Streaming Media Distribution in VANETs

Fabio Soldo, Claudio Casetti, Carla-Fabiana Chiasserini
Politecnico di Torino, Italy
Email: lastname@tlc.polito.it

Pedro Chaparro
Technical University of Catalonia, Spain
Email: pedro.alonso.chaparro@upc.edu

Abstract—Streaming applications will rapidly develop and contribute a significant amount of traffic in the near future. A problem, scarcely addressed so far, is how to support streaming traffic in vehicular networks (VANETs). This problem significantly differs from previous work on broadcast and multicast in ad hoc networks, because of the highly-dynamic topology of VANETs and the strict delay requirements of streaming applications. Our solution, completely relies on inter-vehicular communication for the distribution of multimedia content in an urban environment and it has the following appealing features: (i) it is fully distributed and dynamically adapts to topology changes; (ii) it leverages the characteristics of streaming applications to yield a highly-efficient, cross-layer solution. After optimally dimensioning the network system, we compare the performance of our solution against theoretical results for broadcast capacity in multihop networks.

I. INTRODUCTION

It is commonly acknowledged that mobile ad hoc networks, and, in particular, Vehicular Ad Hoc Networks (VANETs), are ill-suited to support multimedia traffic. Low bandwidth, fleeting connectivity, and highly-dynamic, unpredictable topology are the main shortcomings hindering the support of real-time applications. The variable bit rate (VBR) nature of the traffic, the highly dynamic topology and the strict delay constraints, making no allowance for store-and-forward, pose a different problem from the ones previously addressed in broadcast ad hoc networks.

The issues related to the support of video streaming in VANETs have been previously addressed in [1]–[3]. The network architecture in [1] aims at propagating video streaming through forwarding nodes, in a highway scenario; however, unlike our case, the solution is tailored to traffic delivery from multiple sources to a single receiver. The work in [2] proposes an application-layer approach to deliver live video streaming, by exploiting a cluster-based network topology, in a highway environment. Although clustering is a viable approach for video distribution, the control traffic overhead for creating and maintaining the cluster structure may be significant, especially in highly-mobile scenarios. The study in [3] analyzes a slot-based scheme for delivery of streaming traffic, again, in a highway environment. Finally, relevant to our work is also [4], where some sensor nodes are selected as traffic forwarders and their transmissions are scheduled according to a TDMA scheme.

In this paper, we propose a fully distributed solution called SMUG (Streaming Media Urban Grid) that provides streaming media support in city VANETs. SMUG completely relies on inter-vehicular communication for the distribution of multimedia content, and for the definition of a grid-like structure that is laid over the physical topology of mobile nodes for media distribution. A media stream, generated in a point in space (e.g., at a roadside access point), is fed to SMUG-capable nodes and distributed across the VANET through the distribution structure, which is able to dynamically adapt to the changing topology. A SMUG node may either be (i) an active node, i.e., a node that belongs to the distribution structure and is responsible for forwarding the streaming media in a synchronized fashion, and for independently selecting other nodes as close as possible to ideal grid vertices; (ii) a passive node, i.e., a simple user who is within radio range of an active node and plays back the received streaming media.

Streaming media distribution in SMUG therefore occurs through a mix of local broadcasting from an active node to neighboring passive nodes, and MAC-layer multicasting from one active node to another. An active node exploits a built-in positional device (e.g., a GPS) to dynamically select its next-hop active nodes. GPS signals are additionally used to synchronize all active nodes to a structured TDMA transmission. Graph coloring techniques are used both to dimension the TDMA system and to solve the scheduling problem of neighboring active nodes, thus minimizing the chance of collisions in adjacent areas.

Clearly, the scheduling algorithm may sometimes fail and result in bandwidth waste due to collisions stemming from non-ideal grid selection. To overcome this, SMUG nodes are capable of (i) detecting a collision by means of passive acknowledgments at the MAC layer, (ii) reclaiming part of the wasted bandwidth to solve the contention and (iii) salvaging a failed transmission by sending a reduced amount of information. This approach (hereinafter called “opportunist access”) is particularly amenable to cross-layer interactions between the application and the MAC layer, such as the use of multiple-description video streams allowing a selective discard of descriptors.

Broadcasting in ad hoc networks has been widely studied. However, differently from previous work, we address these issues by offering a fully-distributed solution spanning several layers, from the application to the MAC layer.

II. NETWORK SCENARIO

Consider a VANET deployed in a urban environment. We assume that one or more gateway nodes, either fixed or mobile, provide streaming media to drivers (e.g., audio) and passengers...
Distribution of multimedia content completely relies on inter-vehicular communication; in addition, vehicles may wish to exchange best-effort data traffic in a peer-to-peer fashion: news summaries, public transportation timetables, traffic warnings and so on.

All network nodes are supposed to be equipped with a positioning system, such as GPS, so that they are aware of their location and are synchronized in time. We denote by $R_m$ the maximum node radio range and assume that vehicles keep updated knowledge of who their 1-hop neighbors are by in-band HELLO signaling messages, broadcast periodically. HELLO messages carry node IDs and GPS position.

Streaming media and best-effort traffic (the latter including traffic warnings and so on) are transmitted over a data channel, which is organized according to a TDMA structure. As shown in Fig. 1, the data channel is structured in fixed-length frames of duration $T_F$. Each time frame is further divided into $S$ identical slots, where the proper value of $S$ can be determined using graph-coloring theory.

Multimedia content is assumed to be a video sequence. Here, we consider the video to be encoded into multiple descriptors (namely, 3), although other techniques could be considered as well. Each descriptor is composed of several video frames (e.g., $I$, $B$ or $P$), of different size. The node protocol stack includes a Segmentation and Reassembly (SAR) layer, such that, at the transmitter, each video frame is segmented (if needed) and format into a packet that will cover up to one third of a MAC payload of maximum size. In other words, every MAC packet to be transmitted in a time slot carries three (or parts of three) video frames, each corresponding to a different descriptor. The SAR header also carries frame sequence numbers. At the receiver, the SAR layer reassembles (if necessary and possible) different parts of the same video frame and sends the video frame to the upper layers.

Note that, due to the VBR nature of video traffic, $I$ frames are typically very large, while, e.g., $P$ frames are small. This implies that, when an $I$ frame needs to be transmitted, one or more slots will be filled up; when, instead, $P$ frames are sent, a large portion of the slot will remain free and can thus be reused for best-effort traffic (see Fig. 1).

### III. The SMUG Protocol

The keypoints of the solution we propose are as follows.

- Active (i.e., forwarder) nodes are selected so as to maximize the coverage area of the streaming transmission; this results in a grid-like distribution structure that is laid over the physical topology of mobile nodes. The distribution structure is formed in a distributed manner, and dynamically adapts to topology changes.
- Then, the scheduling of the active nodes transmissions on the data channel is optimized by deriving some results on graph coloring and applying them to the grid-like distribution structure.
- Through passive acknowledgments, active nodes can detect collisions among concurrent streaming transmissions.
- Exploiting the properties of video coding, a single slot on the data channel can accommodate a scheduled transmission of multimedia content, or the contention for a transmission in case a previously scheduled transmission has failed, or both.
- If the multimedia content does not fill up the full slot length, a Contention Period (CP) is allocated in the slot residual part to allow best-effort traffic to be transmitted.

In the following, we provide a detailed description of the SMUG solution.

#### A. Children Selection and Scheduling Algorithm

Here, we describe the creation of the distribution graph and the scheduling of active nodes across it. For ease of description, we will take on the point of view of a generic active node called relay node. We refer to the active node scheduling the relay node and feeding traffic to it as parent node; we then refer to the active nodes being scheduled and fed traffic by the relay node as children nodes. The procedure is recurring, so the parent node and children nodes are expected to behave like a relay node in their own turn.

The scheduling of children nodes by the relay node occurs within a single time frame interval and requires the awareness of the surrounding environment, specifically: (i) the identity of its parent, (ii) the position of its 1-hop neighbors and (iii) the current scheduling status of the time frame (i.e., the slots already scheduled by its relay parent). The identity of children nodes changes in time, therefore their choice is a dynamic process that has to be periodically repeated. The length of the refresh period is chosen as twice the positioning system update period (e.g., 1 s in GPS).

The relay node will therefore partition the surrounding space in four identical sectors at 90 degrees of each other. Ideally, near the center of one of them sits its parent (parent sector). Near the center of the remaining three (children sectors), the relay node will look for nodes who are eligible to be children. The “sector center” in our case is a point on the bisectrix of the angle formed by sector borders, and at a distance that satisfies two conflicting requirements: (i) close enough that radio reception is not impaired and (ii) far enough from the relay node to minimize co-channel interference with other nodes scheduled in the same slot, while maximizing stream spacial advancement.

The orientation of the four-sector space may be chosen in several ways; we verified through simulation that the most efficient choice is the following: the orientation of each sector is deterministically chosen so that its bisectrix points toward one of the four cardinal points. Fig. 2 shows one possible
orientation as seen by an active node (X) in a sample topology. Sectors are tagged using their cardinal orientation (N, E, S and W) with respect to X. X’s task is to select three children, one for each children sector (W, N and E, since sector S is already taken up by its parent node P).

The children selection leverages the availability of a neighbor list stored at each node. For each children sector, X chooses its relative (active) children node among its neighbors, according to the following procedure:

- select the node closer to the bisectrix within the outer rim (shaded area in the figure) enclosed between distance 2/3R and R, where R is the radio range of the relay node; OR
- if no such node can be found, select the node closer to the bisectrix within the inner rim (white area in the figure) as far as 2/3R; OR
- if no such node can be found, declare the sector unschedulable.

Fig. 2 shows the outcome in the sample topology, i.e., the choice of the three children nodes, each tagged as C. Black nodes are passive nodes, i.e., nodes that were unsuitable, or second-choice, to become active nodes according to the above procedure.

It should be observed that the radio range R cannot be deterministically established due to changing propagation conditions. Therefore, a relay node assumes R to be equal to the distance of the farthest one-hop neighbor from which it received a HELLO message.

If no children sector is schedulable, the relay node will just transmit for the benefit of its own neighboring passive nodes, but will refrain from scheduling any children until the children selection is successful on a later attempt. If one or more children are selected, the relay assigns their transmissions to a specific slot according to the “grid coloring” scheme explained in Sec. IV. Then, the scheduling is notified to the children.

B. Scheduled Access for Streaming Media

When transmitting within its allocated slot, the relay node appends the following information to the stream frame it is relaying (it amounts to the MAC-layer header):

- number of bits of the stream frame;
- identity of the three (or fewer) children;
- slot number in which each children is supposed to be scheduled.

No explicit acknowledge is expected from children. Rather, an implicit acknowledge (“passive ACK”) is obtained at the relay node by monitoring transmissions in the slots where its children are supposed to be relaying the content. If children transmissions, relaying the correct packet, are heard, the relaying procedure is considered to be successful. Otherwise, one of the following situations may arise.

1) No transmission is heard during the time slot: the relay node will try an “on-the-fly” rescheduling of the missing child within the same slot (see below, opportunistic access).

2) A garbled transmission is received: the relay node assumes it did not receive it because of concurrent neighboring transmissions that only affected its own reception; the relaying is considered successful and no rescheduling is deemed necessary.

3) A correct transmission is decoded, but its source is not the scheduled child: the scheduled child may have successfully received the transmission and then lost a children conflict. No action is taken since the slot is occupied.

4) A correct transmission by the scheduled child is decoded, but the video frame has a higher sequence number than the one transmitted by the relay: this situation may arise if the child is scheduled by two or more relay nodes at different levels in the distribution tree (hence relaying just one copy, the “newer” one, of the video frame). In this case, the relay node gives up its own parent role toward that child and waits to be scheduled by it in the next time frame, thus receiving the more recent stream feed. Recursively, the more recent copy of the stream will worm its way outward at the expense of the older copy of the stream.

C. Opportunistic Access for Streaming Media

Opportunistic access within a time slot occurs in the following situation: a relay node does not hear one of its children using the slot where it was scheduled for transmission. Likely, a node close to the silent child has collided with the relay node transmission. In this instance, the relay node will try and start a contention procedure to claim the slot and, at the same time, to salvage its latest transmission by sending one or two out of the three (parts of the) multiple-description frames.

The time slot is therefore used as follows:

- a grace period (similar to IEEE 802.11 DIFS) needed to declare the medium as idle;
- an RTS/CTS exchange between the relay node and the silent child to reclaim the slot;
- a leftover period carrying the transmission of one or two (depending on the length of the RTS/CTS exchange) of the multiple descriptions.

The RTS/CTS exchange provides for contending relay nodes to claim the leftover slot through a slotted Aloha procedure. After sending the RTS, if the contending relay node hears a CTS addressed to itself, it will use the leftover slot time to transmit two of three of the descriptions.
s.t. \( C \) distance-R dimension and that neighboring nodes are all at the same
the active node plays a role similar to that of IEEE 802.11e
it accesses the medium as in DCF. It can be observed that
for the transmission of at least one minimum-size data packet,
if the residual slot time at the end of post-backoff is sufficient
first perform a post-backoff as in standard 802.11 DCF and,
the slot occupancy indication. A node willing to transmit must
field. All nodes within the radio range of the transmitter are en-
contention-based access. The sender is assumed to advertise
fill up the time slot with the streaming media available at
at least a single description. Clearly, two or more contending
nodes will keep competing for the same time slot in every
frame, till one of them moves out of range (this is a frequent occurrence due to the highly mobile environment).

D. Contention-based Access for Best-effort Traffic

If a time slot is underutilized, i.e., an active node cannot
fill up the time slot with the streaming media available at
the time of transmission, the residual slot time is allocated to
contention-based access. The sender is assumed to advertise
the slot portion that it is about to use through a MAC header
field. All nodes within the radio range of the transmitter are en-
titled to contend during the CP, provided they correctly decode
the slot occupancy indication. A node willing to transmit must
first perform a post-backoff as in standard 802.11 DCF and, if
the residual slot time at the end of post-backoff is sufficient
for the transmission of at least one minimum-size data packet,
it accesses the medium as in DCF. It can be observed that
the active node plays a role similar to that of IEEE 802.11e
Hybrid Coordinator.

IV. GRID COLORING

We first consider a regular grid topology in the x,y-plane,
i.e., that every node has two neighbors along each spatial dimension and that neighboring nodes are all at the same
distance \( R \) from each other. Also, let the set of radio resources
available in the communication system (namely, the number of time slots within a time frame) be represented as a set
of colors \( C \), and let \( C \) be the set cardinality. Then, a proper
distance-\( k \)-C-coloring (or briefly \((C, k)\)-coloring) of graph \( G \)
is a mapping \( \phi \) from the set \( V \) into the set of available colors
\( C \) s.t. \( \phi(u) \neq \phi(v) \), \( \forall u, v \in V \) connected by a shortest path
of at most \( k \) hops, with \( k > 0 \). Let us denote each color
with a non-negative integer (i.e., \( C = \{0, 1, \ldots, C - 1\} \));
also, let \( Z_C \) be the set of non-negative integers modulo \( C \).
Assume that the grid vertices are assigned colors according to
the following method, that we name constant-step coloring.
Given two adjacent nodes on the x-dimension, \( u_x \) and \( v_x \), we have:
\( \phi(v_x) = \phi(u_x) + a \pmod{C} \), while, for any pair
of adjacent nodes on the y-dimension, \( u_y \) and \( v_y \), we have:
\( \phi(v_y) = \phi(u_y) + b \pmod{C} \), with \( a, b \in Z_C \). As an example,
in Fig. 3 we show a grid-structure topology whose nodes have
been colored using the constant-step method, with \( C = 5 \),
\( a = 1 \), and \( b = 2 \).

Interestingly, We can prove that, at least for the distances
\( k \) of practical interest, the proposed constant-step coloring
optimally solves the distance-\( k \)-coloring problem over a grid
topology, that is, it finds the minimum number of colors
needed to color the grid. Details omitted for the lack of space
can be found in [5].

Now, let us apply the constant-step coloring to the system
dimensioning and the scheduling problem in our VANET. Consider
that the number of colors represents the number of slots, \( S \), within a single time frame, and that active nodes that
are assigned the same color are scheduled for transmission in
the same slot. Clearly, it is of crucial importance to the
network performance to maximize the spatial separation
between vehicles transmitting simultaneously.

In particular, to avoid collisions in our case, the distance
between any two nodes scheduled in the same time slot should
be greater than \( 3R \). Following our approach, we found that the
minimum number of slots (colors) needed to achieve such a
separation is equal to \( 8 \). We therefore consider that each time
frame includes at least \( S = 8 \) slots, and that each active node
schedules its children according to the constant-step coloring.

We point out that, since the vertices of the forwarers grid
are chosen at variable distance depending on propagation con-
ditions and node availability, the number of colors needed to
achieve separation may be higher. However, in our simulations,
\( S = 8 \) proved to be an excellent tradeoff between interference
mitigation and radio resource efficiency, showing that the
theoretical results derived under our admittedly simplistic
assumptions yield a practical payoff.

V. PERFORMANCE EVALUATION

We analyzed the performance of SMUG using the ns-2
simulator on a city section, where nodes move according
to a realistic mobility model [6] and under typical urban
propagation conditions [7]. The vehicular traffic model takes
into account both micro- and macro-mobility interactions. We
derived the first set of results from snapshots of the VANET
at different time instants, thus corresponding to different
network topologies. These experiments allow us to compare
the performance of SMUG against the theoretical results for
broadcast capacity in [8]. The second set of results instead
provides a broader set of statistics (including the PSNR of
received video streams) collected over the whole simulation.

We assume that a fixed gateway node acts as a source of
the streaming media, and that the streaming content has to be
distributed to all vehicles in the network. Performance results
are computed for a single\(^1\) streaming video session issued
by the gateway. The transmission rate on the data channel
is equal to 2 Mb/s, the time frame periodicity is the same
as the one of video frames. We consider a network area size
of 1 km\(^2\), with the gateway node located at the center of the
area, and we set \( R_m=250 \) m (recall, however, that each vehicle
keeps updated knowledge of its actual 1-hop neighbors through
HELLO messages, transmitted on the data channel). The road
topology for the city section is shown in Fig. 4. It includes
several road intersections regulated by traffic lights or stop
signs, where vehicles are allowed to queue up. Each vehicle
in the topology is initially assigned a desired speed, uniformly

\(^1\)We are currently deriving results in the case of multiple streams.

Fig. 3. Example of constant-step coloring, with \( C=5 \), \( a=1 \), \( b=2 \).
chosen in the interval \([10,20]\) m/s. Depending on road traffic conditions and regulations, each vehicles tries to maintain its desired speed, slowing down and picking up speed again if required [6].

A. Comparison against upper bounds to broadcast capacity

We exploit the results in [8], which provide upper bounds to the maximum achievable throughput, when the network topology is represented by an arbitrary connected graph. The maximum achievable throughput is the maximum rate at which all nodes can successfully receive the broadcast transmission. Specifically, we look at two upper bounds defined in the following way:

1) in a topology with a predefined backbone, the ratio of the size of the Maximum Independent Set (MIS) to the number of backbone nodes;

2) in a general topology, the ratio of the size of the MIS to the size of the Minimum Connected Dominating Set (MCDS).

The MIS size represents the maximum number of active nodes whose transmissions can be received by at least one node. For the comparison to be meaningful, we consider ten snapshots of a realistic VANET simulation taken at random times, each corresponding to a different topology. For each topology, we compute the theoretical upper bounds and derive the throughput achieved by SMUG, as the average throughput at receiver nodes. Again, for consistency with theoretical results, the gateway is assumed to broadcast CBR traffic saturating the channel capacity.

In Fig. 5, we compare the two upper bounds with the SMUG throughput. Each comparison was carried out for a different number of nodes \(N\) in the VANET. As noted in [8], we observe that the theoretical broadcast capacity decreases with increasing \(N\), since, due to the node interference, both the size of the MCDS and that of the backbone grow much faster than the size of the MIS. Surprisingly, however, in a VANET with a small number of nodes, this effect is mitigated by the fact that increasing \(N\) corresponds to widening the geographical area covered by the connected network (i.e., the area where the service is provided). Thus, the node density on the actual network area grows more slowly than the number of nodes.

As for the SMUG performance, it should be pointed out that a small number of nodes (namely, 80 nodes) results in fewer nodes that can be selected as active nodes. Also, since SMUG is tested under urban mobility, the choice of active nodes is less than ideal due to vehicles following the road layout. These factors result in an incomplete, askew grid; hence, in a lower throughput for small values of \(N\). When \(N\) increases, instead, SMUG achieves an excellent performance.

B. Video streaming and best-effort data

Here, we use video YUV sequences provided by [10]. Using the technique described in [9], the video sequences are encoded into 3 descriptors in the Quarter Common Intermediate Format (QCIF) resolution\(^\text{2}\), \((176 \times 144)\). Each descriptor is associated to a constant video rate of 8 frames/s, a mean bit rate of 55.3 kb/s, and a peak rate of 259.3 kb/s; this yields a total mean and peak bit rate associated to the video stream equal, respectively, to 166 kb/s and 778 kb/s. Nodes are also assumed to attempt the transmission of best effort data, whenever possible. The size of the SAR layer header is set to 26 bits [20], while the MAC and physical layer headers are of 26 and 24 bytes, respectively. At the MAC layer, we fix the maximum payload size to 1236 bytes and the minimum payload size (useful for best-effort data transfer) to 200 bytes. The RTS and CTS size, as well as all other parameters and timings used for best-effort data transfer, are set to the default values as specified by the 802.11b technology.

Figure 6 plots the results obtained at the node experiencing the best and worst performance, as well as the performance averaged over all network nodes, as functions of the number of nodes. The plot also carries the average throughput for aggregate video and best-effort data traffic, and highlights the efficient bandwidth usage obtained through SMUG.

The good performance is further confirmed by the mean received video PSNR, whose best, average, worst performance are plotted in Fig. 7. Some nodes, namely those whose route keeps them closer to the gateway, achieve excellent performance (close to the ideal, no-loss PSNR equal to 34 dB). The average performance starts at about 6 dB below the ideal mark and closes in as the number of nodes grows.

\(^*\text{Note that typically hand-held wireless devices have a screen size that corresponds to the QCIF video format}\)
VI. Conclusions

We addressed the support of video streaming in VANETs by providing a fully-distributed solution, SMUG, spanning several architectural layers, from the application to the MAC layer. SMUG leverages the properties of video coding to design a collision-resolution mechanism and the characteristics of VBR traffic to efficiently exploit radio resources. It also promotes best-effort traffic exchange in a VANET without infrastructure support. We compared the performance of SMUG against theoretical results for broadcast capacity and we tested it via simulation in a realistic vehicular environment.

ACKNOWLEDGMENTS

This work was supported by Regione Piemonte through the VICSUM project.

REFERENCES


TABLE I
BEST EFFORT TRAFFIC PERFORMANCE FOR DIFFERENT NUMBER OF NODES

<table>
<thead>
<tr>
<th>N</th>
<th>80</th>
<th>160</th>
<th>240</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficiency</td>
<td>0.60</td>
<td>0.58</td>
<td>0.56</td>
</tr>
</tbody>
</table>

In Table I, we report the access efficiency (i.e., the ratio of throughput to available bandwidth) for what concerns best effort traffic. We point out that SMUG allows best effort to pick up what is left by the streaming media and that every active node acts as a “hot spot”. As expected, the access efficiency decreases with the increase of the number of network nodes.