using a pay-per-use model. Such resources cover a big area and demand a considerable amount of electricity for networking devices, cooling technologies, displays, and server farms, among other things [16]. In addition, large cloud loads, especially those powering LLMs, can consume even more energy. However, some companies are taking steps to address this issue by using energy-efficient hardware and shared infrastructure powered by renewable energy sources [17]. For example, HPE's GreenLake for LLMs [18] will be powered by almost 100% renewable energy, with measures such as power management optimisation through HPE software. This will help to reduce the environmental impact of cloud computing while still providing the benefits of large-scale AI training and deployment. However, a trade-off exists between providing a good user experience, protecting user privacy, and reducing energy consumption can be achieved by designing a proper architecture.

6. Conclusions

Our proposed architecture represents a promising approach to addressing the challenges associated with ALPDAs. By allowing users to choose between QoE and sustainability, we provide a flexible solution that can meet the diverse needs and preferences of different users. Future research could further explore the potential of this approach and its implications for developing more intelligent and sustainable ALPDAs.

References

[1] Allouch, M., Azaria, A., & Azoulay, R. (2021). Conversational agents: Goals, technologies, vision and challenges. Sensors, 21(24), 8448
[2] Rediger, A. M. (2016). Always-Listening Technologies: Who Is Listening and What Can Be Done About It. Loy. Consumer L. Rev., 29, 229.
[3] Poushneh, A. (2021). Humanizing voice assistant: The impact of voice assistant personality on consumers' attitudes and behaviors. Journal of Retailing and Consumer Services, 58, 102283.
[4] Mirghaderi, L., Sziron, M., & Hildt, E. (2022). Investigating user perceptions of commercial virtual assistants: A qualitative study. Frontiers in Psychology, 13, 944714.
[5] Tabalba, R., Kirshenbaum, N., Leigh, J., Bhatacharya, A., Johnson, A., Grosso, V., ... & Zellner, M. (2022, July). Articulate+: An Always-Listening Natural Language Interface for Creating Data Visualizations. In Proceedings of the 4th Conference on Conversational User Interfaces (pp. 1-6).
[6] Manikonda, L., Deotale, A., & Kambhampati, S. (2018, December). What's up with privacy? User preferences and privacy concerns in intelligent personal assistants. In Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society (pp. 229-235).
[7] Buteau, E., & Lee, J. (2021). Hey Alexa, why do we use voice assistants? The driving factors of voice assistant technology use. Communication Research Reports, 38(5), 336-345.
[8] Ford, M., & Palmer, W. (2019). Alexa, are you listening to me? An analysis of Alexa voice service network traffic. Personal and ubiquitous computing, 23, 67-79.
[9] Luger, E., & Sellen, A. (2016, May). " Like Having a Really Bad PA" The Gulf between User Expectation and Experience of Conversational Agents. In Proceedings of the 2016 CHI conference on human factors in computing systems (pp. 5286-5297).
[10] Malkin, N., Egelman, S., & Wagner, D. (2019, September). Privacy controls for always-listening devices. In Proceedings of the New Security Paradigms Workshop (pp. 78-91).
[11] Cao, T., He, L., & Niu, F. (2022, December). End-to-end speech topic classification based on pre-trained model Wavlm. In 2022 13th International Symposium on Chinese Spoken Language Processing (ISCSLP) (pp. 369-373). IEEE.
[12] Nagwani, N. K. (2015). Summarizing large text collection using topic modeling and clustering based on MapReduce framework. Journal of Big Data, 2(1), 6.
[13] Duarte, J. M., & Berton, L. (2023). A review of semi-supervised learning for text classification. Artificial Intelligence Review, 1-69.
[14] Wang, R., Ao, J., Zhou, L., Liu, S., Wei, Z., Ko, T., ... & Zhang, Y. (2022, May). Multi-view self-attention based transformer for speaker recognition. In ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 6732-6736). IEEE.
[15] Hassan, F., Sánchez, D., & Domingo-Ferrer, J. (2021). Utility-preserving privacy protection of textual documents via word embeddings. IEEE transactions on knowledge and data engineering, 35(1), 1058-1071.
[16] Bharany, S., Sharma, S., Khalaf, O. I., Abdulsahib, G. M., Al Humaimeedy, A. S., Aldhyani, T. H., ... & Alkahtani, H. (2022). A systematic survey on energy-efficient techniques in sustainable cloud computing. Sustainability, 14(10), 6256.
[17] Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274.
[18] Castro, P., Isahagian, V., Muthusamy, V., & Slominski, A. (2023). Hybrid serverless computing: Opportunities and challenges. Serverless Computing: Principles and Paradigms, 43-77.

Authors Bio-data

Luigi Serreli is a Ph.D. student in Electronic and Computer engineering at the Department of Electrical and Electronics Engineering at the University of Cagliari (Italy). He holds a B.Sc. in Electrical and Electronic Engineering and an M.Sc. in Internet Engineering at the University of Cagliari, discussing the thesis “A generative Adversarial Network (GAN) Fingerprint approach over LTE”.

Michele Nitti is an Assistant Professor at the University of Cagliari, Italy since 2015. He served as a technical program chair for various international conferences (IEEE BMSB 2017, IEEE IoT V&T Summit 2020). Currently, he is a member of the editorial board for the IEEE IoT Journal, Elsevier Computer Networks and co-founder of an academic spin-off (GreenShare s.r.l.) which works in the mobility sector. His main research interests are in architecture and services for the Internet of Things (IoT).

Towards Sustainable Video Streaming on Android-based Devices: Reducing Energy Consumption and Empowering Users

Simone Porcu, Lazizjon Suyunov, and Luigi Atzori
DIEE, University of Cagliari, Italy
{simone.porcu, lazizjon.suyunov, l.atzori}@unica.it

Abstract

Video streaming contributes to a large part of greenhouse gas emissions caused by the extensive use of Information and Communication Technologies (ICT). We describe a possible solution to decrease energy consumption while maintaining acceptable levels of end-user's Quality of Experience (QoE). The solution is an end-device application for monitoring energy consumption and stimulating the user towards a greener behavior.

Keywords: Quality of Experience, Sustainability, Video Streaming, Energy, End-device.

1. Introduction

The widespread adoption of ICTs plays a vital role in various aspects of our modern life; however, their usage comes at a cost in terms of increased energy consumption and CO_2 emissions. Indeed, the ICT sector is estimated at ca. 1.8%–2.8% of global greenhouse gas emissions in 2020 [1]. Moreover, these numbers are expected to rise further with the ongoing digitalisation trend.

Video streaming, in particular, is a prominent contributor to resource demands due to the continuous pursuit of high-quality and high-resolution contents concerning, for example, video-on-demand, video-conferencing, and live streaming applications. Video traffic has increased by 24% in 2022 and now accounts for 65% of all Internet traffic [2]. In particular, many studies identified the end-device as the element of the video streaming chain that consumes the greatest amount of energy [3]-[6]. For instance, one hour of streaming on Netflix consumes an average of 76.9 Wh, of which 70% is used in the device, 25% in the transmission network, and 5% in the data center [5]. Therefore, it is of particular importance to investigate how the energy consumption may be reduced at the end-device side. In this paper, we particularly focus on end-devices (i.e., smartphones) running the Android operating system, which was the leading mobile operating system worldwide in the second quarter of 2023 with a market share of 70.8% [7].

One of the most effective ways to reduce energy consumption of video streaming services is to reduce the video quality, e.g., by reducing the spatial or temporal video resolution [3], [6]. However, such a video quality reduction may result in a negative impact on the end user's Quality of Experience (QoE), which is defined as the "*degree of delight or annoyance of the user of an application or service*" [8]. Accordingly, the reduction of the video quality must be conducted carefully, with a knowledge of the relevant benefits in terms of energy reduction and the impact on the QoE.

To advance in this regard, in this paper we propose a service to monitor battery consumption during video streaming sessions at the user end-device that can be used to investigate the possible strategies to implement greener actions. Specifically, this research focuses on developing an Energy Consumption Monitoring App (ECMA) for Android Smartphone (SP) devices. The primary objective of this tool is to gain insights into the energy consumed by different applications on these devices, to analyze user behavior in terms of application usage, to inform the users about their level of sustainability, and to motivate them to adopt more sustainable behaviors. Ultimately, the goal is to set a green usage mode to adopt an energy-efficient usage of the SP among users.

2. End-device Monitoring Application

We propose an end-device application to monitor and manage data usage and battery consumption during video streaming sessions. The ECMA strives to contribute to a greener, more sustainable digital landscape by bridging the gap between user awareness and energy-efficient video streaming. Unlike traditional monitoring tools, this app provides users with information about the environmental impact of their streaming habits by summarizing the consumed energy and corresponding produced CO_2 after each streaming session. This empowers the users to make more informed choices concerning their video streaming behavior, by encouraging them to consider the energy implications of their actions. Moreover, according to the user authorisations, the application can limit the device's network capabilities to reduce the incoming video data flow. In this way, the users can actively contribute to environmental sustainability by reducing video streaming bitrates. However, it is essential to investigate the impact of video quality reduction on the user's QoE and identify the trade-off between sustainability and perceived QoE.

2.1 ECMA Implementation Details

In this section, we present our methodology to develop the ECMA for Android SP devices. First, we used a combination of Java and Kotlin as programming languages to write the ECMA code. The application user interface rendering was handled using Jetpack Compose[1], while parallel tasks were managed using Kotlin Coroutines[2] and Flows[3]. Regarding the background tasks, they are implemented using Service[4] in collaboration with the Broadcaster Receivers[5], which manage the information exchanging between the Android system and the background routines. The application employs SQLite[6] and JSON[7] for database management, with Room[8] serving as the database helper. To prepare the data to be displayed in the application Activities we used the ViewModel[9] class. To ensure dependency inversion, we integrated Dagger Hilt[10] into the project. The application's navigation system was designed using Voyager[11], with a custom structure. Throughout the development process, we adhered to SOLID, KISS, and DRY principles, while also incorporating clean architecture, clean code practices, and an implementation based on the Model-View-Intent (MVI) pattern. By utilizing Service and Activity, the battery monitor application continuously collects data on battery percentage, as well as a list of used applications and corresponding usage time.

An additional feature of ECMA is the capability to introduce bandwidth limits. This is obtained by implementing a logic layer between the phone and server which can enforce desired custom network rules. Exploring various possibilities, we settled on the idea of creating a virtual private network (VPN), utilizing the developer APIs provided by Android for VPN solutions. Not only do VPNs offer a secure means for devices not physically connected to access the network but these also grant us the ability to modify rules within this virtual layer dynamically. This feature gives us the opportunity to fine-tune bandwidth utilization and then to reduce the energy consumption.

2.2 Monitoring App Energy Consumption

One of the primary goals of the ECMA is to encourage sustainable behaviors in users. By providing feedback on energy-starved apps and their implications on battery life, users may become more conscious of their energy consumption habits. Being informed on the energy consumption, they may be motivated to make informed decisions, such as reducing the usage of energy-demanding apps or finding more energy-efficient alternatives. Once the ECMA has identified the apps responsible for significant energy consumption, it provides feedback to the user, as highlighted in Figure 1. ECMA presents the user with comprehensive insights into their energy

---

[1] https://developer.android.com/jetpack/compose?authuser=2
[2] https://kotlinlang.org/docs/coroutines-overview.html
[3] https://developer.android.com/kotlin/flow
[4] https://developer.android.com/guide/components/services
[5] https://developer.android.com/guide/components/broadcasts
[6] https://developer.android.com/training/data-storage/sqlite
[7] https://developer.android.com/reference/org/json/JSONObject
[8] https://developer.android.com/training/data-storage/room
[9] https://developer.android.com/reference/android/arch/lifecycle/ViewModel?hl=en
[10] https://developer.android.com/training/dependency-injection/hilt-android
[11] https://voyager.adriel.cafe/

consumption behavior. Users can visualize the most energy hungry apps and understand their overall impact on device performance and battery life. In order to provide accurate feedback on which applications starve for energy, the ECMA employs data analysis techniques such as analysis of variance (ANOVA) and Pearson correlation coefficient (PCC) to establish a correlation between battery consumption and application usage. By continuously monitoring the battery drain while the user engages with various applications, the tool can identify which apps have a more substantial impact on energy consumption. The analysis provides insights into the energy-starved apps and the overall energy usage patterns of the user.

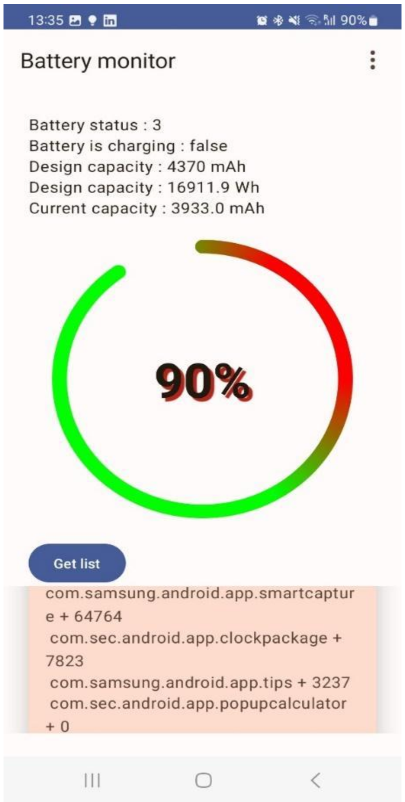
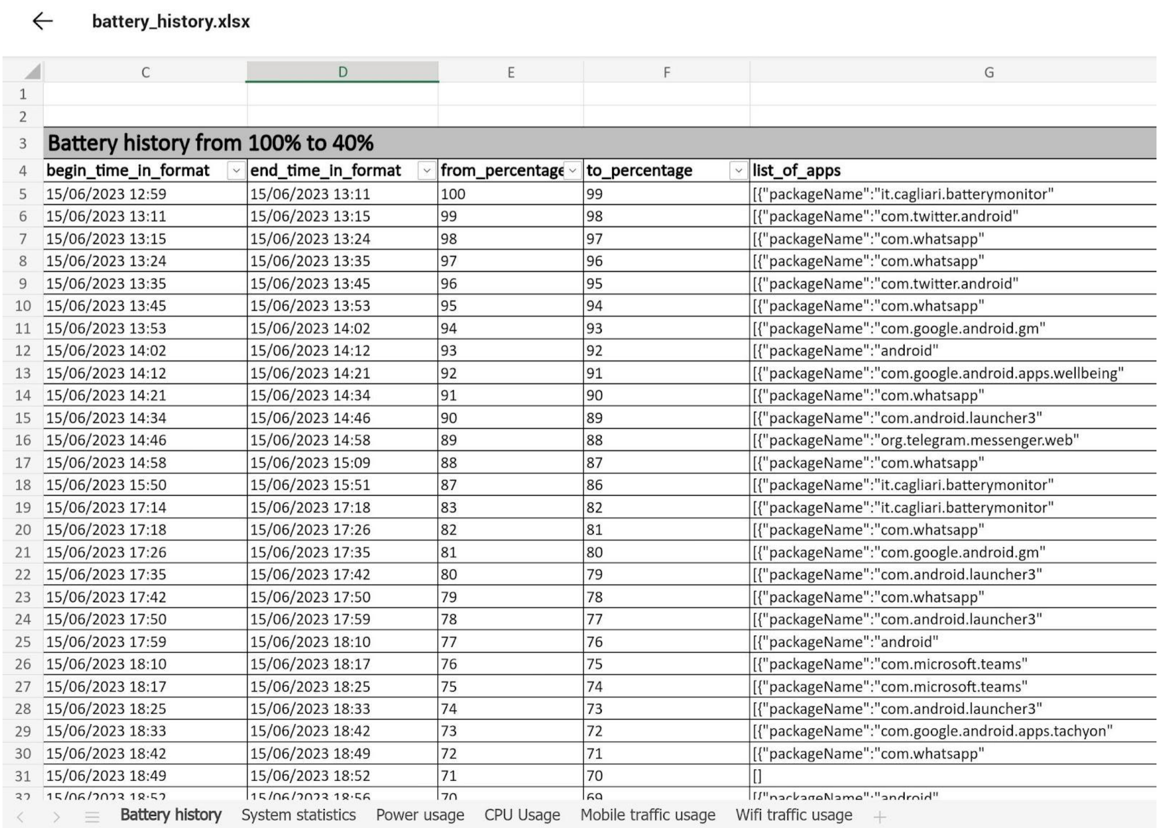


Figure 1: Screenshot of the application that shows the battery percentage and the list of used applications during the last period (when the battery level changed from 91% to the current 90%). The list of applications can be scrolled down and exported.

2.3 ECMA For Research Purposes

Not only does the ECMA focus on short-term monitoring, but it also gathers data over extended periods to identify long-term energy consumption trends. This information can be relevant for research purposes, since a long history consumption can help to create models for predicting energy demands for end-devices. To better understand the correlations between battery consumption and energy-starved applications, we explored the possibility of using Batterystats, a built-in Android framework tool designed to collect battery data on devices. By utilizing Android Debug Bridge (adb), it is possible to extract the collected Batterystats data and transfer it to a development machine for further analysis. Moreover, employing the Android Battery Historian tool, it is possible to transform the extracted data into an HTML visualization, accessible through a Web browser. This visualization provides valuable insights into how various processes consumed battery reserve power and pinpointed specific tasks in our application that could be deferred or removed to optimize battery life.

Figure 2 shows the collection of all the run apps and the information such as time usage, battery consumption and number of iterations of the user with the app since the end-device first run. This information, collected in the long period can produce insightful results for addressing power management. However, it is worth noting that this method requires the physical connection of the end-device to the computer and involves several steps to access and interpret the data. Despite this limitation, the combination of the battery monitoring application information and the adoption of Batterystats and Android Battery Historian can provide insightful results. In fact, collecting the whole battery history data and the data concerning the apps running in the phone, it is possible to find correlations between running apps and the old energy consumptions of the same app. This provides significant information that can help the ECMA algorithm in future applications to provide power consumption predictions to the final user.



battery_history.xlsx

	C	D	E	F	G
1					
2					
3	Battery history from 100% to 40%				
4	begin_time_in_format	end_time_in_format	from_percentage	to_percentage	list_of_apps
5	15/06/2023 12:59	15/06/2023 13:11	100	99	[{"packageName":"it.cagliari.batterymonitor"
6	15/06/2023 13:11	15/06/2023 13:15	99	98	[{"packageName":"com.twitter.android"
7	15/06/2023 13:15	15/06/2023 13:24	98	97	[{"packageName":"com.whatsapp"
8	15/06/2023 13:24	15/06/2023 13:35	97	96	[{"packageName":"com.whatsapp"
9	15/06/2023 13:35	15/06/2023 13:45	96	95	[{"packageName":"com.twitter.android"
10	15/06/2023 13:45	15/06/2023 13:53	95	94	[{"packageName":"com.whatsapp"
11	15/06/2023 13:53	15/06/2023 14:02	94	93	[{"packageName":"com.google.android.gm"
12	15/06/2023 14:02	15/06/2023 14:12	93	92	[{"packageName":"android"
13	15/06/2023 14:12	15/06/2023 14:21	92	91	[{"packageName":"com.google.android.apps.wellbeing"
14	15/06/2023 14:21	15/06/2023 14:34	91	90	[{"packageName":"com.whatsapp"
15	15/06/2023 14:34	15/06/2023 14:46	90	89	[{"packageName":"com.android.launcher3"
16	15/06/2023 14:46	15/06/2023 14:58	89	88	[{"packageName":"org.telegram.messenger.web"
17	15/06/2023 14:58	15/06/2023 15:09	88	87	[{"packageName":"com.whatsapp"
18	15/06/2023 15:50	15/06/2023 15:51	87	86	[{"packageName":"it.cagliari.batterymonitor"
19	15/06/2023 17:14	15/06/2023 17:18	83	82	[{"packageName":"it.cagliari.batterymonitor"
20	15/06/2023 17:18	15/06/2023 17:26	82	81	[{"packageName":"com.whatsapp"
21	15/06/2023 17:26	15/06/2023 17:35	81	80	[{"packageName":"com.google.android.gm"
22	15/06/2023 17:35	15/06/2023 17:42	80	79	[{"packageName":"com.android.launcher3"
23	15/06/2023 17:42	15/06/2023 17:50	79	78	[{"packageName":"com.whatsapp"
24	15/06/2023 17:50	15/06/2023 17:59	78	77	[{"packageName":"com.android.launcher3"
25	15/06/2023 17:59	15/06/2023 18:10	77	76	[{"packageName":"android"
26	15/06/2023 18:10	15/06/2023 18:17	76	75	[{"packageName":"com.microsoft.teams"
27	15/06/2023 18:17	15/06/2023 18:25	75	74	[{"packageName":"com.microsoft.teams"
28	15/06/2023 18:25	15/06/2023 18:33	74	73	[{"packageName":"com.android.launcher3"
29	15/06/2023 18:33	15/06/2023 18:42	73	72	[{"packageName":"com.google.android.apps.tachyon"
30	15/06/2023 18:42	15/06/2023 18:49	72	71	[{"packageName":"com.whatsapp"
31	15/06/2023 18:49	15/06/2023 18:52	71	70	[]
32	15/06/2023 18:52	15/06/2023 18:56	70	69	[{"packageName":"android"

Battery history | System statistics | Power usage | CPU Usage | Mobile traffic usage | Wifi traffic usage

Figure 2: List of applications detected by ECMA.

2.4 Limiting Network Throughput/Bandwidth

In addition to monitoring energy consumption and providing insights into app usage, the ECMA incorporates a feature that allows the users to limit the network bandwidth when using video streaming applications. This approach empowers users to reduce their video streaming sessions' impact on the network, contributing to a more sustainable use of network resources. The Bandwidth management feature of the ECMA offers users the ability to control the network bandwidth allocated to video streaming applications. When a user accesses a video streaming service, the ECMA detects the application and presents an option to limit the network bandwidth dedicated to the streaming session. Users can choose from predefined bandwidth options or set a custom limit based on their preferences.

Video streaming applications are among the most bandwidth-intensive services, and their widespread usage can waste network resources, especially during peak hours. By allowing users to cap the network bandwidth used by video streaming, the ECMA empowers individuals to be responsible consumers of network resources. Limiting bandwidth can help alleviate network congestion and improve overall network performance, benefiting all users sharing the same network infrastructure. Moreover, users can adjust the network bandwidth limits based on their specific needs and network conditions. Reducing the network bandwidth for video streaming sessions has a positive environmental impact as well. Data transmission across networks consumes electricity and generates greenhouse gas emissions. By limiting bandwidth during video streaming, users actively contribute to a reduction in the carbon footprint associated with network operations.

3. Conclusion

Empowering users with tools for monitoring the energy consumption and getting information on the applications that have a significant contribution on this, allows them to take a more environmentally conscious behavior. In this paper, we have presented a tool with this objective: an application to monitor energy consumption on the user end-device that empowers the user to make greener decisions by reducing avoidable high-quality video flows. With this approach, we can pave the way for a sustainable video streaming ecosystem; however, the impact on the user's QoE should be carefully considered. This tool will be made soon available by the authors to the research community as it can help in conducting experiments and novel advanced energy-aware streaming approaches.

Acknowledgement

This work has been partially supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, "Sustainable Mobility Center" (Centro Nazionale per la Mobilità Sostenibile, CNMS, CN_00000023).

References

[1] C. Freitag, M. Berners-Lee, K. Widdicks, B. Knowles, G. S. Blair, and A. Friday, "The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations," Patterns, vol. 2, no. 9, p. 100340, 2021.

[2] Sandvine, "2023 Global Internet Phenomena Report," 2023. [Online]. Available: https://www.sandvine.com/global-internet-phenomena-report-2023.

[3] R. Madlener, S. Sheykhha, and W. Briglauer, "The electricity-and CO2-saving potentials offered by regulation of European video-streaming services," Energy Policy, vol. 161, p. 112716, 2022.

[4] T. Tabata and T. Y. Wang, "Life Cycle Assessment of CO2 Emissions of Online Music and Videos Streaming in Japan," Applied Sciences, vol. 11, no. 9, 2021.

[5] De Bosschere, Koen and Duranton, Marc, "Everything as a service," in HiPEAC vision 2021: high performance embedded architecture and compilation. HiPEAC, 2021, pp. 206–211.

[6] P. Suski, J. Pohl, and V. Frick, "All you can stream: Investigating the role of user behavior for greenhouse gas intensity of video streaming," in Proc. of the 7th Int. Conf. on ICT for Sustainability, 2020, pp. 128–138.

[7] [Online] https://www.statista.com/statistics/272698/global-market-share-held-by-mobile-operating-systems-since-2009/

[8] P. Le Callet, S. Möller, and A. Perkis. (2012) Qualinet White Paper on Definitions of Quality of Experience (2012). European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003), Lausanne, Switzerland, Version 1.2, March 2013.

Authors Bio-data

Simone Porcu is an Assistant Professor at the University of Cagliari (Italy), where he received his M. Sc. degree in Computer Science and his Ph. D. in Electronic and Computer Engineering respectively in 2017 and 2021. He also obtained the Doctor Europaeus title for his research period abroad at the Technische Universitat Berlin (September 2018-March 2019). He is a member of the Net4U Laboratory (https://sites.unica.it/net4u/) of the Department of Electrical and Electronic Engineering (DIEE) at the University of Cagliari from 2017, a member of the Italian University Consortium for Telecommunications (CNIT) from 2018, and a member of the IEEE from 2018. He is Associate Editor for IEEE COMSOC MMTC Communications - Frontiers and participated as Guest Editor for the MDPI Electronics journal. His main research interests include Quality of Experience (QoE), multimedia communications, artificial intelligence, and metaverse applications.

Lazizjon Suyunov is a master's degree student in Transport System Engineering, faculty of Civil and Industrial Engineering, University of Sapienza in 2021-2023 (Italy). He is currently a research fellow at the Department of Electrical and Electronic Engineering (DIEE) at the University of Cagliari. His research interests include integrating transport systems with mobile development, GPS system, AR, VR and mobile development.

Luigi Atzori is Full Professor at the Department of Electrical and Electronic Engineering at the University of Cagliari (Italy) and Research Associate at the Networks for Humans laboratory (Net4U) of CNIT (Consorzio Nazionale Interuniversitario per le Telecomunicazioni). L. Atzori research interests are in multimedia communications and computer networking (wireless and wireline), with emphasis on multimedia QoE, multimedia streaming, NGN service management, service management in wireless sensor networks, architecture and services in the Internet of Things. L. Atzori has been the coordinator of the Marie Curie Initial Training Network on QoE for multimedia services (2015–2018), involving ten European Institutions in Europe and one in South Korea. He has been the associate editor for major journals in the area, including: ACM/Springer Wireless Networks, IEEE IoT Journal, Elsevier Ad Hoc Networks, Advances on Multimedia. He has been the guest editor for the IEEE Communications Magazine, the Springer Monet Journal, the Elsevier Signal Processing: Image Communications Journal.

Sustainability in Industrial Remote Operating

Shirin Rafiei, Kjell Brunnström
RISE Research Institutes of Sweden, Stockholm, Sweden
Dept. of Information Systems and Technology, Mid Sweden University, Sundsvall, Sweden
{shirin.rafiei, kjell.brunnstrom}@ri.se

Abstract

With the advent of technologies in industrial remote controlling (Tele-operation), operators interact with machines from a distance. Such systems improvs personal safety and work environments but there are challenges in conveying accurate on-site information to the remote site, like perceiving (on-site) visual information, which is one of the main inputs for remote operators to interpret real-world events. Providing visual information in remote operating systems needs real-time video streaming which is energy demanding. In this research we investigated the impact of video quality (spatial video resolution) and latency (video buffer size) on user's experience and energy consumption. Overall, there is a trade-off between the user's comfortableness and energy consumption in the lab-based Tele-operation system. We observed the energy consumption through an increase in voltage drop over a fix time period where the current was constant with the highest spatial resolution and the highest video latency while there were no significant differences in user's comfort. We could, thus, encourage users to use and adaptive video streaming considering a tradeoff between acceptable QoE and sustainable video stream choices in Tele-operation.

Keywords: Sustainable video streaming, Teleoperating, Industrial remote operating, Greenhouse gas emissions.

1. Introduction

Attention to energy consumption in industrial Teleoperating is growing. In such semi-autonomous applications, there is a need to understand the trade-off between the experienced video quality and sustainable video streaming [1]. Quality of Experience (QoE) is defined as the "*degree of delight or annoyance*" [2], and it is influenced by many items, including the quality of video streaming offered to users. Bitrate, buffer size, video encoding, communication scenario, and hardware complexity are examples of influencing factors on video streaming and, consequently, on what user's experience [3, 4]. Enabling sustainable video streaming requires sufficient video quality referring to satisfy the user's needs with the lowest use of resources [5]. In this study, we investigated on a laboratory Teleoperating platform and hypothesize that there is a tradeoff between the overall user experience and achieving energy efficiency.

2. Method

In the previous section, we discussed video bitrate and buffer size impacts on electricity consumption. This study investigates the impact of different video configuration (Video quality (Q) and latency (L)) on the user's QoE and energy consumption in video streaming applied for a laboratory Tele-operation system [6]. When users interact with Teleoperation for a long time, overall energy consumption and consequently CO_2 emissions becomes important. In this study, the aim is satisfying the user experience while interacting with Tele-operation systems using video streaming without overestimating the need in the consumption of energy. To study this approach, we considered an experimental study in which four different video configurations are applied in a small lab-based Tele-operated platform (LTP) [6]. The different conditions applied in the video streaming consist of two different levels of video quality and two levels of video latency. The streamed video quality was manipulated by changing the special video resolution parameters (Q1: 720W * 640H, Q2: 480W * 320H) while the frame rate was constant (30 fps). As LTP designed for ultra-low latency real-time video streaming, the latency configuration defined as a base line Latency (L1: 150 ms Glass-to-Glass (G2G) latency) and for the second latency configuration 500 ms was added on top of the default by adding the buffer size in the video streaming pipeline (L2: 650 ms). LTP mimics an industrial teleoperation system. The test set up and procedure is described in [6] in details.

The results used in this research is one part of the main experimental study [6] which is conducted at Perception Lab, RISE Kista, Sweden. Ten users participated in the experiment. First, they read the instructions and signed the consent form and then completed the training session. Before starting the main experiment, they answered to the background questionnaire and then completed the recurring questionnaire during the experiment for each trial. In this study we used just one part of the recurring questionnaire's result when users were asked to rate "*their comfort while interacting with LTP*" (users' comfortableness). The ratings were attained with 5-point Likert scale, and we used numerical interval scales rating from 1 to 5 for the analysis. The test conditions relating the results used in this research are shown in Table 1.

For energy consumption, we used a full charge ($\boldsymbol{V2} = 8.3$ Volt) Ni - MH 4000 mAh battery to launch the whole LTP system for 7 minutes and 30 seconds, while the operating time, type of encoding and the display were constant in the entire experiment. The power supply was connected to a GPU NVIDIA Jetson nano and the toy truck [6], and provided the required electricity for both GPU and the toy truck. To consider the experiment situation equal, users remotely controlled the toy truck using the applied video stream configuration for 1 min and the remained 6 min and 30 s was allocated for video streaming. For analyzing the energy consumption, we calculated the voltage drop in the power supply using SKYRC e680 multi-Chemistry microprocessor control charge. After installing a fully charged battery (8.3 Volt) the specific video configuration was loaded for 7 min and 30 s. Finally, after this limited loading time the battery was disconnected, and we checked the battery voltage ($\boldsymbol{V1}$) with the fixed amount of current (set on: 1.0 A). As all the mentioned set up were fixed during the video streaming (namely, time of operation, encoding, frame rate), the voltage operation just was depended on the type of video streaming configuration (spatial video resolution and/or video latency). As a proxy for the energy consumption, we defined "Voltage drop" parameter ($\Delta\boldsymbol{V}$) relative to energy consumption calculated by Lithium based charger with the constant current set on 1 mA.

$$\Delta\boldsymbol{V} = \|\boldsymbol{V2} - \boldsymbol{V1}\|$$

Table 1: Different test conditions for video streaming in LTP system, applied to this study.

Quality levels	G2G Latency (ms)
Q1: 720 W * 640 H	L1: 150 ms
Q2: 480 W * 320 H	L2: 650 ms

Configuration 1	Configuration 2	Configuration 3	Configuration 4
Q1L1	Q2L1	Q1L2	Q2L2

3. Results and discussion

Error! Reference source not found. shows the voltage drop ($\Delta\boldsymbol{V}$) applied to different video configurations in the limited time (7.5 minutes) while all other video streaming conditions were fixed (encoding, operating time, etc.). The voltage drop for all the applied conditions was about 0.26 Volt but for Q1L2 it increased to 0.3 Volt. **Error! Reference source not found.** presents the Mean Opinion Score (MOS) with 95% confidence intervals for user's comfort at different video quality and video latency. The results show user's comfortableness is reduced under the same latency configuration by degrading the video quality from Q1 to Q2 but the rating to all the applied configuration is between 3 (Fair), and 4 (Good). Regarding the applied One-way Repeated-measures Analysis of Variance (ANOVA) for user's comfort in interacting with LTP, there is no significant differences between different enabled video configuration. Considering the depicted figures, users are satisfied with the applied configuration but the increase on the consumed voltage for Q1L2 and providing the highest level of spatial video quality does not show considerable increase in user's experience (comfortableness) interacting with LTP.

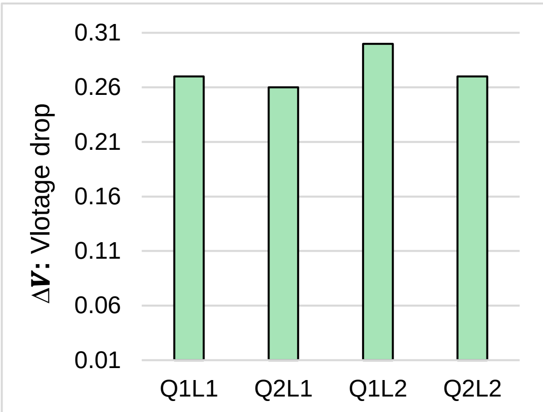


Figure 1: Voltage drop ($\Delta \boldsymbol{V}$) relative to energy consumption calculated by Lithium based charger with the constant current set on 1 mA.

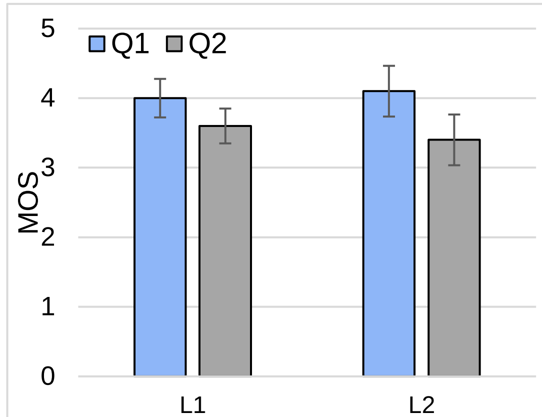


Figure 2: Impact of video latency (buffer size) on MOS for users' comfort (comfortableness) at different video quality.

4. Conclusions

Tele-operating applications are increasing in industrial machinery, requiring visual presentation for remote operators. The study evaluates sustainable video streaming to limit energy consumption and CO_2 emissions, considering user satisfaction with Tele-operating. To address this issue, we presented QoE experiment results of a lab-based Tele-operation, considering the user's comfortableness and the consumed energy while varying the video resolution and video G2G latency. We observed an increase in voltage drop and thus energy using video stream configuration 3 (highest level of spatial resolution and the highest latency level). At the same time, the user's comfortableness is not significantly different. Therefore, when the latency is high, adaptive video streaming (spatial video resolution) would be applicable for remote operators to adjust sustainable video streaming. This study will extend to determine other influencing factors like encoding to optimize the energy consumption level in Tele-operation, so that quality and task performance can be maintained while not consuming more energy than necessary.

References

[1] G. Bingöl, A. Floris, S. Porcu, C. Timmerer and L. Atzori, "Are Quality and Sustainability Reconcilable? A Subjective Study on Video QoE, Luminance and Resolution," 2023 15th International Conference on Quality of Multimedia Experience (QoMEX), Ghent, Belgium, 2023, pp. 19-24.

[2] Perkis, A., Timmerer, C., et al., "QUALINET White Paper on Definitions of Immersive Media Experience (IMEx)", European Network on Quality of Experience in Multimedia Systems and Services, 14th QUALINET meeting (online), May 25, 2020. [Online]: https://arxiv.org/abs/2007.07032.

[3] M. Fiedler, A. Popescu and Y. Yao, "QoE-Aware Sustainable Throughput for Energy-Efficient Video Streaming," 2016 IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom) (BDCloud-SocialCom-SustainCom), Atlanta, GA, USA, 2016, pp. 493-500.

[4] C. Herglotz, W. Robitza, M. Kränzler, A. Kaup and A. Raake, "Modeling of Energy Consumption and Streaming Video QoE using a Crowdsourcing Dataset," 2022 14th International Conference on Quality of Multimedia Experience (QoMEX), Lippstadt, Germany, 2022, pp. 1-6.

[5] Katsenou, A., Wang, X., Bull, D. R., & Schien, D. (2023). Video Quality Sufficiency for Sustainable Video Streaming. Abstract from 15th International Conference on Quality of Multimedia Experience, Ghent, Belgium.

[6] S. Rafiei, C. Singhal, K. Brunnström and M. Sjöström, "Human Interaction in Industrial Tele-Operated Driving:

Laboratory Investigation," 2023 15th International Conference on Quality of Multimedia Experience (QoMEX), Ghent, Belgium, 2023, pp. 91-94.

Authors Bio-data

Shirin Rafiei is a researcher and Ph.D. fellow at RISE Research Institutes of Sweden and Mid Sweden University. Her research interests include Quality of Experience, User Experience, Augmented Telepresence, live video streaming and Augmented Reality.

Kjell Brunnström is a Senior Scientist at RISE Research Institutes of Sweden AB and Adjunct Professor at Mid Sweden University. He is leading development for video quality assessment as Co-chair of the Video Quality Experts Group (VQEG). His research interests are in Quality of Experience for visual media especially immersive media. He is area editor of the Elsevier Journal of Signal Processing: Image Communication and has co-authored > 100 peer-reviewed scientific articles including conference papers.