Energy Aware Video Streaming from Drones with End-to-End Delivery using Packet Trimming

Emre Karakış*, Stuart Clayman†, Mustafa Tüker*, Elif Bozkaya‡, Müge Sayıt*
*International Computer Institute, Ege University, Izmir, Turkey
†Dept. of Electronic Engineering, University College London, London, UK
‡Dept. of Computer Engineering, National Defence University Naval Academy, Istanbul, Turkey

***Abstract*—The end-to-end delivery of video from drones at one edge, to clients at another edge raises a number of architectural and implementation issues. A specific use-case of collecting video from a football field is presented, with a particular focus on the energy used by the drones, when doing positioning and video collection. It is impossible to send continuous video from all of the drones without having an energy management and a recharging process. Therefore an analysis and formulation of the consumed energy for a flight strategy for successful video collection is developed. Results of doing edge processing of the video streaming application, using a packet trimming process and using the Big Packet Protocol (BPP) protocol, demonstrate that the techniques presented provide a low latency, high QoE stream at the client.**

***Index Terms*—Drone, NFV, Energy Models, In-Network Packet Processors, Packet Trimming, Traffic Engineering**

I. Introduction

The emergence of a number of network communication facilities, such as Network Function Virtualization (NFV) and Software Defined Networking (SDN), has allowed new softwarized network scenarios to be developed. The use of cloud computing and the allocation of virtualized functions to various cloud locations has created a flexible and adaptable infrastructure for deployment. There is a much interest in Edge processing, where the utilization of Internet of Things (IoT), and Unmanned Aerial Vehicles (UAVs) / Drones have been a specific application area. Other recent developments have been introduced, including the concepts of in-network packet processing which allows new approaches to be developed for future Internet applications, as well as the concept of packet trimming, which adapts the size of the packets during their transmission over the network, in protocols such as BPP [1].

UAV / Drone networks provide a new approach to video collection and video surveillance applications due to their dynamic nature and controllability. The advantages of low cost, highly mobile, and flexible deployment, mean the drones can capture images and video by monitoring a surveillance area; establish the wireless communication link with a ground control station; and then transmit that data to a ground control station, which is significant in such dynamic scenarios. Traditional stationary video systems can be located to observe a critical region or to monitor activities, including observing the public at the entrance of sports events, on public transportation, and in crowded public areas. However, for occasional events (e.g. football once a week or festivals organised once a year), or for unexpected events (e.g. natural disasters) the use of stationary video systems may not be feasible. Thus, the utilization of drones for video streaming is promising, as they are easily reconfigurable and provide wireless communication link with a ground control station with high flexibility [2], [3].

The combination of all of these elements is a specific motivation in this paper, with a football event as a use-case. We consider an end-to-end deployment of drones with cameras, via a core network, with delivery to a number of clients with various display devices. The drone network at the stadium is connected to one edge network, and then onto a core network, which is connected to another edge network which contains the clients. There is a virtualized network function processing the data from the drones residing at the drone edge. This function collects video data from all of the drones, and then selects which specific video to send to the clients. At the client edge is another virtualized function which collects the video sent in Big Packet Protocol (BPP) packets from the selected drone, and based on the client attributes may undertake some packet trimming. BPP was presented as one of the first protocols for doing packet trimming as a native operation [1]. Using BPP provides a mechanism to trim packets during their journey across the network, and it can be used effectively for in-network computing [4]. Providing end-to-end video streaming services and service differentiation can be done in a more efficient way, by providing low latency and low loss, when using the newly developed technologies of BPP and NFV [5].

In this paper, we present this scenario of end-to-end video transmission of football video between the drones and the clients. We determine the locations of multiple drones to monitor the viewing area, and transmit the video captured by the drone cameras to the virtualized edge server for processing. There is a particular focus on the energy used by the drones, when doing positioning and video collection, as they have limited battery capacity due to battery weight. As such, it is impossible to send continuous video from all of the drones without having an energy management and a recharging process. The video from the drone is Scalable Video Codec (SVC), whose compatibility with BPP has been shown in our previous work [5]. With SVC, an encoded video file consists of several layers, each of them corresponds to a quality alternative.

In this paper, we investigate the optimal deployment of multiple drones to cover the surveillance area of a football field, capture the video from that area, and provide a low latency, high QoE video streaming application using the packet trimming and BPP protocol. The main contributions of this paper are as follows: (i) an analysis and formulation of the consumed energy to determine a flight strategy for successful video collection; and (ii) an approach to providing service differentiation for end-to-end streaming when utilizing the combination of the emerging technologies and protocols, BPP, virtualized edge functions, and SVC. The rest of this paper is organized as follows: in section II, we present the background related to video capturing and streaming via drones and in-network computing with BPP. The details about the system proposed in this work are given given in section III. An experimental performance evaluation of the system, with the QoE results at the clients, is presented in section IV; which is followed by the conclusions in section V.

II. Background

Over the last years, the application scenarios of UAVs have drawn considerable attention from both industrial and military fields, as these aerial platforms can be used in various fields, such as environmental monitoring, disaster assistance, aerial photography, etc. [2]. Recently, the streaming of videos captured by drones has became one of the popular topics because there are a wide range of future network scenarios utilizing video streaming with drones. In [6], the authors proposed a system where a cellular or WiFi connection is selected according to the throughput value for the users stream video by using drones. In [7], a channel allocation model that considers energy consumption and QoE requirements for streaming the videos captured by the drones to a base station. That study focuses on the bandwidth resource allocation problem, rather than end-to-end video streaming to the clients. In [8], the clients signal their bandwidth and delay values to the drone which transfers video to the clients, and the drones use this information to adapt quality. In [9], drones capture different areas in a sport field, where they collect the video on the basis of the signals sent by the motion sensors on sport players. Although these recent studies are related to video streaming from drones, they focus on drone-to-base station or drone-to-client communication, rather then an end-to-end scenario.

For beneficial transmission of video from drones we need a good protocol. RTP is an application layer framing protocol, and provides no improvements over UDP as it is just a payload of UDP packets, so we get packet loss. TCP is reliable, but has high latency. The QUIC protocol, which is basically HTTP over UDP, was introduced to improve the responsiveness of web services. It aims to overcome some issues when using TCP [10]. Many people believe that QUIC will inherently be good for transmitting video, however in [11], they find there is no evidence for any QoE improvement, and in [12], QUIC is still too reliable for real-time video as needed for drones.

BPP was designed as a protocol to be used for low latency/high reliability applications. First introduced in 2018 [1], BPP provides packet trimming capabilities in its design, using a mechanism called Packet Wash, in which chunks in BPP packets are trimmed. This has been shown to reduce latency [13]. Although BPP was devised for sending video, AR, and VR, the first real implementation of using BPP for determining the effects of video transmission with BPP and for streaming video is in [14]. A detailed view of the mechanisms and techniques used for transmitting SVC video across the network, using the packet trimming capabilities of BPP and showing in-network packet adaption, is presented in [5].

The concept of Packet Trimming has a focus as newer hardware has become fast enough to update packets as the cross the network. In [15], packet trimming in the Data Center is presented. In [16], the authors note that trimming the whole payload and just keeping the header can only work well in Data Centers as the trimmed payload can be retransmitted fast enough to the host that trimmed the payload. Therefore, the approach used in data centers does not work across WANs. They suggest that selectively doing packet trimming, rather than the whole payload, can be a less dramatic approach that could work in WANs. To enhance end-to-end transport, a new transport protocol called QUCO [17] was defined. It reacts to congestion by selectively trimming parts of a packet. Their packet trimming scheme reduces the variation in the number of packets going through the network,

In [18], we compare the use of either an ONOS SDN controller or a virtualized edge server as a node for processing BPP packets and doing packet trimming. In that system the server always sends the video at the highest quality, and we show that in-network video quality adaption using a virtualized function provides better performance. We utilize parts of that architecture in this paper, as we also have in-network adaption by trimming the packets of the SVC video at the edge.

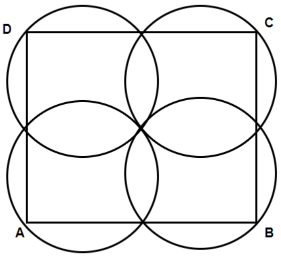
In this work, we focus on end-to-end streaming of the videos from drones, where videos are selected by a Mobile Edge Computing (MEC) server at the drone edge network, and the quality is adapted in the network by another Virtual Network Function (VNF) at the client edge network. This occurs while an energy consumption process manages the position of the drones and sends them to the charging station.

III. Energy Aware End-to-End Video Streaming

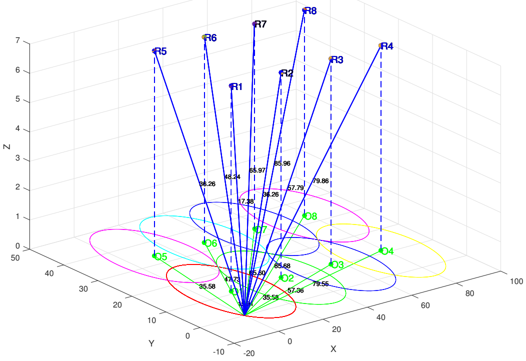
We propose an end-to-end streaming system in which videos are captured by drones and delivered to different clients, by using a protocol supporting packet trimming and utilising in-network computing. New aspects of the system are presented, showing the drone network, the drone placement strategy, the energy management process, and the packet trimming function.

A. Drone Positioning for Capturing Video of a Football Field

In our study, we present a scenario where the videos of a football field are captured by a set of drones and transmitted to a set of users having different characteristics. For such a scenario, how many drones should be used and the positions of each drones should be determined so that drone cameras can capture video while flying over the football field. According to FIFA rules, a football field is a rectangle with a minimum size of $45x90m^2$. In this case, we can view the rectangular area as two distinct squares with sides of 45 meters. The goal is to use as minimum number of drones as possible, while



(a) Circle packing in a square region.



(b) 3D Coordinate Calculations

$$R_1(X_1, Y_1, h) = (11.25, 11.25,\ 7)$$
$$R_2(X_2, Y_2, h) = (33.75, 11.25,\ 7)$$
$$R_3(X_3, Y_3, h) = (56.25, 11.25,\ 7)$$
$$R_4(X_4, Y_4, h) = (78.75, 11.25,\ 7)$$
$$R_5(X_5, Y_5, h) = (11.25, 33.75,\ 7)$$
$$R_6(X_6, Y_6, h) = (33.75, 33.75,\ 7)$$
$$R_7(X_7, Y_7, h) = (56.25, 33.75,\ 7)$$
$$R_8(X_8, Y_8, h) = (78.75, 33.75,\ 7)$$

(c) 3D Positions

Fig. 1: Positioning & Coverage of Drones on Football Field

managing the energy consumption of the drones. In order to solve this, we use the circle packing problem which serves as the foundation for the challenge of calculating the number of drones and their location in coordinates. The challenge of circle packing is finding a way to cover all of the surfaces with the least number of circles. For a rectangle R consisting of the nodes A, B, C, and D, as seen in Fig. 1a, the coverage radius with four drones is shown. The full definition of the equations for these calculations can be found in [19].

B. *Energy Management Process*

In order to provide seamless and continuous video streaming, the drones need to have a recharging scheme and an energy management process. It is impossible to send continuous video from all of the drones without having such energy management. As the drones do not have enough energy to continuously work, they should go to an Energy Supply Station to be recharged, where they are exchanged by replacement drones with a high charge. The deployed drones will go to a specified position and start sending video. For this purpose, the Drone Control Station (Access Point) keeps track of drone position and drone energy levels, using the process shown in Fig. 2a. We describe the two main facets of the *energy consumption model* of the Energy Management Process: (i) the energy consumed when the drone is moving to the designated position, called the *transition energy*, and (ii) the energy consumed when the drone is at the hover position to capture the video, called the *hover energy*.

1) Consumption Model for Drones: Initially, 8 drones are located at the origin point $(0, 0, 0)$, as a start position. In order to calculate the *transition energy*, it is vital to determine the transition time based on the the target distance and drone velocity. The drone placements and discovery of coordinates are based on the mathematical expressions given in [19], with the drone velocities fixed as a constant value of $5m/s$. We can then determine the 3D coordinates of each drone, and the distance from the origin, as shown in Fig. 1b. This calculation is given in Fig. 1c as target point for each drone. According to the distance between the source (Energy Supply Station)

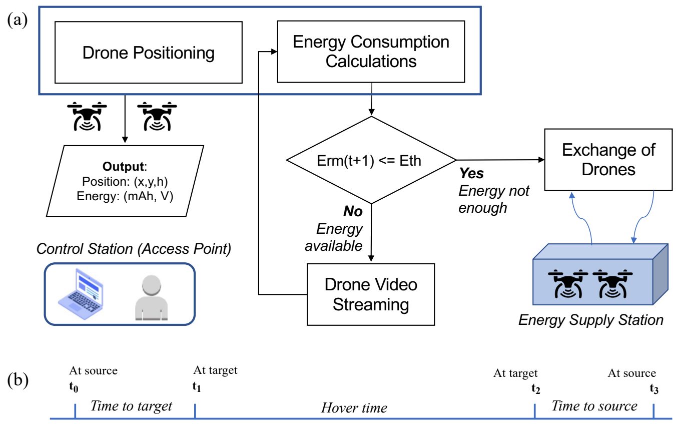


Fig. 2: Drone Energy Life Cycle and Movement Timeline.

and target points and velocity, $t_1 - t_0$ and $t_3 - t_2$ represent the required time to transfer each drone to the target and the source point, respectively, as illustrated in Fig. 2b.

2) Transition Energy Consumption: We define the transition energy to transfer the drones to the designated positions in Eq. 2 [20], based on the mass. There are multiple parameters affecting the drone mass, such as propellers, motors, sensors, GPS, camera weight. Here the weight of the total mass of the deployed drone is assumed to be $2\,kg$. The battery level of the drone $(1Wh = 3600J)$ is expressed in Eq. 1.

$$Battery_{Energy} = Capacity(mAh) * Voltage * 3.6 \quad (1)$$

Energy is calculated for the target and source point transition.

$$E_{transition} = \left(\frac{P_{full} - P_S}{V_{max}} V_d + P_S\right)(t_1 - t_0) \quad (2)$$

where: P_S is the power level in hover position, P_{full} is the hardware power level at full speed, V_d is the velocity during the transition, and V_{max} is the maximum velocity.

3) Hover Energy Consumption: While the drones stay in the hover position to capture the video, the consumed energy is calculated from Eq. 3, as presented in [20], [21].

$$E_{hover} = (T^{3/2}/\sqrt{2n\rho S})(t_2 - t_1) = \big((gm_k)^{3/2}/\sqrt{2n\rho S}\big)(t_2 - t_1) \quad (3)$$

where g represents the gravitational acceleration (m/s^2), m_k represents the drone component mass (kg), n is number of rotors, T is the thrust value, ρ is air density (kg/m^3), which is based on environmental factors such as temperature, humidity and pressure. Under the normal conditions, the air density at sea level is considered as $1.225 kg/m^3$. S is the area of the rotating disk of a rotor (m^2).

From these values, the remaining energy $E_{remaining}$ can be determined in order to decide the recharging status of drones:

$$E_{remaining} = E_{max} - (E_{transition} + E_{hover}) \quad (4)$$

As stated, the drones stay in fixed hover positions for streaming the video, after the positions have been determined. As seen in Fig. 2a, the *Energy Consumption Calculations* are followed according to two main facets of energy consumption model. Here, the transition time from designated position to the origin position, $t_3 - t_2$ depends on the remaining energy. Therefore, the remaining energy must be greater than the transition energy consumed in the interval $t_3 - t_2$ [2].

C. Packet Trimming At the Edge

The drones, which are in a designated position and collecting video, send the video to a MEC server in which the video having the Region of Interest is transmitted to a set of clients by using an object detection software such as YOLO [22]. In this study, there are different types of users requesting to play the video. These users are of different client types, having various devices such as mobile phone, laptop, or TV.

The video is transmitted from the MEC server to the clients using BPP. We assume that the video streaming company and ISP is in cooperation and the bandwidth values of the links that the clients connected to are send to the video streaming company. The video streaming company also has the information about the clients' device characteristics, which require service differentiation. Let the rendering capability and resolution of c^{th} client's device be represented as d_c, which is then transformed by the video streaming company into the bitrate unit, br_c by using a mapping function. The company determines the bitrate of the videos that will be sent to the clients by calculating $min(br_c, abw_c)$, where abw_c represents the available bandwidth of the c^{th} client's connection.

Videos are encoded by using SVC where the bitrate of each layer corresponds to the bitrate determined for the client type. We assume that the core network has enough capacity to send the video with the highest bitrate and the bottleneck of the end-to-end network is the last hop. Therefore, we adapt the quality by using in-network computing at the edge. According to this, the MEC server sends the video with the highest bitrate. At the client edge, a virtualized BPP network function processes the packets and implements packet trimming on the basis of the available bandwidth information of the link that connects the client to network.

IV. Performance Evaluations

In order to measure the performance, we used two different platforms for streaming the video, one for the drone side and one for video transmission and BPP implementation. Drone evaluation and simulation is performed using the Unreal Engine tool and the Microsoft AirSim simulator [23]. To transfer the video from the server to the clients, a Mininet emulator is used. The videos captured by the drones are pre-processed in the server and the video showing the Region of Interest is streamed with the highest quality by the server. The details of the video are as follows:

Video type	H.264	**Frame rate**	20 fps
Resolution	480 x 360	**L0 bitrate**	764.72 Kbps
Frame No	720	**L0+L1 bitrate**	1249.90 Kbps
Duration	36 sec	**L0+L1+L2 bitrate**	3297.51 Kbps

There are different types of clients in the system: mobile, laptop, and TV, and they require different levels of video quality as their device characteristics vary. The BPP process executes as an edge virtualized function, where packets are trimmed according to the client types. As a comparative approach, we also streamed the video using UDP. Although more sophisticated approaches can be used, namely client-side quality adaptation or server-side retransmissions based on client's feedback, here we focus on the comparison of the transport layer protocols without using additional mechanisms.

A. Energy Management and Exchange of Drones

A significant contribution of this work is to track and manage the consumed energy of the drones with a deployment and flight planning strategy, since it is not possible to send continuous video without energy replenishment. This scenario was presented in in Fig. 2. In the *Energy Consumption Calculations*, the Control Station computes the consumed energy for transition from initial point to the target point (given in Fig. 1), and hovering. We then assume that the drone should be recharged when its remaining energy, Eq. 4, reaches the limit for transition time, and hence each drone is directed to the Energy Supply Station for recharging.

In Table I, the energy calculations, described in Section III-B, are shown for 8 drones. As seen in the table, the positions of the drones are calculated and they remain in a fixed position since we aim to cover a football field and capture the video. The coverage radius of each drone on the football field is evaluated as $15.909m$ and each drone can take off up to $7m$ high. Based on the distance to the target point, the Transition Time and Transition Energy are showed in the table. After the drone deployment, the remaining energy and required transition time $(t_3 - t_2)$ are followed to direct the drone to the Energy Supply Station and provide energy replenishment.

B. Evaluations based on QoE Parameters

In the experiments, there are 2 clients of each client type: mobile, laptop, and TV. Due to the Ethernet MTU size, the biggest BPP packet is 1500 bytes. As the video is 3297.51 Kbps, we selected a set of bandwidths which have

Drone	Target Point	X	V_d	Vmax (m/s)	Start Power (P_S)	Max. Power (P_{full})	Start Time (t_0)	Transition Time ($t_1 - t_0$)	Transition Energy
R1	(11.25, 11.25, 7)	17.38 m	5	22.64	529.74	1974.3 W	0	3.48 sec	2953 J
R2	(33.75, 11.25, 7)	36.26 m	5	22.64	529.74	1974.3 W	0	7.25 sec	6153.6 J
R3	(56.25, 11.25, 7)	57.79 m	5	22.64	529.74	1974.3 W	0	11.56 sec	9811 J
R4	(78.75, 11.25, 7)	79.86 m	5	22.64	529.74	1974.3 W	0	16.00 sec	13580.3 J
R5	(11.25, 33.75, 7)	36.26 m	5	22.64	529.74	1974.3 W	0	7.25 sec	6153.6 J
R6	(33.75, 33.75, 7)	48.24 m	5	22.64	529.74	1974.3 W	0	9.65 sec	8190 J
R7	(56.25, 33.75, 7)	65.97 m	5	22.64	529.74	1974.3 W	0	13.19 sec	11195.2 J
R8	(78.75, 33.75, 7)	85.96 m	5	22.64	529.74	1974.3 W	0	17.19 sec	14590.4 J

TABLE I: Power and Energy Consumption of Drones.

proportionally limited values, for the different types of clients. The available bandwidth of the link that connects clients to the network is 0.9 Mbps for the mobile, 1.5 Mbps for the laptop, and 3.5 Mbps for the TV. However, in real life, the BPP packet size, the bitrates of the video layers, and the bandwidth values of the clients connection can be so much higher.

An overall view of the received layers and the lost frame distribution on the client side is given in Fig. 3. This graph shows the percentage of layers that the clients receive, based on their client type and protocol used. Laptop and TV users with BPP never experience frame losses. We observed high level of frame loss in mobile and laptop clients with UDP.

In the experiments, we measure different types of QoE parameters: average bitrate, duration of pauses, and the number of quality switches. We also calculate an overall QoE value indicating the perceptual quality. In Fig. 4a, the average bitrate values received by the clients are given. For both approaches, the average received bitrates increase as the requirements of the clients increase. This is mostly because the available bandwidth for the clients is set according to the device type. Other QoE metrics observed in the clients give better insight about how successful our proposal is in providing service differentiation by considering the device type of the clients.

Duration of pauses is one of the most important QoE parameters that affects negatively perceived quality. These values are presented in Fig. 4b, where the E2E-BPP approach outperforms E2E-UDP transmission. The duration of pauses values are lower for BPP than the UDP approach for each type of client. The number of quality switches is another parameter that affects the perceived quality. The graph given in Fig. 4c shows that E2E-BPP manages to keep the number of quality changes under a certain level for all type of clients, and this number decreases as the device resolution of the clients increases. On the other hand, the clients using UDP transmission experience a high number of quality switches, where the highest number is observed by TV clients.

We also calculate the overall QoE values and present the results in Fig. 4d. QoE is calculated by using a weighted linear function of different QoE metrics proposed in [24]. In the calculations, we normalized the bitrate values according to the client types in order to obtain comparable results among different devices. The results show that all E2E-BPP clients obtain better QoE than the E2E-UDP clients.

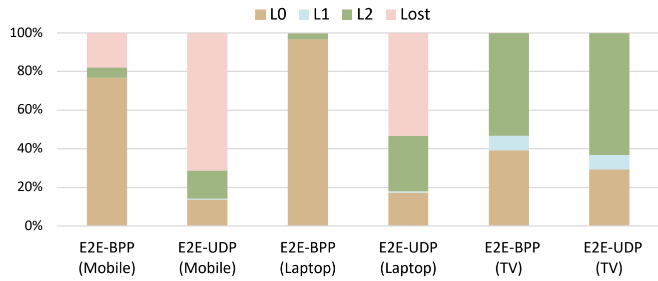


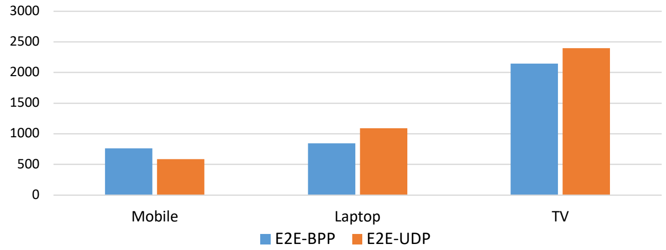
Fig. 3: Layers Received + Loss in the Network

V. Conclusions

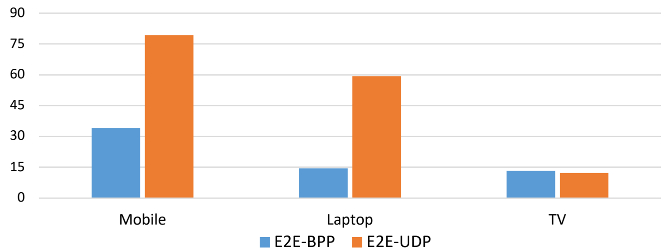
In this paper, we have investigated the problem of end-to-end video delivery, for video streaming applications, from drones to a set of clients having different device types, with a view to find the optimal deployment of the drones over a football field. We have proposed a drone deployment and flight planning strategy which focuses on the energy used by the drones, when doing drone positioning and video collection. We presented an energy consumption model and an energy management process for the drones. Thus we can define the recharging times and direct the drones to an energy supply station for energy replenishment.

The energy management process allows for the drone network to send the video seamlessly, without having a video shortage at the clients, due to the exchange of drones in low energy conditions. In our framework, a virtualized server at the drone edge sends the video with the highest quality in BPP packets. These BPP packets with video are collected by a virtualized BPP processing function that resides on the edge with the clients. The packet trimming operation is done by considering the available bandwidth and the client types. Finally, the results have shown that using the combination of the emerging technologies and protocols, BPP, SVC, and virtualized edge functions, when considering the end-to-end video streaming applications, provides a high QoE performance.

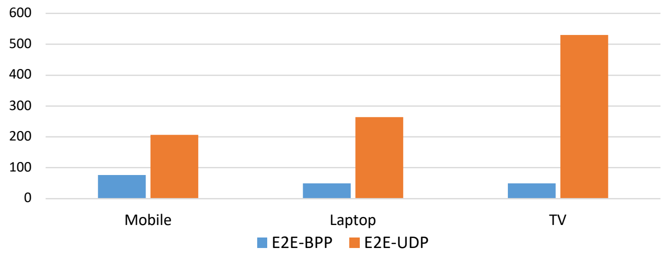
In future work, we will refine the energy consumption model to consider more elaborate aspects and also enhance our energy management process. We will show the effects of energy model on QoE, by considering the bitrates of the SVC layers. We plan to improve the BPP processing function by adding a mechanism that adapts to the number of quality switches.



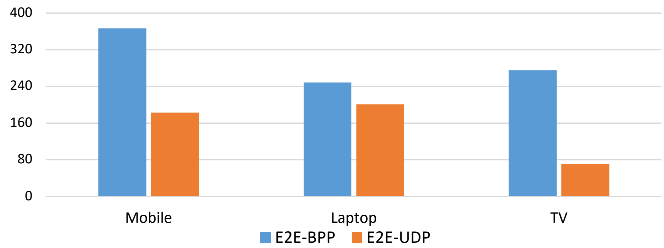
(a) Average Bitrate (in Kbps) *Higher is Better*



(b) Duration of Pauses (in seconds) *Lower is Better*



(c) Quality Switches *Lower is Better*



(d) Overall QoE Results *Higher is Better*

Fig. 4: Quality of Experience Values across the Experiments

Acknowledgment

This work is funded by TUBITAK Electric, Electronic and Informatics Research Group (EEEAG) under grant 121E373. Stuart Clayman is funded by the TUDOR project (UK DCMS).

References

[1] R. Li, A. Clemm, U. Chunduri, L. Dong, and K. Makhijani, "A new framework and protocol for future networking applications," in *NEAT 2018 - Proc. of ACM Workshop on Networking for Emerging Applications and Technologies*, August 2018.

[2] E. Bozkaya, K.-T. Foerster, S. Schmid, and B. Canberk, "Airnet: Energy-aware deployment and scheduling of aerial networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 12 252–12 263, 2020.

[3] A. V. Savkin, C. Huang, and W. Ni, "On-demand deployment of aerial base stations for coverage enhancement in reconfigurable intelligent surface-assisted cellular networks on uneven terrains," *IEEE Communications Letters*, vol. 27, no. 2, pp. 666–670, 2023.

[4] A. Clemm, M. Zhani, and R. Boutaba, "Network management 2030: Operations and control of network 2030 services." *Journal Network System Management*, vol. 28, pp. 721–750, 2020.

[5] S. Clayman and M. Sayıt, "Low Latency Low Loss Media Delivery Utilizing In-Network Packet Wash," *Journal of Network and Systems Management*, vol. 31, no. 1, p. 29, 2023.

[6] C. Singhal and B. N. Chandana, "Aerial-son: Uav-based self-organizing network for video streaming in dense urban scenario," in *2021 Intl. Conf. on COMmunication Systems NETworkS (COMSNETS)*, 2021, pp. 7–12.

[7] C. He, Z. Xie, and C. Tian, "A qoe-oriented uplink allocation for multi-uav video streaming," *Sensors*, vol. 19, no. 15, 2019.

[8] J. Molina, D. Muelas, J. E. L. De Vergara, and J. J. García-Aranda, "Network quality-aware architecture for adaptive video streaming from drones," *IEEE Internet Computing*, vol. 24, no. 1, pp. 5–13, 2020.

[9] X. Wang, A. Chowdhery, and M. Chiang, "Networked drone cameras for sports streaming," in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, 2017, pp. 308–318.

[10] M. Bishop, "Hypertext Transfer Protocol Version 3 (HTTP/3)," Internet Engineering Task Force, Internet-Draft draft-ietf-quic-http-34, Feb. 2021. [Online]. Available: https://datatracker.ietf.org/doc/html/draft-ietf-quic-http-34

[11] M. Seufert, R. Schatz, N. Wehner, and P. Casas, "QUICker or not? -an Empirical Analysis of QUIC vs TCP for Video Streaming QoE Provisioning," in *22nd Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN)*, 2019, pp. 7–12.

[12] M. Palmer, T. Krüger, B. Chandrasekaran, and A. Feldmann, "The QUIC Fix for Optimal Video Streaming," in *Proceedings of the Workshop on the Evolution, Performance, and Interoperability of QUIC*, ser. EPIQ'18. ACM, 2018, p. 43–49.

[13] L. Dong and R. Li, "In-packet network coding for effective packet wash and packet enrichment," in *2019 IEEE Globecom Workshops (GC Wkshps)*, 2019, pp. 1–6.

[14] S. Clayman, M. Tuker, H. Arasan, and M. Sayıt, "Managing Video Processing and Delivery using Big Packet Protocol with SDN Controllers," in *IEEE Conf. on Network Softwarization – Netsoft*, Tokyo, July 2021.

[15] M. Handley, C. Raiciu, A. Agache, A. Voinescu, A. Moore, G. Antichi, and M. Wójcik, "Re-Architecting Datacenter Networks and Stacks for Low Latency and High Performance," in *ACM SIGCOMM*, Aug 2017, pp. 29–42.

[16] C. Westphal, K. Makhijani, and R. Li, "Packet Trimming to Reduce Buffer Sizes and Improve Round-Trip Times - Extended Abstract," in *Buffer Sizing Workshop (BS'19)*. New York, NY: ACM, 2019.

[17] A. Albalawi, H. Yousefi, C. Westphal, K. Makhijani, and J. Garcia-Luna-Aceves, "Enhancing End-to-End Transport with Packet Trimming," in *GLOBECOM 2020 - 2020 IEEE Global Comms Conf*, 2020, pp. 1–7.

[18] S. Clayman, , M. Tüker, E. Karakış, and M. Sayıt, "In-Network Video Quality Adaption using Packet Trimming at the Edge," in *26nd Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN)*, Paris, France, 2023.

[19] A. Heppes and H. Melissen, "Covering a rectangle with equal circles," *Periodica Mathematica Hungarica*, vol. 34, no. 1, pp. 65–81, 1997.

[20] H. Ghazzai, M. B. Ghorbel, A. Kadri, M. J. Hossain, and H. Menouar, "Energy-efficient management of unmanned aerial vehicles for underlay cognitive radio systems," *IEEE Transactions on Green Communications and Networking*, vol. 1, pp. 434–443, 2017.

[21] J. Zhang, J. F. Campbell, D. C. Sweeney II, and A. C. Hupman, "Energy consumption models for delivery drones: A comparison and assessment," *Transportation Research Part D: Transport and Environment*, vol. 90, p. 102668, 2021.

[22] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *CoRR*, vol. abs/1804.02767, 2018. [Online]. Available: http://arxiv.org/abs/1804.02767

[23] Microsoft. (2023, Jan.) Airsim. [Online]. Available: https://www.microsoft.com/en-us/ai/autonomous-systems-project-airsim?activetab=pivot1%3aprimaryr3

[24] X. Yin, A. Jindal, V. Sekar, and B. Sinopoli, "A control-theoretic approach for dynamic adaptive video streaming over http," in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, 2015, pp. 325–338.