Beacon-less Video Streaming Management for VANETs Based on QoE and Link-Quality

Carlos Quadros, Eduardo Cerqueira
Federal University of Pará - Brazil
{quadros, cerqueira}@ufpa.br

Aldri Santos
Federal University of Paraná - Brazil
aldri@ufpr.br

Jaehan Lim, Mario Gerla
UCLA - USA
{ljhar, gerla}@cs.ucla.edu

***Abstract*—Real-time video dissemination over Vehicular Ad hoc Networks (VANETs) is fundamental for many services, e.g., emergency video delivery, road-side video surveillance, and advertisement broadcasting. These applications deal with several challenges due to strict video quality level requirements and highly dynamic topologies. To handle these challenges, geographic receiver-based beacon-less approaches have been proposed as a suitable solution for forwarding video flows in VANETs. In general, the routing decisions are performed only based on network, link, and/or node characteristics, such as link quality and vehicle's location. However, in real situations, due to different requirements and hierarchical structures of multimedia applications, these existent routing decisions are not satisfactory to select the best relay nodes and build up reliable backbones to delivery video content with reduced delay and high Quality of Experience (QoE). This paper introduces the QOe-Driven and LInk-qualiTy rEceiver-based (QOALITE) protocol to allow live video dissemination with QoE assurance in Vehicle-to-Vehicle (V2V) scenarios. QOALITE considers video and QoE-awareness, coupled with location and link quality attributes for relay selection. Simulation results show the benefits of QOALITE when compared to existing work, while achieving multimedia transmission with QoE support and robustness in highway scenarios.**

I. INTRODUCTION

Vehicular Ad Hoc Networks (VANETs) are targeted to support a wide scope of new real-time multimedia services, ranging from on-road safety and security to entertainment video flows. As reported by Cisco, the traffic generated by video-based services will represent over 90% of the global IP data in a few years [1]. In this context, it becomes important to ensure that network resources are efficiently employed to maximize the video quality level from the user's point-of-view, i.e., with higher Quality of Experience (QoE). While the concept of Quality of Service (QoS) is only focused on packet-based resource management and delivery statistics, the perception of videos shared and watched by humans, characterized in terms of QoE, is measured and related to the subjective acceptability of the users.

Despite the potential of VANETs to support high connectivity, the provision and control of on-road real-time video flows is not a straightforward task. The broadcast nature and the highly dynamic network topology of VANETs brings many technical challenges to manage and deliver multimedia content. The routing decisions must adapt to topology changes and be aware of QoE requirements to recover or maintain the video flows with at least a good quality level for humans. Thus, the route decision process for video packets must be done not only based on networking and vehicle parameters, such as loss and position, but also based on application parameters, e.g., frame importance of the videos and the impact of that on the human perception [2].

A key requirement for multimedia content dissemination in VANETs is to decrease the time spent in the relay node selection process. Thus, the building of a forwarding backbone of relay nodes becomes essential. A variety of forwarding geocast solutions has been proposed by considering multimedia delivery [3–5]. Most of them are sender-based approaches and rely on end-to-end routes, which suffer from frequent interruptions in dynamic topologies. Conversely, Receiver-Based (RB) approaches improves the transmission of packets through a distributed hop-by-hop routing decision [6]. The RB backbone shifts the paradigm for the selection of relay nodes and packet forwarding from senders to receivers in a beacon-based or in a beacon-less operational mode. These RB approaches are not static end-to-end routes, allowing the persistence of flows in case of node failures, mobility, and wireless channel variations.

With regard to beacon-less RB concept, forwarding nodes do not need prior knowledge of their neighborhood for relay node's selection, avoiding beacon transmissions, saving bandwidth and media access [7]. Forwarding nodes broadcast packets to their 1-hop neighbors and upon receiving the packets, neighboring vehicles in a contention-based fashion, define timers to relay the packets further [5]. The first node to expire its countdown timer, forwards the received packets. The existing beacon-less RB approaches do not explore multiple criteria to build up routing backbones. Parameters, such as frame type and distortion estimation (caused by probability of occurrence of loss and loss burstiness) [8] must be used as QoE-Indicators to support forwarding decision at routing level [2]. For example, protocols [9] and [5] consider only geographical information for routing decisions. However, due to the unpredictability of the wireless environments, the most distant node might suffer from bad link quality or cannot receive the best frames of a video sequence from the human point-of-view. Therefore, a proposal that considers node, link, network, application, and human perception characteristics to reach better video delivery becomes necessary.

This paper proposes a multi-criteria protocol, that combines location information, current link quality conditions, and QoE-indicators to establish multi-hop backbones for live video dissemination in highway VANETs, named QOe-Driven and LInk-qualiTy rEceiver-based (QOALITE). The proposed protocol creates and controls V2V routes for live video transmissions, reacting well to node failures, and enhancing the user's experience by considering the most important frames of video sequences from the user's point-of-view. QOALITE performs an adaptive use of relay nodes while also improving or at least maintaining the QoE level of the transmitted videos when compared to non-QoE-driven schemes. Thus, QOALITE improves the transmission of on-road videos for medium distances (e.g., 1 to 5 Km) as required for many disaster, surveillance, and even for entertainment in multimedia smart city services.

The remainder of this paper is organized as follows. Section II presents relevant related work. Section III introduces the QOALITE protocol. Section IV describes the test environment, scenario, and simulation results. Finally, conclusions are summarized in Section V.

II. RELATED WORK

In RB approaches, vehicles decide the next relay node through a distributed contention phase, i.e., packets are delayed before sending to the MAC layer. In this context, a beaconless RB backbone-based strategy to create and manage reliable multi-hop routes during the video delivery can be used to reduce the hop-by-hop and packet-by-packet delays. This section presents the main beacon-less RB protocols for backbone discovery and video dissemination in high mobility networks and highway VANETs available in the literature.

Heissenbuttel et al. introduced the idea of Dynamic Forwarding Delay (DFD) for relay node selection in the Beaconless Routing protocol (BLR) [7] over MANETs. In the contention phase, the source node broadcasts data packets for neighboring nodes. The relay node candidates, within the source node's forwarding zone, compute their DFD by using location information before forwarding the received packets. The node closest to the destination generates the shortest DFD and wins the contention phase. By forwarding the packets first, the winner node suppresses the transmission of other nodes and establishes itself as the next forwarding node. BLR operates well in terms of packet delay requirements. However, a basic limitation of this protocol consists of its reliance on a single metric to compute the DFD, reducing the network reliability for long data transmission, such as live video streaming.

Rezende et al. proposed a RB and backbone-based geographic routing approach for V2V video transmissions called VIdeo Reactive Tracking-based UnicaSt Scheme (VIRTUS) [5]. VIRTUS uses a location-based bayesian model for predicting where vehicles are going to move, and thus they can build a backbone by means of a contention-based phase, which considers such predictions. VIRTUS relies on location information and a countdown timer to discover short paths. However, this protocol do not take into account link quality factors and do not combine video-awareness to build reliable backbones, therefore it do not ensure QoE in the received video streaming. Furthermore, it is important to evaluate new video-based routing schemes based on QoE metrics.

Rosário et al. proposed the Link quality and Geographical Opportunistic Routing (LINGO) protocol, which uses multiple criteria to compute the countdown timer [10]. Felice et al. presented a geocast and contention-based protocol for real-time video delivery in VANETs, namely Dynamic Backbone Assisted (DBA) MAC protocol [3]. In DBA-MAC, the formation of the backbone also takes into account several criteria, such as link quality, vehicles location, speed, and direction. DBA-MAC also requires beacons and ACKs to work. Despite LINGO and DBA-MAC consider the link quality for routing decisions, these protocols do not take into account QoE-indicators for relay node selection and backbone maintenance. Further, these approaches use only one sample of packet to define the best forwarder nodes, which can cause false-positive results on the link quality measurement.

From our related work analysis, a beacon-less RB approach is a promising solution for vehicular multimedia applications, since vehicles do not need to proactively broadcast beacon messages to be aware of their neighbors. In addition, the existing RB protocols do not efficiently combine link quality, location information, and QoE-awareness to build reliable backbones. All of these key features are not offered in a unified beacon-less RB approach so far, thus existing proposals lack of robustness and QoE-awareness.

III. THE QOE-DRIVEN AND LINK-QUALITY RECEIVER-BASED PROTOCOL (QOALITE)

This section describes the proposed protocol to manage real-time videos in highway VANETs. QOALITE considers a beacon-less RB approach to select relay nodes and build backbones, while exploring multiple criteria, including link quality, mobility information, and QoE-indicators. QOALITE considers two phases, namely the Distributed Contention-based Forwarding (DCF) and Multi-hop Backbone Forwarding (MBF) phases. In DCF, relay node candidates start the video transmission and compete to choose which nodes will participate in the MBF phase from source to destination. Whereas, MBF provides a contention-free forwarding of the video stream, exploiting the previously built backbone, and allowing dynamic changes to other paths in case of link failures and loss of quality. Each phase will be presented in detail in the following subsections.

We consider a VANET environment where k vehicles (nodes) containing an identifier $(i \in [1, k])$, are moving with a speed s comprised between a minimum (e.g., s_{min}) and a maximum (e.g., s_{max}) speed limit, over a multi-lane highway area. The combination of these nodes create a dynamic graph $G(V, E)$, where vertices $V = \{v_1, v_2, ..., v_k\}$ mean a finite subset of k nodes, and edges $E = \{e_1, e_2, ..., e_m\}$ mean a finite set of asymmetric wireless links between them. We denote a subset $N(v_i) \subset V$ as all 1-hop neighbors within the radio range of a given node v_i. Further, each node v_i has an IEEE 802.11p-compliant radio transceiver, through which it can communicate with $N(v_i)$, a GPS (for location awareness), a multimedia encoder/decoder, and a transmission buffer (*TB*) with a maximum queue capacity (TB_{Max}).

Video sequences consist of media streams with different spatial temporal video features. MPEG standard defines that the Group of Pictures (GoP) is composed of a combination of three frame types, namely I (Intra), P (Predictive), and B (Bidirectionally predictive) frames. The successive frames between two I-frames define a GoP (the typical minimum and maximum values of a GoP in MPEG-4 videos are between 10 and 20). Because of the complex frame structure of MPEG videos, the same degree of packet loss may cause severe

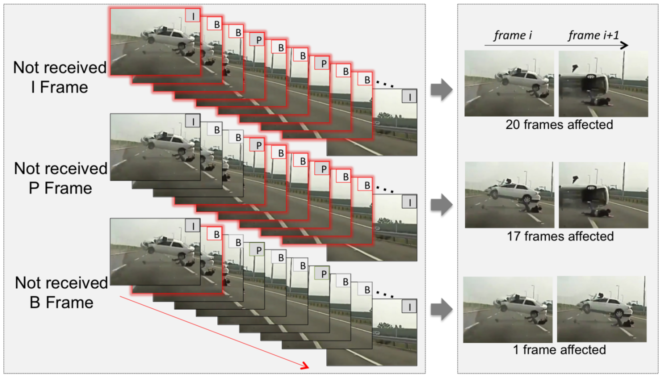


Fig. 1: Different priority of frames

quality degradation or may pass unnoticed, depending on which frame types are affected. Fig. 1 depicts the different degrees of importance for the user's perception in each MPEG frame type for a GoP with size 20. I-frames are the most important ones from the human point-of-view. For a single I or P frame lost, there is an error propagation through frames until the end of the GoP. On the other hand, for a single B frame lost, the impact is not noticed visually, since no others GoP frames are affected.

We represent a video flow $VF = \{p_1, p_2, ..., p_n\}$ as a set of n video packets p. Each p contains, in addition to the data payload, other encoder parameters, such as frame type, Id, length, timestamp, and packet segmentation. To obtain this information, a Deep Packet Inspection (DPI) algorithm enables extraction of the frame type and intra-frame dependency information for each p, since each VF starts with a sequence header followed by a GoP header, and then by one or more coded frames. Whenever a source node (VS) has a VF to send, it first determines the position of the destination node (VD), stores these geographical coordinates along with its own current position in the header of the packet, which contains: $< VS_{Id}, VS_{x,y}, VD_{Id}, VD_{x,y}, p_{information} >$, and triggers the DCF phase by broadcasting video packets to $N(VS)$. We consider that VS receives updates from VD periodically through location schemes, such as those proposed in [7].

A. Distributed Contention-based Forwarding (DCF) phase

The DCF phase employs the task of forwarding the video flows through a contention distributed stage. In this phase, VS starts the video transmission and relay nodes (VR) compete among themselves to choose which nodes will participate in the MBF phase. By using three criteria in the DCF phase, namely location information, current link quality conditions, and QoE-indicators, QOALITE defines the best relay vehicle in each hop. Once VS begins to capture a given VF_i, it starts to transmit the stream to VD in a multi-hop fashion. In other words, VS broadcasts video packets in a *Time Window*, denoted by $W(VF_i) \subset VF_i$, to all the 1-hop neighbors of VS ($N(VS)$). The generic process for DCF phase can be visualized through the Algorithm 1.

Upon receiving a $W(VF_i)$ from a previous transmitting node v_a, the nodes $N(v_a)$ have v_a, VD, and their own position information, extracted from the packet headers. Thus, a given node $v_b \in N(v_a)$ can easily determine when it is located within the set of nodes in the Forwarding Zone (FZ) of v_a ($FZ(v_a) \subset N(v_a)$) (Line 1 of Algorithm 1), that corresponds to an angle α of the line connecting v_a and VD. The concept of FZ becomes important, since it limits the selection of VR to a given sector, avoiding loops and disconnected VRs that are not able to cancel transmissions from they neighbors. We have defined $\alpha = 45^\circ$ to FZ [5].

Algorithm 1 DCF phase

```
    When a given node $v_b \in N(v_a)$ receives broadcasted packets ($W(VF_i) = \sum_{k=1}^{n} p_k$) from a transmitting node $v_a$:
 1: if $v_b \in FZ(v_a)$ then
 2:     if $\exists!\, p_k \in W(VF_i)$ and $p_k \in TB_{v_b}$ then
 3:         Drop $W(VF_i)$ from $TB_{v_b}$
 4:         return
 5:     else
 6:         Compute $FF(v_b)$ (Eq. (7))
 7:         Start $CountdownTimer(v_b)$ (Eq. (1))
 8:         while $CountdownTimer(v_b) \neq 0$ do
 9:             if Overhear $!\, p_k \in W(VF_i)$ then
10:                 Cancel $CountdownTimer(v_b)$
11:                 Drop $W(VF_i)$ from $TB_{v_b}$
12:                 Cancel any subsequent rebroadcast of $p_k \in W(VF_i)$
13:                 return
14:             end if
15:         end while
16:         Broadcasts $W(VF_i)$
17:         $v_b \leftarrow VR_\alpha$
18:     end if
19: else
20:     Drop $W(VF_i)$
21:     return
22: end if
```

If $W(VF_i)$ contains only new received packets (Line 2 of Algorithm 1), nodes located within $FZ(v_a)$ apply the *Fit Function* (FF) (Eq. (7)) prior to relay the packets, conversely nodes outside this area drop the received packets from *TB*. The employ of FF allows VR candidates to mitigate the number of retransmissions inside FZ by choosing only one best relay node, i.e., $VR_\alpha \in FZ(v_a)$. The value of FF $[0, FF_{max}]$ depends on link quality, VR location, and QoE-indicators, as shown in Subsection III-B.

After the calculation of FF, VR candidates must replace the location of v_a by their own locations in the packet header, set a $CountdownTimer$ according to Eq. (1), and wait for the timeout to rebroadcast the buffered packets ($W(VF_i)$). It is easy to see that nodes with higher values of FF are mapped to smaller $CountdownTimer$ sizes, and thus having higher probability to win in the DCF phase.

$$CountdownTimer = CW_{Max} - FF \cdot (CW_{Max} - CW_{Min}) \tag{1}$$

The CW $[CW_{Min}, CW_{Max}]$ is the size of the Contention Window in the 802.11p standard. In a completely distributed manner, the node that generates the smallest $CountdownTimer$, rebroadcasts $W(VF_i)$ first and it is selected as VR_α (lines 16 and 17 of Algorithm 1). As a

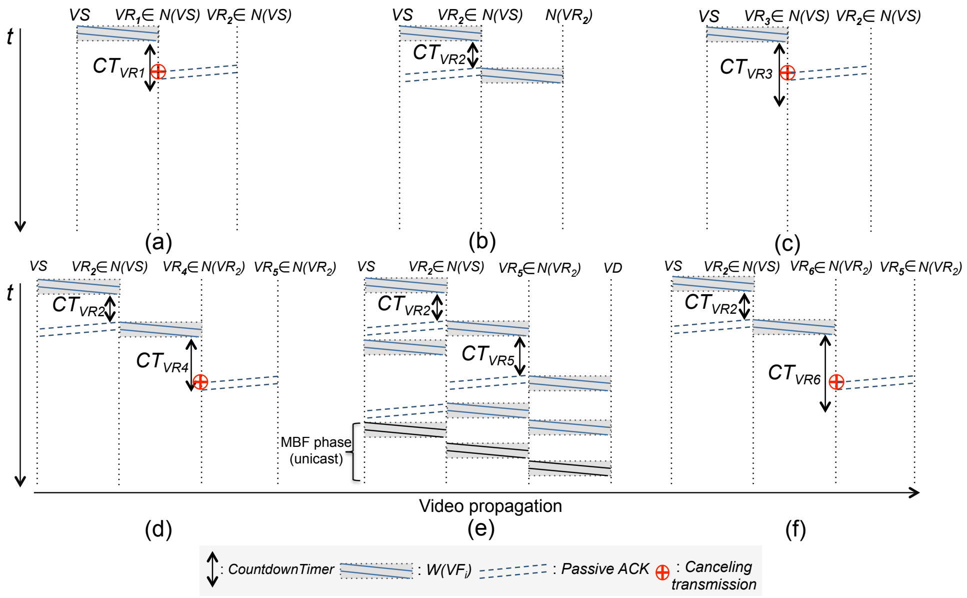


Fig. 2: DCF phase (a), (b), and (c) for the first hop ($N(VS)$), and (d), (e), and (f) for the second hop ($N(VR_2)$)

result, QOALITE provides a reliable route between VS and VD via multiple VR_α. Moreover, as expected in the IEEE 802.11p standard, VR candidates are able to sense the channel during the $CountdownTimer$, so that they will cancel their transmissions attempt in case they overhear transmissions of any $p \in W(VF_i)$ from other nodes. Every node in the $FZ(v_a)$ detects further relaying of $W(VF_i)$ and cancels its $CountdownTimer$ of the same packets. Then, VR_α repeats the same procedure until VF_i reaches VD. Moreover, through a *passive acknowledgment* approach [11], v_a also is able to overhear the further relaying of $W(VF_i)$ and, thus, concluding that it was successfully received by another node in the $FZ(V_a)$, allowing QOALITE to reduce the acknowledgments on the MAC-layer.

Fig. 2 depicts the general overview of the DCF phase of QOALITE. Suppose that $N(VS) \cap FZ(VS) = \{VR_1, VR_2, VR_3\}$, i.e., nodes VR_1, VR_2, and VR_3 are located in the Forwarding Zone of the source vehicle. In this example, VR_2 forwarded $W(VF_i)$ first (Fig. 2.b) and the neighboring nodes (VR_1 and VR_2) overhear this transmission (2.a and 2.c). As a result, these nodes cancel their *CountdownTimers*, delete the buffered packets, and stay in contention mode. For the second hop (Figures 2.d, 2.e, and 2.f), suppose that $N(VR_2) \cap FZ(VR_2) = \{VR_4, VR_5, VR_6\}$, i.e., VR_4, VR_5, and VR_6 are located in the Forwarding Zone of VR_2. Initially, VR_5 wins the contention phase by calculating the smallest *CountdownTimer* and forwards $W(VF_i)$ first. Then, VR_2 uses the transmitted packets as *passive acknowledgement*. Meanwhile, the neighbors of VR_5 (VR_4 and VR_6, from Figures 2.d and 2.f, respectively) overhear VR_5 transmitting $W(VF_i)$, in this way, these nodes cancel their *CountdownTimers*, and stay in contention mode. Finally, VD receives $W(VF_i)$ and VS unicasts subsequent packets to VD by VR_2 and VR_5 i.e., it switches to Multi-hop Backbone Forwarding phase that will be detailed in Subsection III-C.

B. Criteria for forwarding selection

For each VR candidate performing the FF calculation, we have defined three criteria, namely: link quality (C_{LQ}), location information (C_L), and QoE-indicators (C_{QoE}). Thus, QOALITE selects a given VR_α to forward the packets if it has a reliable link, when it is close to VD, and finally, if it has enough QoE-indicators to forward the video flow. Here, we present these three criteria, which are optimized in three different utility functions.

1) Link Quality: QOALITE considers link quality as part of the FF function to define routes, which ensures that the chosen relay node (VR_α) provides multimedia transmission with higher packet delivery guarantees. Therefore, each link e_j has a C_{LQ_j} associated, which represents a single value for the link quality computed at the received side. Through C_{LQ} we attempt to maximize the robustness of the communication among the vehicles. For this reason, we prefer as VR_α of a vehicle v_a, the node receiving the highest value of C_{LQ}, i.e., experiencing the best propagation conditions with vehicle v_a. Eq. (2) expresses C_{LQ} for a given node v_b receiving broadcasted packets ($W(VF_i)$) from a transmitting node v_a:

$$C_{LQ}^{v_b}(W(VF_i)) = \frac{\left(\sum_{i=0}^{n} \frac{P_{RX}^{v_a}(i)}{i}\right) - RS_{thr}}{|RS_{thr}|} \tag{2}$$

where $P_{RX}^{v_a}$ is the power (in dBm) of each message received in $W(VF_i)$ from vehicle v_a and RS_{thr} is the Receiver Sensitivity (in dBm) of its wireless interface. In contrast to previous works that use only one packet or beacon message to calculate link quality, QOALITE computes an average of the perceived signal strength coming from a set of packets in $W(VF_i)$. This way, our proposal establishes more reliable measurements, avoiding false outliers.

2) QoE-Indicators: As seen previously, a MPEG video sequence is composed of I, P, and B-frames, each one having different degrees of importance for the user's perception. The loss of one I-frame causes severe video distortion and error propagation through the other frames within a GoP (Fig. 1). In a typical MPEG-4 video with 24fps, if packets from only one I-frame are lost, the video will be degraded for 0.75s or more. Also, the loss of P-frames at the beginning of a GoP causes a higher video distortion than loss at the end of a GoP. On the other hand, the loss of B-frames only affects the video quality of that particular frame and does not impact heavily on the user's perception. By considering the importance of each video frame, as well as the P-frame position within the GoP, QOALITE prioritizes frames with a greater impact on the average video distortion σ_s^2, opposed to those with lower impact on the perceived QoE. Therefore, it assigns different weights to slices s of packets belonging to each frame as modeled by Eq. (3):

$$\sigma_s^2 \propto \begin{cases} \frac{\alpha_1(R_M - R_I)}{R_M} & \text{if s} \in \text{I-frame} \\ \frac{\alpha_2}{(2^{T-1}-1)R_M} \sum_{i=1}^{T-1} 2^{T-1-i}(R_M - R_{P_i}) & \text{if s} \in \text{P-frame} \\ \frac{\alpha_3(R_M - R_B)}{R_M} & \text{if s} \in \text{B-frame} \end{cases} \tag{3}$$

Where $T-1$ is the number of P-frames per GoP, R_I, R_{P_i}, and R_B means the I, P (with position i in the GoP), and B-frame received rate in $W(VF_i)$, respectively, and R_M is the maximum data-rate supported by the radio transceiver of each vehicle. For instance, for a DCMA-86P2 802.11p WiFi card we have: R_M = 6 Mbps if $P_{rx} >$ -93dbm [12]. The parameters α_1, α_2, and α_3 are weighting factors, where $\sum_{i=1}^{3} \alpha_i = 1$.

The distortion model proposed by Shu Tao [8] considers the impact caused by the loss of single slices of a frame from a video stream. Thus, for a given video frame structure and a probability of occurrence of loss, Eq. (4) defines an overall distortion value for the whole stream:

$$\overline{D} = sL \cdot \overline{n}P_e \cdot \sigma_s^2 \cdot \left(\frac{\gamma^{-t+1} - (T+1)\gamma + T}{T(1-\gamma)^2} \right) \tag{4}$$

Where s and L are the number of slices per video packet and the number of packets per frame, respectively obtained from the standard video codec configurations and packetization. The parameter $\overline{n}$ represents the loss burstiness ($1.06 \geq \overline{n} \geq 1$ for Bernoulli losses, depending on the aggressiveness of the burst errors). The attenuation factor γ ($\gamma < 1$) accounts for the effect of spatial filtering, and varies as a function of the video characteristics and decoder processing. P_e is the probability of loss events (of any length) in the video stream. Both, γ and P_e are given by the effect of the loss pattern experienced by the video stream and the codec's error concealment mechanism. Finally, the Mean Square Error (MSE) distortion $\overline{D}$ provides a QoE-estimate by using a non-linear relation that measures the video quality level by comparing distortions caused by slice losses [8], according to each frame type, denoted at Eq. (5):

$$C_{QoE}^{v_b}(W(VF_i)) = \frac{1}{1 + exp(b_1 \cdot 10 \cdot log_{10}(255^2/\overline{D}) - b_2)} \tag{5}$$

Where, the parameter b_1 determines the slope of the QoE mapping curve, whereas b_2 establishes the central point. By considering 40dB as the highest video quality, and the lowest video quality for values below 20dB, the values of b_1 and b_2 are given by 0.5 and 30, respectively. Based on the average distortion caused by losses in the different frame types in $W(VF_i)$ it is possible to relay node candidates to compute a higher FF to the nodes receiving the most important packets.

3) Location Information: This criterion attempts to minimize the number of hops, since usually longer routes reduce the packet delivery ratio. For this reason, it is selected the VR_α closer to the VD. Hence, VR with small geographical distance towards VD generate higher FF values. The C_L for a given v_b is computed according to Eq. (6):

$$C_L^{v_b}(W(VF_i)) = \frac{E\left[d(v_a, v_b)\right]}{R} \tag{6}$$

Where $E\left[d(v_a, v_b)\right]$ is the geographical average distance between vehicles v_a and v_b calculated through the GPS information piggybacked at the packet headers, and R is the maximum transmitting range of a vehicle v.

Upon computing C_{LQ}, C_{QoE}, and C_L, each RF has an array D $= \{C_{LQ}, C_{QoE}, C_L\}$, containing each one of the previous calculated criteria ($|D| = 3$). As a result, considering the different weights $\omega_p | \sum_{p=1}^{|D|} \omega_p = 1$, Eq. (7) calculates FF by multiplying the values d_p in D and the weights of evaluation criteria, similarly to others multi-criteria approaches [10]:

$$FF = \sum_{p=1}^{|D|} (d_p \times \omega_p) \tag{7}$$

C. Multi-hop Backbone Forwarding (MBF)

As video transmission are often long (e.g., 20 s), whenever a route from VS to VD is established, nodes transmit video packets explicitly addressed to VD without any additional delay and in a unicast fashion (Fig. 2.e). Thereby, QOALITE mitigates additional delays, interferences, and duplicate packets derived by DCF phase by introducing the Multi-hop Backbone Forwarding (MBF) phase. During the transmission, the video content should be delivered even in presence of node failures or channel variations. These issues cause interruptions, being undesirable for the user's experience. QOALITE detects route failures, providing a smoother route management. In particular, QOALITE considers that every node that composes a route $P(VS, VD)$ should perceive whether it is still a reliable or valid route to transmit packets. This is achieved by receiving reply messages. We define a control packet, called Peer Quality Message (PQM), which contains the link quality (LQ) and $\overline{D}$ perceived by each forwarder. Thus, if a given RF_α^2 receives a video flow from RF_α^1, it must compute LQ and $\overline{D}$ perceived in each $W(VF)$ and send a $PQM_{W(VF)}$ to RF_α^1.

Any built backbone returns to the DCF phase, when it detects that the link quality falls below a predefined link quality threshold (LQ_{thr}) or a video distortion threshold ($\overline{D}_{thr}$). In addition, any node that composes a transmitting backbone considers that the route is not valid anymore, as long as it does not receive any reply message from its previous RF_α node within a certain period of time, i.e., time-out $= 0.5s$. Hence, it returns to the DCF phase to re-establish a new backbone.

IV. PERFORMANCE EVALUATION

This section shows the methodology and metrics used to evaluate the quality level of transmitted video flows, where the performance on QOALITE is compared to main related work. In order to establish a relevant scenario, we have considered a 10 Km portion of the San Diego Freeway from OpenStreetMap, which was imported into SUMO (Simulation of Urban MObility). It allows us to reproduce the desired vehicle movements, with realistic behavior and V2V interactions according to empirical data. Each video flow VF must be received by the VD at a distance lower than 1 km from the VS, providing a hop limit, as proposed in [3].

Also, aiming for realistic results, we have used EvalVid - A Video Quality Evaluation Tool-set that allow us evaluating the video streaming quality. In this way, we have conducted the experiments by transmitting real MPEG-4 video sequences (720 x 480 pixels), available in [13], with 768 kbps and 24 fps, duration varying from 10s to 60s, and internal GoP structure configured as two B-frames for each P-frame. The videos and the road/vehicle characteristics were integrated into Network Simulator version 2.33. More details about the simulation parameters are shown in Table I.

TABLE I: Simulation parameter values

Parameter	Value
Scenario (San Diego Freeway)	12Km - 5 lanes
Speed Range	20 to 30 m/s
Radio Range	300
TB_{Max}	30 pkts
Avg. vehicle density (veh/km)	50 to 200
Simulation Time	700s
MAC Layer	IEEE 802.11p DCMA-86P2
Bandwidth	6Mb/s
Atenuation Model	Nakagami Dist.

Each plotted result in the graphs is the average of results generated from 35 simulations: we diversified the vehicle density levels on the roads (50, 100, 150, and 200 nodes/km), and also the included videos (1 per simulation) with typical Internet-based GoP sizes (14 and 20). In our simulations, VS sends the video sequence at any time after the initial 100s and before the last 100s of the simulation. Finally, after receiving the frames, the decoder uses Frame-Copy as the error concealment method. The confidence interval was calculated with a 95% confidence level.

To demonstrate the impact of QOALITE in delivering QoE-aware video flows in VANETs, we use BLR [7], VIRTUS [5], and DBA-MAC [3] for comparison. The BLR and VIRTUS protocols use distance and location based forwarding mechanisms. For BLR the farther away the node, the shorter the $CountdownTimer$. VIRTUS builds a backbone through a bayesian model for determining where vehicles are going to be. Thus, VIRTUS take into account distance to destination, direction, and speed of the vehicles. Regarding DBA-MAC, it takes into account location information and link quality for the FF calculation. These protocols have been improved with respect to their originals, thus all of them are aware of the desired message propagation direction. Moreover, we adjusted them with the MBF phase in order to mitigate the forwarding contention timer: once a vehicle successfully transmits video packets, its timer for the next packets will be minimum. Further, DBA-MAC was extended to a beacon-less version. We introduced these improvements because the standard protocols, as they were, did not represent a fair comparison.

The I, P, and B-frame weights (α_1, α_2, and α_3) affect the QOALITE performance. We have conducted independent empirical evaluations and we concluded that $\alpha_1 = 0.65$, $\alpha_2 = 0.3$, and $\alpha_3 = 0.05$ give the best C_{QoE} results. Furthermore, the weights for each criterion ω_1, ω_2, and ω_3 were fixed in 0.4, 0.25, and 0.35, respectively. This is because QOALITE achieved the best trade-off between lowest number of hops to VD together with reliable links and enough QoE-indicators to forward the packets with an acceptable video quality. In addition, we set CW_{Max} to 100 ms, $W(VF)$ to 80 ms, LQ_{thr} to 0.6, and $\overline{D}_{thr}$ to 0.75.

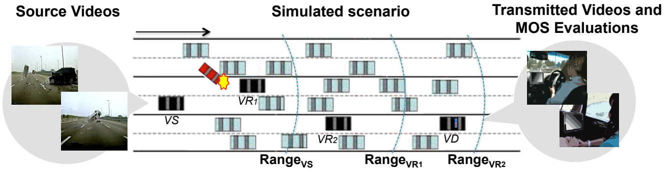


Fig. 3: Evaluated environment

The impact of the proposed solutions were measured by the following QoS metrics: Packet Delivery Rate (PDR) and End-to-end Delay. Since measuring the human experience while watching video sequences is a key requirement for our work, measurements were carried out with Structural SIMilarity (SSIM) and Mean Opinion Score (MOS). SSIM is a well-known objective QoE metric, which measures the structural distortion of the video to obtain a better correlation with the user's subjective impression. For the subjective experiments (MOS), an Android application [14] was used (following the ITU-R recommendations) to playback the transmitted videos and collect their evaluations. We used the Single Stimulus (SS) method of ITU-R BT.500 in the tests, and 35 subjects with ages ranging from 18 to 40 were invited to participate in the process. After watching a video, the viewer assess the video quality level by selecting a score ranging from 1 to 10, where 1 means poor quality and 10 means excellent quality. The distorted videos were played on a Samsung Galaxy Tab 3 with a 8.4-inch display that better represents a vehicle context and conditions,e.g., a display placed on the back seat of a car (Fig. 3). These metrics (SSIM and MOS) determine the behavior of QOALITE regarding DBA-MAC, VIRTUS, and BLR. It also presents the results with the variations in number of vehicles, video codification, and the quality of the video delivery.

Regarding network performance, PDR and average delay, Figures 4 and 5 show the performance results for the four simulated protocols. From Fig. 4, QOALITE and DBA-MAC notably outperform VIRTUS and BLR protocol in terms of received rate. On average, DBA-MAC increases the received rate by 10.8% and 7.6% compared to BLR and VIRTUS, respectively, while QOALITE achieves delivery rates of 1.7%, 12.5%, and 9.9% compared to DBA-MAC, BLR, and VIRTUS, respectively. The highest probability of reception occurs due to selection criteria of QOALITE and DBA-MAC. For these protocols, the forwarding vehicles are (potentially) close vehicles that experience more stable link connectivity. Thus, BLR can face several broken link situations, especially when there are few available vehicles in the neighborhood. VIRTUS reaches a PDR slightly higher in comparison to BLR, i.e., around 3.8%. This occurs due to the fact that VIRTUS predicts a contact time between neighbor nodes. Nevertheless, this scheme does not consider link quality, and so, even for nodes with greater contact time, there may still be interference or fading.

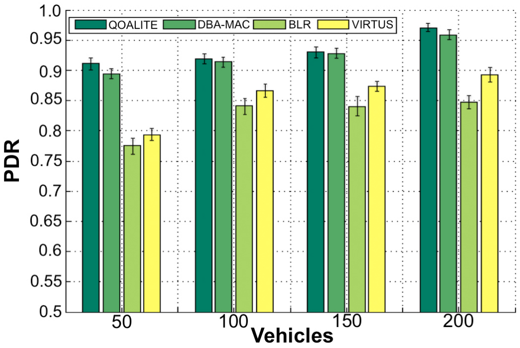


Fig. 4: Average Packet Delivery Rate

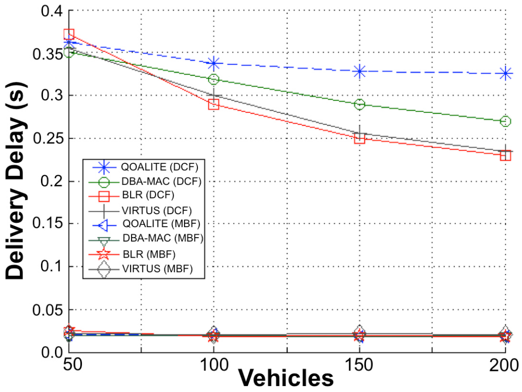


Fig. 5: Average Delay

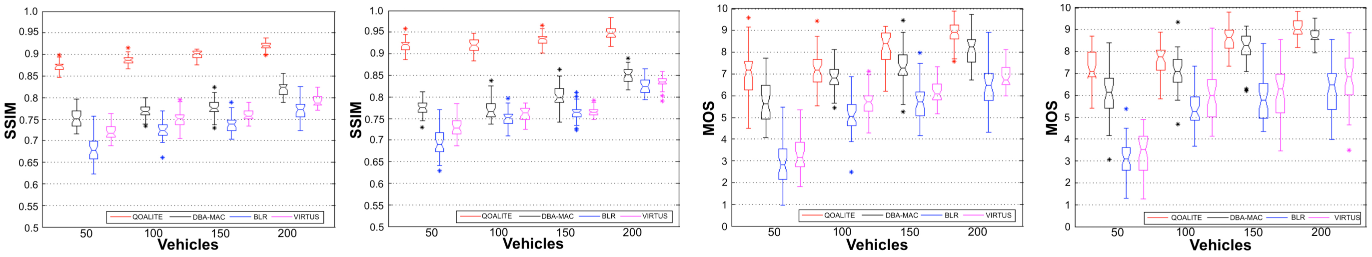
The impact of the DCF and MBF phases are significant on the delivery delay over the transmission (Fig. 5), since the MBF phase allows a great reduction of the delivery delay with the contention-free forwarding. We consider the average delay required by video packets from a $W(VF)$ to be transmitted in a route between VS and VD with 1 Km of length. The BLR protocol at DCF phase, in a low-density scenario, experiences a superior delivery delay, despite the reduced number of hops. This problem is caused by the settings of the *CountdownTimer*, that is adjusted to the CW_{Min} only for relay node candidates located at the maximum transmission range from the sender. Thus, in a low-density scenario, it may frequently happen that the (farthest) forwarding vehicle uses a *CountdownTimer* higher than CW_{Min}, for each hop, resulting in a higher end-to-end delay. Conversely, in high dense scenarios, despite more vehicles competing for getting the channel access, the farthest vehicle is more likely placed near the transmission range border, and thus it sets its *CountdownTimer* close to the CW_{Min} value. Hence, the delivery delay of the protocols at DCF phase, decreases when the number of vehicles increases.

Also, with respect to the delivery delay of each protocol at the DCF phase, it is slightly different mainly when BLR is compared with QOALITE and DBA-MAC. The delay reduction provided by the BLR protocol over QOALITE and DBA-MAC is around 22.1% and 11.7% respectively for 150 nodes, and 25.1% and 12.8% for 200 nodes in the network. As mentioned previously, the QOALITE protocol, according to its forwarding criteria, provides more effort to deliver flows with better quality, this could mean forwarding streams to alternatives nodes with increasing transmission durations. On the other hand, BLR just tries to minimize the number of hops and will not make any effort. This might result in longer delays and path lengths for QOALITE. However, delay levels are much lower than one second, which are negligible even in video applications and are significantly lower than requirements of 4 to 5 seconds defined by CISCO [15].

QoS-based metrics are not enough to measure the video quality level. Thus, to understand the impact of the QoE-Indicators criterion, the results in Fig. 6 present the SSIM and MOS metrics. SSIM values range from 0 to 1, where a higher value means better video quality. In Figures 6.a and 6.b, with 14 and 20 frames by GoP, QOALITE keeps the SSIM values around 0.91 and 0.94, respectively. An average increase of 14.5% compared to DBA-MAC, 17.3% compared to VIRTUS, and 19.8% compared to BLR. It presents more deeply results than those obtained in the Fig. 4 and shows significant benefits to user's experience. This occurred because QOALITE can estimate when the quality of the transmitting flow decreases based on the different received frame types, codec configurations, and losses, allowing vehicles, to switch to others nodes, before increasing damage on the flow quality. For instance, when a given part of video flow is successfully received for a VR_i candidate in $|W(VF_i)|$ milliseconds, and since the spatial distribution of vehicles that does not change very quickly in a short period of time (e.g., 3s), it is likely that VR_i, continue to receive successfully a greater number of packets until a new route becomes necessary. Thus, QOALITE provides a trade-off between hop-length and video quality.

With respect to real-time video assessment, sometimes the correlation between the SSIM results and subjective scores does not have a high accuracy. Thus, it is fundamental to have MOS experiments in order to really understand the video quality level according to human perception [2]. The results in the Figures 6.c and 6.d reinforce the results in the Figures 6.a and 6.b, respectively. Besides that, these results demonstrate that the QoE-indicators metric can be successfully extended from the distortion model introduced by Shu Tao in [8] and performs well when employed in QOALITE. Figures 6.c and 6.d present the average MOS scores for all protocols when the number of vehicles varies and confirms that QOALITE allows the delivery of live video sequences with a good or excellent quality in networks with 50 and 100 vehicles.

We randomly selected one sample frame from one of the transmitted videos [13], aiming to give the reader an idea of the user's point-of-view, as illustrated in Fig. 7. The frame # 328 is the moment when a person is thrown out of the vehicle, which can be very important for rescue teams. The transmitted sequences using QOALITE have low distortion compared to the same frame sent using BLR, VIRTUS, and DBA-MAC. This is because QOALITE establishes a reliable backbone, and, according to QoE-Indicators criteria, QOALITE prioritizes



(a) Metric: SSIM; GOP size: 14 (b) Metric: SSIM; GOP size: 20 (c) Metric: MOS; GOP size: 14 (d) Metric: MOS; GOP size: 20

Fig. 6: Metrics SSIM and MOS for all protocols with different GoP lengths

(a) QOALITE (b) DBA-MAC (c) VIRTUS (d) BLR

Fig. 7: Comparison sample for the different simulated protocols at frame 328

nodes receiving the first slices (packets) of the GoP. It becomes important, since in accident videos, losing the first slice in a frame typically causes more distortion than losing the second slice, since there is typically more motion in the top half of its frames [8]. Apart from the distortions on the frame transmitted via BLR and VIRTUS, the person does not appear clearly, compared to the frame transmitted via QOALITE and DBA-MAC. This may impair the action of rescuers.

V. CONCLUSION

This paper introduced QOALITE to enable an efficient real-time video delivery with QoE-awareness in VANETs. QOALITE aims to share videos with a better quality level than existing works, it employs QoE-indicators criterion that support the selection of the best next hop and switches to other routes as soon as lower quality is identified. Results highlight the performance and QoE-awareness support of QOALITE by measuring the video quality levels when the number of vehicles and video features varies. By creating a backbone and according to its forwarding criteria, QOALITE provides a greater support to a self-organized delivery of streams with a higher quality from the user's point-of-view. This could mean forwarding of streams to alternative nodes (increasing transmission durations), but nonetheless, still are insignificant to real-time video requirements. In future works, we will perform a study taking into account multiple flows, so that, the time window can be dynamically adjusted to avoid collisions.

ACKNOWLEDGMENT

This work was funded by the National Council for Scientific and Technological Development (CNPq), Coordination for the Improvement of Higher Education Personnel (CAPES), and Fundação Amazônia Paraense de Amparo à Pesquisa (FAPESPA)

REFERENCES

[1] C. V. N. Index, "Global mobile data traffic forecast update, 2012–2017, cisco white paper, feb. 6, 2013."

[2] E. Aguiar, A. Riker, E. Cerqueira, A. Abelém, M. Mu, M. Curado, and S. Zeadally, "A real-time video quality estimator for emerging wireless multimedia systems," *Wireless Networks*, pp. 1–18, 2014.

[3] M. Di Felice, L. Bedogni, and L. Bononi, "Group communication on highways: An evaluation study of geocast protocols and applications," *Ad Hoc Networks*, vol. 11, no. 3, pp. 818–832, 2013.

[4] C. Xu, F. Zhao, J. Guan, H. Zhang, and G.-M. Muntean, "QoE-driven user-centric VoD services in urban multihomed P2P-based vehicular networks," *Vehicular Technology, IEEE Transactions on*, vol. 62, no. 5, pp. 2273–2289, 2013.

[5] C. Rezende, H. S. Ramos, R. W. Pazzi, A. Boukerche, A. C. Frery, and A. A. F. Loureiro, "Virtus: A resilient location-aware video unicast scheme for vehicular networks," in *Communications (ICC), 2012 IEEE International Conference on*. IEEE, 2012, pp. 698–702.

[6] C.-J. Hsu, H.-I. Liu, and W. K. Seah, "Opportunistic routing–a review and the challenges ahead," *Computer Networks*, vol. 55, no. 15, pp. 3592–3603, 2011.

[7] T. Braun, M. Heissenbüttel, and T. Roth, "Performance of the beaconless routing protocol in realistic scenarios," *Ad Hoc Networks*, vol. 8, no. 1, pp. 96–107, 2010.

[8] S. Tao, J. Apostolopoulos, and R. Guérin, "Real-time monitoring of video quality in ip networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 16, no. 5, pp. 1052–1065, 2008.

[9] T. Aguilar, S.-J. Syue, V. Gauthier, H. Afifi, and C.-L. Wang, "Coopgeo: A beaconless geographic cross-layer protocol for cooperative wireless ad hoc networks," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 8, pp. 2554–2565, 2011.

[10] D. Rosário, Z. Zhao, A. Santos, T. Braun, and E. Cerqueira, "A beaconless opportunistic routing based on a cross-layer approach for efficient video dissemination in mobile multimedia iot applications," *Computer Communications*, vol. 45, pp. 21–31, 2014.

[11] M. Heissenbüttel, T. Braun, T. Bernoulli, and M. Wälchli, "Blr: beaconless routing algorithm for mobile ad hoc networks," *Computer communications*, vol. 27, no. 11, pp. 1076–1086, 2004.

[12] D. Unex, "86p2: industrial grade, high power 5.86˜ 5.92 ghz wifi mini-pci module for 802.11 p/dsrc application, ar5414a-b2b," *Online: http://www. unex. com. tw/product/dcma-86p2*.

[13] "Videos used in the MOS experiments ('truck accident on highway' and 'car accident on a freeway') encoded with GoPs size 14 and 20. URL https://www.youtube.com/channel/uculjslvbpejblar6isvxbja/videos."

[14] "VEMOS application Vs. 4.0. URL http://gercom.ufpa.br/downloads/multimedia."

[15] C. Hatting *et al.*, *End-to-end qos network design*. Cisco Press, 2005.