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

MMTC Communications - Frontiers

Vol. 18, No. 3, May 2023

CONTENTS

SPECIAL ISSUE ON Sustainable multimedia communications and services-I

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 first paper visually illustrates the user sufficiency sub-category by analysing quality-energy curves. By quantifying the trade-offs between user experience and energy consumption, this study enables the measurement and implementation of actionable interventions, facilitating the assessment of potential carbon reductions.

In the second paper, the authors summarise their most recent concepts and discoveries concerning two fundamental video communication services components: encoding and decoding. Various methodologies are thoroughly discussed for assessing and mitigating the energy consumption of diverse video encoder and decoder implementations.

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.

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.

Video Quality Sufficiency for Sustainable Video Streaming

Angeliki Katsenou, Xinyi Wang, David Bull, and Daniel Schien
University of Bristol, UK
{Angeliki.Katsenou, Xinyi.Wang, Dave.Bull, Daniel.Schien}@bristol.ac.uk

Abstract

With the effects of climate change globally manifesting, all sectors of the economy and society are aiming to reduce their carbon emissions. For the ICT sector the concept of sufficiency has been proposed as a paradigm to go beyond efficiency improvements towards an absolute reduction of carbon emissions. Yet little is known about the practical implementation of sufficiency and its carbon reduction potential. We are providing a concrete example of the user sufficiency sub-category by considering the quality-energy curves, in order to make trade-offs between user experience and energy consumption measurable and actionable, and allowing to quantify the carbon reduction potential from this intervention.

Keywords: QoE, Sufficiency, Sustainability, Video Streaming, Energy.

1. Introduction

Growing attention has been given to the carbon footprint of video streaming. The reason is that video comprises the vast volume of streams data [1]. Popular examples of exchanged video data are on-demand streaming of movies (Netflix, Apple TV, HBO, Amazon Prime, etc.), live video conferencing, and collaborative online workspaces (Zoom, Webex, MS Teams, etc.). Furthermore, advances in powerful and affordable mobile devices and cloud computing techniques, have facilitated consumers to create and share live or on-demand short User-Generated Content (UGC) clips over social media/sharing platforms (Instagram, Facebook, TikTok, YouTube, etc.). The increased volume of exchanged video data and environmental awareness are interconnected [2] suggesting that efficiency must improve as technology usage increases if sustainability targets are to be met. In this paper, we provide a position of our future research on how we could drive video streaming by taking into consideration the quality of experience and the required energy expenditure.

2. Related work

Innovation and optimisation of video codecs can affect the energy consumption in all parts of the streaming pipeline, namely the encoder. Previous researches have first focused on the energy profiling of video encoding and decoding for different standardised technologies, such as H.265, VVC, and AV1 [3]–[6], either using software estimators of power such as Intel's RAPL [7] or hardware-based power meters, such as Tektronix PA1000. Moreover, research has focused also assessing the energy consumed at decoding on different end user devices, e.g., phones, laptops, displays [6], [8].

Further to energy profiling, Schien *et al.* [9] and Herglotz *et al.* [10], have identified potential solutions to increase energy efficiency. Schien *et al.* [9] have evaluated a short-term intervention that would impose a spatial resolution restriction in video streaming from Full High Definition (FHD) to High Definition (HD) and found minimal short-term savings. Herglotz *et al.* [10] have studied the effect of video codec switching, from VP9 to H.264 and vice versa, and switching device, from laptop to PC and the respective combinations with codecs. They found that energy saving can indeed occur with the right combinations. Yet methods to reliably estimate the long-term carbon reduction potential from interventions on the video streaming pipeline are currently lacking. While technical solutions can increase the efficiency of services, they are embedded and constrained by boundaries set within a socio-technical context.

3. QoE Sufficiency for Sustainable Video Streaming

In response to the lack of literature studies, the concept of "sufficiency" has been proposed to refer to strategies that directly aim for absolute impact reductions from lowering production and consumption [11]. Although sufficiency and related to this, sobriety is increasingly advocated, their actual potential to affect meaningful carbon reduction has not been explored. User sufficiency in video streaming refers to satisfying user quality (objective) needs or (subjective) expectations with the lowest use of resources in shared infrastructure (data centres and networks) and on user devices. Quality of experience is also further constrained by technical capabilities. For example, users with a low bandwidth network (e.g., less than 2 Mbps) or older screens (less than FHD resolution displays) might be limited and expect Video Multi-Method Assessment Fusion (VMAF) [12] scores between 70-80[1]. While a user with a 4K TV and fibre broadband is able (and expect) VMAF scores of 90 and higher.

Additionally, it is important to consider that the quality-energy (QE) relation like the quality-rate (QR) relation is content-dependent. Therefore, those two curves, QE and QR, are not always aligned. This is illustrated in the example QR and QE curves of Fig. 1. The two subfigures show the QR and QE curves for two sequences from the YouTube-UGC dataset [13], encoded at nine quantisation points (QP), namely $QP = \{12,18,\ldots,60\}$, with SVT-AV1 [14] codec with medium preset. The energy reported in these graphs was measured during encoding and decoding using a Tektronix power meter on an Intel(R) Core (TM) i9-7900X CPU @ 3.30 GHz with 46 GB RAM. The quality values are represented by the VMAF metric [12], that is used by many video providers as it is better aligned with visual human perception. In this example of Fig. 1, it can be seen that for the Animation-646f sequence the curvature and saturation point of the QE and QR curves are almost aligned, while this is not the case for the Television-5278 sequence. In the latter case, the curvature of the energy expenditure occurs at a lower QP value. Similar patterns can be observed in many other types of content and confirm that the energy optimisation of video encoding and streaming has to take into consideration the content characteristics.

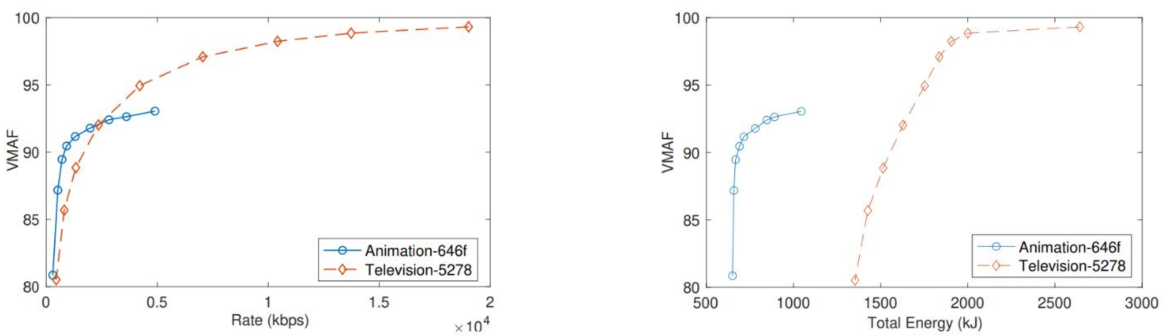

Figure 1: Examples of VMAF-Rate and VMAF-Energy curves using SVT-AV1. These examples demonstrate a different quality-energy tradeoff that cannot be directly aligned with the VMAF-Rate tradeoff. Both sequences are FHD, with 30 fps framerate for Animation-646f and 25 fps for Television-5278.

4. Working Towards Sustainable Video Streaming

Following the discussion above, our position is that we need to explore the use of the QE curve towards greater user sufficiency. This can further be used straightforwardly to drive adaptive bitrate streaming. This means that content-gnostic adaptive bitrate streaming ladders constructed by sampling the quality-rate Pareto front, as in our previous work [15], need to be re-designed by taking into account the quality-energy trade-offs. It also enables estimating the absolute carbon reduction potential from quality constraints [9]. Finally, the Pareto-curve can be used to guide the implementation of user sufficiency in design. For example, behaviour change might be supported by allowing users to choose desired quality levels for varying consumption contexts and provide information on carbon savings alongside such choice. For our future work, we will explore this dimension and will propose a framework that takes into account the quality-rate curve characteristics for an end-to-end optimisation.

[1] VMAF assumes values in the range 0-100 with 100 representing excellent quality.

Acknowledgement

This work has been supported by the Creative Bristol+Bath Cluster and UKRI MyWorld Strength in Places Programme (SIPF00006/1).

References

[1] Sanvdine, "The global internet phenomena report" 2023. [Online]. Available: https://www.sandvine.com/global-internet-phenomena-report-2023-download?submissionGuid=bd3de666-249c-4927-9bab-b84c3577b2c9.

[2] Greenpeace, "Clicking Clean: A Guide to Building the Green Internet," 2015. [Online]. Available: https://www.greenpeace.org/static/planet4-international-stateless/2015/05/153e0823-2015clickingclean.pdf.

[3] A. Katsenou, J. Mao, and I. Mavromatis, "Energy-Rate-Quality Tradeoffs of State-of-the-Art Video Codecs," in Picture Coding Symposium, 2022.

[4] C. Herglotz, A. Heindel, and A. Kaup, "Decoding-energy-rate-distortion optimization for video coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 1, pp. 171–182, 2019.

[5] F. Pakdaman, M. A. Adelimanesh, M. Gabbouj, and M. R. Hashemi, "Complexity Analysis Of Next-Generation VVC Encoding And Decoding," in Proc. of 2020 IEEE International Conference on Image Processing (ICIP), 2020, pp. 3134–3138.

[6] M. Kränzler, C. Herglotz, and A. Kaup, "Energy Efficient Video Decoding for VVC Using a Greedy Strategy Based Design Space Exploration," IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 7, pp. 4696–4709, 2022.

[7] K. N. Khan, M. Hirki, T. Niemi, J. K. Nurminen, and Z. Ou, "RAPL in Action: Experiences in Using RAPL for Power Measurements," ACM Trans. Model. Perform. Eval. Comput. Syst., vol. 3, no. 2, 2018.

[8] C. Herglotz, W. Robitza, M. Kränzler, 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), 2022, pp. 1–6.

[9] D. Schien, P. Shabajee, H. Akyol, L. Benson, and A. Katsenou, "Help, I shrunk my savings! Assessing the Carbon Reduction Potential for Video Streaming from Short-Term Coding Changes," pre-print.

[10] 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), 2023, pp. 79–82.

[11] T. Santarius, J. C. Bieser, V. Frick, M. Höjer, M. Gossen, L. M. Hilty, E. Kern, J. Pohl, F. Rohde, and S. Lange, "Digital sufficiency: conceptual considerations for ICTs on a finite planet," Annales des Telecommunications/Annals of Telecommunications, 2022.

[12] Z. Li, A. Aaron, I. Katsavounidis, A. Moorthy, and M. Manohara, "The NETFLIX Tech Blog: Toward a Practical Perceptual Video Quality Metric," 2016 [Online]. Available: https://netflixtechblog.com/toward-a-practical-perceptual-video-quality-metric-653f208b9652.

[13] YouTube, "UGC Dataset: A large dataset containing YouTube User Generated Content intended for video compression and quality assessment research," https://media.withyoutube.com/, 2021.

[14] F. Kossentini, H. Guermazi, N. Mahdi, C. Nouira, A. Naghdinezhad, H. Tmar, O. Khlif, P. Worth, F. Ben Amara, "The SVT-AV1 Encoder: Overview, Features and Speed-Quality Tradeoffs," in Proc. of Applications of Digital Image Processing XLIII, vol. 11510. International Society for Optics and Photonics, 2020, p. 1151021.

[15] A. V. Katsenou, J. Sole, and D. R. Bull, "Efficient Bitrate Ladder Construction for Content-Optimized Adaptive Video Streaming," IEEE Open Journal of Signal Processing, vol. 2, pp. 496–511, 2021.

Authors Bio-data

Angeliki V. Katsenou received a Diploma in Electrical and Computer engineering and an MSc degree in Signal and Image Processing from the University of Patras, Greece, respectively, and a PhD degree from the Department of Computer Science and Engineering, University of Ioannina, Greece. From 2015-2022, she was a Senior Research Fellow with the Visual Information Lab, University of Bristol, U.K. In 2022, she was an Assistant Professor in the Electronics and Electrical Engineering Department, at Trinity College Dublin, Ireland, and since 2023 she is a Senior Lecturer in Networked Media at the University of Bristol. Her main research interests include topics around the video streaming pipeline: acquisition, analysis, compression, and communication, as well as adding the human in the loop to understand perception and quality of experience. Other side research activities include data and biomedical engineering. She was awarded a Leverhulme Early-Career Fellowship in 2017 and has experience in several EC-funded, UKRI projects and direct collaborations with the video technology industry.

Xinyi Wang received the B.Sc (Eng) with First Class Honours in Telecommunications Engineering and Management from the joint program between Beijing University of Posts and Telecommunications and Queen Mary University of London, Beijing, China, in 2020. She also received the M.Sc. degree with Distinction in Image and Video Communication and Signal Processing from the University of Bristol, Bristol, U.K., in 2021. Since 2022, she has been a PhD student in Electrical and Electronic Engineering at the Visual Information (VI) Laboratory, University of Bristol, Bristol, UK. Her PhD is sponsored by UKRI MyWorld Strength in Places Programme (SIPF00006/1) Scholarship, under the supervision of Dr Angeliki Katsenou and Prof David Bull. Her main research interests include perceptual video analysis, image or video quality assessment, computer vision, and sustainable ICTs.

David R. Bull (M'94-SM'07-F'12) received the B.Sc. degree from the University of Exeter, Exeter, U.K., in 1980; the M.Sc. degree from University of Manchester, Manchester, U.K., in 1983; and the Ph.D. degree from the University of Cardiff, Cardiff, U.K., in 1988. Dr Bull has previously been a Systems Engineer with Rolls Royce, Bristol, U.K. and a Lecturer at the University of Wales, Cardiff, U.K. He joined the University of Bristol in 1993 and is currently its Chair of Signal Processing and Director of Bristol Vision Institute. In 2001, he co-founded a university spin-off company, ProVision Communication Technologies Ltd., specializing in wireless video technology. He has authored over 450 papers on the topics of image and video communications and analysis for wireless, Internet and broadcast applications, together with numerous patents, several of which have been exploited commercially. He has received two IET Premium awards for his work. He is the author of three books, and has delivered numerous invited/keynote lectures and tutorials. Dr. Bull is a fellow of the Institution of Engineering and Technology.

Daniel Schein received the Ph.D. degree in Computer Science from the University of Bristol, Bristol, U.K. in 2015, and the M.Sc. degree in Computer Science from Technical University Berlin, Berlin, Germany. He is currently a Senior Lecturer in Computer Systems in the Department of Computer Science at the University of Bristol. His research aims at improving our understanding and mitigating the environmental impact of energy consumption in the built environment with a focus on information and communication technologies (ICT).

Energy Efficiency in Video Compression

Christian Herglotz, Matthias Kränzler, Geetha Ramasubbu, Lena Eichermüller, André Kaup
Multimedia Communications and Signal Processing
Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
{firstname.lastname}@fau.de

Abstract

In the past decade, the increasing use of video communication services has led to video communications contributing substantially to greenhouse gas emissions nowadays. As a consequence, researchers started investigating the reasons for the high energy consumption and are coming up with solutions for significant energy reduction. In this paper, we review our latest ideas and findings with respect to two important parts in video communication services, namely video encoding and video decoding, which are both performed on billions of devices worldwide.

Keywords: Energy, Video, Encoding, Decoding.

1. Introduction

Today, online video applications are ubiquitous. Online video services provide diverse applications such as video sharing, teleconferencing, video streaming, or Internet protocol television. Furthermore, end-user devices such as smartphones, tablet PCs, desktop PCs, and TV sets are easily accessible such that billions of people make use of this technology everyday.

Studies revealed that the worldwide video communication infrastructure nowadays contributes substantially to greenhouse gas emissions [1]. A main reason is the large number of end-user devices, which cause greenhouse gas emissions in device production and during the use of video services. Another reason is the encoding of videos, which, depending on the encoding setup, can lead to substantial energy consumption on the provider side [2].

In this work, as illustrated in Fig. 1, we summarize our research concerning the energy consumption of encoding and decoding videos in Sections 2 and 3. We will show that these two components, which are essential in video communication pipelines, inherit large potentials for saving energy.

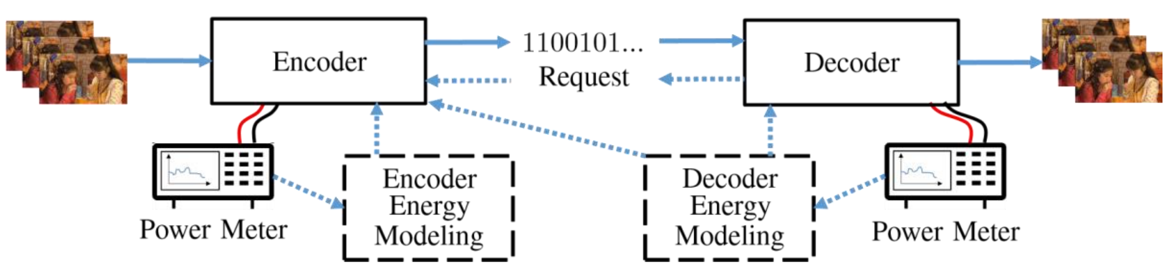

Figure 1: Energy measurement, modeling, and optimization for video encoding (left) and decoding (right). The solid arrows indicate the classic information pipeline in video communication systems and the dotted arrows the opportunities to achieve energy reductions on the encoder and the decoder side.

2. Video Encoding

It is well known that video encoding is much more complex than video decoding. The reason is that usually, practical encoders need to test a high number of coding modes and select the mode leading to the lowest costs in terms of rate and distortion. In contrast, the decoder only has to process the mode selected by the encoder, such that the overall complexity is much lower.

Therefore, it is worthwhile to investigate opportunities to reduce the complexity and thus the energy consumption of the encoding process. A great deal of research targets efficient implementations or fast encoder decisions [3]–[7]. Also, most modern encoder implementations provide so-called presets, which can be chosen to find a trade-off between encoding time and the desired compression performance [8].

A methodological problem in these approaches is that they target a low complexity or a low processing time. Although the complexity and the processing time are highly correlated with the actual energy consumption [9], minimizing them does not ensure that the actual energy consumption is reduced. Consequently, we propose to focus on the energy consumption of video encoding. We develop measurement procedures, perform detailed analyses on measurement results, and optimize encoders with the explicit goal of reducing the encoding energy.

A corresponding measurement setup and first results on the encoding energy consumption and its relation with the encoding time is shown in [9] for the High-Efficiency Video Coding (HEVC) standard [10]. The corresponding part in Fig. 1 is the encoder's power meter and the encoder energy modeling. We found that next to the actual encoding time, which can approximate the encoding energy with a mean relative error of roughly 8% for the x265 encoder implementation [11], the encoding time of the fastest preset is a good estimator for the encoding energy of other presets (mean relative error of roughly 11% across all presets). Hence, the very lightweight and low-energy preset can provide a good estimate on the energy consumption at a high-energy preset, which can help developers in selecting an appropriate preset and thus save energy.

In a similar direction, we developed a method to perform processing time control during encoding AV1 bit streams [12]. It allows the encoder to reach a user-defined encoding time with a high precision [13] and is indicated by the dashed arrow from the encoder energy modeling to the encoder in Fig. 1. The method performs adaptive preset switching during the encoding process and switches the preset on a frame-by-frame basis. The desired target time was reached with a mean relative error of 9%. In the future, it is planned to extend this approach towards encoding energy control such that users of the encoder can set the desired energy consumption of the process.

On top of this, there is a high potential for further improvements. For example, concepts such as rate-complexity-distortion optimization [14] or encoding time optimizations [3] could be adapted to encoding energy. Also, the space of encoder configuration capabilities using video parameters such as frame rate or resolution and the enabling and disabling of coding tools [15] could be further investigated to optimize the energy consumption of encoders.

3. Video Decoding

Substantial work has also been done in optimizing video decoders. Similar to video encoding, a lot of work targeted to reduce the decoding complexity or the decoder processing time [16]–[18]. Also, a common approach is to develop dedicated hardware chips that consume less energy than classic software decoders on a CPU [19]–[22].

Also in this case, we propose to focus on the actual energy consumption of the decoding process. In this direction, similar to encoding, we developed dedicated measurement methods and analyzed the energy consumption of existing software as well as hardware decoders in detail [15], [23], [24]. Analyzing the measurement results, we find that the decoding energy can be modeled using a wide variety of different classifiers such as high-level video features [25], the decoding time [26], and bit stream features [23], [27], which was validated for multiple different video codecs [28]. With the help of this knowledge, we are able to develop various methods to reduce the decoding energy consumption.

A powerful method to reduce the decoding energy is teaching encoders to generate decoding-energy-saving bitstreams (dashed arrow from decoder energy modeling to the encoder in Fig. 1). In this direction, the classic rate-distortion optimization (RDO) process [29], which selects the best coding mode during encoding, is extended to decoding-energy-rate-distortion optimization (DERDO) [30]. We find that it leads to maximum decoding energy savings between 30% and 40% for the HEVC video codec. To achieve this, we exploit a bit-stream-feature-based energy estimation model which calculates the estimated decoding energy during encoder runtime. This information then helps to select coding modes leading to a low decoding energy.

For the successor codec Versatile Video Coding (VVC) [31], we developed a different method which enables and disables coding tools during encoding [15]. We performed extensive measurements on the impact of disabling coding tools on rate-distortion performance and decoding energy consumption and derived optimized encoder configurations [32]. We find that by disabling computationally complex coding tools, decoding energy savings of more than 40% can be achieved.

Another application for bit-stream-feature-based models is the use in metadata of the video bit stream. In the context of the Green Metadata standard [33], we developed syntax elements for a supplemental enhancement information (SEI) message, which can be used by the decoder device to reduce its power consumption when using software decoding [34]. The principle is as follows:

During encoding, the encoder collects statistics on the use of certain coding tools in a defined segment (e.g., a number of frames) of the final bit stream. These statistics are transmitted to the receiver as side information. With the help of these statistics, which are transmitted prior to the actual video information, the decoder can estimate the processing time or processing energy for decoding the corresponding video segment (dashed arrow from the decoder energy model to the decoder in Fig. 1). In order to keep the constraint that the video segment needs to be decoded in real-time, the decoder can then calculate an available CPU time budget for the segment. If the time budget exceeds the estimated processing time, the device can reduce the CPU's frequency using techniques such as dynamic voltage and frequency scaling (DVFS). This will ultimately lead to a reduced power consumption [35].

Another direction targeting energy-efficient video decoding deals with spatio-temporal scaling of the video before compression. In terms of rate-distortion performance, it is well known that at lower bitrates, it can be beneficial to reduce the spatial resolution of a video to increase the compression performance [36]. In our works, we evaluate the impact of reducing the video's spatial and temporal resolution on the decoder's energy consumption. Concerning the spatial resolution, we find that for HD sequences, it is highly beneficial for the smartphone's power consumption to spatially downscale the video [37]. Power savings up to 15% are reported. For temporal downscaling, which corresponds to a reduction of the frame rate, the effectiveness of the scaling process highly depends on the content of the sequence [38]. As such, temporal downscaling is beneficial for highly static videos, however it can be detrimental in sequences containing complex motion. Power savings up to 80% are observed at the same visual quality. In the Green Metadata standard [33], such rescaling can be requested by the decoder (dashed arrow from the decoder to the encoder in Fig. 1).

In future work, we will apply these concepts to other video codecs and combine them to foster further energy savings. Also, more research can be performed on specialized hardware decoder implementations, which are used on most modern portable devices.

4. Conclusion

In conclusion, video encoding and video decoding are essential processes in online video systems that consume a significant amount of energy. We discussed various approaches to analyze and reduce this energy consumption of various video encoder and decoder implementations. In future work, we will investigate further opportunities for energy reduction and extend our research towards upcoming technologies such as neural-network-based image and video compression and the joint optimization of encoding and decoding energy.

References

[1] The Shift Project, "Climate crisis: The unsustainable use of online video," Tech. Rep., 2019.
[2] C. Herglotz, M. Kränzler, R. Schober, and A. Kaup, "Sweet streams are made of this: The system engineer's view on energy efficiency in video communications," IEEE Circuits and Systems Magazine, vol. 23, no. 1, pp. 57–77, 2023.
[3] J. Vanne, M. Viitanen, and T.D. Hamalainen, "Efficient mode decision schemes for HEVC inter prediction," IEEE Transactions on Circuits and Systems for Video Technology, vol. 24, no. 9, pp. 1579–1593, Sep 2014.

[4] A. Tissier, A. Mercat, T. Amestoy, W. Hamidouche, J. Vanne, and D. Menard, "Complexity reduction opportunities in the future vvc intra encoder," in Proc. IEEE 21st International Workshop on Multimedia Signal Processing (MMSP). IEEE, 2019, pp. 1–6.

[5] M. B. A. Khernache, Y. Benmoussa, J. Boukhobza, and D. Menard, "Hevc hardware vs software decoding: An objective energy consumption analysis and comparison," Journal of Systems Architecture, vol. 115, 2021.

[6] J. Brandenburg, A. Wieckowski, T. Hinz, A. Henkel, V. George, I. Zupancic, C. Stoffers, B. Bross, H. Schwarz, and D. Marpe, "Towards fast and efficient vvc encoding," in Proc. 22nd International Workshop on Multimedia Signal Processing (MMSP), 2020, pp. 1–6.

[7] A. Wieckowski, B. Bross, and D. Marpe, "Fast partitioning strategies for vvc and their implementation in an open optimized encoder," in Proc. Picture Coding Symposium (PCS). IEEE, 2021, pp. 1–5.

[8] P.-H. Wu, V. Kondratenko, and I. Katsavounidis, "Fast encoding parameter selection for convex hull video encoding," in Proc. Applications of Digital Image Processing XLIII, Andrew G. Tescher and Touradj Ebrahimi, Eds. International Society for Optics and Photonics, 2020, vol. 11510, pp. 181 – 194, SPIE.

[9] G. Ramasubbu, A. Kaup, and C. Herglotz, "Modeling the HEVC encoding energy using the encoder processsing time," in Proc. IEEE International Conference on Image Processing (ICIP), Bordeaux, France, 2022.

[10] G.J. Sullivan, J. Ohm, Woo-Jin Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1649 –1668, Dec 2012.

[11] "x265: H.265 / HEVC video encoder application library," https://www.videolan.org/developers/x265.html, accessed 2023-07.

[12] Faouzi Kossentini, Hassen Guermazi, Nader Mahdi, Chekib Nouira, Amir Naghdinezhad, Hassene Tmar, Omar Khlif, Phoenix Worth, and Foued Ben Amara, "The SVT-AV1 encoder: overview, features and speed-quality tradeoffs," in Applications of Digital Image Processing XLIII. International Society for Optics and Photonics, 2020, vol. 11510, p. 1151021.

[13] L. Eichermüller, G. Chaudhari, I. Katsavounidis, Z. Lei, H. Tmar, A. Kaup, and C. Herglotz, "Encoder complexity control in svt-av1 by speed-adaptive preset switching," in accepted for IEEE International Conference on Image Processing (ICIP), 2023.

[14] X. Li, M. Wien, and J.-R. Ohm, "Rate-complexity-distortion optimization for hybrid video coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 21, no. 7, pp. 957 –970, Jul 2011.

[15] M. Kränzler, C. Herglotz, and A. Kaup, "Energy efficient video decoding for VVC using a greedy strategy based design space exploration," IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 7, pp. 4696–4709, 2022.

[16] Y.-H. Chen and V. Sze, "A 2014 mbin/s deeply pipelined CABAC decoder for HEVC," in Proc. IEEE International Conference on Image Processing (ICIP), Paris, France, Oct 2014, pp. 1967–1971.

[17] A. Wieckowski, G. Hege, C. Bartnik, C. Lehmann, C. Stoffers, B. Bross, and D. Marpe, "Towards a live software decoder implementation for the upcoming versatile video coding (vvc) codec," in Proc. IEEE International Conference on Image Processing (ICIP), 2020, pp. 3124–3128.

[18] R. Yang, M. Xu, Z. Wang, Y. Duan, and X. Tao, "Saliency-guided complexity control for HEVC decoding," IEEE Transactions on Broadcasting, vol. 64, no. 4, pp. 865–882, Dec. 2018.

[19] C.-T. Huang, M. Tikekar, C. Juvekar, V. Sze, and A. Chandrakasan, "A 249mpixel/s HEVC video-decoder chip for quad full HD applications," in IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC), San Francisco, USA, Feb 2013, pp. 162–163.

[20] D. Zhou, S. Wang, H. Sun, J. Zhou, J. Zhu, Y. Zhao, J. Zhou, S. Zhang, S. Kimura, T. Yoshimura, and S. Goto, "An 8k H.265/HEVC video decoder chip with a new system pipeline design," IEEE Journal of Solid-State Circuits, vol. 52, no. 99, pp. 113–126, 2016.

[21] D. Engelhardt, J. Hahlbeck, J. Möller, and B. Stabernack, "FPGA implementation of a full HD real-time HEVC main profile decoder," IEEE Transactions on Consumer Electronics, vol. 60, no. 3, pp. 476–484, August 2014.

[22] S. Park and K. Ryoo, "The hardware design of effective SAO for HEVC decoder," in Proc. IEEE 2nd Global Conference on Consumer Electronics (GCCE), Las Vegas, USA, Oct 2013, pp. 303–304.

[23] C. Herglotz, D. Springer, M. Reichenbach, B. Stabernack, and A. Kaup, “Modeling the energy consumption of the HEVC decoding process,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 1, pp. 217–229, Jan. 2018.

[24] C. Herglotz and A. Kaup, “Decoding energy estimation of an HEVC hardware decoder,” in Proc. International Symposium on Circuits and Systems (ISCAS), Firenze, Italy, May 2018, pp. 1–5.

[25] C. Herglotz and A. Kaup, “Estimating the HEVC decoding energy using high-level video features,” in Proc. 23rd European Signal Processing Conference (EUSIPCO), Nice, France, Aug 2015, pp. 1591–1595.

[26] C. Herglotz, E. Walencik, and A. Kaup, “Estimating the HEVC decoding energy using the decoder processing time,” in Proc. International Symposium on Circuits and Systems (ISCAS), Lisbon, Portugal, May 2015, pp. 513–516.

[27] M. Kränzler, C. Herglotz, and A. Kaup, “Decoding energy modeling for versatile video coding,” in Proc. International Conference on Image Processing (ICIP), 10 2020, pp. 1–5.

[28] C. Herglotz, Y. Wen, B. Dai, M. Kränzler, and A. Kaup, “A bitstream feature based model for video decoding energy estimation,” in Proc. Picture Coding Symposium (PCS), Nürnberg, Germany, Dec 2016.

[29] G.-J. Sullivan and T. Wiegand, “Rate-distortion optimization for video compression,” IEEE Signal Processing Magazine, vol. 15, no. 6, pp. 74 –90, Nov 1998.

[30] C. Herglotz, A. Heindel, and A. Kaup, “Decoding-energy-rate-distortion optimization for video coding,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 1, pp. 172–181, Jan. 2019.

[31] B. Bross, J. Chen, J.-R. Ohm, G. J. Sullivan, and Y.-K. Wang, “Developments in international video coding standardization after AVC, with an overview of versatile video coding (VVC),” Proceedings of the IEEE, vol. 109, no. 9, pp. 1463–1493, 2021.

[32] M. Kränzler, A. Wieckowski, G. Ramasubbu, B. Bross, A. Kaup, D. Marpe, and C. Herglotz, “Optimized decoding-energy-aware encoding in practical vvc implementations,” in Proc. IEEE International Conference on Image Processing (ICIP). IEEE, 2022, pp. 1536–1540.

[33] Information technology - MPEG Systems Technologies - Part 11: Energy-Efficient Media Consumption (Green Metadata) - Third Edition, ISO/IEC 23001-11, ISO/IEC JTC 1/SC 29, 2023.

[34] C. Herglotz, M. Kränzler, A. Kaup, E. Francois, M. Radosavljevic, E. Reinhard, X. Ducloux, D. Menard, Y. He, M. Coban, D. Rusanovskyy, and M. Karczewicz, “Proposed update of green mpeg on complexity metrics for decoder-power reduction for VVC,” Mpeg m57977, ISO/IEC JTC 1/sc 29/WG 3, 2021.

[35] F. C. Fernandes, X. Ducloux, Z. Ma, E. Faramarzi, P. Gendron, and J. Wen, “The green metadata standard for energy-efficient video consumption,” IEEE MultiMedia, vol. 22, no. 1, pp. 80–87, 2015.

[36] J. Katsenou, A. V .and Sole and D. R. Bull, “Content-gnostic bitrate ladder prediction for adaptive video streaming,” in Proc. Picture Coding Symposium (PCS). IEEE, 2019, pp. 1–5.

[37] C. Herglotz, A. Kaup, S. Coulombe, and S. Vakili, “Power-efficient video streaming on mobile devices using optimal spatial scaling,” in Proc. IEEE International Conference on Consumer Electronics (ICCE), Berlin, Germany, Sep. 2019.

[38] C. Herglotz, M. Kränzler, R. Ludwig, and A. Kaup, “Video decoding energy reduction using temporal-domain filtering,” in Proc. Green Multimedia Systems Workshop (GMSys) on the ACM Multimedia Systems Conference (ACM MMSYS), 2023.

Authors Bio-data

Christian Herglotz received the Dipl.-Ing. in electrical engineering and information technology in 2011 and the Dipl.-Wirt. Ing. in business administration and economics in 2012, both from Rheinisch-Westfälische Technische Hochschule (RWTH) Aachen University, Germany. Since 2012 he has been a Research Scientist with the Chair of Multimedia Communications and Signal Processing, Friedrich-Alexander University Erlangen-Nürnberg (FAU), Germany, where he received his Dr.-Ing. degree in 2017. In 2018 and 2019, he worked as a PostDoc-Fellow at École de technologie supèrieure in collaboration with Summit Tech Multimedia, Montrèal, Canada on energy efficient VR technologies. Since 2019, he is with Friedrich-Alexander University Erlangen-Nürnberg as a senior scientist. His current research interests include energy efficient video communications and video coding. Since 2020, he is with the Visual Signal Processing and Communications Technical Committee of the IEEE Circuits and Systems Society. Since 2023, he serves as an associate editor for the Transactions on Circuits and Systems for Video Technology (T-CSVT).

Matthias Kränzler received the B.Sc. and M.Sc. degrees in electrical engineering and information technology from the Friedrich Alexander Universität-Erlangen-Nürnberg (FAU), Germany, in 2017 and 2019, respectively. During his studies, he worked as a student assistant on modeling and optimizing the energy demand of video decoders from 2015 to 2019. Since 2019, he has been a Research Scientist with the Chair of Multimedia Communications and Signal Processing, Friedrich-Alexander Universität Erlangen-Nürnberg (FAU), Germany. In 2023, he worked as a PhD intern at Apple in Cupertino, USA. His research interests include video coding and solutions for energy efficient video communication.

Geetha Ramasubbu received the B.E. degree in electronics and Communication Engineering from Anna University, Chennai, India, in 2015 and the M.Sc. degree in Communications and Multimedia Engineering from Friedrich-Alexander University Erlangen-Nürnberg (FAU), Germany, in 2020, respectively. Since 2021, she has been a Research Assistant with the Chair of Multimedia Communications and Signal Processing FAU. Her research interests include video encoding and solutions for energy-efficient video encoding.

Lena Eichermüller (Student Member, IEEE) was born in Eschenbach, Germany, in 1996. She received the B.Sc. and M.Sc. in computational engineering, from the Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany, in 2019 and 2022, respectively. In 2022, she joined the Chair of Multimedia Communications and Signal Processing, Friedrich-Alexander-Universität Erlangen-Nürnberg, where she is currently working towards her Ph.D. Her research interests include energy efficient video communications and video coding.

Andrè Kaup (Fellow, IEEE) received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineer-ing from RWTH Aachen University, Aachen, Germany, in 1989 and 1995, respectively. He joined Siemens Corporate Technology, Munich, Germany, in 1995, and became the Head of the Mobile Applications and Services Group in 1999. Since 2001, he has been a Full Professor and the Head of the Chair of Multimedia Communications and Signal Processing at Friedrich-Alexander University Erlangen-Nürnberg (FAU), Germany. From 2005 to 2007 he was Vice Speaker of the DFG Collaborative Research Center 603. From 2015 to 2017, he served as the Head of the Department of Electrical Engineering and Vice Dean of the Faculty of Engineering at FAU. He has authored around 450 journal and conference papers and has over 120 patents granted or pending. His research interests include image and video signal processing and coding, and multimedia communication. Dr. Kaup is a member of the IEEE Image, Video, and Multidimensional Signal Processing Technical Committee and a member of the Scientific Advisory Board of the German VDE/ITG. He is an IEEE Fellow and a member of the Bavarian Academy of Sciences. He is a member of the Editorial Board of the IEEE Circuits and Systems Magazine. He was a Siemens Inventor of the Year 1998 and obtained the 1999 ITG Award. He received several IEEE best paper awards, including the Paul Dan Cristea Special Award in 2013, and his group won the Grand Video Compression Challenge from the Picture Coding Symposium 2013. The Faculty of Engineering with FAU and the State of Bavaria honored him with Teaching Awards, in 2015 and 2020, respectively. He served as an Associate Editor of the IEEE Transactions on Circuits and Systems for Video Technology. He was a Guest Editor of the IEEE Journal of Selected Topics in Signal Processing.

Towards Low-Latency and Energy-Efficient Hybrid P2P-CDN Live Video Streaming

Reza Farahani, Christian Timmerer, and Hermann Hellwagner
Institute of Information Technology, Alpen-Adria-Universität Klagenfurt, Austria
{firstname.lastname}@aau.at

Abstract

Streaming segmented videos over the Hypertext Transfer Protocol (HTTP) is an increasingly popular approach in both live and video-on-demand (VoD) applications. However, designing a scalable and adaptable framework that reduces servers' energy consumption and supports low latency and high quality services, particularly for live video streaming scenarios, is still challenging for Over-The-Top (OTT) service providers. To address such challenges, this paper introduces a new hybrid P2P-CDN framework that leverages new networking and computing paradigms, i.e., Network Function Virtualization (NFV) and edge computing for live video streaming. The proposed framework introduces a multi-layer architecture and a tree of possible actions therein (an action tree), taking into account all available resources from peers, edge, and CDN servers to efficiently distribute video fetching and transcoding tasks across a hybrid P2P-CDN network, consequently enhancing the users' latency and video quality. We also discuss our testbed designed to validate the framework and compare it with baseline methods. The experimental results indicate that the proposed framework improves user Quality of Experience (QoE), reduces client serving latency, and improves edge server energy consumption compared to baseline approaches.

Keywords: HAS; DASH; Edge Computing; NFV; CDN; P2P; Low Latency; QoE; Video Transcoding; Energy Efficiency.

1. Introduction

A. Motivation

The rise of innovative video streaming technologies, the evolution of networking paradigms, and the growing population of users who prefer to stream online video content instead of traditional TV have collectively established video as the predominant traffic on the Internet [1]. According to the Cisco Annual Internet Report, video traffic will make up more than 60% of the entire IP network traffic by 2023, where live video streaming has experienced remarkable popularity among all types of video traffic, constituting approximately 17% of the total video traffic [2]. HTTP Adaptive Streaming (HAS) delivery systems, such as those based on the MPEG Dynamic Adaptive Streaming over HTTP (DASH) [3] standard or Apple HTTP Live Streaming (HLS) [4], have emerged as the dominant technologies utilized by OTT service providers like Facebook, YouTube, and Twitch for delivering live video streaming [5]. In HAS-based streaming, a video is divided into small segments of fixed duration. Each segment then is encoded at multiple qualities or bitrates, known as representations. HAS clients then employ an adaptive bitrate algorithm (ABR), which enables them to adaptively download representations from media servers (i.e., Content Delivery Network (CDN) servers) considering bandwidth and playout buffer conditions. Although integrating CDN services alongside HAS-based systems has represented a significant advancement in video delivery, the substantial surge in demand for high quality and low-latency live video poses numerous challenges for OTT providers. For instance, increasing the number of video users in popular events such as football matches can overload CDN servers. This situation may cause OTT services to deliver unsatisfactory quality and latency to end users. Consequently, OTT providers put significant pressure on backhaul networks as a bottleneck of video delivery systems [6].

Recent studies have demonstrated that incorporating clients' capabilities within a Peer-to-Peer (P2P) network to establish hybrid P2P-CDN video delivery systems effectively tackles the aforementioned challenges while offering several benefits. These advantages include mitigating network congestion, enhancing streaming stability, and lowering overall delivery costs [7], [8]. Considering these benefits, many companies, such as Peer5 and Livepeer, have adopted the utilization of peer-assisted networks, employing promising networking protocols like WebRTC to accomplish the aforementioned goals. Anjum *et al.* [9] indicate that current hybrid P2P-CDN live streaming systems do not use the full potential of peers to deliver high-quality and low-latency live streaming. Hence, the primary motivation of our research is to develop a hybrid P2P-CDN live streaming system that accomplishes the following key goals: (i) use both the computing and bandwidth capabilities offered by the P2P network, (ii) utilize modern networking paradigms such as NFV and edge computing to propose a virtual P2P-CDN tracker server at the edge of the network, and (iii) ensure the satisfaction of HAS client requests with improved QoE and modest latency.

B. Related Work

The Internet Engineering Task Force (IETF) introduced a peer assistance standard called Application-Layer Traffic Optimization (ALTO) [10]. ALTO enables P2P applications to receive abstract maps of network information, such as information from CDNs. Having such information allows ALTO to optimize the utilization of network resources and efficiently deliver traffic while maintaining optimal application performance. However, the ALTO protocol still has drawbacks, and its usage is limited [11]. As another well-known peer assistance product, Akamai proposed the NetSession Interface [12] that supports peer and CDN cooperation. However, it forces users to install extra software on their devices. Ma *et al.* [13] introduced machine learning-based approaches for hybrid P2P-CDN systems enabling their trackers for peer selection. However, their system does not employ edge-supported methods nor use peers' computational resources. Multiple works like [14], [15] customized HAS players with a prefetching module to propose such a hybrid system to reduce CDN bandwidth usage and transmission costs. Our previous works [16]–[20] proposed edge- and SDN- (Software-Defined Networking) assisted video streaming frameworks without considering P2P capability and mainly focusing on VoD scenarios. In other works [21], [22], we proposed a hybrid P2P-CDN architecture for low latency live video streaming without implementation or evaluation of the energy consumption of the tracker server. However, the framework introduced in this paper is different from all the aforementioned approaches since it: *(i)* uses all feasible resources of peers, i.e., storage, bandwidth, and computation, *(ii)* introduces a virtualized tracker server (VTS), deployable on any networking appliance, *(iii)* is compatible with any ABR algorithm due to not forcing clients to modify their codes or installing additional software, and *(iv)* investigates the energy consumption at the edge.

C. Contributions

To tackle these challenges mentioned in the previous sections, in this paper, we leverage HAS, P2P, CDN, NFV, and edge computing technologies to present a new hybrid P2P-CDN framework for live video streaming. Our primary objective is to enhance the serving latency and QoE for HAS clients, taking into account different resource limitations. To do that, we design an action tree including all possible actions for serving clients' requests employed by Virtual Tracker Servers (VTSs) at the edge of a P2P-CDN network. To validate the practical deployment of our solution, we implement the proposed approach and analyze its performance through experiments conducted in a cloud-based testbed, including 350 DASH clients. The experimental results demonstrate high users' QoE, low latency, and low edge server energy consumption compared with selected baseline approaches.

2. System Design

The proposed architecture, including four core layers, is illustrated in Fig. 1.

1) Media Origin Layer: This layer first encodes the raw live video sequences, then packages them into DASH format before storing them on the origin server. Note that this layer has the ability to package the encoded video in other formats, such as HLS or Common Media Application Format (CMAF).
2) CDN Layer: This layer includes a group of CDN servers, which can be OTT servers or a purchased service from CDN providers. Each server stores various parts of video sequences. Moreover, CDN

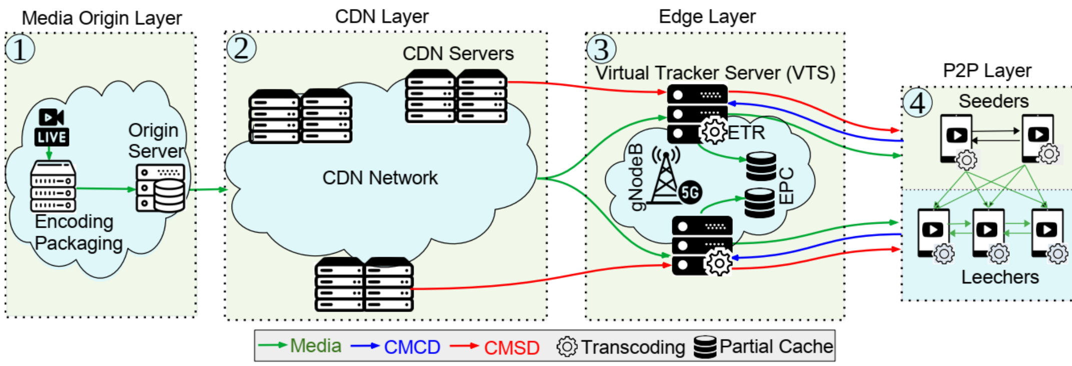

Figure 1: Proposed multi-layer architecture.

servers employ Common Media Server Data (CMSD) [23], [24] messages to periodically inform the edge layer about their cache occupancy.

3) Edge Layer: This layer includes virtualized edge components called Virtual Tracker Servers (VTSs), which are placed close to base stations (e.g., gNodeB in 5G). VTS servers are equipped with partial cache and video transcoding functions to serve client requests directly from cached qualities or to construct requested qualities from existing higher qualities, respectively. In the proposed system, clients' requests are directed to a VTS server; the VTS then considers information received from the clients as CMCD messages and servers' CMSD messages, besides other monitored information, such as available bandwidth, peers' computational and power resources, and peers' joining/leaving times. It finally employs an action tree (Fig. 2) to decide from where (i.e., adjacent peers, VTS, CDN servers, or origin server) and using which approach (i.e., fetch or transcode) to respond to the requested quality level. The designed action tree proposes the following potential actions that can be employed during the decision-making process (action numbering as in the figure):
   (1) Use the P2P network and transmit the requested quality directly from an adjacent peer with maximum stability (i.e., the least recent joining time).
   (2) Transcode the requested quality from a higher quality at the most stable adjacent peer and transmit it through the P2P network.
   (3) Fetch the requested quality directly from the VTS server.
   (4) Transcode the requested quality from a higher quality at the VTS.
   (5) Fetch a higher quality from a CDN server with maximum available bandwidth and transcode it at the VTS.
   (6) Fetch the requested quality directly from a CDN server with maximum available bandwidth.
   (7) Fetch the requested quality from the origin server.

4) P2P Layer: This layer utilizes all feasible peers' idle resources, i.e., bandwidth, storage, and computation, to offer services like "distributed video transcoding". The P2P network is based on the tree-mesh structure and includes two types of peers: Seeders and Leechers. In this scheme, seeders' requests can be served by all nodes (i.e., CDNs, origin, edge, or other seeders) except leechers, while all nodes can serve leechers' requests. In the proposed system, peers periodically communicate their cache occupancies to the edge layer using Common Media Client Data (CMCD) [23] messages and, in turn, receive updates from the edge layer through CMSD messages.

3. Evaluation Setup and Results

A. *Evaluation Setup*

To evaluate the performance of the proposed system in a realistic large-scale environment, we select InternetMCI [25] as a real backbone network topology and instantiate a cloud-based testbed on the CloudLab [26]

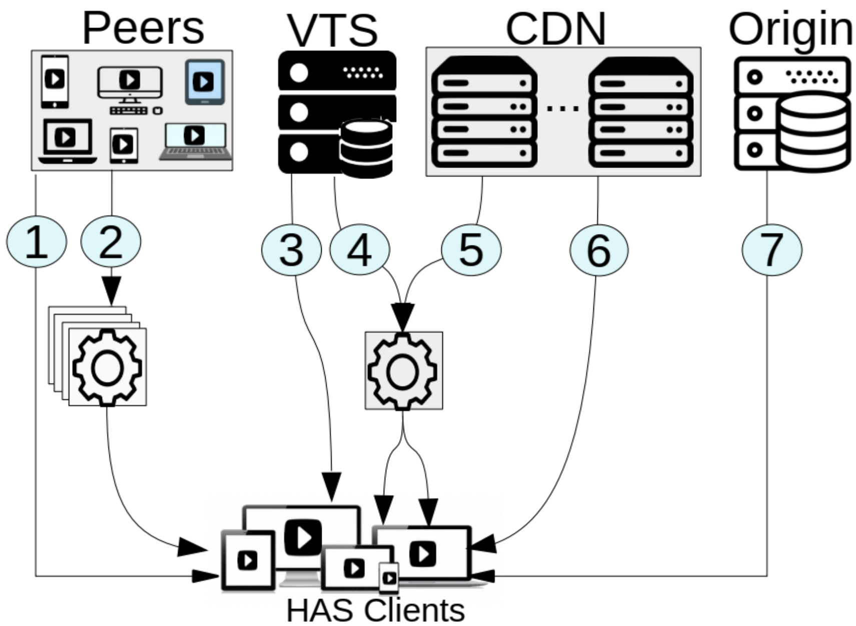

Figure 2: Proposed action tree.

environment. This testbed contains 375 elements, including 350 AStream [27] DASH players (seven groups of 50 peers), running SQUAD [28] and BOLA [29] as hybrid and buffer-based ABR algorithms, respectively. Five Apache HTTP servers (i.e., four CDN servers and an origin server), 19 OpenFlow (OF) backbone switches, 45 backbone layer-2 links, and a VTS server are used in our testbed. FFmpeg and FFmpegKit4 are used to measure the segment transcoding time on the VTS and peers, respectively. Moreover, iPhone 11 (Apple A13 Bionic, iOS 15.3), a Xiaomi Mi11 (Snapdragon 888, Android 11), and a PC (Apple M1, MacOS 12.0.1) are used to measure the transcoding time on the heterogeneous P2P network. Power consumption is measured by device tools, e.g., Android Energy Profiler and Android Battery Manager. The CodeCarbon [30] project measures power consumption for PC-type clients. Furthermore, we use a standard QoE model [31] to evaluate the system's performance regarding objective QoE. Table I shows the details of other configuration parameters.

B. Evaluation Results

Executing computationally-intensive tasks, such as video transcoding, on peers must be fast enough to avoid imposing significant delays on the system and should not excessively drain the peers' batteries. Otherwise, clients' requests may be served by other actions, leading to network and edge server congestion. Thus, the first experiment aims to determine the power consumption (i.e., power values are given as kWh$\times 10^{-3}$ and percentage of battery usage) and the percentage of battery usage when various operations are concurrently executed on a single peer, such as playing video I and transcoding video II concurrently for a five-minute video (equivalent to 150 segments). The results, as presented in Table II, reveal that the transcoding (TR) operation consistently consumes the most power while playing back a video (PLY) incurs significantly lower battery usage. Furthermore, combining the PLY and TR tasks does not impose a substantial burden on the peers' batteries when compared to the energy consumed by individually upscaling a video.

In the next experiment, we assess the transcoding times on peers and explore the impact of peer transcoding on users' perceptual quality. We employ the Video Multi-method Assessment Fusion (VMAF) [33] scores to evaluate the perceptual quality for a three-minute video. As depicted in Table III, the transcoding process for the entire video takes 1.28 to 20.44 seconds (0.014 to 0.22 seconds per segment) on PC peers and 5.85 to 74.91 seconds (0.065 to 0.8 seconds per segment) on mobile peers.

In the final experiment, we run testbed experiments to analyze the performance of the proposed framework compared to the following baseline methods:

1) Non Hybrid (NOH): The NOH system is a regular CDN-based streaming approach that does not incorporate any P2P support.
2) Simple Edge-enabled Hybrid (SEH): The SEH system utilizes a basic VTS server without caching and transcoding capabilities. In this system, peers can only be served through one of the actions 1, 6, or 7 (Fig. 2).
3) Non Transcoding-enabled Hybrid (NTH): As common in similar works, the NTH-based system has caching capability but lacks transcoding capability. In this approach, peers are only served through one of the actions 1, 3, 6, or 7 (Fig. 2).
4) Edge Caching/Transcoding Hybrid (ECT): An ECT-enabled system does not involve transcoding at the peer side. In this approach, requests can be served through all actions except action 2.

We first calculate the average serving latency (in seconds) for all clients, encompassing both transmission latency and computational latency, as an integral component of the end-to-end (E2E) delay. As illustrated in Fig. 3(a), the proposed system exhibits the best performance for both ABR algorithms concerning this metric. By efficiently downloading requested segments through the most suitable actions with shortened serving time (combining computation and transmission), and from nodes with available resources in terms of bandwidth and computation, the serving time is significantly reduced.

While improving latency is crucial, it is equally important to have a comprehensive standard model to analyze the system's behavior concerning the enhancement of users' QoE. For this purpose, we employ the standard objective QoE model P.1203 [31]. As depicted in Fig. 3(b), the proposed system exhibits superior performance for both ABR algorithms in this aspect as well. We further analyze the system's performance in terms of edge energy consumption (measured in kWh) required for running edge transcoding. As evident from Fig 3(c), the proposed system outperforms the ECT system (the only transcoding-enabled baseline system) and shows reduced energy consumption. This is primarily because of serving requested representations through *(i)* distributed caching, i.e., the representations cached in all layers of the network (CDN, P2P, and edge), and *(ii)* distributed transcoding, the representations transcoded by transcoding-enabled peers, which allows our system to transcode fewer segments at the edge.

Table I: Experimental setup.

Cache Replacement Policy	Least Recently Used (LRU)
Maximum Size of CDN Caches	40% of the video dataset
Maximum Size of VTS Caches	5% of the video dataset
Maximum Size of Peer Caches	5 segments
Number of Live Channels	5 channels
Segment Duration	2 s
Bitrate Ladder [32]	{(0.089,320p), (0.262,480p), (0.791,720p), (2.4,1080p), (4.2,1080p)}[Mbps, content resolution]
Links' Bandwidth (CDNs to VTS)	100 Mbps
Links' Bandwidth (Origin to VTS)	50 Mbps
Channel Access Probability Model	Zipf ($\alpha = 0.7$)
Monitoring Interval	1 s

Table II: Power consumption and battery usage of different operations for a 5-min. video on peers.

Operation	TR: 791k→262k		TR: 4219k→2484k	
	Power (kWh$\times 10^{-3}$)	Battery	Power (kWh$\times 10^{-3}$)	Battery
PLY	119	0.39%	152	0.49%
TR	130	0.42%	352	1.14%
TR + PLY	237	0.77%	479	1.55%

Table III: Transcoding times and VMAF scores for a 3-min. video on different peers.

Input BR (bps)	Target BR (bps)	Client PC		Client Mobile	
		Time (s)	VMAF	Time (s)	VMAF
4219k	89k	4.31	15.38	15.98	13.75
4219k	262k	5.33	44.61	18.32	42.13
4219k	791k	11.74	76.21	39.28	73.14
4219k	2484k	20.44	93.33	74.91	91.53
2484k	89k	3.80	14.35	16.55	13.01
2484k	262k	4.83	42.27	18.82	40.02
2484k	791k	11.36	71.56	39.76	69.06
791k	89k	2.05	12.21	10.43	11.24
791k	262k	3.35	36.33	14.81	34.76
262k	89k	1.28	11.01	5.85	10.32

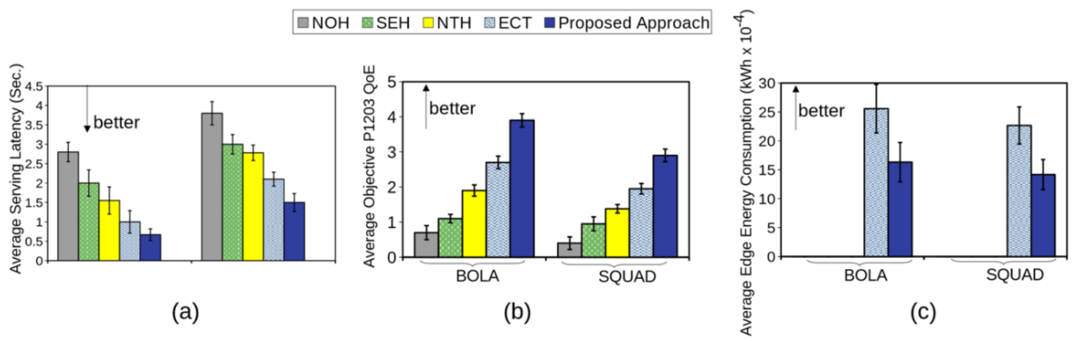

Figure 3: Performance of the proposed framework compared with baseline methods in terms of (a) average serving latency, (b) average objective P1203 QoE, (c) average edge server energy consumption for 350 clients running the BOLA and SQUAD ABR algorithms.

4. Conclusion

This paper presented a novel hybrid P2P-CDN live video delivery system, incorporating modern networking paradigms such as NFV and edge computing, along with distributed video transcoding. The system introduced a multi-layer architecture and an action tree, encompassing all possible actions to serve HAS clients from different nodes (i.e., peers, CDNs, edge, origin server), ensuring satisfactory users' QoE and acceptable latency. To evaluate the system's performance, a cloud-based testbed with 350 clients was established, and multiple experiments were conducted. The experimental findings confirmed the superiority of the proposed approach in terms of users' QoE, serving latency, and edge server's energy consumption compared to competitive schemes.

References

[1] Sandvine, "The Global Internet Phenomena Report," White Paper, January 2023. [Online]. Available: https://www.sandvine.com/phenomena

[2] V. Cisco, "Cisco Visual Networking Index: Forecast and Trends, 2018–2023," White Paper, 2018. [Online].Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf

[3] ISO/IEC. 2019, "Information Technology — Dynamic Adaptive Streaming over HTTP (DASH) — Part 1: Media Presentation Description and Segment Formats," International Organization for Standardization, International Standard 23009-1:2019, Dec. 2019.

[4] R. Pantos and W. May, "HTTP Live Streaming," RFC 8216, 2017. [Online]. Available: https://rfc-editor.org/rfc/rfc8216.txt

[5] A. Bentaleb, B. Taani, A. C. Begen, C. Timmerer, and R. Zimmermann, "A Survey on Bitrate Adaptation Schemes for Streaming Media over HTTP," IEEE Communications Surveys & Tutorials, 2018.

[6] A. A. Barakabitze, N. Barman, A. Ahmad, S. Zadtootaghaj, L. Sun, M. G. Martini, and L. Atzori, "QoE Management of Multimedia Streaming Services in Future Networks: A Tutorial and Survey," IEEE Communications Surveys & Tutorials, 2019.

[7] N.-N. Dao, A.-T. Tran, N. H. Tu, T. T. Thanh, V. N. Q. Bao, and S. Cho, "A Contemporary Survey on Live Video Streaming from a Computation-Driven Perspective," ACM Computing Surveys, 2022.

[8] R. Farahani, "CDN and SDN Support and Player Interaction for HTTP Adaptive Video Streaming," in 12th ACM Multimedia Systems Conf., 2021.

[9] N. Anjum, D. Karamshuk, M. Shikh-Bahaei, and N. Sastry, "Survey on Peer-Assisted Content Delivery Networks," Computer Networks, 2017.

[10] R. Alimi, R. Penno, Y. Yang, S. Kiesel, S. Previdi, W. Roome, S. Shalunov, and R. Woundy, "Application-Layer Traffic Optimization (ALTO) Protocol," RFC 7285, 2014.

[11] S. Ellouze, B. Mathieu, and T. Lemlouma, "A Bidirectional Network Collaboration Interface for CDNs and Clouds Services Traffic Optimization," in IEEE Int'l. Conf. on Communications (ICC), 2013.

[12] M. Zhao, P. Aditya, A. Chen, Y. Lin, A. Haeberlen, P. Druschel, B. Maggs, B. Wishon, and M. Ponec, "Peer-Assisted Content Distribution in Akamai NetSession," in Internet Measurement Conf., 2013.

[13] Z. Ma, S. Roubia, F. Giroire, and G. Urvoy-Keller, "When Locality is Not Enough: Boosting Peer Selection of Hybrid CDN-P2P Live Streaming Systems using Machine Learning," in Network Traffic Measurement and Analysis Conf. (IFIP TMA), 2021.

[14] T. T. T. Ha, J. Kim, and J. Nam, "Design and Deployment of Low-Delay Hybrid CDN–P2P Architecture for Live Video Streaming Over the Web," Wireless Personal Communications, 2017.

[15] H. Yousef, J. Le Feuvre, P.-L. Ageneau, and A. Storelli, "Enabling Adaptive Bitrate Algorithms in Hybrid CDN/P2P Networks," in 11th ACM Multimedia Systems Conf., 2020.

[16] R. Farahani, F. Tashtarian, A. Erfanian, C. Timmerer, M. Ghanbari, and H. Hellwagner, "ES-HAS: An Edge- and SDN-Assisted Framework for HTTP Adaptive Video Streaming," in 31st ACM NOSSDAV Workshop, 2021.

[17] R. Farahani, F. Tashtarian, H. Amirpour, C. Timmerer, M. Ghanbari, and H. Hellwagner, "CSDN: CDN-Aware QoE Optimization in SDN-Assisted HTTP Adaptive Video Streaming," in 46th IEEE Conf. on Local Computer Networks (LCN), 2021.

[18] R. Farahani, F. Tashtarian, C. Timmerer, M. Ghanbari, and H. Hellwagner, "LEADER: A Collaborative Edge- and SDN-Assisted Framework for HTTP Adaptive Video Streaming," in IEEE Int'l. Conf. on Communications (ICC), 2022.

[19] R. Farahani, M. Shojafar, C. Timmerer, F. Tashtarian, M. Ghanbari, and H. Hellwagner, "ARARAT: A Collaborative Edge-Assisted Framework for HTTP Adaptive Video Streaming," IEEE Trans. on Network and Service Management, 2022.

[20] R. Farahani, A. Bentaleb, C. Timmerer, M. Shojafar, R. Prodan, and H. Hellwagner, "SARENA: SFC-Enabled Architecture for Adaptive Video Streaming Applications," in IEEE Int'l. Conf. on Communications (ICC), 2023.

[21] R. Farahani, H. Amirpour, F. Tashtarian, A. Bentaleb, C. Timmerer, H. Hellwagner, and R. Zimmermann, "RICHTER: Hybrid P2P-CDN Architecture for Low Latency Live Video Streaming," ACM Mile-High Video (MHV), 2022.

[22] R. Farahani, A. Bentaleb, E. C¸etinkaya, C. Timmerer, R. Zimmermann, and H. Hellwagner, "Hybrid P2P-CDN Architecture for Live Video Streaming: An Online Learning Approach," in IEEE Global Communications Conf. (GLOBECOM), 2022.

[23] CTA-5004, "Web Application Video Ecosystem—-Common Media Client Data," 2020. [Online]. Available: https://cdn.cta.tech/cta/media/media/resources/standards/pdfs/cta-5004-final.pdf

[24] CTA-5006, "Common Media Server Data," 2022. [Online]. Available: https://github.com/cta-wave/common-media-server-data

[25] http://www.topology-zoo.org/dataset.html, accessed: 2023-07-26.

[26] R. Ricci, E. Eide, and C. Team, "Introducing CloudLab: Scientific Infrastructure for Advancing Cloud Architectures and Applications," login: The Magazine of USENIX & SAGE, 2014.

[27] "AStream: A Rate Adaptation Model for DASH," accessed: 2023-07-20. [Online]. Available: https://github.com/pari685/AStream

[28] C. Wang, A. Rizk, and M. Zink, "SQUAD: A spectrum-based quality adaptation for dynamic adaptive streaming over HTTP," in 7th ACM Conf. on Multimedia Systems, 2016.

[29] K. Spiteri, R. Urgaonkar, and R. K. Sitaraman, "BOLA: Near-optimal Bitrate Adaptation for Online Videos," in 35th IEEE Int'l. Conf. on Computer Communications (INFOCOM), 2016.

[30] "CodeCarbon: Estimate and Track Carbon Emissions from Machine Learning Computing," accessed: 2023-07-26. [Online]. Available: https://github.com/mlco2/codecarbon

[31] "ITU-T. Rec. P.1203. Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport - video quality estimation module." [Online]. Available: http://handle.itu.int/11.1002/ps/P1203-01

[32] S. Lederer, C. Muller, and C. Timmerer, "Dynamic Adaptive Streaming over HTTP Dataset," in¨ 3rd ACM Multimedia Systems Conf., 2012.

[33] Z. Li, A. Aaron, I. Katsavounidis, A. Moorthy, and M. Manohara, "Toward a Practical Perceptual Video Quality Metric," [Online] https://netflixtechblog.com/toward-a-practical-perceptual-video-quality-metric-653f208b9652, 2016, accessed: 2022-12-25.

Authors Bio-data

Reza Farahani is a last-year Ph.D. candidate at the Institute of Information Technology (ITEC), Alpen Adria-Universität Klagenfurt (AAU). Currently, he is working on the Horizon Europe GraphMassivizer project. During the period from October 2019 to February 2023, he was involved in the ATHENA project, funded by the Christian Doppler Forschungsgesellschaft and the industry partner Bitmovin GmbH. From November 2022 to January 2023, he was a visiting scholar at the 5G & 6G Innovation Centre (5GIC & 6GIC), Institute for Communication Systems (ICS), at the University of Surrey, UK. He received his B.Sc. and M.Sc. degrees in computer engineering, in 2014 and 2019, from the University of Isfahan, Isfahan, IRAN, and the University of Tehran, Tehran, IRAN, respectively. He also has been working in the computer networks field in different roles, e.g., Cisco network engineer, network protocol designer, network programmer, and Cisco instructor (R&S, SP) for over six years. His research interests are Multimedia Communication, Video Networking, Serverless Computing, Cloud-Edge Continuum, Network Softwarization/Virtualization, Distributed Systems, Mathematics Optimization, and Distributed Learning approaches. Further information at https://www.rezafarahani.me/.

Christian Timmerer (M'08-SM'16) is a full professor of computer science at Alpen-Adria-Universität Klagenfurt (AAU), Institute of Information Technology (ITEC), and he is the director of the Christian Doppler (CD) Laboratory ATHENA (https://athena.itec.aau.at/). His research interests include multimedia systems, immersive multimedia communication, streaming, adaptation, and quality of experience where he co-authored eight patents and more than 300 articles. He was the general chair of WIAMIS 2008, QoMEX 2013, MMSys 2016, and PV 2018 and has participated in several EC-funded projects, notably DANAE, ENTHRONE, P2P-Next, ALICANTE, SocialSensor, COST IC1003 QUALINET, ICoSOLE, and SPIRIT. He also participated in ISO/MPEG work for several years, notably in MPEG-21, MPEG-M, MPEG-V, and MPEG-DASH, where he served as standard editor. In 2012 he cofounded Bitmovin (http://www.bitmovin.com/) to provide professional services around MPEG-DASH, where he holds the position of the Chief Innovation Officer (CIO) – Head of Research and Standardization. Further information at http://timmerer.com.

Hermann Hellwagner is a full professor of computer science at Alpen-Adria-Universität Klagenfurt (AAU) where he leads the research group Multimedia Communication (MMC) in the Institute of Information Technology (ITEC). Earlier, he held positions as an associate professor at the University of Technology in Munich (TUM) and as a senior researcher at Siemens Corporate Research in Munich. His current research areas are distributed multimedia systems, multimedia communication and adaptation, QoS/QoE, and communication in multi-UAV networks. He has published widely on parallel computer architecture, parallel programming, and multimedia communication and adaptation. He is a senior member of the IEEE and of the ACM. He was a member of the Scientific Board of the Austrian Science Fund (FWF) (2005-2016) and an FWF Vice President (2013–2016). Currently, he is a member of the CD Senate of Christian Doppler Forschungsgesellschaft (CDG). Further information at https://www.itec.aau.at/∼hellwagn/.