Wireless Networks (2021) 27:2559–2575
https://doi.org/10.1007/s11276-021-02606-1

Real-time video frame differentiation in multihomed VANETs

Rui Lopes[1] · Miguel Luís[2,3] · Susana Sargento[1]

Accepted: 15 March 2021 / Published online: 27 March 2021

Abstract

Providing high quality video transmission in VANETs is very challenging due to the highly dynamic, unpredictable topology, and low bandwidth characteristics. In this article, we design a system able to optimally transmit RTP video streams in an IP-based multihomed VANET. By splitting the video through its different frame types in the array of multihomed paths from the center node of the network to the clients, the system can then send critical frames, depending on the used coding standard, through a more reliable path, in order to improve the video performance even if the quality is degraded due to bad signal reception. Two different content-based multihomed video distribution schemes have been proposed: linear selection, where the system is able to select more reliable paths for higher prioritized segments; and adaptive selection, where the paths are sensed to assess their congestion level in a real-time approach. Through real deployment of these approaches in a real vehicular scenario, with mobility, handovers and multihoming, the proposed approaches achieved a strong decrease in the loss percentage, with a maximum of approximately 60.4%, greatly improving the video quality while on the move.

Keywords IP-based mobility · Multihoming · Frame type differentiation · Video transmission

1 Introduction

Content distribution is an area of fundamental importance in days to come, with a significant increase in the number of content access and share in the Internet. Lately, most of these shared contents are made via video, some of them transmitting live.

With the deployment of smart cities, vehicular ad-hoc networks (VANETs) then enter into the picture, where spontaneous creation of wireless networks for data exchange are made on the vehicles' domain. Here, users can send and receive content, as they are part of a continuous service provider inserted on an ad-hoc mesh network. Some of this content can be video, especially live video streaming. Nevertheless, video transmissions can also be done as dissemination of a video emergency alert on a car accident happening some kilometers ahead, or an image of the road in front of an obstacle (such as a truck on course).

As it is commonly acknowledged, VANETs are ill-suited to support streaming video traffic: their highly dynamic, unpredictable topology and low bandwidth are the main issues hindering multimedia applications [2]. This poses a problem to the delivery of real-time content, while we try to proceed according to the users' strict criteria of both quality of experience (QoE) and quality of service (QoS).

In this article, we propose an approach that is able to tolerate and guarantee proper behavior of multimedia transport protocols, such as RTP, within a mobility and multihomed network. A running mobility protocol, with multihoming, whilst it tries to take advantage of a more available connection (given the redundancy of connections), it can have a load balancing feature that splits the traffic amongst the array of available connections to points of attachment, with which paths are completed from a

✉ Rui Lopes
ruieduardo.fa.lopes@ua.pt

Miguel Luís
nmal@av.it.pt

Susana Sargento
susana@ua.pt

[1] Instituto de Telecomunicações, Universidade de Aveiro, 3810-193 Aveiro, Portugal

[2] Instituto de Telecomunicações, 3810-193 Aveiro, Portugal

[3] ISEL - Instituto Superior de Engenharia de Lisboa, 1959-007 Lisbon, Portugal

mobility manager (MM) to the clients, independently of the structure or semantics of the transmitted packets. If the transmitted content is video, independently of the applied coding, blindly dividing its stream over multiple paths could lead to undetermined segments of video (or fragments of segments) to be lost, if, for some reason, an active channel does not have or cannot ensure ideal conditions, suddenly dropping packets (or simply accumulating too many packets on its queues).

By analyzing communication channels between nodes and assuming environments where mobility and multihoming are available, with the same and different technologies, we propose to split video transmissions on a real-time paradigm, according to the transmitted type of frames and the characteristics of the available networks and technologies. Two different content-based multihomed video distribution schemes are proposed: linear selection, where the system is able to select more reliable paths for higher prioritized segments; and adaptive selection, where the paths are sensed to asses their congestion level in a real-time approach. Through real deployment of these approaches in a real vehicular scenario, with mobility, handovers and multihoming, the proposed approaches achieved a strong decrease in the loss percentage, with a maximum of approximately 60.4% reduction, greatly improving the video quality while on the move. This approach therefore results in granting successful delivery of MPEG type frames, with the possibility of applying redundancy factors dynamically to each transmission channel.

This article extends the work in [15], by the same authors, by adding an updated and extensive related work, a detailed explanation about the proposed video multihoming strategies and an exhaustive performance evaluation analysis based on real world scenarios. This work positions itself in a category of server-oriented solutions, in comparison with the other related work as described in the following section. Whilst the most of the state-of-the-art reduces the issues with the QoE, this work tries to study approaches to the delivery of specific video frame types by a network perspective, identifying the packets containing video frame fragments (and its respective frame types) and dispatching them to specific roadside units (RSUs) towards an on-board unit (OBU) where the requesting client is waiting for the content to appear. With a first goal of diminishing the loss frame-related packet ratio, we allow clients to perceive the content movement even when video coding errors occur.

In summary, the main contributions of this article are as follows:

- An assessment of the video quality in a mobile and multihoming approach in real vehicular scenarios;
- An architecture to determine the priority of the video frames and assign the preferred technology to the specific quality;
- Two different content-based multihomed video distribution algorithms that are able to select more reliable paths for higher prioritized segments;
- Seamless integration of the approaches with a real mobility- and multihomed-enabled platform;
- Real deployment, test and performance evaluation in a real vehicular scenario.

The remainder of this article is structured as follows. Related work is described in Sect. 2. Section 3 showcases a variety of definitions and use-cases relevant to the understanding of this work. Section 4 introduces our proposed algorithms for a video segment distribution in the mobility network, ending with some topics on its implementation. Section 5 showcases the evaluation results, followed by a discussion. This article is concluded in Sect. 6.

2 Related work

Over the past few years, several solutions have been proposed to study and enable video streaming in wireless networks within multi-homing contexts. These proposed solutions can be organized into different categories according to their chosen abstraction layer.

Starting from the physical layer of the standard communication model, some research groups studied ways to optimize video transmissions by considering factors such as energy consumption while servicing in a multi-homing context, in a heterogeneous wireless access medium. In [11], the authors proposed a framework that allows mobile terminals to determine the transmission power for the used radio interfaces and selectively drop packets under energy constraints. Simultaneously, the most valuable packets are assigned to different radio interfaces, minimizing the video quality distortion. The same authors, in [10], describe an energy management sub-system, allowing mobile terminals to support a sustainable multihoming video transmission in a fading channel. In [9], complementing this work, it is determined the maximum video quality lower bound over the entire video transmission duration, with some statistical guarantees.

More recently, in [1], Agarwal et al. introduced a new paradigm called cognitive multi-homing, where a cognitive radio-enabled base station transmits to users simultaneously over the licensed cellular bands. The aim is to minimize the transmission cost while meeting the users' energy and received video quality constraints, while it is adjusted the sensing duration, the transmission rate over the cognitive radio networks, and the network selection for retransmission of lost packets.

Another perspective takes insights directly from an application abstraction layer, where the quality of experience is estimated through the quality of the received video or ratio on its noise, further adjusting or segmenting the transmission of video frames in the network. In [12], Wu et al. proposed the creation of a framework that deliberately splits large-size video frames into smaller sub-frames. After this action, the sub-frames are then dispatched, individually, onto a specific wireless network to the multi-homed client, moderated by proper scheduling that diminishes the frame-level delay and maintains video quality in terms of peak signal-to-noise ratio (PSNR).

Specifically, with low-delay H.264 video coding, in [22], the authors apply such logic to the large-size I-frames over the lossy channels within a delay constraint, proposing a delay stringent coded transmission (ASCOT) framework, designed to deliver low-delay encoded high-definition streaming while applying coding. Here, as the P-frames are significantly smaller in size than the I-frames, they are protected and scheduled in groups, reducing the frame-level delay, consequently enhancing the streaming video quality.

Despite large-size video frames in specific, most of the solutions for video streaming considering multipath scenarios forward packets via different paths in a content-agnostic way, disallowing effective leverage of scarce wireless resources. In [23], Wu et al. address the creation of a content-aware concurrent multipath transfer, that is, frame-level scheduling based on estimated video parameters and feedback channel status. Deng et al. [6], proposed a quality of experience prediction model to prefer the QoE values of different rate allocation scenarios as an optimization problem of maximizing the perceived video quality at the receiver. Similarly, in [13], Kim et al. proposed a solution to send TCP packets belonging to a video transmission along in a set of different paths, according to an optimization of the quality of experience at the user level, evaluated through a mean opinion score (MOS) value.

In a different abstraction level, Zhang et al. [26] presented a region-of-interest (ROI) video transmission in heterogeneous wireless networks with multi-homed terminals. On a ROI-coded mobile video stream, only the most essential items of a video transmission are deeply encoded, such as background encoded with a larger quantization parameter. A channel monitor keeps track of the status of each communication path available, and sends feedback signals to the streaming controller for packet-scheduling control, with a deep learning method to predict channel selection. In this architecture, a scheduling approach based on a network and rate-distortion model is made, maintaining the playback fluency by improving the quality (as estimated by the PSNR) and balancing end-to-end delay.

Without going deeply within both ends of the standardized communication layered model, a mid-term approach can also be chosen. In [18], Nightingale et al. introduced an implementation of multipath video streaming, taking into consideration a Network Mobility (NEMO)-enabled mobile network and the usage of the scalability features of the H.264 scalable video coding (SVC) standard addendum. Here, the scheduling of RTP packets is performed after identifying each frame type, trying to reduce the path switching frequency, and improving playback fluency. Also, at a packet level, in [3], the authors describe a multipath MPEG video streaming solution able to prevent the loss of video frames, ensuring their playback deadlines by scheduling the packets containing high priority (I and P) video frames on the lowest-loss-rate set of paths.

Several other solutions on this mid-term approach have already been studied by the international community for several years. Such solutions can be categorized in two different groups: *server-oriented solutions*, as proposals that server-wisely try to accommodate a good and reliable transmission to a client on a mobility network, ensuring the delivery of all contents; and *receiver-oriented solutions*, as applications that allow the mobility network clients to manage and organize themselves to receive the externally transmitted content promptly, providing specific feedback to the server. Currently, the leading solutions regarding server-oriented architectures use erasure coding (EC) to ensure the delivery of real-time transport protocol (RTP) packets, in a new protocol named EC-RTP [7, 16], in which borderline components translate incoming RTP packets to EC-RTP, applying erasure coding to them. This creates a new protocol domain where RTP packets are handled differently inside the mobility network. Then, as packets exit the mobility units, another component reorders the received packets, re-translating them as RTP to the client.

Another solution, proposed in 2014 by Yao et al. [24], describes, specifically for an H.264 video stream, an IPB-frame adaptive mapping-mechanism, using IEEE 802.11e standard. This standard implements the requirements of QoS classification on the entering video transmissions, by creating different access categorys (ACs) and then classifying different video frames to them, according to their priority. This solution, despite its features, needs the entering video to be subject to a transcoding procedure if any of the AC queues is full, in order to change the group-of-pictures (GoP) length or the amount of B-frames being transmitted.

Finally, still regarding server-based solutions, Chau et al. [5], in 2018, implemented a cooperative multicast of scalable video using network coding, to ensure the delivery of content, over relay-based cellular networks, guaranteeing a smooth playback. Being a good candidate for using in 5G networks, this work features a cooperative mechanism which overcomes the performance degradation of wireless broadcast due to interferences and path losses, whilst it uses a network coding technique to ensure correct multimedia content delivery.

Regarding receiver-oriented solutions, a more complex architecture can be found in [14], where one can create a collaborative network—named cooperative video streaming (CVS),—allowing clients to receive help downloading content from external networks, through known neighbors, depending on their roles in a group context. Creating the concept of an helper (who retrieves data from a remote server), a forwarder (who forwards data from a helper), and a requester (who wants the data), all nodes work together to retrieve video content. This work has been extended by Yaqub et al. [25] with a pre-selection of the neighbors, exploiting multiple on-road characteristics.

Video data is naturally large, where packet collision becomes a very common issue in high density video transmissions. As the number of vehicles increase, it becomes more probable that broadcast storms do happen, which will rebound as packets rebroadcasts the same content. For this reason, Naeimipoor et al. [17] developed a hybrid video dissemination protocol (HIVE) that deploys a receiver-based relay node selection in addition to a medium access control (MAC) congestion control mechanism. In parallel, an erasure coding is applied to the transmitted video packets, ensuring the content delivery.

In 2019, Immich et al. [8] have developed a real-time live transmission mechanism using high efficiency video coding (HEVC) streams, which incorporates an ant colony optimization scheme to dynamically allocate a precise amount of redundancy into the content delivery system. This incorporation allows a system to use a combination of forward error correction techniques and unequal error protection to prevent high packet losses, ensuring the correct delivery of multimedia content from a source to the requesting client.

Considering the set of previous work in the literature, there are several proposals that consider the multipath split of the video taking the content into account. However, to the best of our knowledge, we are not aware of proposals that consider the split of the video through its different frame types and importance in a multihomed mobility-enabled vehicular network.

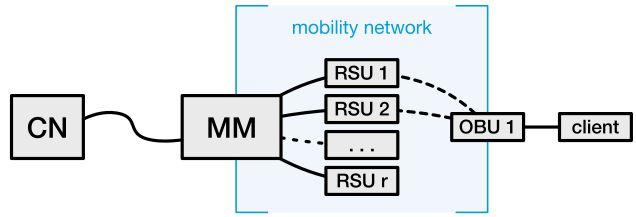


Fig. 1 Mobility network topology

3 Transmitting video with multihomed VANETs

This section introduces definitions and use cases for the transmission of video in multihomed VANETs.

3.1 Definitions

This study considers a VANET with mobility running an IP-based mobility protocol with the support of multihoming. Such a network has a simple representation depicted in Fig. 1, whose components are described as follows:

- An *MM* is the central node of the mobility network, being the one that manages the communications on the static-mobility border;
- A *RSU* is a service provider gateway, subject to the MM, that has a proper technology interface allowing clients, on OBUs to connect;
- An *on-board unit (OBU)* is the interface of a mobility network client, positioned in a vehicle, which connects to multiple RSUs, as it passes by them;
- A *corresponding node (CN)* is an outsider node from the mobility network, to whom a mobility client wants to have content access.

The mobility protocol must be executed in different instances in the MM, in the RSUs, and the OBUs, so that each one can share information about the available interfaces, services, users' requests, and traffic flows.

As soon as an OBU is able to connect to two or more RSUs through different technologies (IEEE 802.11p/WAVE[1] and Wi-Fi, for instance), the OBU will send a request for multihoming to the MM, and the latter can evaluate a percentage of traffic flows to apply from each RSU to the OBU, granting the best load balancing possible for the traffic that flows to the OBU [4], taking into account the delay of the traffic flows. This percentage, the multihoming rule, is a dynamic value which adjusts itself according to the network and traffic conditions. This procedure, on the OBU, is started by a connection manager, responsible for the selection of the best connections available in a given time instant.

[1] https://standards.ieee.org/standard/802_11p-2010.html.

In terms of video transmission on a network, in this work we decided to explore the implementation of H.264 [21], currently one of the most used formats for the distribution and manipulation of video content. Moreover, it supports a network abstract layer signaling information about video content on a packet-level.

Since video is a continuous stream of images semantically related to each other, the number of changing pixels (picture elements) can sometimes be considered irrelevant, since their number is very low. For this reason, MPEG has three types of frames to compress data, which are the following:

- *I-frames* a frame that is compressed solely based on the information contained in a frame; here there is no reference to any of the other video frames before or after this frame; its name I stands for intra;
- *P-frames* a frame that has been compressed using the data contained in itself and data from the closest proceeding I or P frame; here, P stands for predicted;
- *B-frames* a frame that has been compressed using the data contained from the closest proceeding I or P, and the closest following I or P frame; here, B stands for bidirectional.

In fact, as with H.264, the granularity of prediction types works on a slice-level, which is a spatially distinct region of a frame that is encoded separately from any other region in the same frame. Consequently, the standard I-, P- and B-frames MPEG types are now to be mentioned as I-, P- and B-slices.

The encoded video sequences, to be transmitted over a network communication channel, need to be encapsulated in multiple packets. MPEG's H.264 extension also implements a new concept which eases the transmission of video frames under a network context. By the usage of small units of data called network abstract layer (NAL) units, the coded video data can be accessed in a network layer, where identification of a content type is made, disregarding content itself.

In MPEG, to allow varying connection qualities such as in mobile nodes, it was developed a coding standard that allows graceful degradation in lossy transmission environments: the SVC. Being appended to the MPEG-4 as an extension [20], the SVC will liberate different spatial and time resolutions so that receivers do not miss the contents, independently of their capabilities. With these coding procedures, a video segment translates itself onto a set of layers, being one of them—the critical one to be received—named *base layer*, and the others—supplemental—named *enhancement layers*. These layers are arrayed in the following manner: in the start of the stream, a base layer is placed, with which a device can play the content at its minimum spatial resolution; following it, several enhancement layers allow the device to increase such spatial resolution if those could be retrieved. On a time axis, a similar situation can occur, where a base layer marks the minimum frame rate to watch content and, by further collecting enhancement layers, a higher frame rate could be achieved.

3.2 Use-cases

In related works, such as [7, 16], the proposed approaches were designed taking in consideration some applications that are useful to specific activities. We enumerate the following use-cases that will also be considered in our work:

- *Video graceful degradation* If the network traffic increases and receiving nodes need to perform handovers between different RSUs with multihoming while on a highway, then the clients should be able to recognize the transmitted content, since the local mobility anchor (LMA) will send the critical segments of video through the more reliable path directed to them. This way, the client's user will experience some graceful degradation of the video quality when the reception quality gets worse, while being able to capture the content.
- *Backup of distribution of emergent video content* In emergency scenarios, it is quite useful to have a backup solution if a video transmission is required for a team to get their eyes on the scene in a sudden way. With graceful video degradation, as mentioned in the last point, this transmission can be degraded by the reception quality getting worse, but the content will still be understandable.

4 Proposed video multihoming

This section describes the proposed approach for video multihoming to support both use-cases described earlier.

4.1 Concept and architecture

Our proposed solution starts by identifying the semantics of the contexts in the transmitted data chunks and then, by identifying the beginning and end marks of each video segment, it groups them as a unit. Having groups of packets already identified as distinct sections or components of a video stream, a load balancing mechanism can now work intelligently with such chunks, where losses will not degrade the quality of experience as much as the original blind balancing method would be.

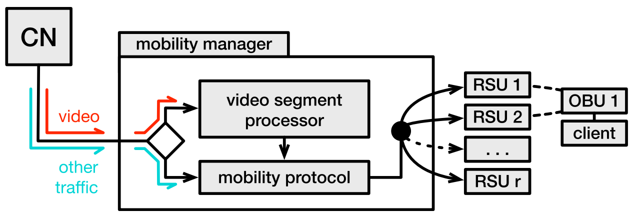


Fig. 2 Cooperative module with a mobility protocol execution

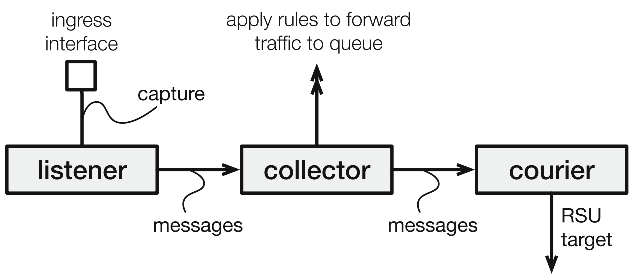


Fig. 3 Components of the video segment processor top module

In this work, we focus on video transmissions in the downlink direction, that is, from an external network (a CN) to a mobile client in the VANET. Having this in mind, the ideal place to insert a coordinator module, as a component able to perform such identifications and load balancing terms, is side-by-side with a mobility manager, a module working as a helper to the mobility protocol running daemon.

As depicted in Fig. 2, by inserting a cooperative module with the protocol in use, we can now identify video streams entering the mobility network, monitoring and grouping sets of packets as they belong to a single video segment, such as a frame, or a scalability enhancement or base layer.

In Fig. 2 there is a significant component which is the core function of this work: the *video segment processor*. This module is able to identify, collect, and send hints to the mobility protocol about to which RSU a frame/layer should be sent. Such a component will have the following structure, as shown in Fig. 3.

Attached to the ingress network interface, where video stream packets enter the mobility manager, a *listener* module is responsible for identifying if a received packet is video-related or not. For this purpose, a capture must be in place, able to identify video-related packet headers. When such element is detected, then a recognition of its frame/layer type should be made by decoding its NAL unit type information, sending then an internal message to the collector module, informing that the current received packet belongs to a given frame/layer and stream.

Following the listener module, the *collector* module will act depending on the messages received from the latter, and create ways for another module (the courier) to dispatch information to the protocol in use. By receiving messages from the listener, the collector must forward the packets to a specific queue, in a way that the mobility protocol can work with such packets, independently of the other non-video traffic.

Next, the *courier* receives the collector's messages warning the presence of a new frame/layer on the queue, and keeps track of them. As soon as a frame/layer is received entirely by the MM, this module will then interact with the current protocol execution instance, informing that a new specific segment has reached the mobility manager and it must be distributed in a specific path towards the targeted client.

4.2 RSU targeting algorithms

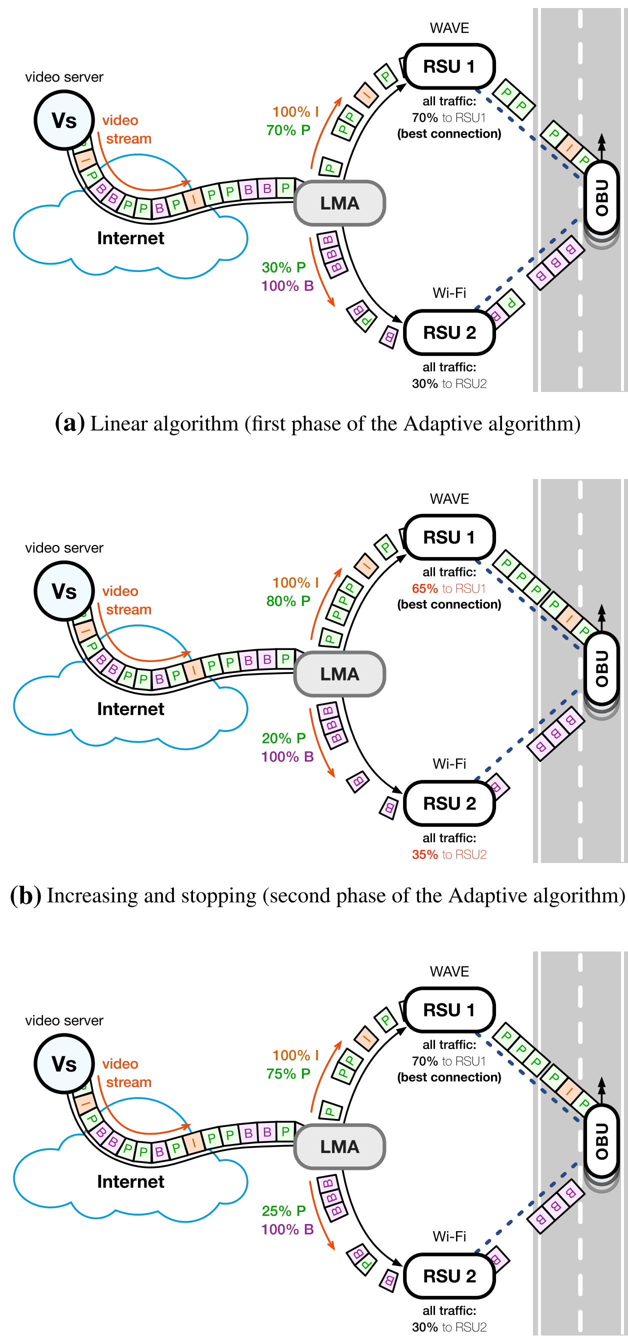
We consider a specific case of NEMO proxy mobile IPv6 (N-PMIPv6) implemented as a mobility protocol, where a local mobility anchor (LMA) is the machine with the MM responsabilities. Since the courier module needs to select a path towards the client to forward a given video segment, it must also receive the information of the N-PMIPv6 caches, where a list of the connected RSUs establishing a path towards a client's OBU exists, as well as their network access technologies and the assigned multihoming flow percentage to each RSU. Moreover, with such data and having in consideration that different video segments can have more or less priority than others (depending on the encoding), the courier must be able to select more reliable paths–paths whose availability is the highest of the set of available connections–for higher prioritized segments, leaving the other segments to be sent by the remainder RSUs. This method is referred to as *linear selection*, and it is described in algorithm 1, where S is the set of available connections towards a client, and F is a given frame.

As an example, considering an H.264 over RTP video stream, then the courier will select I-frames to be transmitted over the most reliable path, P-frames to be partially transmitted in the same path (in a multihoming rule percentage relation) and others, along with B-frames. Figure 4a depicts such a scenario.

Algorithm 1 Linear strategy—Stateless logic

function LINEARSTRATEGY(S, F)
  ; If there is no connected RSUs...
  if S **is** $\emptyset$ **then**
    pass
  ; If there is one connected RSU...
  if $|S|$ **is** 1 **then**
    return $S[0]$.GETADDRESS()
  ; Get list per available RSU technology t.
  $S_{\mathrm{WAVE}}, S_{\mathrm{WiFi}}, \ldots, S_t \leftarrow S$.GETVECBYTECH()
  ; If this frame F is high priority...
  if F.ISHIGHPRIORITY() **then**
    ; and if there is no WAVE RSU
    if S_{WAVE} **is** $\emptyset$ **then**
      ; Return the best non-WAVE RSU.
      return GETBESTON($S_{\mathrm{WiFi}} \cup S_t$)
    else
      ; Return the best WAVE RSU.
      return GETBESTON(S_{WAVE})
  ; If this frame F is medium priority...
  else if F.ISMEDIUMPRIORITY() **then**
    $r_{\mathrm{WAVE}} \leftarrow$ GETTOTALRULEDIVISION(S_{WAVE})
    $c_{\mathrm{WAVE}} \leftarrow$ GETBESTON(S_{WAVE})
    $c_{\mathrm{WiFi}\cup t} \leftarrow$ GETBESTON($S_{\mathrm{WiFi}\cup t}$)
    ; $B(n, p)$ is the binomial distribution.
    $\alpha \leftarrow$ GETRANDOMVALUE($B(2, r_{\mathrm{WAVE}})$)
    ; Return the best WAVE or non-WAVE RSU, by r_{WAVE}.
    return $c_{\mathrm{WAVE}}, c_{\mathrm{WiFi}\cup t}[\alpha]$
  ; Otherwise, if this frame F is low priority...
  else
    ; If there is no non-WAVE RSU...
    if $(S_{\mathrm{WiFi}} \cup S_t)$ **is** $\emptyset$ **then**
      ; If there is only one WAVE RSU...
      if $|S_{\mathrm{WAVE}}|$ **is** 1 **then**
        ; Return the best WAVE RSU.
        return GETBESTON(S_{WAVE})
      ; If there are more then one WAVE RSU...
      else
        ; Return the second best WAVE RSU.
        return GETSECONDBESTON(S_{WAVE})
    ; If there are non-WAVE RSUs...
    else
      ; Return the best non-WAVE RSU.
      return GETBESTON($S_{\mathrm{WiFi}} \cup S_t$)

The second strategy, *adaptive selection*, will make use of the cooperation with the N-PMIPv6 program execution, using the changes of the multihoming rule evaluation to flag when to stop adjusting the percentages of packets on the most reliable connection. Starting with the linear selection, the most relevant segments are forwarded via the most reliable connection and the rest through the other connections (Fig. 4a). Now, if the selector starts to increase the percentage of the next most relevant segments to the most reliable connection (Fig. 4b), it will allow to use more of the most reliable connection with a video stream that is more error-prone, relatively to other types of traffic, in a user experience perspective. This increase in such percentage must occur until the rule changes when the mobility manager senses that there is a possibility of congestion in a path (Fig. 4c). When this happens, the exceeded amount of traffic should be repositioned, in order for the rule to reset its old value, where the mobility network will assume no further modifications until a new topology change occurs.



(a) Linear algorithm (first phase of the Adaptive algorithm)

(b) Increasing and stopping (second phase of the Adaptive algorithm)

(c) Rollback (third phase of the Adaptive algorithm)

Fig. 4 Linear and adaptive selection example with a H.264 over RTP video stream. The linear selection is depicted in subfigure (**a**), while **a–c** depict the set of states of the adaptive selection strategy

As an example, if we consider an SVC video stream, then the courier will select the base layers to be sent over the most reliable path, and the enhancement layers to be preferably sent through the same, that is, starting to send in multiple paths (as in the linear strategy) and incrementing the number of packets via the best path, until the multihoming rule changes, noticing the possibility of a congested connection. This strategy is described in the algorithm 2, where S is the set of available connections towards a client, F is a given frame, and q is an algorithm state ($q^\star$ is its future state).

Algorithm 2 Adaptative strategy—Stateful logic

```
function ADAPTATIVESTRATEGY(S, F, q)
    ; Update the status of the system.
    q* ← q
    ; If there is no connected RSUs...
    if S is ∅ then
        pass
    ; If there is one connected RSU...
    if |S| is 1 then
        q*.LASTFRAMETYPE ← F.TYPE()
        return S[0].GETADDRESS(), q*
    ; Get list per available RSU technology t.
    S_WAVE, S_WiFi, ..., S_t ← S.GETVECBYTECH()
    ; If this frame F is high priority...
    if F.ISHIGHPRIORITY() then
        increment q.HIGHPRIORITYFRAMES
        q*.LASTFRAMETYPE ← HIGHPRIORITY
        if S_WAVE is ∅ then
            ; Return the best non-WAVE RSU.
            return GETBESTON(S_WiFi ∪ S_t), q*
        else
            ; Return the best WAVE RSU.
            return GETBESTON(S_WAVE), q*
    ; If this frame F is medium priority...
    else if F.ISMEDIUMPRIORITY() then
        increment q.MEDIUMPRIORITYFRAMES
        q*.LASTFRAMETYPE ← MEDIUMPRIORITY
        r_WAVE ← GETTOTALRULEDIVISION(S_WAVE)
        c_WAVE ← GETBESTON(S_WAVE)
        γ ← THRESHOLDOFDIFFERENCE
        ω ← FLOWDIVISIONRULEHASCHANGED(q, q*, γ)
        if r_WAVE is within a reasonable range then
            q*.WAVEPERCENTAGE ← ω × γ + r_WAVE
        α ← RAND(B(2, q*.WAVEPERCENTAGE()))
        return c_WAVE, c_WiFi∪t[α], q*
    ; Otherwise, if this frame F is low priority...
    else
        increment q.LOWPRIORITYFRAMES
        q*.LASTFRAMETYPE ← LOWPRIORITY
        ; If there is no non-WAVE RSU...
        if (S_WiFi ∪ S_t) is ∅ then
            ; If there is only one WAVE RSU...
            if |S_WAVE| is 1 then
                ; Return the best WAVE RSU.
                return GETBESTON(S_WAVE), q*
            ; If there are more than one WAVE RSU...
            else
                ; Return the second best WAVE RSU.
                return GETSECONDBESTON(S_WAVE), q*
        else
            ; Return the best non-WAVE RSU.
            return GETBESTON(S_WiFi ∪ S_t), q*
```

4.3 Implementation

The current deployment architecture is depicted in Fig. 5. Packets are sensed by the listener-collector-courier triple and, if a video stream is detected, all traffic with such destination address and port is forwarded to the video queue using `iptables`. The N-PMIPv6 executes a finite state machine which uses `libnfqueue` library on callback functions to retrieve packets from the queue. As an extension, a video queue callback is implemented to manage stream-related packets, whose RSU target is retrieved from the courier module.

In terms of video codec, following the work of Lee et al. and Yaqub et al. [14, 25], the H.264 extension to the MPEG-4 standard is chosen to work with. Although it is complex and extensive, it is the most common video codec. In this work, this codec will be transmitted through the mobility network, inside RTP packets, as our base transport protocol.

4.4 Considerations on different access technologies

The described and implemented scenarios in this work refer the N-PMIPv6 protocol as our base mobility protocol to manage the connections between mobile nodes within a mobile network using IEEE 802.11p/WAVE access technology. Despite these considerations, this solution is able to be implemented with other technologies, and in other topologies that manage nodes and routes in a mobility scenario.

To adapt this solution to the mobility context, some requirements need to be accomplished: (1) the mobility protocol or another entity being responsible by estimating a load balancing factor must be able to provide these values to an external module (our courier submodule); (2) from the array of different network access technologies, one of them should be considered as prioritized in face of the others, as we consider the WAVE connection in our scenarios; finally, (3) the video segment processor module needs to be implemented in a machine where the load balancing feature is running, usually in the static-mobility domains border.

As for nowadays, 5G networks are getting down the spotlight, and this work is compatible with its usage, since its architecture is agnostic to the network access technologies. In fact, if the array of connections to RSUs is made with the 5G network (as with the RSU-OBU connection), this work still applies, as long as the previous requirements are accomplished. Figure 6 depicts an example showing the usage of the video segment processor in cooperation with a load balancing feature using a 5G and Cellular Vehicle-to-Everything (C-V2X) infrastructure.

5 Results and evaluation

This section tests the performance of the proposed approaches in a mobile environment. The tests are split into three different phases, as described below:

1. *Phase 1* transmission of different resolutions of video through the mobility network, without movement;
2. *Phase 2* transmission of different resolutions of video through the mobility network, considering simulated movement;
3. *Phase 3* transmission of different resolutions of video through the mobility network, with real movement in a real vehicular scenario.

In these phases, a number of different spatial video resolutions are inserted onto the mobility network via a video server machine located outside the network. The different video resolutions are presented in Table 1.

The usage of multiple video resolutions aims to test the system behavior with different volumes of data traversing the network channels, to evaluate the system with low and high congested transmission mediums.

The streamed video is a cut of the first 3 min of the short "Big Buck Bunny", downgraded from 1080p to 720p, 360p, and 240p, with 82.4 MB, 44.4 MB, 21.8 MB, and 14.3 MB, respectively.

5.1 Testbed characterization

The evaluation tests are performed using the topology in Fig. 1. Considering only two RSUs, four different computer systems are in use, with two different wireless network technologies, as described in Table 2.

In terms of network access technologies, the tests are performed with both WAVE and Wi-Fi technologies. More particularly, one of the RSUs allows a connection with WAVE, and the other with Wi-Fi. To those, an OBU carrying both technologies is also considered, to which a computer serving as a client, is connected via Wi-Fi. Table 3 describes the characteristics of the network cards used in these tests.

5.2 Discussion of results

As mentioned before, the tests are conducted in three different phases: a first one, referred as *static tests*, which will not have mobility; a second one that emulates handovers, referred as *handover tests*; and a third one with handovers in a real vehicular scenario, *outside tests*.

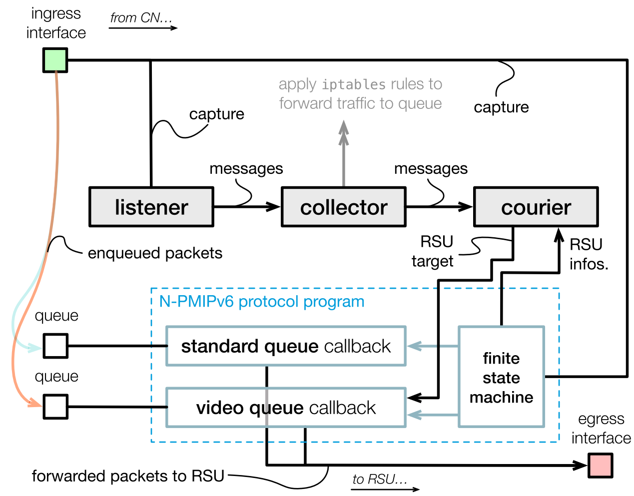


Fig. 5 Implemented global architecture

The 3 strategies for video distribution are tested in the different scenarios: the current mobility network distribution which does not take into account the video content (denoted as *baseline*); the linear strategy, denoted simply as *linear*; and the adaptive strategy, denoted as *adaptive*. The presented results are a mean of the 5 runs with, when possible, a 95% confidence interval.

5.2.1 First phase: results of static tests

In the first phase, the tests are conducted in a laboratory environment, simulating conditions where, in the 3 min of the video transmission, the vehicle was stopped and with access to both WAVE and Wi-Fi technologies.

We consider the several video resolutions. Figure 7 shows the loss percentage in the 240p video resolution with the different video distribution strategies.

At first sight, we can argue that the presence of a video frame distribution algorithm enhances the quality of reception of video streams in the mobility network's clients: the loss percentages dropped approximately 87.5% from the baseline tests to both linear and adaptive ones. With a video frame distribution algorithm enabled, the probability of complete frames reaching the client's machine is more significant, once the distribution mechanism performs its action frame-wise, not packet-wise (or fragment-wise).

In this video transmission, the packetization in NAL units will create an average of two packets (fragments) per I- and P-frame. With this in mind, and considering the absence of a video frame distribution algorithm, the fragments can follow different paths towards the client, reaching out-of-order with higher probability. With a video frame distribution enabled, as the fragments of a single frame are always dispatched in a group (as a unit) towards the client, such scenario is avoided.

Since the most significant footprint frame-type-wise on an H.264 encoded video is made out of P-frames, Fig. 8 shows us the division applied to them along the transmission of video, in comparison to the division that was initially provided by the genetic algorithm of the baseline mobility protocol. We observe that the percentage of P-frames heading towards the WAVE-enabled RSU is smaller than the given multi-homing rule in the baseline experiments. This reflected a more prominent usage of Wi-Fi simultaneously with the WAVE connection, leading to more significant losses on the client machine, since the Wi-Fi technology has a medium access control protocol enabled, opposite to the handshake-less WAVE medium, as we are using it.

Since the average number of video fragments per frame is two in these video resolution conditions, the linear algorithm shows that a bigger percentage of packets traversing the WAVE technology is similar to the given by the multi-homing rule. In the adaptive strategy, as it tries to use more of the WAVE connection, increasing the usage percentage of this technology until a new multi-homing rule is evaluated, the registered percentage of P-frames packets via WAVE is bigger than the multi-homing rule. In both cases, the number of losses becomes low, since most of the packets are being distributed in semantically equal groups and diminish the probability of out-of-order receptions.

With the 360p video resolution, the obtained results are presented in Fig. 9. Similarly as before, we can argue that the presence of a video frame distribution algorithm enhances the quality of reception of video streams in the mobility network's clients. This is supported by the results, which show that the loss percentages have now dropped approximately 63.2% from the baseline to the linear strategy, and 56.2% to the adaptive strategy. Here we can also conclude that none of the connections from the RSUs to the OBU can be considered congested. From the last results to these, the only thing that has changed was the size of the frames, which consequently will provide us with more fragments per frame (about three to four fragments per frame).

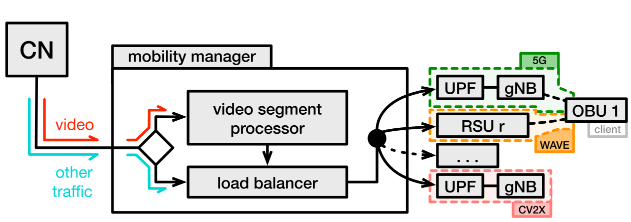


Fig. 6 Execution example with different network access technologies

Table 1 Video resolutions by value, width and height

Value	Width	Height
240p	352	240
360p	480	360
720p	1280	720
1080p	1920	1080

Table 2 Testbed nodes technical specifications

	CN	LMA (virtualized)
Processor	Intel Core i5, 2.3GHz	AMD E1-1200, 1.40 GHz
Memory	8 GB	1 GB
Storage	256 GB	16 GB
OS	macOS Catalina	Ubuntu 16.04.3 LTS

	RSU/OBU	Client
Processor	AMD Geode, 500 MHz	Intel Core i7, 2.6 GHz
Memory	64 MB	16 GB
Storage	128 MB	256 GB
OS	VeniamOS (OpenWRT-based)	Ubuntu 18.04.2 LTS

In Fig. 10, one can observe the distribution of the P-frame packets in comparison with the baseline multi-homing rule. In this case, the results are similar to the previous one. Still, in both linear and adaptive algorithms, the losses have reduced, while the percentage of P-frames have now intersected the baseline multi-homing rule, since its variation has now reached higher ranges. However, still, the percentage of P-frames packets in the adaptive algorithm has been increased in comparison with its assigned multi-homing rule.

The results of the 720p video resolution are observed in Fig. 11. We observe that the losses have increased, and reduce 26.3% from the baseline strategy to the linear, and of 38.9% from the baseline to the adaptive strategy.

In Fig. 12 the loss percentage is now larger, but the percentage of P-frame packets is lower than in the baseline multi-homing rule. Now, as the number of fragments per frame is larger, the number of P-frame packets is more approximate to the multi-homing rule in the linear algorithm, and failing to be less than the average of the multi-homing rule in the adaptive case. This can be explained by the fact that the genetic algorithm recalculated the multi-homing rule more frequently, not giving the adaptive algorithm the opportunity to increase the WAVE usage percentage.

The results of the video transmission with a resolution of 1080p are shown in Fig. 13, where we also show the P-frame packet distribution. The network starts to reflect

Table 3 Testbed network access technology hardware specifications

	RSU/OBU
WAVE interface	Mini-PCI 802.11p-compliant wireless interface with the Atheros AR5414 chipset, controlled by an `ath5k` driver
Wi-Fi interface	Wi-Fi module compliant with IEEE 802.11a/b/g
	Client
Wi-Fi interface	Intel Dual Band Wireless-AC 7625, module compliant with IEEE 802.11ac

losses due to the high bitrate, which turns the medium significantly congested. In this case, there are no gains in using a video frame distributing algorithm, since the nature of the network is the one not allowing frames to arrive at the client. Although possible to transmit 1080p video streams with the chosen technologies, in our laboratories the WAVE transmission rate is limited by its hardware. As a consequence, and to grant WAVE as the best connection available, Wi-Fi has been reduced in its transmission rates.

Even not registering gains, we show that the P-frame packets distribution is similar to the other cases, showing that the transmission problems are due to the streaming being too demanding, causing a bottleneck scenario.

5.2.2 Second phase: results of handover tests

The second phase of tests is also conducted in a laboratory environment, simulating conditions where the vehicle is moving from a WAVE-only area to a common WAVE and Wi-Fi area, then leaving to an exclusive Wi-Fi area. These handovers are performed on a rate of 1 handover per minute, as depicted in Fig. 14. Due to the results in the last phase, we excluded the video resolutions of 720p and 1080p, focusing our analysis exclusively on the 240p and 360p resolutions.

Figure 15(a) shows results of the entire 3 min of transmission, covering the length of time that the client was receiving packets that only passed through WAVE, followed by multi-homing and, later, by a time range where only the Wi-Fi connection was available; the second will only show the results of the second minute of the test, only concerning multi-homing. The handover tests could not be done in a perfectly controlled environment. As a consequence to that, the OBU has connected itself to the WAVE-enabled RSU through the time it should only remain connected to the Wi-Fi-enabled RSU, avoiding a multi-homing scenario.

In the total time of the experiment, we observe that, in the linear video frame distribution scheme, the percentage of P-frame packets going through the WAVE connection is larger than in the baseline multi-homing rule; however, observing the second minute in detail, we can find a very similar plot to what we have obtained in the first phase tests. Reconsidering that, a video frame is split onto an average of two packets, and the percentage of P-frames going through the WAVE connection is higher than in the baseline multi-homing rule.

In the 360p video resolution, Fig. 16, the loss percentage, although very close to the baseline conditions, is lower using both linear and adaptive algorithms of video frame

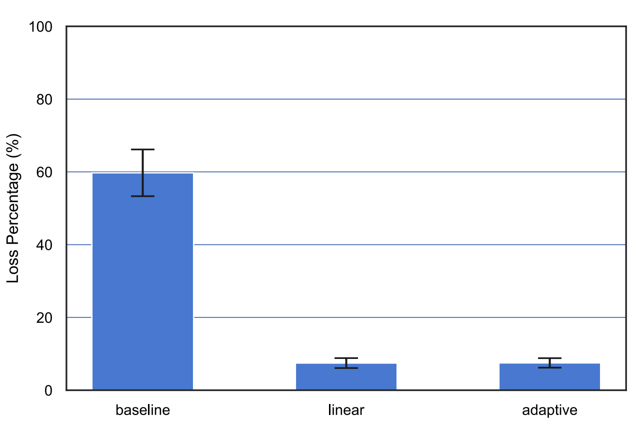


Fig. 7 Loss percentage on 240p video resolution by strategy in static tests

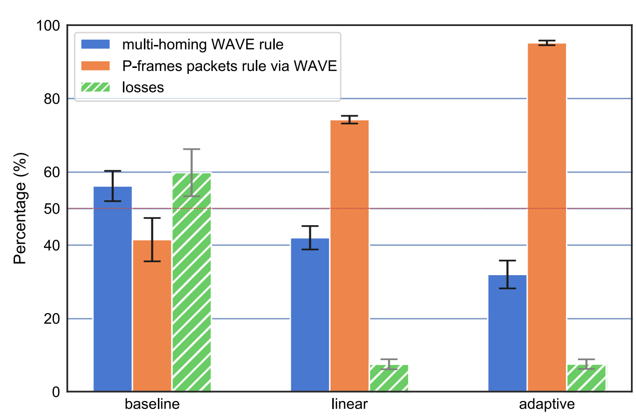


Fig. 8 P-frame packets distribution in comparison with the baseline multi-homing rule in static 240p experiments

distribution. In both linear and adaptive results, we can notice a reduction of the loss percentage given by an increase of percentage in the distribution of P-frame packets to the WAVE connection, allowing more complete frames to arrive in-order on the client.

5.2.3 Third phase: results of outside tests

Finally, to conclude our tests, some of them are executed on an outside environment, where real-world conditions are introduced to our system evaluation. In this scenario, only the 360p video resolution is chosen to be tested out, since it is the highest definition we got with positive results in the two first phases.

To perform this test, a car drives within a straight line with a distance of 100 m. The car started its course moving from far, but within the range of the WAVE-enabled RSU, stopping for 30 s in front of the RSU. Then, it moved again entering the common WAVE-Wi-Fi range at approximately 1 min of the experiment, stopping for another 30 s in-between both RSUs, getting the best signal reception from both. After this time, it moves again in a straight line, crossing the 2 min barrier, now, until the side of the Wi-Fi-enabled RSU, where only a Wi-Fi connection was made available. With this, the car stopped until the end of the test. A representation of this test is depicted in Fig. 17.

The results are depicted in Fig. 18. In terms of the baseline behavior of the mobility network, the percentage of P-frame packets transmitted over WAVE was smaller than the baseline multi-homing rule. Even so, the loss percentage can be considered small, which can be justified by the fact that the WAVE connection remained active for 2/3 of the total transmission time, as well as Wi-Fi's, reducing the losses. Moreover, in each of these 2/3 of the time for each network access technology, at least 1/3 of the time was spent connected to a single technology, where losses could only happen due to failures on the routing

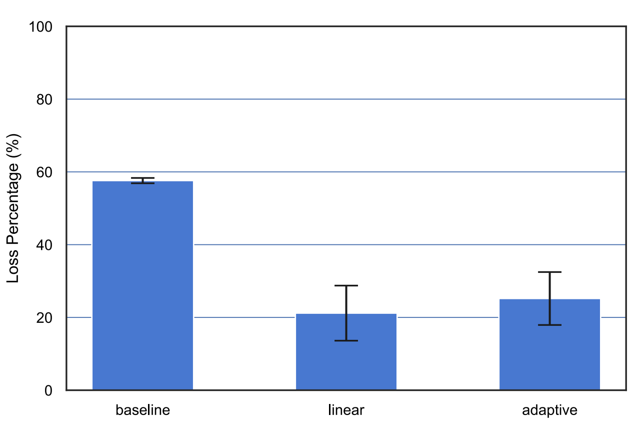


Fig. 9 Loss percentage on 360p video resolution by strategy in static tests

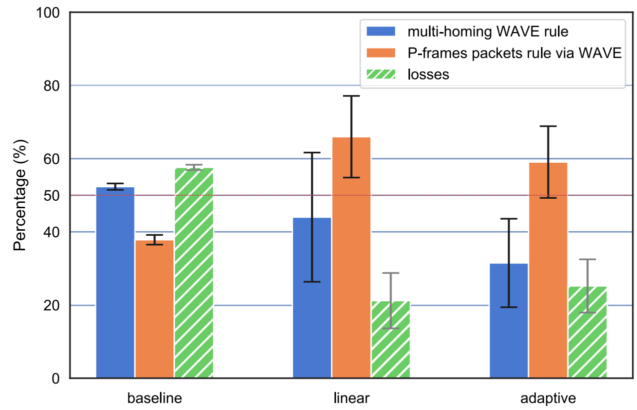


Fig. 10 P-frame packets distribution in comparison with the baseline multi-homing rule in static 360p experiments

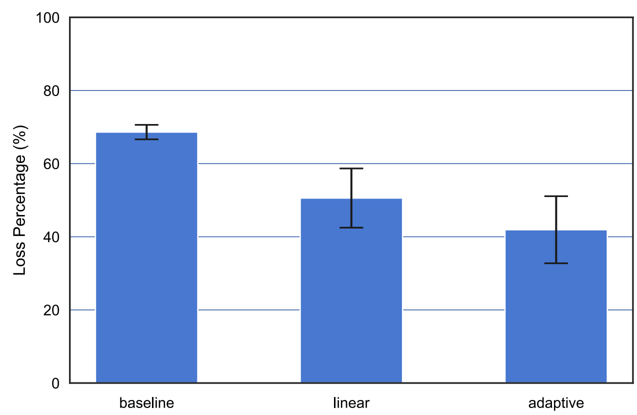


Fig. 11 Loss percentage on 720p video resolution by strategy in static tests

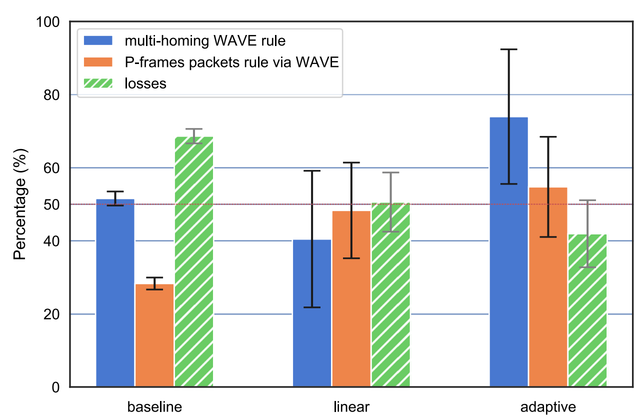


Fig. 12 P-frame packets distribution in comparison with baseline multi-homing rule in static 720p experiments

procedures of the mobility protocol or physical disturbances in the medium channels.

Relative to the algorithms, the percentage of P-frame packets was always bigger than the multi-homing rule. However, in none of the cases, the adaptive algorithm overpassed the linear algorithm on such percentage. This

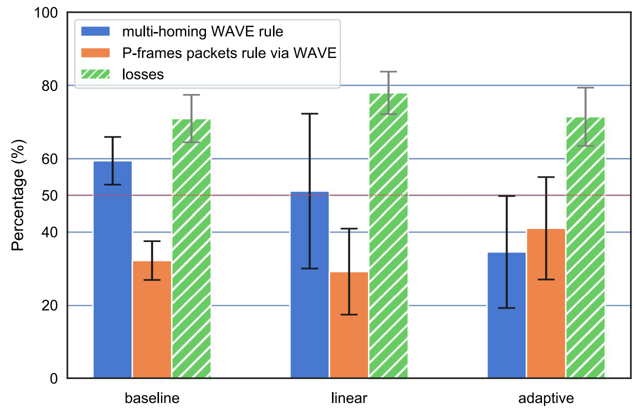


Fig. 13 P-frame packets distribution in comparison with baseline multi-homing rule in static 1080p experiments

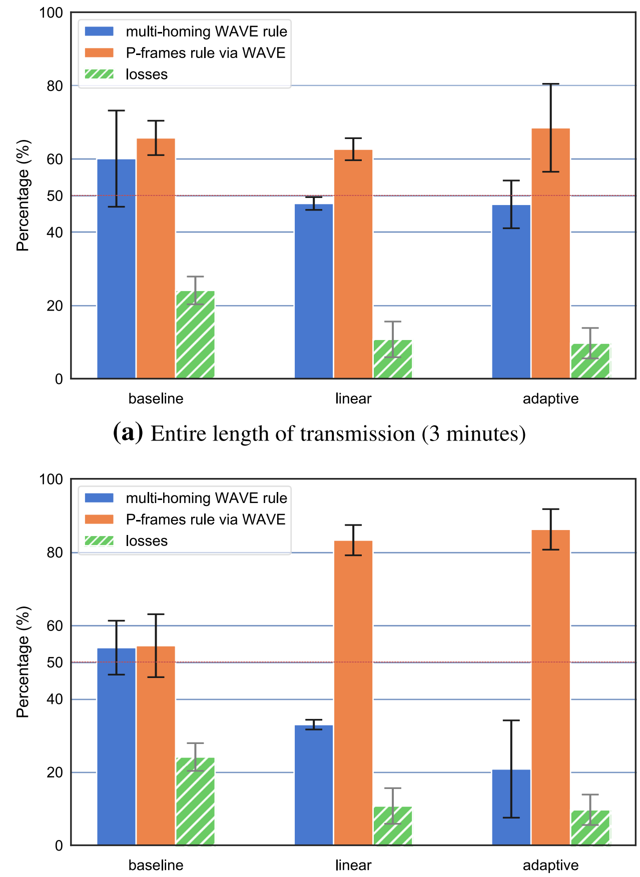
can be justified by the multi-homing rule being recalculated with a higher frequency than usual, not allowing the adaptive strategy to increment the percentage assigned to the WAVE communication channel.

5.3 Assessment of the video quality

To validate the results in terms of the visualization in the client side of the mobility network, in this section we characterize the effects of the applied video content scheduling in the QoE, estimating a (MOS) for each RSU targeting algorithms.

Figure 19 presents the MOS evaluation of the 3 approaches, using the MOS calculation in [19], where the mean opinion score is obtained relatively to variations of the video bitrate, screen size and video size ratios, frames per second, and the amount of buffer stalls. This MOS formula has been validated against the evaluation of the quality perception with real users.

This data, gathered in chunks of 1 s each in the tests where the handovers were present, is collectively shown in Fig. 19, where one can verify that the proposed strategies had a positive consequence on the QoE in both linear and adaptive algorithms, in comparison with the baseline. However, within those algorithms, we cannot conclude on the advantage between them, as already preceived in the aforementioned analysis.



(a) Entire length of transmission (3 minutes)

(b) Second minute of transmission

Fig. 15 P-frame packets distribution in comparison with baseline multi-homing rule in handover 240p experiments

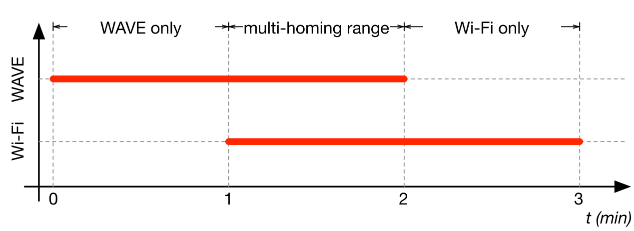
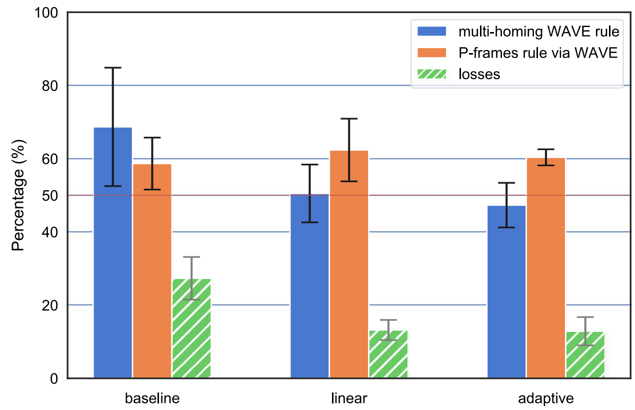


Fig. 14 Representation of the handover tests, on a time axis

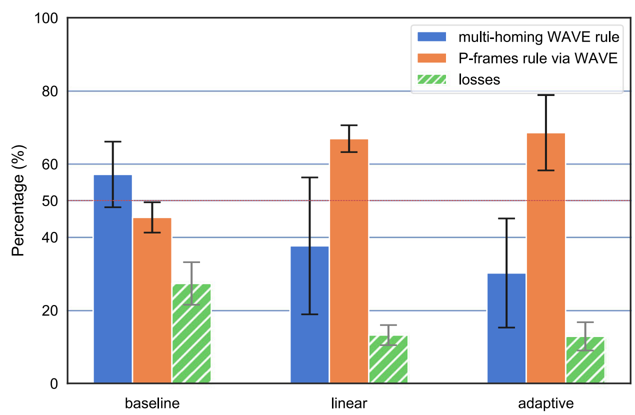
As the MOS scale provides us with a subjective evaluation of the video visualization quality, we complemented our analysis with the estimation of both peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) values.

Starting with the PSNR, in Fig. 20 we observe that the introduced noise, relatively to the video that was served, is higher in the baseline, than in both linear and adaptive algorithms. In fact, right in the beginning of the video, where the higher motion of the section is registered (where a full update of frame pixels are made with residuals during 8 s, as it is depicted in Fig. 21, being the more error-prone section of the 3 min video), the baseline has a downward variation of values of the PSNR, translating in a higher noise, in opposition to the linear and adaptive, which behaved better in these scenarios.

As expected, the estimated SSIM also positions the adaptive and linear algorithm applications above the baseline experiments. Here, the value of SSIM, as depicted in Fig. 22, shows that while the baseline has an index of approximately 0.497, the linear conquered an index of 0.651, and the adaptive an index of 0.715. This evaluation



(a) Entire length of transmission (3 minutes)



(b) Second minute of transmission

Fig. 16 P-frame packets distribution in comparison with baseline multi-homing rule in handover 360p experiments

allows to confirm that the adaptive and linear approaches have a higher ability to maintain the structure of the video after the distribution of frames by its type, through different RSUs.

Comparing the linear and the adaptive algorithms, we observe that the adaptive has slightly better results. In fact, in the video length section from the start until the depicted frame in Fig. 21 in 00:13, the adaptive algorithm also reached a SSIM of almost 1.

6 Conclusions and future work

The transmission of video in a VANET in an always moving environment is a very challenging task to perform. In this article we proposed an approach to transmit H.264 video content diminishing the loss percentages, while multi-homing is enabled. By distributing different and complete MPEG frames (according to its type) through specific RSUs to whom a client's OBU is connected, in a linear and adaptive approaches, we achieved a strong

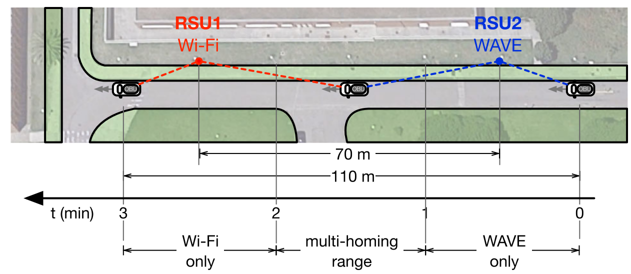
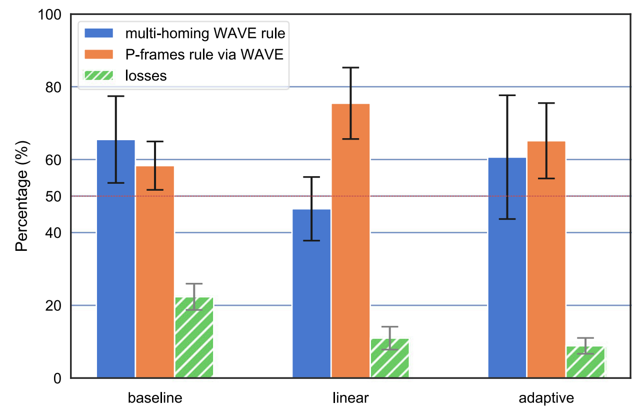
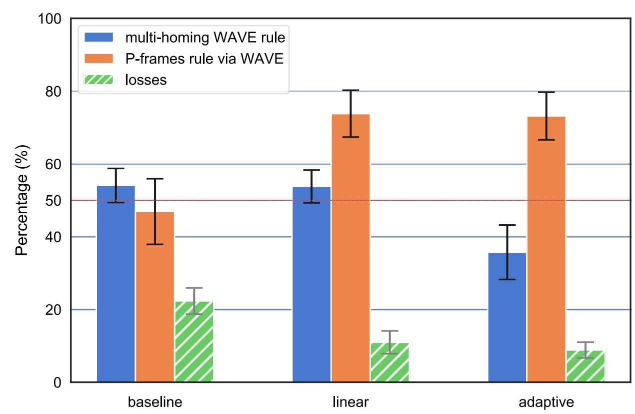


Fig. 17 Representation of the outside tests, on a time axis



(a) Entire length of transmission (3 minutes)



(b) Second minute of transmission

Fig. 18 P-frame packets distribution in comparison with baseline multi-homing rule in outside 360p experiments

decrease in the loss percentage, with a maximum of approximately 60.4%. When considering the subjective quality of the video, through MOS, PSNR and SSIM, we confirm that the proposed approaches linear and adaptive, outperform the baseline, with slightly better results to the adaptive approach. In our approach, according to the technologies available and their performance characteristics, we determine in real-time the best video content to the sent through each technology as the car moves in the road.

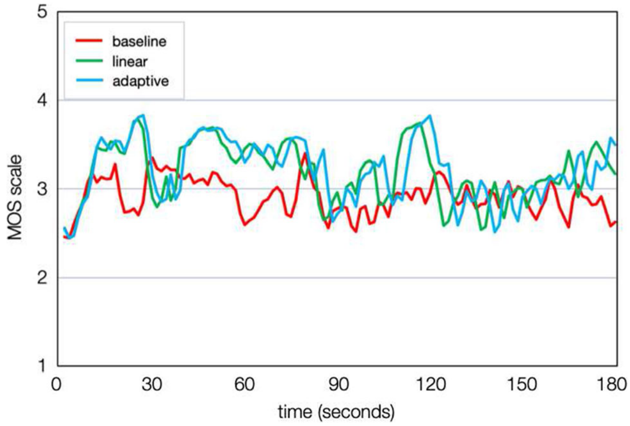


Fig. 19 MOS scale on handover tests, within the 3 min of transmission

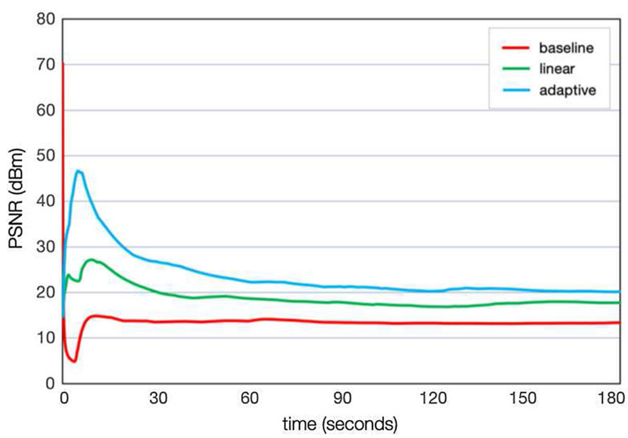


Fig. 20 PSNR scale on handover tests, within the 3 min of transmission

In terms of future work, we will follow-up this work in different directions:

1. Implementation of a video segment processor with the detection of SVC's base and enhancement layers—with scalability in mind, the mobile network's client could always obtain a visualization of the video quality, with a time, spatial, and quality resolution dependable of the amount of enhancement layers it can get from the network, for each video frame;
2. Active video transcoding in the LMA according to the communication channel's current throughput towards a client—by having an active transcoder positioned in the LMA, every video stream entering the mobile network could be transcoded, in order to produce a different amount of I-frames, or adjusting other encoding features, depending on the communication channels' characteristics towards a client;
3. Extension of the current system to a cooperative video streaming paradigm—with a cooperative video streaming, other network nodes, of similar nature such as the client or the OBU, located in the vicinities, could create a group and help a requester node to get the content, following the work in [14], but in a multi-homing-enabled context;
4. Implementation of coding techniques to ensure the delivery of video packets on the mobility network—to ensure that the packets are completely delivered to the client, a coding feature must be implemented and some derivative of Luby transform codes could be integrated onto our solution, allowing the distribution of video frames to be redundantly sent when the connection channel's quality is medium to low, guaranteeing the reception and further visualization of data on the client-side.

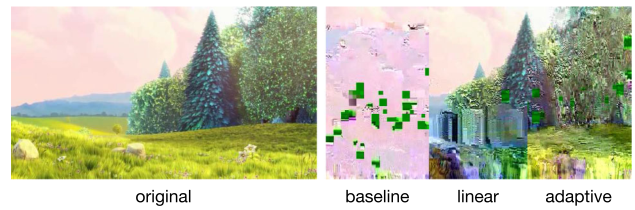


Fig. 21 Frame of "Big Bucks Bunny", in 00:13, as viewed with the application of the baseline, linear, and adaptive algorithms, in comparison with the original image

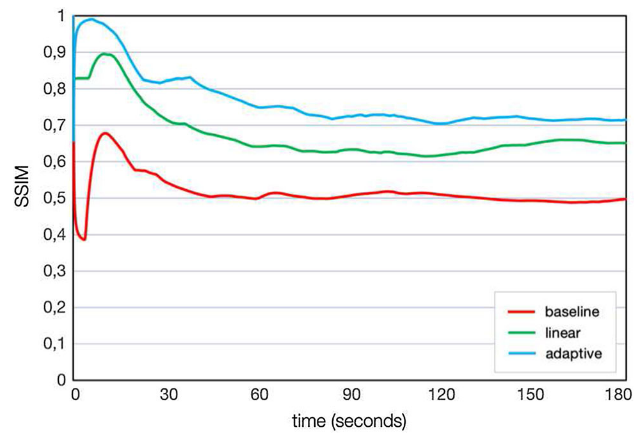


Fig. 22 SSIM scale on handover tests, within the 3 min of transmission

Acknowledgements This work is supported by the European Regional Development Fund (FEDER), through the Competitiveness and Internationalization Operational Programme (COMPETE 2020) of the Portugal 2020 framework [Project 5G-PERFECTA with Nr. 038190 (POCI-01-0247-FEDER-038190)].

References

1. Agarwal, S., & De, S. (2016). Cognitive multihoming system for energy and cost aware video transmission. *IEEE Transactions on Cognitive Communications and Networking, 2*(3), 316–329.
2. Al-Sultan, S., Al-Doori, M. M., Al-Bayatti, A. H., & Zedan, H. (2014). A comprehensive survey on vehicular Ad Hoc network. *Journal of Network and Computer Applications, 37*(1), 380–392. https://doi.org/10.1016/j.jnca.2013.02.036.
3. Allen, L. R., Falowo, O. E., & Chan, H. A. (2013). Mpeg video streaming solution for multihomed-terminals in heterogeneous wireless networks. In *2013 IEEE 10th consumer communications and networking conference (CCNC)* (pp. 677–682).
4. Capela, N., & Sargento, S. (2015). An intelligent and optimized multihoming approach in real and heterogeneous environments. *Wireless Networks*. https://doi.org/10.1007/s11276-015-0896-1.
5. Chau, P., Shin, J., & Jeong, J. P. (2018). Efficient scalable video multicast based on network-coded communication. *Wireless Networks, 24*(5), 1561–1574. https://doi.org/10.1007/s11276-016-1409-6.
6. Deng, Z., Liu, Y., Liu, J., Zhou, X., & Ci, S. (2017). Qoe-oriented rate allocation for multipath high-definition video streaming over heterogeneous wireless access networks. *IEEE Systems Journal, 11*(4), 2524–2535.
7. Fang, Z. (2014). RTP compatible: Two models of video streaming over VANETs. M.sc. thesis, University of Ottawa, Ottawa.
8. Immich, R., Cerqueira, E., & Curado, M. (2019). Efficient high-resolution video delivery over VANETs. *Wireless Networks, 25*(5), 2587–2602. https://doi.org/10.1007/s11276-018-1687-2.
9. Ismail, M., & Zhuang, W. (2013). Statistical QoS guarantee for wireless multi-homing video transmission. In *2013 IEEE global communications conference (GLOBECOM)* (pp. 4615–4620).
10. Ismail, M., & Zhuang, W. (2014). Mobile terminal energy management for sustainable multi-homing video transmission. *IEEE Transactions on Wireless Communications, 13*(8), 4616–4627.
11. Ismail, M., Zhuang, W., & Elhedhli, S. (2013). Energy and content aware multi-homing video transmission in heterogeneous networks. *IEEE Transactions on Wireless Communications, 12*(7), 3600–3610.
12. Jiyan, W., Jingqi, Y., Xiaokun, W., & Junliang, C. (2013). A low latency scheduling approach for high definition video streaming over heterogeneous wireless networks. In *2013 IEEE global communications conference (GLOBECOM)* (pp. 1723–1729).
13. Kim, D., & Park, S. (2015). Adaptive video packet transmission for guaranteeing video quality on mobile devices. In *2015 IEEE 5th international conference on consumer electronics—Berlin (ICCE-Berlin)* (pp. 456–458).
14. Lee, C. H., Huang, C. M., Yang, C. C., & Lin, H. Y. (2014). The K-hop cooperative video streaming protocol using H.264/SVC over the hybrid vehicular networks. *IEEE Transactions on Mobile Computing*. https://doi.org/10.1109/TMC.2013.153.
15. Lopes, R., Luís, M., & Sargento, S. (2019). Real-time video transmission in multihomed vehicular networks. In *2019 IEEE vehicular networking conference (VNC)*.
16. Mammeri, A., Boukerche, A., & Fang, Z. (2016). Video streaming over vehicular ad hoc networks using erasure coding. *IEEE Systems Journal, 10*(2), 785–796. https://doi.org/10.1109/JSYST.2015.2455813.
17. Naeimipoor, F., & Boukerche, A. (2014). A hybrid video dissemination protocol for VANETs. In *2014 IEEE international conference on communications (ICC)* (pp. 112–117). IEEE. https://doi.org/10.1109/ICC.2014.6883304.
18. Nightingale, J., Wang, Q., & Grecos, C. (2010). Optimised transmission of h.264 scalable video streams over multiple paths in mobile networks. *IEEE Transactions on Consumer Electronics, 56*(4), 2161–2169.
19. Salvador, A., Nogueira, J., & Sargento, S. (2015). QoE Assessment of HTTP adaptive video streaming. In: S. Mumtaz, J. Rodriguez, M. Katz, C. Wang, & A. Nascimento (Eds.) Wireless internet (vol. 146, pp. 235–242). Springer, Cham. https://doi.org/10.1007/978-3-319-18802-7_32. http://link.springer.com/10.1007/978-3-319-18802-7_32. Series Title: Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering.
20. Schwarz, H., Marpe, D., & Wiegand, T. (2007). Overview of the scalable video coding extension of the H.264/AVC standard. In *IEEE transactions on circuits and systems for video technology*. https://doi.org/10.1109/TCSVT.2007.905532.
21. Wiegand, T., Sullivan, G. J., Member, S., Bjøntegaard, G., Luthra, A., & Member, S. (2003). Overview of the H.264 video coding standard. *IEEE Signal Processing Magazine, 13*(7), 560–576. https://doi.org/10.1109/TCSVT.2003.815165.
22. Wu, J., Yuen, C., Cheung, N., & Chen, J. (2016). Delay-constrained high definition video transmission in heterogeneous wireless networks with multi-homed terminals. *IEEE Transactions on Mobile Computing, 15*(3), 641–655.
23. Wu, J., Yuen, C., Wang, M., & Chen, J. (2016). Content-aware concurrent multipath transfer for high-definition video streaming over heterogeneous wireless networks. *IEEE Transactions on Parallel and Distributed Systems, 27*(3), 710–723.
24. Yao, X. W., & Wang, W. I. (2014). IPB-frame adaptive mapping mechanism for video transmission over IEEE 802.11e WLANs. *ACM SIGCOMM Computer Communication Review, 44*(2), 6–12.
25. Yaqub, M. A., Ahmed, S. H., & Kim, D. (2018). Asking neighbors a favor: Cooperative video retrieval using cellular networks in VANETs. *Vehicular Communications, 12,* 39–49. https://doi.org/10.1016/J.VEHCOM.2017.12.002.
26. Zhang, Z., Jing, T., Han, J., Xu, Y., Li, X., & Gao, M. (2017). Roi-based video transmission in heterogeneous wireless networks with multi-homed terminals. *IEEE Access, 5,* 26328–26339.

Rui Lopes received the M.Sc. degree in Computer and Telematics Engineering from the University of Aveiro, Portugal, in 2019. He is currently pursuing the Ph.D. degree in Computer Engineering from the University of Aveiro, Portugal. He is Teaching Assistant in programming courses at the Department of Electronics, Telecommunications and Informatics, University of Aveiro, Portugal. He joined the Instituto de Telecomunicações, Aveiro in 2016, with research interests including vehicular ad-hoc networks and video transmissions over mobile networks.

Miguel Luís received the M.Sc. and Ph.D. degrees in Electrical and Computer Engineering from the Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal, in 2009 and 2015, respectively. He is an Adjunct Professor in Instituto Superior de Engenharia de Lisboa (ISEL) and Researcher with Instituto de Telecomunicações (IT), and has been involved in several national and European research projects targeting new communications for mobile networks. Currently, he is the coordinator of two national funded projects ''InfoCent-IoT: Efficient information centric networks for IoT infrastructures'' and ''MH-SDVanet: Multihomed Software Defined Vehicular Networks " and he contributes to several other research projects such as Aveiro STEAM City (EU-UIA program), SNOB-5G (FCT-MIT program) and 5G-Perfecta (Celtic-P2020), to name a few. Miguel has published more than 60 scientific works, including 3 book chapters and 25 publications in peer-reviewed international journals. His research interests include medium access control for wireless systems, routing and dissemination mechanisms for mobile networks and management, orchestration and softwarization of future networks.

Susana Sargento is a Full Professor in the University of Aveiro and a senior researcher in the Institute of Telecommunications, where she is leading the Network Architectures and Protocols (NAP) group (https://www.it.pt/Groups/Index/62). She received her Ph.D. in 2003 in Electrical Engineering in the University of Aveiro, being a visiting student at Rice University in 2000 and 2001. She joined the Department of Computer Science of the University of Porto between 2002 and 2004, and she was a Guest Faculty of the Department of Electrical and Computer Engineering from Carnegie Mellon University, USA, in August 2008, where she performed Faculty Exchange in 2010/2011. Since 2002 she has been leading many national and international projects, and worked closely with telecom operators and OEMs. She has been involved in several FP7 projects (4WARD, Euro-NF, C-Cast, WIP, Daidalos, C-Mobile), EU Coordinated Support Action 2012-316296 "FUTURE-CITIES", EU Horizon 2020 5GinFire, national projects, and CMU-Portugal projects (S2MovingCity, DRIVE-IN with the Carnegie Melon University) and MIT-Portugal Snob5G project. She has been TPC-Chair and organized several international conferences and workshops, such as ACM MobiCom, IEEE Globecom and IEEE ICC. She has also been a reviewer of numerous international conferences and journals, such as IEEE Wireless Communications, IEEE Networks, IEEE Communications. She was the founder of Veniam (https://www.veniam.com), which builds a seamless low-cost vehicle-based internet infrastructure, and she was the winner of the 2016 EU Prize for Women Innovators. Susana is also the co-coordinator of the national initiative of digital competences in the research axis (INCoDe.2030, https://www.incode2030.gov.pt/), belongs to the evaluation committee of the Fundo200M (https://www.200m.pt) government co-investment and funding), and she is one of the Scientific Directors of CMU-Portugal Programme (https://www.cmuportugal.org/). Her main research interests are in the areas of self-organized networks, in ad-hoc and vehicular network mechanisms and protocols, such as routing, mobility, security and delay-tolerant mechanisms, resource management, and content distribution networks. She regularly acts as an Expert for European Research Programmes.