Autonomous Robots (2020) 44:75–91
https://doi.org/10.1007/s10514-019-09851-6

Live multicast video streaming from drones: an experimental study

Raheeb Muzaffar[1] · **Evşen Yanmaz**[1,2] · **Christian Raffelsberger**[1] · **Christian Bettstetter**[1,3] · **Andrea Cavallaro**[4]

Received: 16 August 2018 / Accepted: 6 April 2019 / Published online: 27 April 2019

Abstract
We present and evaluate a multicast framework for point-to-multipoint and multipoint-to-point-to-multipoint video streaming that is applicable if both source and receiver nodes are mobile. Receiver nodes can join a multicast group by selecting a particular video stream and are dynamically elected as designated nodes based on their signal quality to provide feedback about packet reception. We evaluate the proposed application-layer rate-adaptive multicast video streaming over an aerial ad-hoc network that uses IEEE 802.11, a desirable protocol that, however, does not support a reliable multicast mechanism due to its inability to provide feedback from the receivers. Our rate-adaptive approach outperforms legacy multicast in terms of goodput, delay, and packet loss. Moreover, we obtain a gain in video quality (PSNR) of 30% for point-to-multipoint and of 20% for multipoint-to-point-to-multipoint streaming.

Keywords Multicast video streaming · Rate-adaptation · IEEE 802.11 · Drones

1 Introduction

Mobile aerial devices commonly referred to as "drones" range in size, endurance, maneuverability, flight time, and payload capacity (Hanscom and Bedford 2013). Small drones can be used for diverse applications such as search and rescue, surveillance, and transportation (Hayat et al. 2016). Cameras mounted on multiple drones can capture aerial views of several areas of an extended region. Multipoint-to-point-to-multipoint video streaming from the air is therefore important, among other application scenarios, for first responders to obtain a simultaneous overview of different observed areas.

Many applications would benefit from video streaming from multiple drones (Chmaj and Selvaraj 2015; Gupte et al. 2012; Quaritsch et al. 2008). When fixed communication infrastructure is unavailable and a quick communication setup is required, the problem of reliably streaming videos arises as multicast from multiple drones to mobile receivers over IEEE 802.11, which is a widely supported protocol by

This is one of the several papers published in *Autonomous Robots* comprising Special Issue on Robot Communication Challenges: Real-World Problems, Systems, and Methods.

This work is funded by the security research programme KIRAS of the Federal Ministry for Transport, Innovation, and Technology (bmvit), Austria under the grant agreement n. 854747 (WatchDog) and ICT of the future programme (4th call) of the Austrian Research Promotion Agency (FFG) under grant agreement n. 855468 (ForestiMate). The support of the UK EPSRC project NCNR (EP/R02572X/1) and EACEA Agency of the European Commission under EMJD ICE FPA n. 2010-0012 is also acknowledged. The work of Evsen Yanmaz was performed while being employed at Lakeside Labs GmbH, Klagenfurt, Austria.

✉ Raheeb Muzaffar
muzaffar@lakeside-labs.at

Evşen Yanmaz
eyanmaz@alumni.cmu.edu

Christian Raffelsberger
raffelsberger@lakeside-labs.at

Christian Bettstetter
Christian.Bettstetter@aau.at

Andrea Cavallaro
a.cavallaro@qmul.ac.uk

[1] Lakeside Labs GmbH, Klagenfurt, Austria

[2] Ozyegin University, Istanbul, Turkey

[3] Institute of Networked and Embedded Systems, Alpen-Adria-Universität Klagenfurt, Klagenfurt, Austria

[4] Centre for Intelligent Sensing, Queen Mary University of London, London, UK

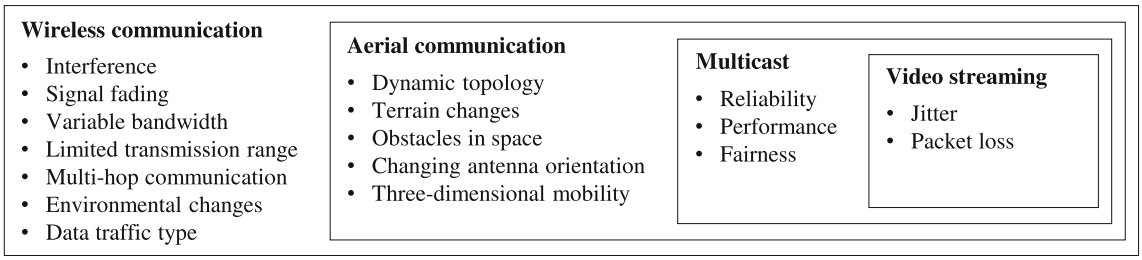


Fig. 1 IEEE 802.11 communication and video streaming challenges in aerial networks

drone platforms and operates in the license-free spectrum (Hayat et al. 2016; Fu et al. 2007; Andre et al. 2014).

Multicast video streaming can enable large-scale situational awareness by simultaneously transmitting identical data to multiple spatially distributed receivers (users). However, multicast bears the challenges of reliability, fairness, performance, and jitter (Choi et al. 2007; Vella and Zammit 2013; Zhu et al. 2004; Nafaa 2007). Wireless communication in various mobility scenarios and diverse environmental conditions pose several challenges related to ground and aerial transmission, multicast, and video streaming (see Fig. 1). Signal fading, attenuation, and interference affect communication and cause temporal link failures and packet losses (Lindeberg et al. 2011). Mobility further reduces performance due to changes in the network topology, requiring routing and multi-hop communication (Crow et al. 1997; Takai et al. 2001). Moreover, wireless communication is limited to its transmission range, which can further be affected by the terrain, environmental changes, and obstacles, thus resulting in higher and bursty bit errors (Bekmezci et al. 2013; Su et al. 2016). Other factors that influence the wireless link quality include the antenna characteristics and orientation that generate different radiation patterns (Yanmaz et al. 2011, 2013; Ahmed et al. 2011).

Multicast packets in 802.11 are sent as broadcast and decoded by the recipients of a multicast group. Since acknowledgements (ACK) of multicast packets from multiple receivers cause overhead (in addition to the issue of scheduling and synchronization), multicast packets are not acknowledged (Maraslis et al. 2012). Therefore, a source is unaware of the reception of sent packets and is unable to detect and retransmit lost packets (Dujovne and Turletti 2006). Similarly, the lack of feedback in 802.11 multicast does not allow a source to adapt the transmission rate when link conditions vary. This limitation may cause some nodes to suffer from network congestion while other network resources are underutilized. Fairness can be achieved through rate adaptation to meet the requirements and reception condition of the receivers. In addition to this, the legacy 802.11a/b/g use the lowest bit rates (1, 2, or 6 Mbit/s) for multicast traffic. The performance of the receiver nodes degrades with the use of lowest bit rates considering conditions when they may afford better rates and increase network performance (Choi et al. 2007). Finally, reducing jitter and long packet delays is crucial for video streaming (Lindeberg et al. 2011). Quality of experience (QoE) support is required for a smooth video reception in live video multicast streaming, which is hindered by the aforementioned challenges (Zhu et al. 2004; Nafaa 2007).

These challenges are addressed by several approaches, including 802.11aa amendments (Maraslis et al. 2012; Salvador et al. 2013), promiscuous reception of unicast transmission (PRUT), polling-based (PB) schemes, and leader-based (LB) protocols (Vella and Zammit 2013). However, these approaches require modifications of the medium access control (MAC) layer and are not readily usable. Moreover, these approaches have not yet been validated for a highly mobile environment requiring high data rates, such as streaming videos from drones to mobile receivers.

In this paper, we describe and evaluate for the first time with an aerial testbed our application-layer rate-adaptive multicast video streaming framework (Muzaffar et al. 2016a, b). This framework addresses the multicast challenges of reliability, performance, and fairness as well as the delay and packet loss issues associated to video streaming. Our application-layer framework is *suitable for any 802.11 wireless ad-hoc setup and does not require any modifications of the MAC layer*. MAC layer modification schemes are complicated to configure with most drone platforms and are undesirable for the targeted applications such as search and rescue, forest fire monitoring, and surveillance due to configuration and compatible deployment issues. The proposed application-layer framework is a ready-to-go approach suitable for most off-the-shelf drone platforms.

We build an IEEE 802.11a ad-hoc network among AscTec Pelican quadrotor drones (Ascending Technologies http://www.asctec.de/en/uav-uas-drones-rpas-roav/asctec-pelican) and the ground receivers to evaluate our framework. Specifically, we test point-to-multipoint and multipoint-to-point-to-multipoint video streaming: a team of drones stream videos of different observed areas as unicast to a video multicast gateway (VMG) that transcodes the received videos and multicasts them to multiple ground receivers in a multipoint-to-point-to-multipoint fashion (see Fig. 2). The ground

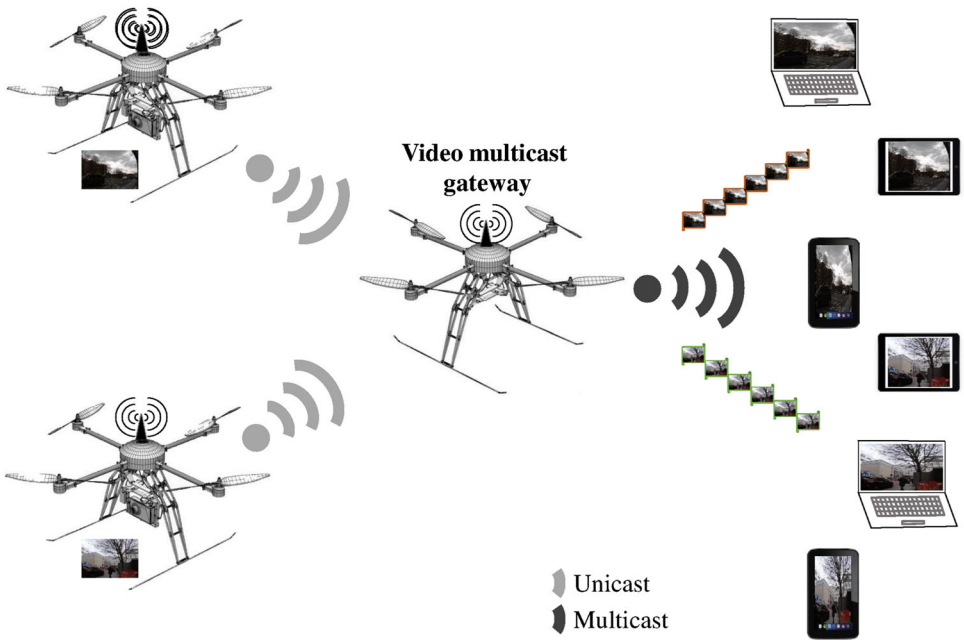


Fig. 2 Two-hop multicast video streaming from drones. Drones unicast videos of their respective observed areas to the video multicast gateway that streams videos as multicast to the ground receivers

receivers can be mobile but in our experiments are stationary at a minimum distance of 40 m from a source drone. Since the drones fly at a speed of approximately 5 m/s, the signal strength rapidly decreases causing packet and frame losses that in turn result in blurred and distorted videos. Application-layer feedback about packet reception from multiple nodes is used for retransmission of lost packets, which improves reliability. Similarly, the feedback is used for rate adaptation, which enhances performance and fairness. In particular, we adapt the link transmission rate (PHY rate), the video encoding rate, and the video frame rate with changing reception conditions of the receivers to address multicasting and video streaming challenges in mobile environments. We refer to the combination and adaption of the encoding rate, link rate, and frame rate parameters as the *ELF rate adaptation*. We experimentally validate that our rate-adaptive multicast video streaming approach outperforms the legacy 802.11 multicast in terms of goodput, packet loss, delay, and video quality using air-ground communication links through high-speed drones.

The remainder of the paper is organized as follows. Section 2 gives an overview of 802.11aa amendments and 802.11 multicast approaches. Section 3 details on the proposed approach and explains the extension to our previous work. Section 4 presents the experimental testbed and Sect. 5 discusses the results. Finally, Sect. 6 concludes the paper.

2 Related work

In this section, we discuss solutions for wireless multicast transmission and analyze their advantages and shortcomings for video streaming. Approaches for reliably multicasting data over wireless local area networks (WLANs) include 802.11aa amendments, PRUT, PB schemes, and LB protocols.

The IEEE 802.11aa Task Group is working toward improving multicast transmission over WLANs through the Group Addressed Transmission Service (GATS), since the legacy 802.11 multicast is unreliable and uses the basic fixed transmission rate (1, 2, or 6 Mbit/s) Maraslis et al. (2012), Salvador et al. (2013). GATS specifies three schemes, namely the directed multicast service (DMS), the groupcast with retries (GCR), unsolicited retries (GCR-UR), and the GCR BlockACK (GCR-BA).

The DMS sends multicast packets as unicast to all recipients of the group while each recipient follows the normal ACK and retransmit policy. This solution provides reliability at the cost of greater overhead and is not suitable for large networks. The GCR-UR approach multicasts packets repeatedly to gain reliability through redundancy. The number of times packets are sent is undefined and is dependent on the implementation and application requirements. An ACK mechanism is not used which makes the method less reliable than DMS, but scalable with reduced overhead. However, multiple transmissions of the packets received successfully waste network resources causing additional delays for video streaming. Lastly, GCR-BA sends multiple multicast frames to poll one or more receivers for a block acknowledgment. Frames not received correctly can then be retransmitted. The performance of the scheme depends on the configuration of the block parameters, application requirements, and the network density (Salvador et al. 2013). GCR-BA offers a good tradeoff between reliability, scalability, and the overhead of

receiving individual ACKs from the chosen number of recipients and retransmission of lost packets (Maraslis et al. 2012; Salvador et al. 2013).

A comparison of the GATS and an enhancement that selects the best GATS service for a given scenario is suggested in Banchs et al. (2014). The performance of the schemes is evaluated using throughput, reliability, and video frame error rate measurements by varying the number of receiver and sender nodes. The legacy multicast service performs worst in all scenarios, while the DMS performs best with one multicast receiver. The GCR-BA performance is best when the number of receivers is small, i.e. between 2 and 7, while the GCR-UR performs better with a large number of receivers, i.e. more than 22. Thus, there is no single scheme that is best suited for all scenarios and a selection of a scheme based on the given scenario could result in an improved network performance. However, the results may differ under diverse network conditions and when mobility is considered.

Multicast approaches such as promiscuous reception of unicast transmission (PRUT), polling-based (PB) schemes, and leader-based (LB) protocols were surveyed by Vella and Zammit (2013). PRUT sends multicast data to a selected receiver using its unicast address while other members of the multicast group listen in promiscuous mode (Ge and McKinley 2002; Chandra et al. 2009). This improves the quality of the received video stream at the selected receiver due to the request to send/clear to send signaling solving the hidden node problem and MAC retransmission mechanism (Ge and McKinley 2002; Chandra et al. 2009; Tourrilhes 1998). Other members also benefit by improving reliability if they experience correlated packet losses as experienced by the unicast receiver. The drawback of this scheme is that each multicast member of the group is to be configured with the MAC and IP addresses of the unicast receiver for promiscuous reception. Apart from compromising security, since all members receive data directed to the selected unicast node, multicast receivers will experience total packet loss in case the unicast receiver fails or leaves the multicast group (Vella and Zammit 2013). DirCast (Chandra et al. 2009) handles failure of the unicast receiver by re-evaluating the choice every 30 s or whenever a new node joins the multicast group and reassigns the responsibility when required. Furthermore, proactive forward error correction (FEC) is applied that adds redundancy, varying with the amount of packet loss, to enable receivers to recover from erroneous packets. However, this adds an overhead of 10 kbit/s per client (Chandra et al. 2009).

The PB schemes sequentially inquire from each member of the multicast group about a packet reception through a request for ACK (RAK) control frame. A packet not acknowledged is retransmitted until an ACK from all members is received (Paris et al. 2013; Piamrat et al. 2009). The QoE-based dynamic rate adaptive multicast (Q-DRAM) concerns rate adaptation based on the implementation of pseudo-subjective quality assessment (PSQA) at the multicast receivers reducing the frequency of feedback (Piamrat et al. 2009). Although PB schemes guarantee reliability through polling and retransmission, they consume additional network resources and are thus inefficient for video streaming applications that require adherence to stringent packet delays (Vella and Zammit 2013).

LB protocols select a leader node on behalf of the multicast group that is responsible for sending ACKs for received packets. Non-leader nodes can seek retransmission through negative ACKs (NAKs). The selection of the leader node can be based on different criteria, e.g. the one with the strongest/weakest link with the source, or the node that joins the group first (Kuri and Kasera 1999). The leader election protocol (LEP) dynamically selects the receiver with the worst channel conditions as the leader through a modified internet group management protocol (IGMP) membership reports where a duplicated bit is reserved to identify if a leader election is required (Thierry and Yongho 2006). However, leader selection with the worst channel conditions increases the chance for the leader to frequently disassociate itself from the source resulting in recurring leader elections (Vella and Zammit 2013). In Li and Herfet (2009), when a node sends a request to join a group, the source checks in a table if a group leader exists and replies accordingly to the new node. If, however, the leader leaves without an announcement, the second node that joined is activated as a leader. This is determined when no feedback is received after several retransmissions.

LB multicast with auto rate fallback (LM-ARF) uses feedback from the leader and non-leader nodes to adapt the contention window and the PHY rate (Choi et al. 2007). LB-ARF is based on commercial ARF rate adaptation where the source increases the PHY rate to the next (higher) level upon ten consecutive ACKs, while it decreases the rate to the previous (lower) level upon two consecutive NAKs (Kamerman and Monteban 1997; Biaz and Wu 2008). Robust rate adaptive multicast (RRAM) extends LB-ARF by allowing non-leader nodes to oppose adaptation to a higher PHY rate (Thierry and Yongho 2006). The source sends a modified MAC frame to inform all members of the rate change. Non-leaders can oppose the adaptation if the current signal-to-interference-noise-ratio (SINR) does not correspond to a pre-defined SINR-PHY table. SNR-based auto rate for multicast (SRAM) uses the SNR of the receivers to select the PHY rate according to the receiver that experiences the worst channel conditions. SRAM focuses on the received video quality such that the corresponding selected PHY rate maintains the video peak-signal-to-noise-ratio (PSNR) above 30 dB (Park et al. 2006).

LB protocols suffer in terms of scalability, since considering retransmission upon a NAK of non-leaders consumes additional channel time. Furthermore, if the resignation of a leader node is unnoticed, the source keeps on sending retrans-

Table 1 Comparison of 802.11 multicast schemes (this table is an extension of those in Muzaffar et al. 2016a, Muzaffar et al. 2016b)

Refs.	Scheme	Changes to MAC?	Reliability	Scalability	Rate adaptation	Evaluation	Multicast groups	Hops
Paris et al. (2013)	Polling based	Yes	None	Low	Joint reception correlation	Testbed	1	1
Piamrat et al. (2009)	Polling based	Yes	None	Low	User experience in time	Simulation	1	1
Choi et al. (2007)	Leader based	Yes	High	Low	Auto rate fallback	Simulation	1	1
Park et al. (2006)	Leader based	Yes	None	High	Beacon Signal	Simulation	1	1
Maraslis et al. (2012), Salvador et al. (2013)	Directed multicast	Yes	High	Low	None	Testbed	1	1
	Unsolicited retries	Yes	Depends on no. of retries	High	None	Testbed	1	1
	Block ACK	Yes	Depends on block size	Depends on block size	None	Testbed	1	1
Ours	Dynamic leader	No	Medium	High	RTP packet feedback	Aerial testbed	1 & 2	1 & 2

missions since no ACK is received from the leader (Kuri and Kasera 1999; Li and Herfet 2008). Other group members will eventually time out and start the process of subscribing to the group afresh, thereafter, a new leader will be elected. Moreover, existing approaches require non-negligible modifications to the MAC layer, which makes their deployment difficult and time consuming. Besides MAC modifications, a drawback for leader based schemes with worst channel conditions is the fact that performance and fairness challenges are not addressed. Additionally, their performance in real scenarios with mobile and drone networks is unclear.

Table 1 compares existing approaches with *ELF rate adaptation*. The *ELF rate adaptation* approach proposes a dynamic leader scheme that has some advantages, compared to the aforementioned leader based approaches. In our dynamic leader scheme some nodes act as a backup of the leader node that can take over to provide feedback in case the leader node stops responding. A backup node is then immediately assigned the role of leader dynamically, thus preventing leader outages. Furthermore, to improve packet delivery rate, non-leader nodes can ask for retransmission with NACKs, as detailed in the next section.

3 Multipoint-to-point-to-multipoint video streaming

We use the real-time transport protocol (RTP) to stream videos and the RTP control protocol (RTCP) to signal packet information against each transmitted RTP packet. Drones stream their videos as unicast transmission to a video multicast gateway (VMG), which receives RTP packets from the source drones and provides feedback through an application-layer acknowledgment, either `AL-ACK` for packet reception, or a negative acknowledgment `AL-NAK` for retransmission of lost packets (see Fig. 3). The maximum number of retransmissions depends on the buffer size, that in our case, is set to store for a maximum duration of 500 ms. Thus, a retransmission usually happens only once. Source drones regulate the *PHY rate*, *video encoding rate*, and *video frame rate* based on the received feedback.

One can argue that the unicast 802.11 MAC layer offers an efficient PHY rate adaptation with the changing channel conditions. The 802.11 standard, although, does not specify any particular algorithm, minstrel, a rate control algorithm (LinuxWireless Project https://wireless.wiki.kernel.org/en/developers/do-cumentation/mac80211/ratecontrol/minstrel), is used by default in Linux-based systems (Yin et al. 2011). However, throughput variations in highly mobile networks can be significant since adaptations are made on the basis of MAC layer frame transmissions in the 802.11a standard. Additionally, probe frames are frequently used to determine if the transmission rate can be adjusted. Thus, application-

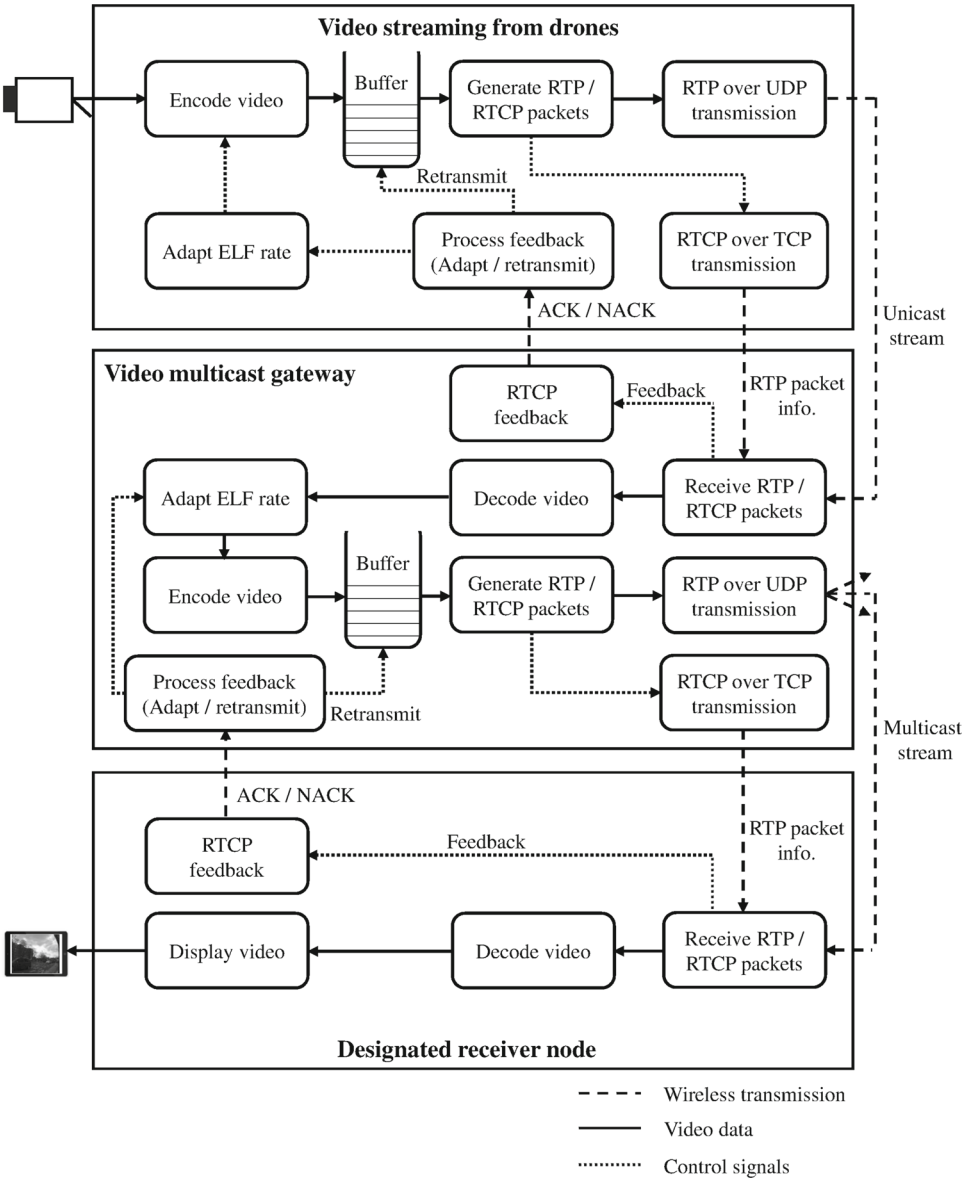


Fig. 3 Source drones unicast real-time video streams over the wireless medium as RTP packets to the VMG. The VMG multicasts received video streams to multiple ground receivers and creates multicast groups based on selected videos for high-quality stream. RTCP signaling provides RTP packet information to the VMG from drones and to the receiver nodes from the VMG. Feedback about packet reception or packet loss is provided by the designated nodes to the VMG and from the VMG to the drones. Lost packets are retransmitted and the *video encoding*, *link* (PHY), and *frame* rates (ELF) are adapted based on the received feedback

layer adaptive solutions are more feasible for streaming applications in mobile environments (Kacianka and Hellwagner 2015; Kofler et al. 2011).

Our scheme comprises several functions. The role assignment function, described in Sect. 3.1, elects members of the multicast group to provide feedback on packet reception. The feedback and retransmission processes are explained in Sect. 3.2, while Sect. 3.3 details the rate adaptation mechanism. The deregistration process is covered in Sect. 3.4, and an overview of the point-to-multipoint video streaming is presented in Sect. 3.5.

3.1 Role assignment

The VMG multicasts video streams to the ground receivers as low-quality overview videos. The low-quality videos are by default sent at a constant video encoding rate of 64 kbit/s. This design parameter can be adjusted based on the application requirements.

The ground receivers can select a particular video to be received in high quality. The selection of a video at the ground receiver triggers the creation of a multicast group at the VMG. The VMG assigns roles to the members of the multicast group as *designated nodes* based on their signal quality from the VMG. *Designated nodes* are responsible to provide feedback about packet reception to the VMG.

The role assignment procedure is defined in Algorithm 1. As for the notation, $\mathcal{M} = \{\mathcal{M}_1, \ldots, \mathcal{M}_{n_1}\}$ defines the multicast groups; P is the primary designated node; $S = \{S_1, \ldots, S_{n_2}\}$ are the secondary designated nodes S_i, ordered by signal strength; $B = \{B_1, \ldots, B_{n_3}\}$ are the best effort nodes B_i, ordered by signal strength; and V is a new mobile

receiver node, R_X is the RSS of node X, and card(.) is the cardinality of a set.

A new receiver node V requests to join a multicast group $\mathcal{M}_i$ by selecting an overview video to receive its high-quality stream. Initially, the first node that joins the multicast group becomes the *primary* designated node, P. As the cardinality of the multicast group, $card(\mathcal{M}_i)$, increases, the node with the highest signal quality becomes P. After P, members with good signal quality become part of the set of *secondary* designated nodes, S, and act as a backup of P. Roles are assigned hierarchically based on signal strength to minimize feedback delays and channel contention time. Fewer than half of the nodes of the multicast group are assigned *designated node* roles. However, this parameter is a tradeoff between scalability and reliability and can be adjusted based on the application requirements. Nodes that do not provide feedback become part of the set of *best effort* nodes, B, and receive videos on a best effort basis.

Algorithm 1: Role assignment

Input: V, R_V, R_P, R_{S_i}, R_{B_i}, card(S), card($\mathcal{M}_i$).
Output: Role is assigned to V joining multicast group $\mathcal{M}_i$.

```
if card(M_i) > 0 then                          // If M_i has one or more members
    if card(S) > 0 then                        // S is not an empty set
        M_i = M_i ∪ {V}                        // V is added to M_i
        if (card(S)+1)/card(M_i) > 0.5 then    // Fewer than half nodes are assigned roles
            if R_V > R_P then                  // Signal quality of V is better than P
                B = B ∪ {S_n2}                 // Nodes change their set
                S = S \{S_n2}
                S = S ∪ {P}; P = V
            else if R_V > R_Si then            // Signal quality of V is better than S
                B = B ∪ {S_n2}                 // Weakest node of S moves to B and V joins S
                S = S \{S_n2}; S = S ∪ {V}
            else
                B = B ∪ {V}                    // V joins B
            end
        else if R_V > R_Si then                // Signal quality of V is better than S
            B = B ∪ {S_n2}                     // Weakest node of S moves to B and V joins S
            S = S \{S_n2}; S = S ∪ {V}
        else
            B = B ∪ {V}                        // V joins B
        end
    else if card(M_i) ≥ 2 then                 // If S is empty and M_i has two or more members
        M_i = M_i ∪ {V}                        // V is added to M_i
        if R_V > R_B1 then                     // Signal quality of V is better than B_1
            S_1 = V                            // V is assigned the role S
        else
            S = {B_1}
            B = B \{B_1}; B = {V}              // Member of B moves to S and V is added to B
        end
    else
        M_i = M_i ∪ {V}; B_1 = V               // V joins M_i as a member of B
    end
else
    M_i = {V}; P = V                           // V joins M_i and is assigned the role P
end
```

We used the role assignment process to select designated nodes, based on the signal quality, to provide feedback about packet reception. As described in Sect. 2, this is advantageous compared to leader-based approaches that select the first node joining the multicast group as the leader (Kuri and Kasera 1999; Li and Herfet 2008) or the node with the worst channel conditions (Thierry and Yongho 2006; Li and Herfet 2009). In these approaches, if the current leader leaves the group, a new leader election process starts or the node that joins the multicast group immediately after will act as the leader. However, such a scheme is not suitable for highly mobile devices since leader elections are more frequent and there exists no leader to respond to packets received during the election process (Vella and Zammit 2013). Our approach improves the design of the LB approaches through the role assignment process where S nodes act as a backup of P to avoid retransmissions while assigning the role of P to another group member, as explained in Sect. 3.2.

3.2 Feedback and retransmission

Our approach is similar to the leader-based schemes, described in Sect. 2. However, unlike the leader-based

schemes that use feedback of a leader node for rate adaptation, we allow feedback from multiple receivers (*designated nodes*). P is mainly responsible to acknowledge received packets through AL-ACK, while either P or a node S_i from S can request a retransmission through an AL-NAK. AL-ACKs are also used to indicate a possible change in the receivers' reception condition to the VMG. In particular, if AL-ACKs from P to the VMG for two consecutive video packets are not received by a node S_i but this node received the corresponding video packets, it will send an AL-ACK to the VMG. To support this function, all nodes have to listen in promiscuous mode.

Our approach has similarities with the 802.11aa GCR-BA scheme, since we use RTP packets for video transmission, whereby an RTP packet may span over a variable number of MAC frames. Similar to a block of frames in 802.11aa GCR-BA scheme, we provide feedback against each RTP packet. However, we use an application-layer approach that does not use MAC layer ACKs and hence does not require any MAC layer modifications.

The signal quality of the receivers is assessed through the *probe group* function by the VMG, to dynamically assign the feedback responsibility to the group members, accounting for the change in reception conditions due to mobility. Thus, a new node joining the group can be assigned the responsibility of P if the signal strength of the new node is higher than the one of P (i.e. $R_V > R_P$). Similarly, a new node can be assigned the role of S_i, if its signal strength is better than the one of the existing S_i (i.e. $R_V > R_{S_i}$). Likewise, in case two consecutive AL-NAKs are received from P while an AL-ACK is received from any S_i for the same RTP packet, the *probe group* function assesses signal quality of the group members and if required assigns the role of P to another group member that has the strongest link with the VMG. The *probe group* function also invokes rearrangement of the group members in S and B based on their signal quality.

3.3 ELF rate adaptation

The goal of the rate adaptation function is to adjust the *video encoding*, *link transmission* (*PHY*), and *frame* (*ELF*) rates based on the AL-ACKs and AL-NAKs from the designated nodes. To address multicast challenges of performance and fairness (see Fig. 3), the *PHY rates* are adapted as the receivers reception conditions vary. This exploits maximal possible network capacity, compared to the legacy multicast that uses the lowest (1, 2, or 6 Mbit/s) *PHY* rates. Although our rate adaptation mechanism is applicable to all IEEE 802.11 variants that support ad-hoc networking, we use 802.11a with *PHY rates* of 6, 9, 12, 18, 24, 36, 48, and 54 Mbit/s. The *PHY rate* is initially set to 54 Mbit/s to utilize maximal link capacity. It is reduced sequentially upon a signal loss to widen the communication range. Ideally, the required link capacity is proportional to the encoding rate applied to the video stream with some added overhead. However, due to unstable wireless link conditions and communication challenges, the required link capacity may not be maintained. Thus, the adaptation scheme also regulates the *video encoding rate* and the *frame rate*, based on varying link conditions to reduce packet losses.

The *ELF rate adaptation* is implemented using GStreamer libraries (Gstreamer framework https://gstreamer.freedesktop.org). GstRtpBin is configured to generate RTCP sender reports (SR). The receiver receives RTCP packets containing SR against each RTP packet sent. If a packet is not received, the feedback information is communicated back to the source using rtpjitterbuffer. rtpjitterbuffer will wait until the next RTP packet is expected and will generate a NACK otherwise.

The rate adaptation works as follows and is applied by the encoder to the next group of pictures (GoP). GoP may contain several frames depending upon the channel conditions. The parameters and values were chosen after extensive experimental validations (Muzaffar et al. 2016b; Kacianka and Hellwagner 2015):

- The PHY rate is increased upon ten consecutive AL-ACKs and is decreased upon a signal loss, i.e. when the feedback response is not received from any of the designated nodes.
- The source (drones and VMG) initially starts with a video encoding rate of 512 kbit/s and increases it by 5 % to a maximum of 8192 kbit/s upon receiving an AL-ACK. It is decreased with the same rate of 5 % upon receiving three consecutive AL-NAKs down to a minimum of 128 kbit/s.
- A packet is retransmitted when an AL-NAK is received by the source (drone or VMG).
- The frame rate is initially set to 25 frames/s and is reduced by one upon three consecutive AL-NAKs down to a minimum of 10 frames/s. The frame rate is decreased only if the video encoding rate is smaller than or equal to 256 kbit/s. The frame rate is increased by a frame upon reception of an AL-ACK, up to a maximum of 25 frames/s.

3.4 Deregistration

A multicast group member can initiate the *deregistration* process by switching to another multicast group or by leaving the group. If the node leaving the group belongs to the set of *designated nodes*, the *probe group* function is initiated to reassign roles if required. If a node relents itself from a multicast group without initiating the *deregistration* process, VMG from the *probe group* function when initiated, will get to know that the node is no longer part of the group. If the P node relents itself without initiating the *deregis-*

tration process, the VMG will initiate the *probe group* and *role assignment* function, since it will no longer receive the AL-ACKs from P and will select a new P.

3.5 Point-to-multipoint

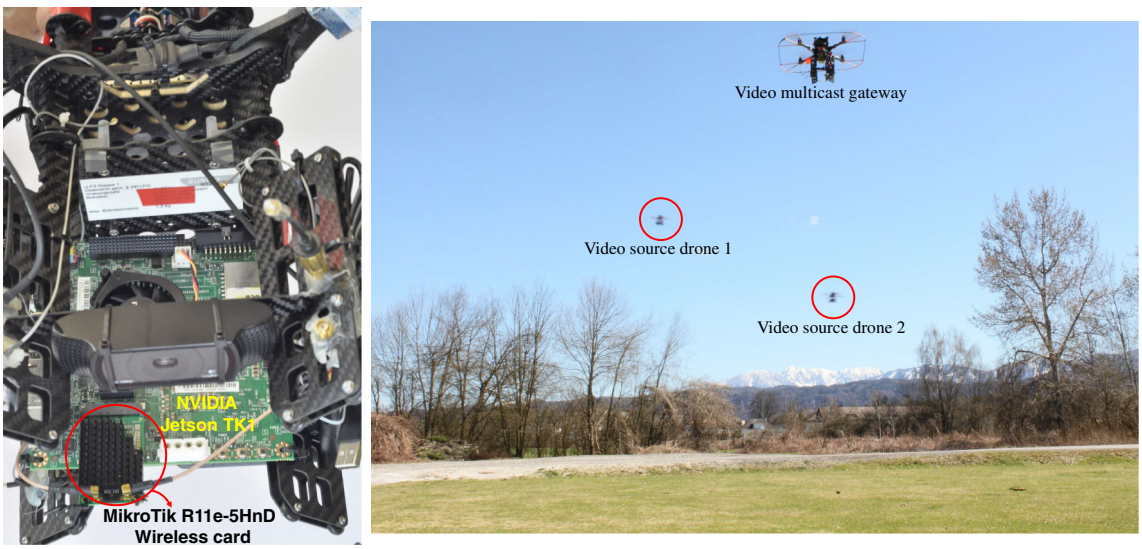
In the point-to-multipoint scenario, live video is streamed from a source drone to multiple ground receivers as multicast. As an extension to Muzaffar et al. (2016b), in this paper, we present an experimental evaluation through an aerial testbed setup. Thus, the point-to-multipoint configuration is a single-hop video streaming i.e. from a drone to multiple receivers, and is without the VMG as in Fig. 3. Members of the multicast group are assigned roles as *designated nodes* based on their signal quality from the source drone instead of the VMG as presented in Sect. 3.1. *Designated nodes* are made responsible to provide feedback for packet reception as presented in Sect. 3.2 and the *ELF rates* are adapted as described in Sect. 3.3. Multicast group members can initiate the *deregistration process* as stated in Sect. 3.4.

4 Testbed

We use MikroTik R11e-5HnD high-power wireless cards to build an 802.11a ad-hoc network among drones and ground receivers. These cards support 802.11a/n with a maximum transmission power of 27 dBm at 6 Mbit/s with a receiver sensitivity of -96 dBm and a maximum transmission power of 24 dBm at 54 Mbit/s with a receiver sensitivity of -76 dBm (MikroTik https://www.mikrotik-store.eu/de/mikrotik-r11e-5hnd). Due to limited flight time and to access our rate-adaptive approach for the given scenarios, we configured the modules with a reduced transmission power of 14 dBm for all experiments. Logitech C920 cameras capture videos up to full HD 1080p video quality at 30 frames/s with H.264 video compression. Processing, video streaming, and video reception is performed by NVIDIA Jetson TK1 boards (NVIDIA http://www.nvidia.com/object/jetson-tk1-embedded-dev-kit.html), with quad-core 2.3 GHz ARM Cortex-A15 CPU and an energy consumption of $1-5$ W.

We mounted the NVIDIA Jetson TK1 board and the Logitech C920 camera on AscTec Pelican drones (Ascending Technologies http://www.asctec.de/en/uav-uas-drones-rpas-roav/asctec-pelican) (Fig. 4a). The source drones are set to fly at distances from 40 to 140 m from the VMG (Fig. 4b). The receiver nodes remain static and are placed at a distance of 5 m from each other (Fig. 5). The VMG is set to hover at a distance of 40 m from the N_1/P nodes to maintain connectivity between the source and the ground receivers. Similarly, in case of point-to-multipoint scenario, the source drone (D_3) is set to fly at distances from 40 to 140 m from the N_1/P node. All drones fly at an altitude of 50 m.

We evaluate our framework with a drone source and static ground receivers (N_1/P, N_2/S_1, and N_3/B_1) in the point-to-multipoint scenario. In the case of the multipoint-to-point-to-multipoint scenario, two source drones (D_1 and D_2), a VMG (D_3), and four ground receivers are used. Three ground receivers ($N_1/P^{\mathcal{M}_1}$, $N_2/S_1^{\mathcal{M}_1}$, and $N_3/B_1^{\mathcal{M}_1}$) receive video for the multicast group $\mathcal{M}_1$ and one receiver ($N_4/P^{\mathcal{M}_2}$) receives video for $\mathcal{M}_2$. All receivers are approximately one meter above the ground.



(a) **(b)**

Fig. 4 **a** NVIDIA Jetson TK1 board and Logitech C920 camera mounted over the AscTec Pelican drone, **b** two video source drones flying away from the hovering video multicast gateway drone in the multipoint-to-point-to-multipoint scenario

We analyze the effect of mobility from the high-speed source drones on the received video stream quality. We manually configure the node closest to the source drone as P, the second closest as S_1, and the farthest as B_1. We compare our rate-adaptive approach with the fixed transmission rate of 6 Mbit/s and constant video encoding rate of 128 kbit/s and 256 kbit/s. Our rate-adaptive approach can adjust the transmission rate from 54 to 6 Mbit/s, video encoding rate from 512 up to 8192 kbit/s as an upper bound and down to 128 kbit/s as a lower bound, and frame rate from 25 to 10 frames/s based on the received feedback from the designated nodes.

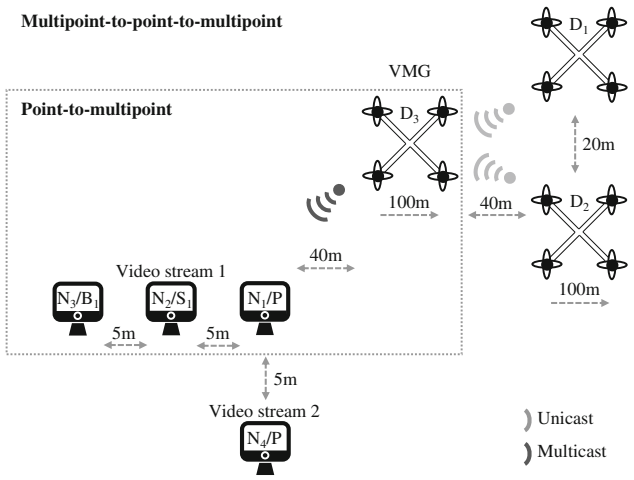


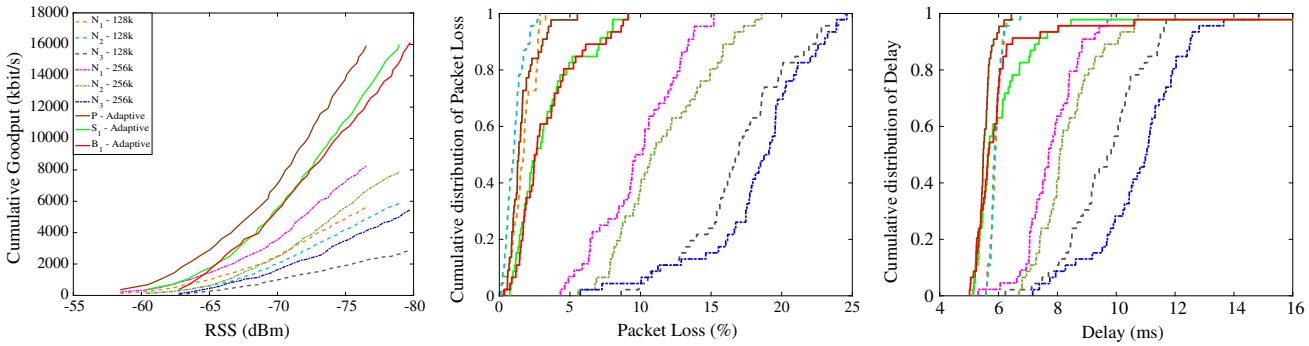
Fig. 5 Experimental testbed setup for point-to-multipoint and multipoint-to-point-to-multipoint scenarios. In the point-to-multipoint scenario, three ground receiver nodes N_1/P, N_2/S_1, and N_3/B_1 are statically placed 5 m apart from each other. The source drone, D_3 is set to fly at distances from 40 m to 140 m from N_1/P. In the multipoint-to-point-to-multipoint scenario, source drones D_1 and D_2 stream live videos to D_3 as unicast. D_3 acts as the video multicast gateway (VMG) and multicasts the received videos to the ground receivers. VMG hovers at a distance of 40 m from the receiver N_1/P while D_1 and D_2 are set to fly at distances from 40 m to 140 m from D_3/VMG. All drones fly at 50 m above the ground level

5 Performance evaluation

We evaluate the performance of ELF rate adaptation in terms of goodput, packet loss, and delay. Video quality is evaluated in terms of PSNR, no-reference perceptual blur metric (Crete et al. 2007), and video multi-method assessment fusion (VMAF) developed by Netflix (https://github.com/Netflix/vmaf).

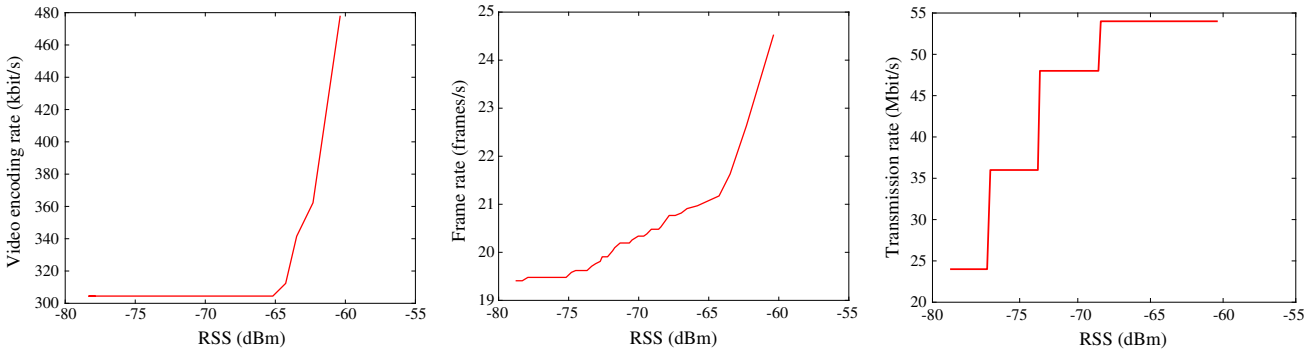
5.1 Point-to-multipoint

Figure 6 shows the mean values from five experimental runs for the point-to-multipoint scenario. The *cumulative goodput* (Fig. 6a) is calculated by adding the received bytes of the video stream as the drone flies away from the receivers. The received signal strength, RSS decibel-milliwatts (dBm), decreases as the distance between the source and the receivers increase. We observe improved goodput for the rate-adaptive approach compared to the fixed transmission rate due to the adaptive and higher video encoding rate. The goodput for fixed transmission rate remains lower due to a lower video encoding rate. It can be observed that receiver N_1 in the case of 128 kbit/s encoding rate and N_3 in the case of 256 kbit/s encoding rate undergo the highest packet loss (Fig. 6b) resulting in the lowest cumulative goodput among all other receivers. Due to a high packet loss, the delay between packet reception for N_1 in the case of 128 kbit/s encoding rate and N_3 in the case of 256 kbit/s encoding rate is also large (Fig. 6c). Our rate-adaptive approach outperforms the fixed transmission schemes in *fairness* since the packet loss and delay for all the three receivers remain low: in 90% of the cases the packet loss for the rate-adaptive receivers P, S_1, and B_1 remains under 5%. Thus, the effect of retransmission of lost packets for which the feedback is provided by the designated receivers, is also evident on B_1. The packet loss and delay for B_1 are similar to S_1, although B_1 only acts as a best effort node and is the farthest from the source.



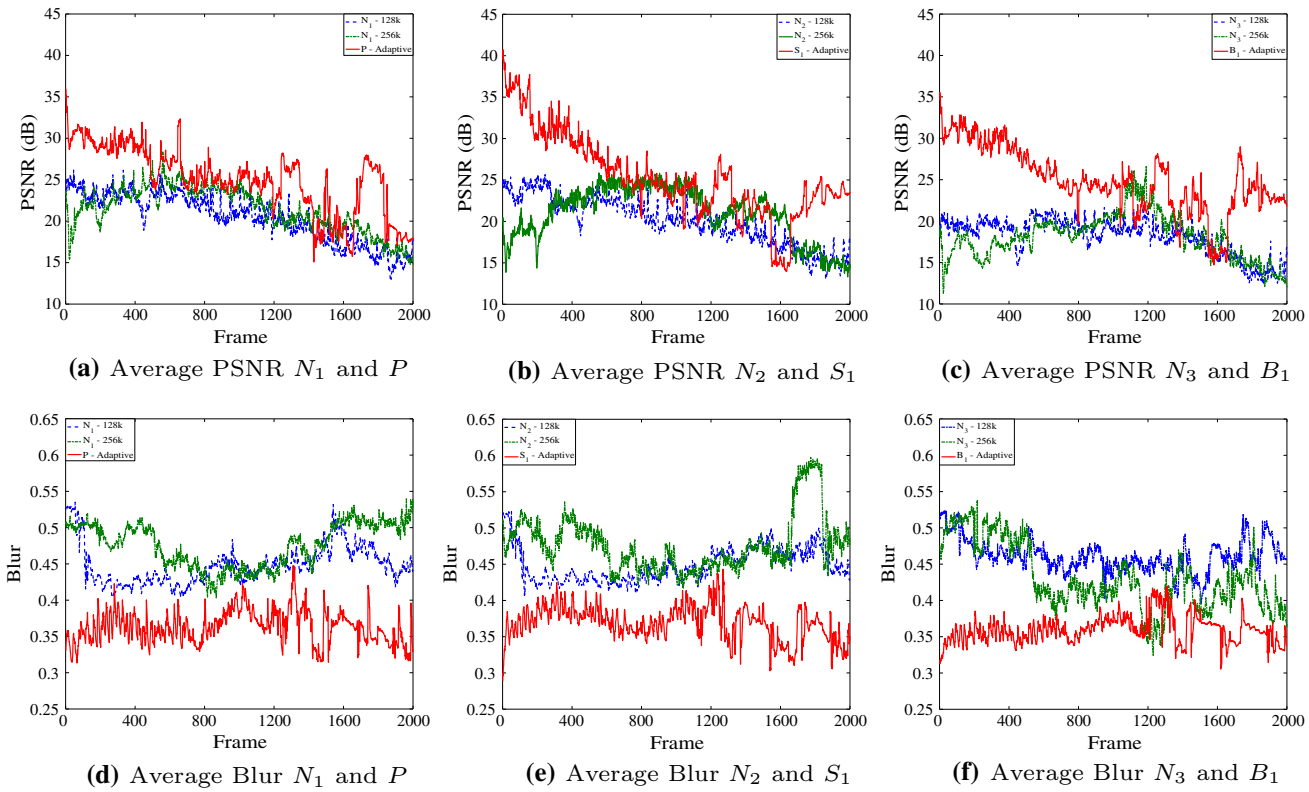
(a) Goodput **(b)** Packet loss **(c)** Delay

Fig. 6 Performance of multicast frame transmission as the source drone flies away from the receivers



(a) Video encoding rate **(b)** Frame rate **(c)** Transmission rate

Fig. 7 Rate adaptation based on the feedback from the designated nodes with decreasing signal strength



(a) Average PSNR N_1 and P **(b)** Average PSNR N_2 and S_1 **(c)** Average PSNR N_3 and B_1

(d) Average Blur N_1 and P **(e)** Average Blur N_2 and S_1 **(f)** Average Blur N_3 and B_1

Fig. 8 Average video quality in PSNR and Blur metrics for point-multipoint scenario. The results are averaged over five experiments

The performance in terms of delay and packet loss when the video is streamed at 256 kbit/s encoding rate with 6 Mbit/s transmission rate is inferior to our rate-adaptive approach due to the lack of feedback and retransmission mechanism, and consequently the lack of a rate adaptation mechanism. Based on the feedback received from the designated nodes, the source adapts the video encoding rate from 478 kbit/s to 302 kbit/s (Fig. 7a), frame rate from 25 frames/s to 19 frames/s (Fig. 7b), and the transmission rate from 54 Mbit/s to 24 Mbit/s (Fig. 7c) as the source drone moves away from the receivers, to address the challenges of performance and fairness. Because of this dynamic rate adaptation, on average, 65% improvement in goodput over 128 kbit/s encoding rate and 50% improvement over 256 kbit/s encoding rate, is observed. Similarly, using adaptive transmission rate, on average, resulted in 20% and 50% improvement in delay over 128 kbit/s encoding rate and 256 kbit/s encoding rate, respectively.

We also evaluated the video quality using the PSNR and blur metrics (Fig. 8). Typically, the average PSNR value for a frame transmitted wirelessly ranges from 20 to 45 dB, and the video quality is considered excellent when the average PSNR is above 37 dB (Gross et al. 2004). The blur metric

is a no-reference evaluation method that detects loss of spatial details by predicting blur annoyance (Crete et al. 2007). The blur metric ranges between 0 (sharp) and 1 (blurry). Furthermore, we evaluated video quality with VMAF, a perceptual assessment algorithm that is considered to be the closest to subjective visual perception. VMAF is a full-reference metric that provides values between 0 (worst) and 100 (best). However, experimental evaluations using legacy multicast service shows that a VMAF score well below 50 is observed for a number of static stations even though the packet reception rates are well above 99% (Gringoli et al. 2018).

In all our experiments, the video transmission starts before the drone takes off and stops after landing. We consider frames from 501 till 2500 to remove the effect of takeoff and only consider frames received until the drone reaches its maximum distance of 140 m. Since the takeoff and flight times vary, the comparison on video quality may not be in perfect synchrony. The PSNR and blur metrics are calculated as a mean of five experimental runs for 2000 frames. All frames are numbered by the source, thus it is possible to detect missing frames. A missing frame is compared with the last received frame to maintain comparison position of the transmitted and received frames. The PSNR traces of the received video stream for nodes N_1 and P, N_2 and S_1, and N_3 and B_1 are presented in Fig. 8a–c, respectively. A similarity in the PSNR traces for nodes P, S_1, and B_1 can be observed. This is due to the fact that the rate adaptation is mainly triggered by the feedback from P and then from S_1. The RSS fluctuates heavily due to high mobility of the drone and increasing distances from the receivers. This causes transmission errors requiring retransmission and adaptation. The rate adaptation needs some time to react to the changing conditions. Hence the video quality degrades before the adaptation aids the improvement of the PSNR, especially between frames 1200 and 2000. In other cases, for receivers N_1, N_2, and N_3, we observe a decreasing or stable trend in the PSNR traces caused by packet loss and lack of adaptation and retransmission mechanism. Compared to legacy multicast, the proposed rate-adaptive approach on average shows an improvement of 30% on video quality for all receivers.

The blur metric traces of the received video stream for nodes N_1 and P, N_2 and S_1, and N_3 and B_1 are presented in Fig. 8d–f, respectively: the video received using the adaptive approach is, on average, 30% sharper compared to the legacy multicast approach at all receivers.

Table 2 presents the VMAF score for N_1 and P, N_2 and S_1, and N_3 and B_1 as a mean of video streams received with five experimental runs. The VMAF score using the adaptive approach, on average, shows 30% improvement compared to the legacy multicast approach at all receivers.

Table 2 VMAF performance for point-to-multipoint scenario

Scheme	N_1/P	N_2/S_1	N_3/B_1
128	17.7	18.2	18.3
256	15.3	16.2	17.8
Adaptive	25.8	25.9	25.0

The results are averaged over five experiments

5.2 Multipoint-to-point-to-multipoint

When we analyze the results of the multipoint-to-point-to-multipoint communications, there is no direct communication between the video sources (drones D_1 and D_2) with the ground receivers (N_1/P^{M_1}, $N_2/S_1^{M_1}$, $N_3/B_1^{M_1}$, and N_4/P^{M_2}). D_1 and D_2 send their video streams ($\mathcal{M}_1$ and $\mathcal{M}_2$) to the VMG drone (D_3).

We separately compare the unicast results from the sources D_1 and D_2 to the VMG (D_3) and the multicast results from VMG (D_3) to the ground receivers. Figure 9 shows the mean values from five experimental runs. We observe an improved goodput at VMG for the two video streams ($\mathcal{M}_1$ and $\mathcal{M}_2$) with the rate-adaptive approach, compared to the fixed transmission rate. This is due to the higher encoding rate, retransmission of lost packets, and rate adaptation (Fig. 9a).

Similarly, improved goodput is observed at the receivers with the rate-adaptive approach (Fig. 9b). On average, 80% improvement in goodput over 128 kbit/s encoding rate and 40% improvement over 256 kbit/s encoding rate is observed. An improved performance of our rate-adaptive approach is also noticeable from the comparison presented for packet loss (Fig. 9c, d) and delay (Fig. 9e, f). However, comparing the results from the point-to-multipoint scenario, we can observe that the goodput of the multipoint-to-point-to-multipoint scenario is considerably lower than the goodput of the point-to-multipoint scenario, although the received signal quality at the VMG (D_3) is higher than the signal quality at the receivers of the point-to-multipoint scenario. The signal quality is better due to the aerial communication between the video source drones and the VMG, where ground reflections are not present. Nevertheless, the goodput is lower at the VMG, since it serves as a gateway for retransmitting the received packets to the ground receivers and hence has to act as a receiver and transmitter at the same time which reduces its reception time. Thus, this dual role of the VMG also creates a bottleneck for the ground receivers, resulting in higher delay and packet loss, which decrease the overall goodput.

The adaptation of the video encoding, frame, and transmission rates of the video streams ($\mathcal{M}_1$ and $\mathcal{M}_2$) at the VMG are presented in Fig. 10a–c, respectively. $\mathcal{M}_1$ adapts the video encoding rate from 396 to 166 kbit/s, frame rate

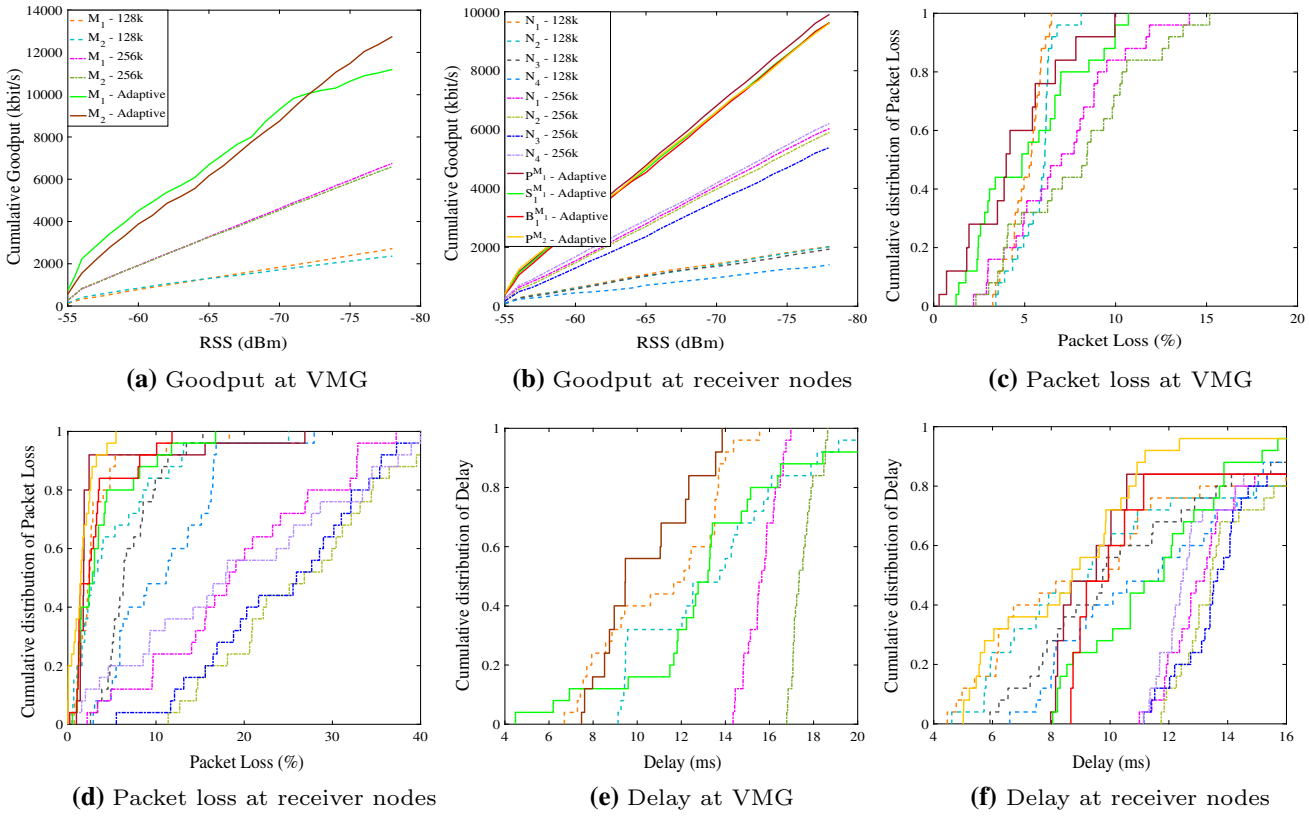


Fig. 9 Performance of multicast frame transmission as the source drones fly away from the video multicast gateway (VMG). The VMG sends the received video streams to the ground receivers

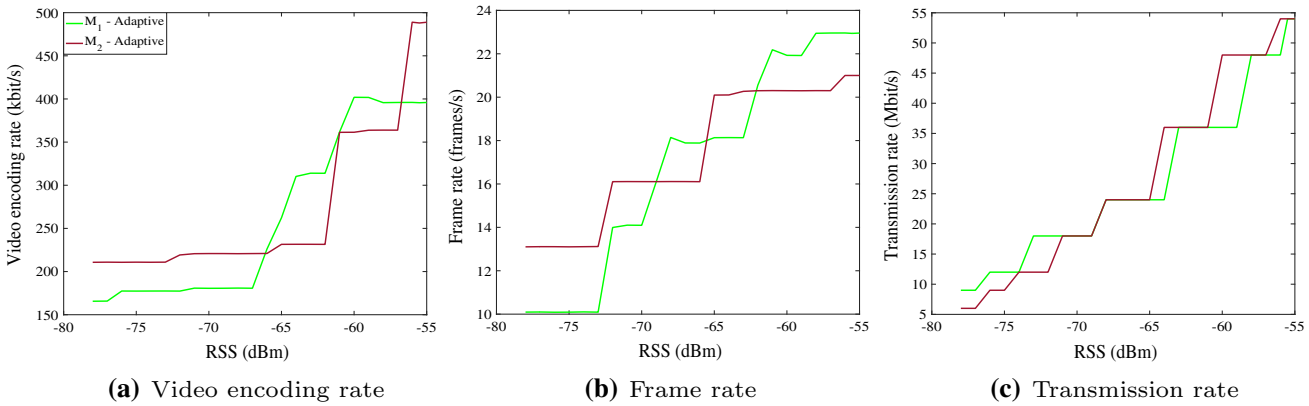
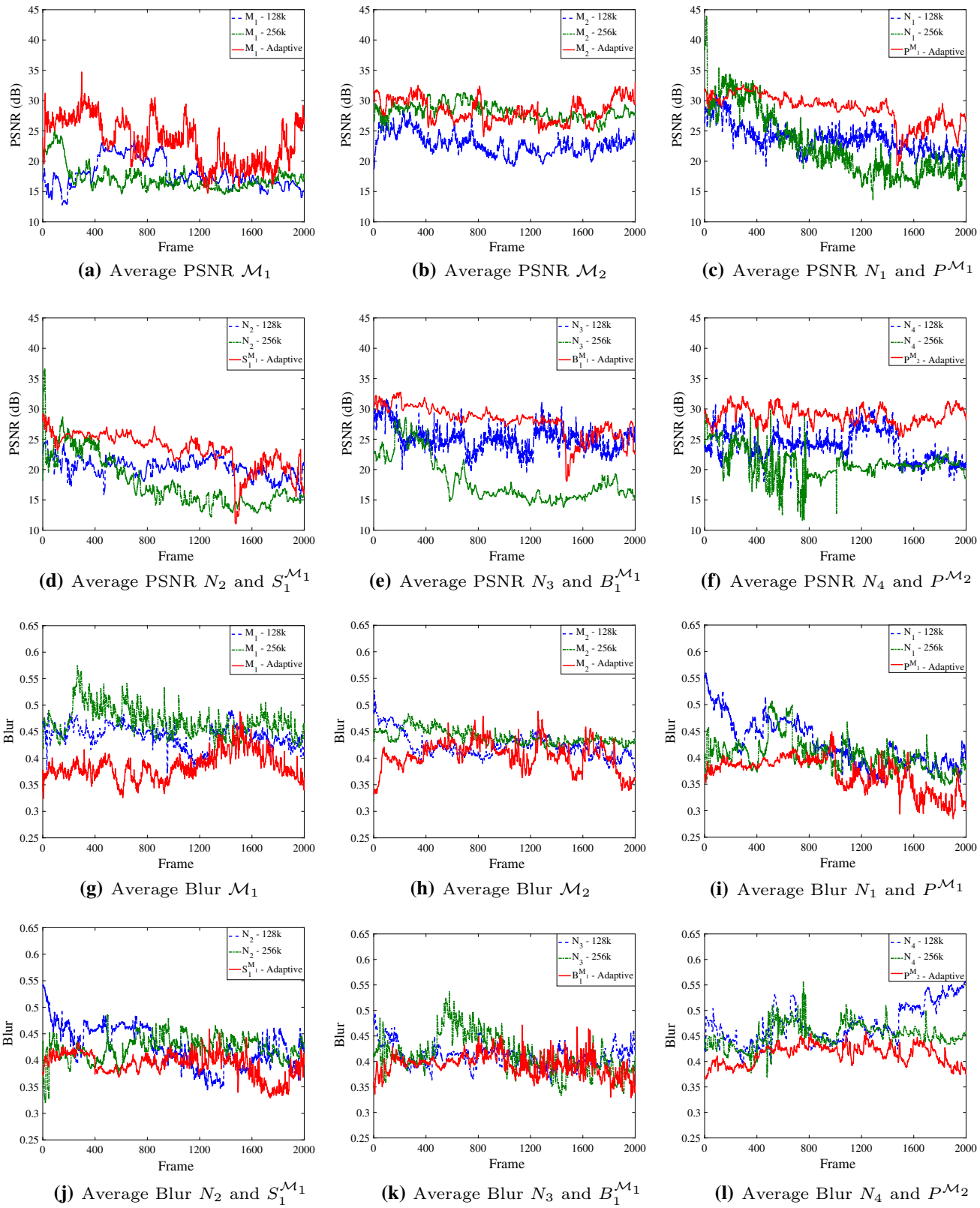


Fig. 10 Rate adaptation based on the feedback from the video multicast gateway with decreasing signal strength

from 23 to 10 frames/s, and the transmission rate from 54 to 9 Mbit/s, while $\mathcal{M}_2$ adapts the video encoding rate from 488 to 210 kbit/s, frame rate from 21 to 13 frames/s, and the transmission rate from 54 to 6 Mbit/s. At a particular time instant, the VMG selects the lower transmission rate required by $\mathcal{M}_1$ or $\mathcal{M}_2$. The rate is adapted based on the feedback from the VMG to the source drones as D_1 and D_2 move away from D_3 and the signal strength decreases. At the ground receivers, the transmission rate remains at the maximum since the VMG hovers at a distance of 40 m and the signal strength almost remains constant. The video encoding rate and the frame rate is adapted similarly to the rate adapted by the VMG.

Figure 11 shows the video quality for the multipoint-to-point-to-multipoint scenario. The PSNR traces of the received videos for $\mathcal{M}_1$, $\mathcal{M}_2$, N_1 and P^{M_1}, N_2 and $S_1^{M_1}$, N_3 and $B_1^{M_1}$, and N_4 and P^{M_2} are presented in Fig. 11a–f, respectively. We observe that, despite using higher and adaptive video encoding rate, the received video quality is better than legacy multicast, because of fewer packet losses, retransmission, and rate adaptation. Compared to legacy mul-



(a) Average PSNR $\mathcal{M}_1$ **(b)** Average PSNR $\mathcal{M}_2$ **(c)** Average PSNR N_1 and $P^{\mathcal{M}_1}$

(d) Average PSNR N_2 and $S_1^{\mathcal{M}_1}$ **(e)** Average PSNR N_3 and $B_1^{\mathcal{M}_1}$ **(f)** Average PSNR N_4 and $P^{\mathcal{M}_2}$

(g) Average Blur $\mathcal{M}_1$ **(h)** Average Blur $\mathcal{M}_2$ **(i)** Average Blur N_1 and $P^{\mathcal{M}_1}$

(j) Average Blur N_2 and $S_1^{\mathcal{M}_1}$ **(k)** Average Blur N_3 and $B_1^{\mathcal{M}_1}$ **(l)** Average Blur N_4 and $P^{\mathcal{M}_2}$

Fig. 11 Average video quality in PSNR and Blur metrics for multipoint-to-point-multipoint scenario. The results are averaged over five experiments

Table 3 VMAF performance for multipoint-to-point-to-multipoint scenario

Scheme	$\mathcal{M}_1$	$\mathcal{M}_2$	N_1/P^{M_1}	$N_2/S_1^{M_1}$	$N_3/B_1^{M_1}$	N_4/P^{M_2}
128	17.0	16.9	19.0	20.4	19.7	18.5
256	17.3	22.6	10.9	10.3	7.9	10.4
Adaptive	34.9	37.3	33.3	32.0	31.1	35.9

The results are averaged over five experiments

ticast, the ELF rate-adaptive approach shows, on average, an improvement of 16% over 128 kbit/s encoding rate and 22% improvement over 256 kbit/s encoding rate on the video quality for all receivers.

The blur metric traces for $\mathcal{M}_1$, $\mathcal{M}_2$, N_1 and P^{M_1}, N_2 and $S_1^{M_1}$, N_3 and $B_1^{M_1}$, and N_4 and P^{M_2} are presented in Fig. 11g–l, respectively. Similar to the PSNR results, the blur metric traces show that using the rate-adaptive approach the received video quality remains sharper, with an improvement of 10% compared to the legacy multicast approach.

Table 3 presents the VMAF score for $\mathcal{M}_1$, $\mathcal{M}_2$, N_1 and P^{M_1}, N_2 and $S_1^{M_1}$, N_3 and $B_1^{M_1}$, and N_4 and P^{M_2} as a mean of video streams received with five experimental runs. The VMAF score using the adaptive approach, on average, shows 50% improvement compared to the legacy multicast approach at all receivers.

6 Conclusion

We validated the first application-layer ELF rate-adaptive multicast video streaming framework on an aerial testbed that addresses multicast challenges of reliability, performance, and fairness and video streaming challenges of delay and packet loss. The proposed approach is suitable for all 802.11 variants that support ad-hoc networks and, unlike existing rate-adaptive multicast approaches, does not require any modifications of the MAC layer.

The experimental evaluations are based on single-hop (point-to-multipoint) and two-hop (multipoint-to-point-to-multipoint) scenarios over multiple experimental runs. Roles are assigned to multicast group members as designated nodes to provide feedback on packet reception, and can be switched between members based on their signal quality to cater for mobility. The ELF rate-adaptive approach reacts upon receiving feedback from designated nodes by adapting the link transmission, video encoding, and frame rates. Reliability is achieved by retransmissions of lost packets, resulting in fewer packet losses. Performance and fairness are achieved through ELF rate adaptation that leads to a superior video quality.

In the point-to-multipoint scenario, more than 50% gain in goodput is observed that resulted in 30% gain in the PSNR of the received videos. Similarly, in the multipoint-to-point-to-multipoint scenario more than 35% gain in goodput is observed that resulted in 20% gain in the PSNR of the received videos.

Acknowledgements The authors would like to thank Arke Vogell and Micha Rappaport for acting as pilots for the test scenarios.

References

Ahmed, N., Kanhere, S.S., & Jha, S. (2011). Link characterization for aerial wireless sensor networks. In *Proceedings of IEEE Global Communications Conference (GLOBECOM)*.

Andre, T., Hummel, K. A., Schoellig, A. P., Yanmaz, E., Asadpour, M., Bettstetter, C., et al. (2014). Application-driven design of aerial communication networks. *IEEE Communications Magazine*, *52*(5), 129–137.

Ascending Technologies (n.d.) AscTec Pelican. http://www.asctec.de/en/uav-uas-drones-rpas-roav/asctec-pelican, last accessed Aug 2018

Banchs, A., de la Oliva, A., Eznarriaga, L., Kowalski, D. R., & Serrano, P. (2014). Performance analysis and algorithm selection for reliable multicast in IEEE 802.11aa wireless LAN. *IEEE Transactions on Vehicular Technology*, *63*(8), 3875–3891.

Bekmezci, I., Sahingoz, O. K., & Temel, S. (2013). Flying ad-hoc networks (FANETs): A survey. *Elsevier Ad-hoc Networks*, *11*(3), 1254–1270.

Biaz, S., & Wu, S. (2008). Rate adaptation algorithms for IEEE 802.11 networks: A survey and comparison. In *Proceedings of IEEE Symposium on Computers and Communications (ISCC)*.

Chandra, R., Karanth, S., Moscibroda, T., Navda, V., Padhye, J., Ramjee, R., & Ravindranath, L. (2009). Dircast: A practical and efficient Wi-Fi multicast system. In *Proceedings of IEEE International Conference on Network Protocols (ICNP)*.

Chmaj, G., & Selvaraj, H. (2015). Distributed processing applications for UAV/drones: A survey. In *Proceedings of Springer International Conference on Systems Engineering (ICSENG)*.

Choi, S., Choi, N., Seok, Y., Kwon, T., & Choi, Y. (2007). Leader-based rate adaptive multicasting for wireless LANs. In *Proceedings of IEEE Global Communications Conference (GLOBECOM)*.

Crete, F., Dolmiere, T., Ladret, P., & Nicolas, M. (2007). The blur effect: perception and estimation with a new no-reference perceptual blur metric. In *Proceedings of SPIE Human Vision and Electronic Imaging (HVEI)*.

Crow, B. P., Widjaja, I., Kim, J. G., & Sakai, P. T. (1997). IEEE 802.11 wireless local area networks. *IEEE Communications Magazine*, *35*(9), 116–126.

Dujovne, D., & Turletti, T. (2006). Multicast in 802.11 WLANs: An experimental study. In *Proceedings of ACM Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*.

Fu, X., Ma, W., & Zhang, Q. (2007). The IEEE 802.16 and 802.11 a coexistence in the license-exempt band. In *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC)*.

Ge, P., & McKinley, P.K. (2002). Comparisons of error control techniques for wireless video multicasting. In *Proceedings of*

IEEE International Performance, Computing, and Communications Conference (IPCCC).

Gringoli, F., Serrano, P., Ucar, I., Facchi, N., & Azcorra, A. (2018). Experimental QoE evaluation of multicast video delivery over IEEE 802.11aa WLANs. *IEEE Transactions on Mobile Computing*. https://doi.org/10.1109/TMC.2018.2876000.

Gross, J., Klaue, J., Karl, H., & Wolisz, A. (2004). Cross-layer optimization of OFDM transmission systems for MPEG-4 video streaming. *Elsevier Computer communications*, *27*(11), 1044–1055.

Gupte, S., Mohandas, P.I.T., & Conrad, J.M. (2012). A survey of quadrotor unmanned aerial vehicles. In *Proceedings of IEEE Southeastcon*.

Hanscom, A.F.B., & Bedford, M. (2013). Unmanned aircraft system (UAS) service demand 2015-2035: Literature review and projections of future usage. Tech. rep., U.S. department of transportation, John A. Volpe national transportation systems center, https://fas.org/irp/program/collect/service.pdf, last accessed Aug 2018.

Hayat, S., Yanmaz, E., & Muzaffar, R. (2016). Survey on unmanned aerial vehicle networks for civil applications: A communications viewpoint. *IEEE Communications Surveys & Tutorials*, *18*(4), 2624–2661.

Kacianka, S., & Hellwagner, H. (2015). Adaptive video streaming for UAV networks. In *Proceedings of ACM International Workshop on Mobile Video (MoVid)*.

Kamerman, A., & Monteban, L. (1997). WaveLAN-II: A high-performance wireless LAN for the unlicensed band. *Bell Labs Technical Journal*, *2*(3), 118–133.

Kofler, I., Kuschnig, R., & Hellwagner, H. (2011). In-network adaptation of H.264/SVC for HD video streaming over 802.11g networks. In *Proceedings of International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV)*.

Kuri, J., & Kasera, S.K. (1999). Reliable multicast in multi-access wireless LANs. In *Proceedings of IEEE Conference on Computer Communications (INFOCOM*.

Li, Z., & Herfet, T. (2008). Beacon-driven leader based protocol over a GE channel for MAC layer multicast error control. *International Journal of Communications, Network and System Sciences*, *1*(2), 144–153.

Li, Z., & Herfet, T. (2009). MAC layer multicast error control for IPTV in wireless LANs. *IEEE Transactions on Broadcasting*, *55*(2), 353–362.

Lindeberg, M., Kristiansen, S., Plagemann, T., & Goebel, V. (2011). Challenges and techniques for video streaming over mobile ad-hoc networks. *Springer Multimedia Systems*, *17*(1), 51–82.

LinuxWireless Project (n.d.) Minstrel Specification. https://wireless.wiki.kernel.org/en/developers/do-cumentation/mac80211/ratecontrol/minstrel, last accessed Aug 2018.

Maraslis, K., Chatzimisios, P., & Boucouvalas, A. (2012). IEEE 802.11aa: Improvements on video transmission over wireless LANs. In *Proceedings of IEEE International Conference on Communications (ICC)*.

MikroTik (n.d.) MikroTik - R11e-5HnD. https://www.mikrotik-store.eu/de/mikrotik-r11e-5hnd, last accessed Aug 2018.

Muzaffar, R., Vukadinovic, V., & Cavallaro, A. (2016a). Rate-adaptive multicast video streaming from teams of micro aerial vehicles. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*.

Muzaffar, R., Yanmaz, E., Bettstetter, C., & Cavallaro, A. (2016b). Application-layer rate-adaptive multicast video streaming over 802.11 for mobile devices. In *Proceedings of ACM Multimedia (ACMMM)*.

Nafaa, A. (2007). Provisioning of multimedia services in 802.11-based networks: Facts and challenges. *IEEE Wireless Communications*, *14*(5), 106–112.

Netflix (n.d.) Video multi-method assessment fusion. https://github.com/Netflix/vmaf, last accessed Aug 2018.

NVIDIA (n.d.) NVIDIA Jetson TK1 developer kit. http://www.nvidia.com/object/jetson-tk1-embedded-dev-kit.html, last accessed Aug 2018.

Paris, S., Facchi, N., Gringoli, F., & Capone, A. (2013). An innovative rate adaptation algorithm for multicast transmissions in wireless LANs. In *Proceedings of IEEE Vehicular Technology Conference (VTC)*.

Park, Y., Seok, Y., Choi, N., Choi, Y., & Bonnin, J.M. (2006). Rate-adaptive multimedia multicasting over IEEE 802.11 wireless LANs. In *Proceedings of IEEE Consumer Communications and Networking Conference (CCNC)*.

Piamrat, K., Ksentini, A., Bonnin, J.M., & Viho, C. (2009). Q-DRAM: QoE-based dynamic rate adaptation mechanism for multicast in wireless networks. In *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM)*.

Quaritsch, M., Stojanovski, E., Bettstetter, C., Friedrich, G., Hellwagner, H., Rinner, B., Hofbaur, M., & Shah, M. (2008). Collaborative microdrones: Applications and research challenges. In *Proceedings of ACM Autonomics*.

Salvador, P., Cominardi, L., Gringoli, F., & Serrano, P. (2013). A first implementation and evaluation of the IEEE 802.11aa group addressed transmission service. *ACM SIGCOMM Computer Communication Review*, *44*(1), 35–41.

Su, G. M., Su, X., Bai, Y., Wang, M., Vasilakos, A. V., & Wang, H. (2016). QoE in video streaming over wireless networks: Perspectives and research challenges. *Springer Wireless Networks*, *22*(5), 1571–1593.

Takai, M., Martin, J., & Bagrodia, R. (2001). Effects of wireless physical layer modeling in mobile ad-hoc networks. In *Proceedings of ACM Mobile Ad-hoc Networking and Computing (MobiHoc)*.

Team G (n.d.) Gstreamer framework. https://gstreamer.freedesktop.org/, last accessed Aug 2018.

Thierry, T., & Yongho, S. (2006). Practical rate-adaptive multicast schemes for multimedia over IEEE 802.11 WLANs. inria research report rr-5993. https://hal.inria.fr/inria-00104699, last accessed Aug 2018.

Tourrilhes, J. (1998). Robust broadcast: Improving the reliability of broadcast transmissions on CSMA/CA. In *Proceedings of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*.

Vella, J. M., & Zammit, S. (2013). A survey of multicasting over wireless access networks. *IEEE Communications Surveys & Tutorials*, *15*(2), 718–753.

Yanmaz, E., Kuschnig, R., & Bettstetter, C. (2011). Channel measurements over 802.11a-based UAV-to-ground links. In *IEEE Global Telecommunications Conference (GLOBECOM Wkshps)*.

Yanmaz, E., Kuschnig, R., & Bettstetter, C. (2013). Achieving air-ground communications in 802.11 networks with three-dimensional aerial mobility. In *Proceedings of IEEE Conference on Computer Communications (INFOCOM)*.

Yin, W., Hu, P., Indulska, J., & Bialkowski, K. (2011). Performance of mac80211 rate control mechanisms. In *Proceedings of ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*.

Zhu, H., Li, M., Chlamtac, I., & Prabhakaran, B. (2004). A survey of quality of service in IEEE 802.11 networks. *IEEE Wireless Communications*, *11*(4), 6–14.

Raheeb Muzaffar received the bachelor's and master's degrees in computer science from Bahria University, Pakistan, in 2001 and 2002, respectively, and the M.S. degree in information technology from National University of Sciences and Technology, Pakistan, in 2006. He worked for six years with various technical organizations post his MS degree. From 2012 to 2016, he worked as researcher with Alpen-Adria-Universität Klagenfurt, Austria, and Queen Mary University of London, U.K., under the Erasmus Mundus Joint Doctorate Program and received his doctorate in 2016. Since 2016, has been working with Lakeside Labs GmbH, Klagenfurt Austria. His research interests include wireless multimedia communication for drones, wireless communication and networked systems, and analysis of mission oriented drone applications.

Evşen Yanmaz received the B.S. degree in electrical and electronics engineering from Bogazici University, in 2000, the M.S. degree in electrical engineering from SUNY at Buffalo, in 2002, and the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, in 2005. She is an Assistant Professor in Ozyegin University, Istanbul, Turkey. She was a Senior Researcher and a Project Leader with Lakeside Labs, Klagenfurt, Austria, was a Senior Researcher with the NES Institute, University of Klagenfurt, a Post-Doctoral Researcher with Los Alamos National Laboratory, a Researcher with Carnegie Mellon University, and an Intern with Nokia Research Center, USA. She has published over 50 peer-reviewed works focusing on the fields of telecommunications, networking, self-organization, and path planning. Her research interests include self-organizing networks, cooperative networks, and coordination of airborne and ground sensor networks.

Christian Raffelsberger received the M.Sc. and the Ph.D. degrees in computer science from the Alpen-Adria-Universität Klagenfurt, Austria, in 2010 and 2015, respectively. From 2010 to 2015 he was part of the Multimedia Communications group at AAU Klagenfurt, where he worked as project assistant in several international research projects. Since 2015 has been a senior researcher at the Lakeside Labs GmbH, Klagenfurt, Austria. His research interests include wireless and delay-/disruption-tolerant networking, and multimedia communication.

Christian Bettstetter received a Dipl.-Ing. degree in 1998 and Dr.-Ing. degree (summa cum laude) in 2004, both in electrical and information engineering from Technische Universität München (TUM), Munich, Germany. He was a research and teaching staff member at the Institute of Communication Networks, TUM, until 2003. From 2003 to 2005, he was a senior researcher with DOCOMO Euro-Labs. He has been a professor at the University of Klagenfurt, Austria, since 2005, and founding director of the Institute of Networked and Embedded Systems since 2007. He is also the founding scientific director of Lakeside Labs, a research company on self-organizing networked systems.

Andrea Cavallaro is Professor and the Director of the Centre for Intelligent Sensing at Queen Mary University of London, UK. He is also a Turing Fellow at the Alan Turing Institute, the UK National Institute for Data Science and Artificial Intelligence. He received his Ph.D. in Electrical Engineering from the Swiss Federal Institute of Technology (EPFL), Lausanne, in 2002. He was a Research Fellow with British Telecommunications in 2004/2005. Prof. Cavallaro is Senior Area Editor for the IEEE Transactions on Image Processing; and Associate Editor for the IEEE Transactions on Circuits and Systems for Video Technology and IEEE Multimedia. He is vice chair of the IEEE Signal Processing Society, Image, Video, and Multidimensional Signal Processing Technical Committee and an elected member of the IEEE Video Signal Processing and Communication Technical Committee.