Martín-Sacristán, D.; Roger, S.; Garcia-Roger, D.; Monserrat Del Río, JF.; Spapis, P.; Zhou, C.; Kaloxylos, A. (2021). Low-Latency Infrastructure-Based Cellular V2V Communications for Multi-Operator Environments With Regional Split. IEEE Transactions on Intelligent Transportation Systems. 22(2):1052-1067. https://doi.org/10.1109/TITS.2019.2962097

Low-Latency Infrastructure-Based Cellular V2V Communications for Multi-Operator Environments with Regional Split

David Martín-Sacristán, Sandra Roger, *Senior Member, IEEE,* David Garcia-Roger,
Jose F. Monserrat, *Senior Member, IEEE,* Panagiotis Spapis, Chan Zhou, Alexandros Kaloxylos, *Senior Member, IEEE*

***Abstract*—Mobile network operators are interested in providing Vehicle-to-Vehicle (V2V) communication services using their cellular infrastructure. Regional split of operators is one possible approach to support multi-operator infrastructure-based cellular V2V communication. In this approach, a geographical area is divided into non-overlapping regions, each one served by a unique operator. Its main drawback is the communication interruption motivated by the inter-operator handover in border areas, which prevents the fulfillment of the maximum end-to-end (E2E) latency requirements of fifth generation (5G) V2V services related to autonomous driving. In this work, we enable a fast inter-operator handover based on the pre-registration of the users on multiple operators, which substantially reduces the handover time to guarantee maximum E2E latency values of** 100 **ms in non-congested scenarios. To further reduce the latency of time-critical services to always less than** 70 **ms, even with the handover interruption time, while providing a latency around** 20 **ms in the majority of locations, we propose to complement the former technique with a mobile edge computing approach. Our proposal consists in the localization of application servers and broadcasting entities in all the base stations, to avoid the communication through the core network, together with the use of a new set of nodes in the base stations of cross-border areas called inter-operator relays, to minimize the communication latency between operators. Based on analytic and simulation results, it is demonstrated that the proposed techniques are effective to support low-latency infrastructure-based cellular V2V communications in multi-operator environments with regional split.**

***Index Terms*—Low-latency, multi-operator, multi-PLMN, inter-operator handover, infrastructure-based, cellular V2V, regional split.**

I. INTRODUCTION

Cellular standards have recently included new functionalities to support vehicular communications. With that aim, the Third Generation Partnership Project (3GPP) has considered two communication modes for Long Term Evolution (LTE): direct Vehicle-to-Vehicle (V2V) communications and infrastructure-based V2V communications. The direct communication mode is also known as PC5-based V2V due to the use of a new sidelink communications interface between User Equipments (UEs) referred to as *PC5 interface*. The infrastructure-based communication mode is known as Uu-based V2V, and is based on the so-called *Uu interface*, which is the conventional LTE radio interface for uplink and downlink communications between UEs and the Base Stations (BSs) of the Radio Access Network (RAN). In particular, the broadcasting capabilities of the LTE Multimedia Broadcast Multicast Service (MBMS) have proven useful for classical V2V services such as Cooperative Awareness Message (CAM) and Decentralized Environmental Notification Message (DENM) delivery [1], [2], [3]. However, the requirements guaranteed by prior art are often insufficient for more challenging fifth generation (5G) V2V applications such as autonomous driving, where an exchange of time-critical messages is key in guaranteeing traffic safety [4], [5]. As defined in [5], autonomous driving services require End-to-End (E2E) delays of 100, 50, 25, 20, 10 or 3 ms, depending on the service, and very high reliability (e.g., 99.999%). For instance, Advanced Driver-Assistance Systems (ADAS) are an example of safety applications demanding extremely low latency. In fact, pre-crash sensing warning messages should be disseminated to neighboring nodes across different Mobile Network Operators (MNOs) within 20 ms [4], [6], which presents challenging implications for 5G V2V. This is particularly important because the system should be able to sense and warn the driver in a very short time so as to reduce injuries to motor vehicle occupants.

Comparing the two 3GPP communication modes, MNOs are interested in providing infrastructure-based cellular V2V to increase their revenues by leveraging their network deployments and available licensed spectrum. Besides, infrastructure-based cellular V2V is also interesting for the vehicle manufacturers given that they already have equipment on board most of their new vehicle series with Uu-based communication capabilities. At last, we consider that 3GPP V2V communication solutions are also appealing for the users due to the higher quality of service guarantees that can be achieved compared to other free direct communication solutions based on the use of unlicensed spectrum. As already mentioned, this high quality of service in terms of high reliability and low latency is of paramount importance for some services, and hence the economic factor

D. Martín-Sacristán, D. Garcia-Roger and J. F. Monserrat are with the iTEAM research institute, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain, e-mail: {damargan, dagarro, jomondel}@iteam.upv.es.

S. Roger is with the Computer Science Department, Universitat de València, Av. de la Universitat s/n, 46100 Burjassot, Spain, e-mail: sandra.roger@uv.es.

P. Spapis and C. Zhou are with Huawei Technologies, German Research Center, 80992 Munich, Germany, e-mail:{panagiotis.spapis, chan.zhou}@huawei.com.

A. Kaloxylos is with the Department of Informatics and Telecommunications, University of Peloponnese, 22100, Greece, e-mail: kaloxyl@uop.gr.

is deemed less decisive in the technology selection. All things considered, we regard justified the focus of this paper on the 3GPP infrastructure-based V2V communication mode.

A. Background

Many cellular V2V related works assume all end devices operating under the same operator but, as identified in [7], this assumption is not realistic. Instead, a multi-operator scenario needs to be considered as very likely for V2V communication.

One potential approach to support V2V communication in a region with multiple operators is to split that region into disjoint areas, where each area is served by a single operator. The regions could be distributed among operators either with direct interactions or via the intervention of the Intelligent Transportation Systems (ITS) regulators. Fig. 1 illustrates this multi-operator scenario showing a road with two areas whose border is represented by a gray line. All the vehicles located in the left area are served by a different operator (OpA) from the vehicles in the right area (OpB), and connected to their respective Core Networks (CNs) (CN OpA and CN OpB). Note that, an operator (like OpB) may have a BS in OpA's region, but that BS would not be involved in V2V communication services.

This scenario provides several benefits. First, it simplifies the multi-operator environment by transforming it into multiple single-operator regions, circumscribing the challenges of multi-operator scenarios to just the borders between regions. Second, the communication between vehicles will be in many cases intra-operator. This fact implies a general reduction of the communication latency, given that the communication among network entities of different operators is avoided. In addition, having all vehicles under the same operator domain enables the use of local breakout schemes to further reduce the delay [8]. Third, it does not require any use of common spectrum resources for V2V services as proposed in [9], whose common management could be very complex [10], and whose business case raises questions about the long run if one virtual operator is supposed to manage this spectrum pool [11]. Rather, the regional split is a good business model that distributes the profit opportunities fairly among the operators. Finally, the split is scalable and cost efficient since it supports an undefined number of operators with only one Radio Frequency (RF) chain in the UEs. Conversely, other solutions based on multi-connectivity may require terminals with multiple subscriber identity modules and RF chains [12], with an increased terminal cost.

All the advantages offered by the regional split have made this multi-operator solution one of the best options [13], even though it is not free of technical problems. The main drawback of the regional split among operators appears in the border area where the vehicles have to perform the inter-operator handover. Typically, this process creates an interruption in the communication that can have a detrimental effect on the quality of the services provided, since some packets may be delayed or lost. In order to alleviate this effect, the borders should be selected in areas where the transmission of messages requiring high reliability and/or low-latency is not frequent (e.g. crossroads should be avoided). For this reason, the simulation results that are presented in Section V deal with a highway in an otherwise sparse road density area, where two operators agree to draw the borders that will divide their regions. However, even such intelligent selection of the borders does not fully solve the problem, as it will be later discussed.

B. Main goal and key contributions

The idea of reducing E2E latency is a target pursued by previous attempts in the V2V research literature, even recent ones as [14] or [15]. However, the main goal of this paper is more specifically to reduce the latency of inter-operator communications in multi-operator environments with regional split, especially for time-critical messages, by providing a solution valid for devices with only one RF chain. The problem addressed is currently unsolved in the literature and the present paper explains the details of a novel solution which can be summarized in the following contributions.

The first contribution of this paper is a detailed analysis of the E2E latency for Uu-based V2V communications in multi-operator environments considering both intra and inter-operator communications. This analysis includes also the assessment of the impact of inter-operator handovers. We consider that this thorough analysis based on realistic assumptions could be very useful for the research community.

Then, two complementary solutions to minimize the inter-operator communication latency are proposed:

- First, we enable a fast inter-operator handover based on the pre-registration of the users on multiple operators well in advance of reaching the border area between the regions allocated to different operators in a regional split of operators. This paper proposes new network entities and the signaling required for that pre-registration. The fast inter-operator handover enables E2E latency values lower than 100 ms, and hence it is valid for all V2V services defined by 3GPP in [4] except for pre-crash sensing warning, which requires 20 ms.
- Second, to further reduce the E2E latency of time-critical services we propose to complement the former technique with a solution that reduces the length of the communication path for both intra-operator and inter-operator communications of time-critical messages. Specifically, we propose a Multi-access Edge Computing (MEC) approach [16] in which all the BSs are equipped with (i) a local application server for a fast processing of time-critical messages sent in uplink by vehicles, and (ii) a local broadcasting system for a fast local distribution of time-critical messages in downlink, whereas BSs located in the border area of the regional operator split are also equipped with (iii) a new node called inter-operator relay. This node monitors messages broadcasted by other BSs and sends them to the local application server of its own BS for its local distribution. As a result of the shortening of the communication paths, this technique is able to reduce the maximum E2E latency of time-critical messages to less than 20 ms in the majority of locations and to less than 70 ms in the worst-case scenario.

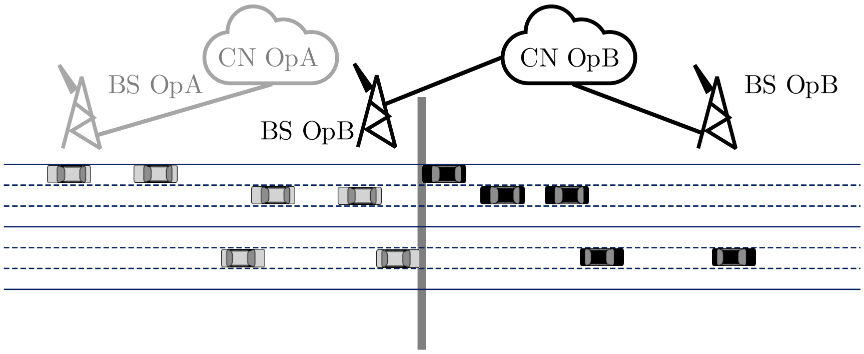

Fig. 1: Illustration of baseline multi-operator communication scenarios with two regional network operators.

C. Paper structure

The remainder of this paper is structured as follows. Section II describes the limitations of the conventional cellular V2V solutions to provide low-latency communications in multi-operator environments with regional split, including a detailed analysis of the E2E latencies. Section III presents the fast inter-operator handover mechanism proposed in this paper, while Section IV details our proposal to use MEC and inter-operator relaying. Section V describes the simulation environment used for the performance evaluation. After that, simulation results are discussed in Section VI, whereas the main conclusions are drawn in Section VII.

II. CONVENTIONAL APPROACH

As mentioned in the introduction, besides being a good business model, the regional split of operators simplifies the management of multi-operator environments by circumscribing the problems to the border area among the regions allocated to different operators.

In order to understand the problems related to the E2E latency in the border areas, it is necessary to describe first several assumptions concerning the network architecture of each operator and how the operators are interconnected, as well as to present and decompose the user plane delay involved in both intra-operator and inter-operator communication.

A. Network architecture and data paths

Fig. 2 shows a simplified view of the network architecture considered for a two-operator communication scenario based on LTE MBMS. The architecture of each operator is the basic LTE architecture, where ITS back-end servers are also considered. The UE connects to the RAN formed by BSs through the radio interface. The BSs are connected to the Serving Gateway, which is the CN access point for user data. In the CN, the Packet Data Network (PDN) Gateway provides connectivity to external packet data networks and between BSs. The Broadcast Multicast Service Center (BM-SC) and MBMS Gateway are MBMS-related entities used for broadcasting or multicasting of data from content servers to the UEs. The MBMS Gateway routes the MBMS information from the CN to the BSs where that information has to be distributed.

In Fig. 2, each operator has its own ITS server that manages the ITS sessions of the vehicles connected to it and a location database continuously updated using messages sent by the vehicles. The ITS servers of different operators can communicate with each other with different aims, such as to exchange vehicle-originated messages or the location of vehicles. To support such exchange, inter-server links are established as shown in Fig. 2.

Most of ITS messages generated by the UEs need to reach the server to be distributed to their intended receivers. In fact, the common procedure to distribute those messages starts with a user sending a message to its ITS server in uplink. Once received, the ITS server updates its location database with the information of the message, and, based on the location information, the server decides to which receivers the message should be sent. These receivers could be, e.g., those located closer than a certain relevance distance from the message originator. After that, the message is sent in downlink to the identified receivers. Therefore, for those messages, two communication paths can be clearly distinguished. In the uplink path, to reach the ITS server, the messages follow the path *Sender UE→BS→Serving-Gateway→PDN-Gateway→ITS server*. In the downlink path, to reach the UEs from the ITS server, both unicast and multicast can be used. In practice, single-cell multicast transmission, Single-Cell Point-To-Multipoint (SC-PTM) in LTE, is considered as the most attractive option by the industry due to its higher capacity compared to the other options [17]. In SC-PTM, the path followed by the messages in downlink is *ITS server→BM-SC→MBMS-Gateway→BS→Destination UE*. In case of inter-operator communication, i.e. when the sender UE and the destination UE are served by different operators, the messages are transmitted from the ITS server in the operator of the sender UE to the ITS server in the operator of the destination UE through an inter-ITS-servers link. Fig. 2 shows an example of both an intra-operator and an inter-operator distribution of an ITS message. Specifically, the dashed lines represent a message sent from UE #1 in OpA reaching UE #2 and UE #3 also in OpA, and reaching UE #4 in OpB.

B. Latency assumptions

The main delay components in the V2V communication are indicated in Fig. 2. Processing delays are shown next to their

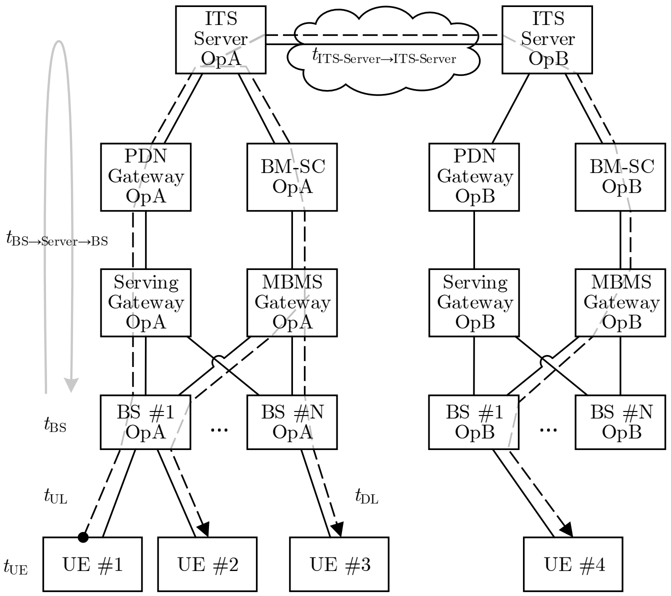

Fig. 2: Network architecture, message transmission latency components and exemplary message path from one transmitter to three receivers including intra-operator and inter-operator communication.

corresponding network element, while the communication delays are indicated next to the links connecting architectural blocks.

In [7], [18] and [19], some reference values for the delay components presented in Fig. 2 can be found. In particular, the back-haul delay due to communication between two BSs through a back-end ITS server, denoted by $t_{\text{BS}\rightarrow\text{Server}\rightarrow\text{BS}}$, is assumed to be 20 ms. The same value is used for unicast and MBMS transmissions. The processing delays in UE, t_{UE}, and in BS, t_{BS}, are assumed to be different in transmission ($t_{\text{UE-tx}}$ and $t_{\text{BS-tx}}$) and reception ($t_{\text{UE-rx}}$ and $t_{\text{BS-rx}}$) as elaborated in [18] and [19]. A typical value for the processing time in the lower layers is 1 ms for transmission and 1.5 ms in reception. In the UE reception, an additional processing time in higher layers of 3 ms is assumed as in [7] and [19].

The exact delays for transmission over the radio interface in uplink, t_{UL}, and in downlink, t_{DL}, depend on the scheduling decisions made by the BSs. Without scheduling delays, the minimum delay would be 1 ms in downlink, and 9 ms in uplink, assuming that the UE does not have pre-allocated resources and has to request them to transmit. The downlink delay of 1 ms corresponds with the duration of an LTE subframe. In uplink, the 9 ms can be achieved if the frequency used for sending the scheduling requests is the highest one (one request per subframe).

It is worth noting that there is an additional latency component in uplink and downlink transmission over the air due to the strict timing of the transmissions, the beginning of which must correspond to the beginning of a subframe. This component, known as subframe alignment delay, has then a value between 0 and 1 ms.

In the inter-operator communications, one additional delay component must be considered: the transmission delay between ITS-servers, $t_{\text{ITS-Server}\rightarrow\text{ITS-Server}}$. There is not a typical value reported for that delay component in the literature. However, in [20], round-trip delay values between service provider premises located in the same region are around 40 ms in developed regions. The round-trip delay is the time to transmit an IP packet from a source to a destination and receive a reply packet from the destination to the source. Therefore, we have assumed a transmission delay between servers of 20 ms (half the round-trip delay). Table I summarizes the above mentioned user plane delay components and the values considered in this work.

TABLE I: Assumed values for user plane delay components

Component	Value
t_{UL}	≥ 9 ms
t_{DL}	≥ 1 ms
t_{UE}	1 ms in transmission ($t_{\text{UE-tx}}$) 4.5 ms in reception ($t_{\text{UE-rx}}$)
t_{BS}	1 ms in transmission ($t_{\text{BS-tx}}$) 1.5 ms in reception ($t_{\text{BS-rx}}$)
$t_{\text{BS}\rightarrow\text{ITS-Server}\rightarrow\text{BS}}$	20 ms
$t_{\text{ITS-Server}\rightarrow\text{ITS-Server}}$	20 ms

C. Latency analysis without inter-operator handover

Adding all the processing and transmission delays experienced along the path from a user to the ITS server and along the path from the server to a user, the total delay for a user-to-user communication is, at least, 38 ms in intra-operator communications ($t_{\text{UE-tx}}+t_{\text{UL}}+t_{\text{BS-rx}}+t_{\text{BS}\rightarrow\text{Server}\rightarrow\text{BS}}+t_{\text{BS-tx}}+t_{\text{DL}}+t_{\text{UE-rx}}$), and 58 ms in inter-operator communications. Due to subframe alignment before uplink and downlink transmissions over the air, the above mentioned latency values could reach 40 ms and 60 ms, respectively. Additionally, these values may increase due to scheduling delays in uplink and downlink. In any case, in absence of congestion, with reasonable scheduling delays, the typical 100 ms maximum E2E latency requirement of CAM messages is fulfilled.

D. Implications of the inter-operator handover

In the previous analysis we have assumed that the UEs remain served by the same operator during the whole communication process. However, in case of splitting a region between different operators, UEs perform an inter-operator handover when they pass from one region to another, with the following implications:

1) Need for message duplication: Currently, to complete the inter-operator handover, vehicles with only one RF chain have to detach from one operator and then attach to the other one. This means experiencing a service interruption that takes more than 300 ms, which is the typical duration reported in [21] for network attachment.

If there were no inter-operator handovers during the transmission of messages, the ITS messages could be distributed only once at each operator and they could reach all the UEs. However, with inter-operator handovers, some UEs may detach from one operator before the ITS message is distributed by this operator, and attach to the other operator when the ITS message has already been distributed by the second operator. Therefore, to avoid the loss of messages, the ITS messages must be sent at least twice by each operator. The first transmission should be performed as soon as possible to minimize latency. The second one should have the proper timing to ensure the reception of the message in due time for those UEs that have performed an inter-operator handover and have attached to one operator before receiving the message from the operator they detached from. Note that, as shown in Fig. 2, the inter-operator handover solution is implemented at the application layer, being the ITS servers of the MNOs the entities distributing the information required.

2) Worst-case scenario: It arises when a UE that has just started an inter-operator handover generates an ITS message, and a destination UE in another operator starts an inter-operator handover just when the message is about to be sent to that UE by its serving BS. In that case, the communication delay would be, at least, equal to the inter-operator communication delay calculated above plus two times the inter-operator handover interruption time, i.e. 658 ms without subframe alignment delays or up to 660 ms with subframe alignment delays, which is unacceptable even for classical ITS services like the distribution of CAMs. Finally, note that as a by-product of the pre-registration, IP addresses and other network configuration parameters are already assigned to the UE upon pre-registration and are no longer a contributing factor to the handover latency.

3) Location information: Another implication of the inter-operator handover is that some changes could be needed in the way the ITS server determines in which BSs it has to distribute the messages. The main problem is that, in case of inter-operator handover, the location database of an ITS server may not have the location information of the UEs that have just performed an inter-operator handover. In that case, the distribution of the messages should occur in all the cells whose serving area overlaps with the border between regions served by different operators, to ensure that all the UEs receive the message. Another option is the addition of some inter-server signaling that, in case of being as fast as the inter-operator handover process, would allow the distribution of messages only based on the use of the location databases. In fact, in case of having this signaling, the additional transmissions required to reach the users after inter-operator handover could be sent in unicast just to those users, thus, the efficiency of the transmission compared to multicast transmission over multiple cells would increase.

III. FAST INTER-OPERATOR HANDOVER

As analyzed in the previous section, the multi-operator scenario with regional split of operators may involve unacceptable communication delays in border areas caused by inter-operator handovers. As mentioned above, the main delay component in an inter-operator handover is the delay of network attachment to the destination operator.

As a solution to achieve a fast inter-operator handover, this paper proposes to have a vehicle registered not only in the operator actually providing it a service, but also to have the vehicle pre-registered in the multiple operators to which the vehicle is likely to perform an inter-operator handover in the near future. This pre-registration speeds-up the future inter-operator handovers by avoiding their main delay component. Although pre-attached in multiple networks, note that a device is not connected to all of them. In fact, given that the aim is to support users with only one available RF unit, this work assumes that UEs can be registered in multiple operators but are active only in one of them at a given time. As a result, a device is considered to be in connected state in only one of the operators while being in idle mode in the others. By connected state we mean a state similar to the 3GPP Radio Resource Control (RRC) CONNECTED mode, in which the device may have access to resources for transmission of data and the network knows with high accuracy the position of the device. By idle state we mean a state similar to the 3GPP RRC IDLE mode, in which the user does not have access to the resources of the network and its position is known in terms of a broader location area. Note that, being the geographical area split into disjoint regions assigned to the different MNOs (as it will be developed later in Section IV), the vehicle would select the suitable operator to be on connected state depending on the MNO assigned to its geographical location.

Nowadays, most mobile devices attach and operate under one single operator, or Public Land Mobile Network (PLMN). For this reason, during power-on they search and select the operator with whom they have a subscription to attach to [22], known as home PLMN (HPLMN) [23]. In cases where their operator is not present in an area (e.g., usually when the device is abroad), the end device performs roaming and attaches to a visited PLMN (VPLMN). In what follows, we use the same notation as 3GPP, assuming that the HPLMN is the PLMN with whom the device has a subscription, or a preferred subscription. The following subsections detail the proposal of this paper for fast inter-operator handover.

A. Functionalities and logical entities

To perform the pre-registration, we propose three new functionalities running in three generic logical network entities. These entities may correspond to some existing LTE or 5G entities as specified in [24] and [25], respectively. First, we use a Mobility Server, whose role could be played by the LTE Mobility Management Entity, or by the Access and Mobility Management Function and the Session Management Function in 5G. Some of the functionalities of the Mobility Server would be the selection of data gateway after initial attach, bearer activation and deactivation, and authentication of the user via interaction with subscriber servers. Second, we consider the use of a Subscriber Server, which is located in the HPLMN and is able to request to other operators the attachment of a UE. The Subscriber Server role can be played by the Home Subscriber Server in LTE networks or by the Authentication Server Function and the User Data Management in 5G. The Subscriber Server contains a database with user-related and subscription-related information for mobility management, call and session establishment, and authorization. Third, we

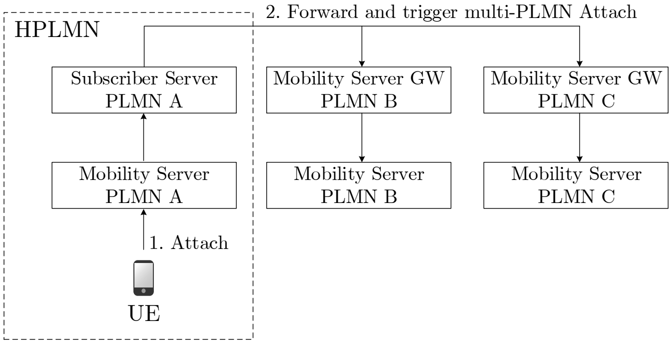

Fig. 3: Example of automatic attachment to three operators.

use a Mobility Server Gateway in each operator to receive the attachment messages from the Subscriber Server of the HPLMN. The reason for using such an entity is that it enables the topological information of the VPLMNs to be invisible to the HPLMN and thus it protects the operators' topology from being exposed. If obfuscating the network topology is not required, then the Mobility Server Gateway may be skipped.

B. New signaling

Fig. 3 illustrates the case of a vehicle performing a typical attachment to its HPLMN. The Subscriber Server of the HPLMN triggers through the CN the attachment (i.e., registration of the same vehicle) in all other available operators (operators B and C in Fig. 3).

In Fig. 4 we detail the particular signaling for the multi-operator attach process in the case of a vehicle performing an attachment to its HPLMN. The core of this process follows the principles of the LTE attach process [22]. However, as we explain in the next paragraphs, to support a multi-attach process, new information elements and messages are needed.

Initially, the UE starts the Attach procedure with a transmission to the BS of an Attach Request message including a device identifier, among other parameters. The UE will indicate that this is a "multi-PLMN" attach message and it also may transmit its current location (e.g., in terms of GPS coordinates). Based on this information, the BS forwards the Attach message to the Mobility Server. The next step for the UE is to be authenticated. After authentication, the establishment of bearers and the configuration of network components will take place. Also, a location update will be carried out, where the Subscriber Server will be notified not only about the new location of the UE and the selected Mobility Server, but also that these actions refer to a "multi-PLMN" attach and that the UE needs to be attached to all other available PLMNs. Thus, the Subscriber Server will send a Remote Attach Request to the Mobility Server Gateway of the other networks, informing that the specific device needs to attach also to them (VPLMNs).

The Mobility Server Gateway, upon reception of such a message, will select the most appropriate Mobility Server, which in turn will select the most appropriate BS (e.g., both choices select the available entities in the specific PLMN) and also configure any needed bearers from the gateways (i.e., router or similar) to the BS. Furthermore, it will perform any required configuration to the BS (e.g., denote that the UE is attached and in idle condition). When these steps are completed, the Mobility Server of the VPLMN will forward to the Mobility Server Gateway and, subsequently, to the HPLMN, a Remote Attach Complete message with VPLMN configuration information to be used by the UE when it will switch to this VPLMN, as well as the spatial coordinates for which this VPLMN should be used by the UE. The configuration information may include data such as temporary identifiers to be used when in VPLMN or power control configuration parameters, for example. The information collected by all VPLMNs is eventually sent as a Non Access Stratum signaling to the UE (message "Multi-PLMN attach information"). Essentially, this information would contain data similar to those transferred today with the use of an RRC CONNECTION SETUP message. At this point of the specific example, the UE is attached to all available PLMNs (Fig. 4 denotes only one VPLMN but multiple VPLMNs may exist).

In the example, the multi-attach process starts with the establishment of a connection between a device and its HPLMN. If the connection is with a VPLMN, the Update Location message shown in Fig. 4 should be sent from the Mobility Server of the VPLMN to the Subscriber Server of the HPLMN, traversing the Mobility Server Gateways of those PLMNs. The rest of the process would be identical to that shown in Fig. 4.

It is also worth noting that the signaling in Fig. 4 presents a case in which the UE sends the Attach Request to a BS in one PLMN, after synchronizing and reading some parameters of that BS, and the UE ends up being connected in this PLMN. However, in general, the UE may send the Attach Request in one PLMN and end up being connected to another PLMN, while being in idle mode in the former. In that case, this situation has to be signaled to the UE to force it to synchronize with a BS of the destination PLMN. This can be done with a message sent to the UE from the BS that received the Attach Request. The transmission of this message would be requested to the BS by the Mobility Server after the reception of an Update Location or Remote Attach Request indicating that the UE should be in idle, and not in connected mode, in this PLMN.

Another important aspect is that the multi-registration may not only occur in the initial attach but also after that point. During normal operation, the UE may send location updates to the CN. The Mobility Server of the PLMN in which the UE is in connected state would translate such location updates into Update Location messages sent to the Subscriber Server of the HPLMN. As a result of a location update, the Subscriber Server may trigger the attachment of the UE to a new PLMN or the detachment of the UE from a PLMN depending on whether those PLMNs are needed or not in the new location. Fig. 5 shows an example in which, initially, a UE is connected in VPLMN1 and idle in VPLMN2. At a certain point in time, the UE sends a location update to the Subscriber Server, which then decides that the UE no longer needs to be attached to VPLMN2. Therefore, a Remote Detach Request is sent to the Mobility Server of VPLMN2. The Mobility Server, after receiving this request, removes the

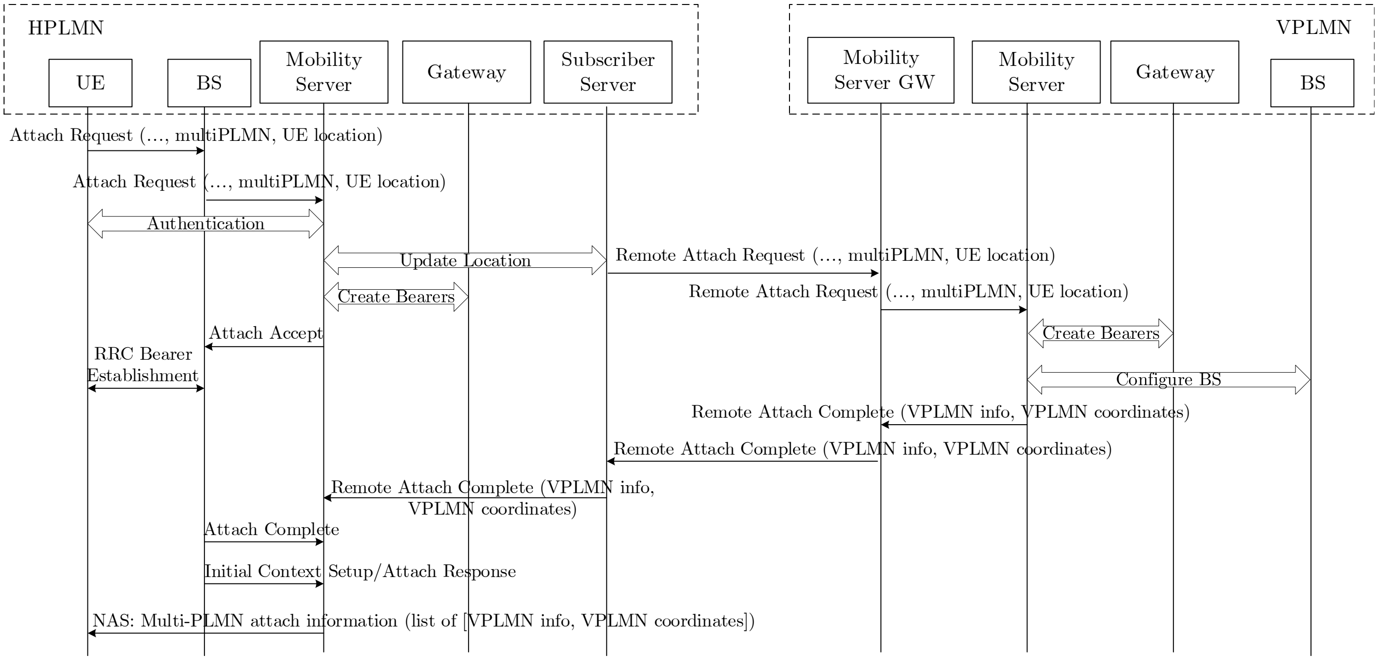

Fig. 4: Multi-attach process in a multi-operator domain.

radio bearers established between the Gateway and some BS, and undoes the BS configuration performed during attachment. The completion of these actions is notified to the Subscriber Server with a Remote Detach Complete message.

C. Signaling overhead and management implications

The pre-registration which is the basis of the fast inter-operator handover involves an additional signaling overhead in the system. However, this extra overhead occurs only in the core network, which is not commonly the most congested part of the network, and does not affect data communication over the air. Note further that pre-registration is only carried out in specific and infrequent situations (e.g. during location update). Concerning the management cost, we consider that once the new entities/functions are introduced in the core of the system, no additional efforts are required by the system manager. Note that the attach/detach processes defined in this paper are not dependent on the actions of the management system. Therefore, there is a low impact of our proposal on the regular operation of the system.

D. Latency analysis

Thanks to the pre-registration to multiple operators, an inter-operator handover is basically equal to a typical intra-operator handover. Accordingly, the interruption time for the fast inter-operator handover is reduced from the 300 ms of a typical inter-operator handover to lower values that might be the experienced ones in an intra-operator handover. Specifically, we assume an interruption of 20 ms for intra-operator handover, which is a conservative value compared to that reported by [18] and close to the values reported by [26]. Therefore, following the explanation in Section II, the E2E communication delay in a worst case non-congested scenario would be 100 ms, instead of the 660 ms reported in Section II-D for typical inter-operator handovers with subframe alignment delays. Therefore, the fast inter-operator handover is valid to support the maximum E2E latencies for typical ITS services, which are usually around 100 ms. However, in case of scheduling delays due to, e.g., a temporary congestion, the maximum E2E latency could be exceeded.

IV. MEC AND INTER-OPERATOR RELAYING

Based on the previous latency analysis, the proposed fast inter-operator handover enables the support of conventional ITS services using cellular V2V in scenarios with regional split of operators, as long as scheduling delays are avoided and the CN back-haul latencies are kept within the limits assumed. However, it would be desirable an additional reduction of the communication latency to increase the guarantees of fulfilling the requirements of conventional services while approaching to the needs of more time-critical services. In order to further reduce the latency of V2V communication, we propose a MEC solution to bring some entities currently found beyond the RAN closer to the users. To this end, we consider local break-out schemes [7], which enable the localization of ITS servers in the BSs instead of in remote locations. Our scheme also assumes the localization of MBMS functions to run broadcasting functions directly in the BSs [7][27] instead of being placed in CN entities. This MEC solution would be deployed in all the BSs to obtain a system-wide latency reduction. We propose that the local entities would be used for the dissemination of time-critical messages while non-time-critical messages would still use the remote entities.

Although the use of localized servers and functions shortens the data path traversed by time-critical messages, it can

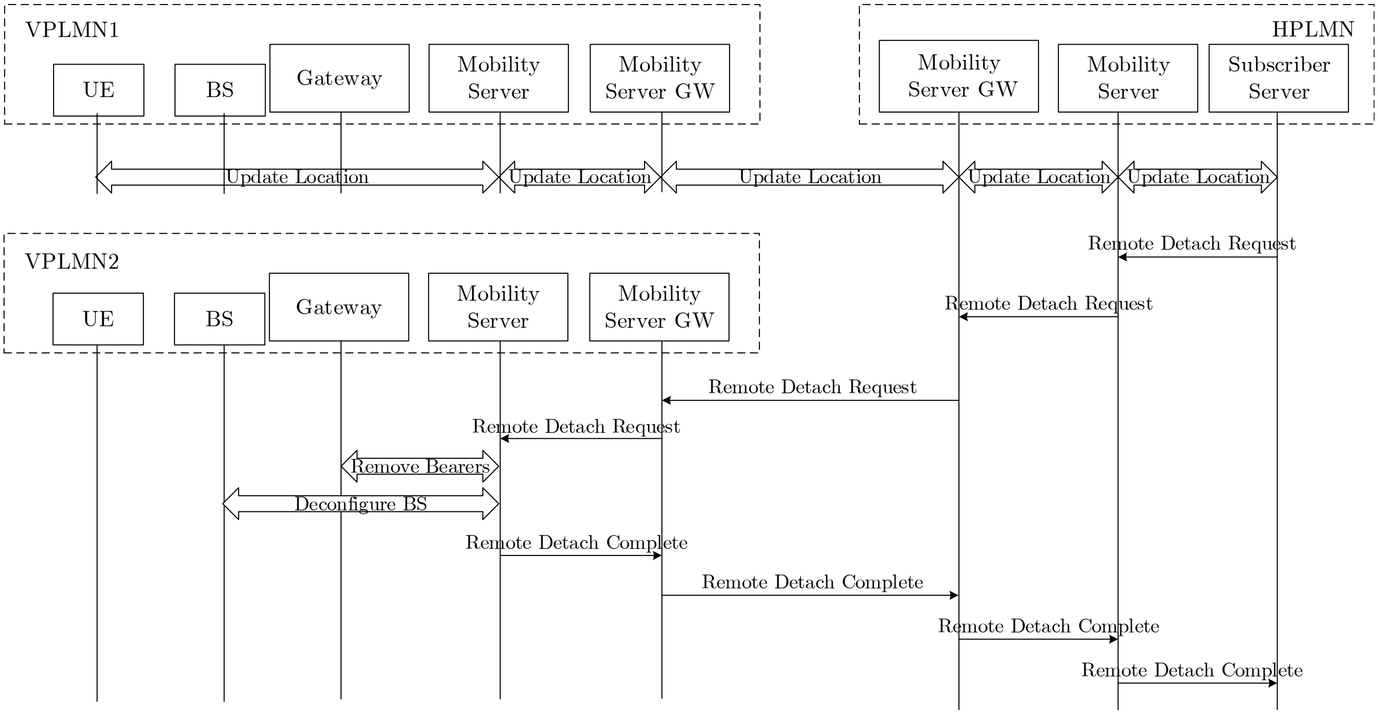

Fig. 5: Detach from one VPLMN triggered by a Location Update.

compromise service continuity in some cases, as raised in [7]. For instance, when a user changes its serving BS, the ITS message may not be available in the destination BS, due to the use of local servers. To ensure the service continuity within each region of the regional split, inter-MEC communication through, e.g., high speed links between neighbour BSs of the same operator could be used, as we suggested in [15]. However, in border areas we have BSs of different operators and it could be unfeasible to set direct links between them. For example, if the operators want to protect their topology from being exposed. Therefore, to overcome the continuity problem in border areas we propose a solution, complementary to the fast inter-operator handover, that includes the definition of a new inter-operator relay. This entity is associated with a BS serving the border area between regional operators and should have a low-latency link with the associated BS (e.g., a fiber-optic connection).

The inter-operator relay monitors the broadcast transmissions of the operators other than the operator of its associated BS. When the inter-operator relay receives a time-critical message, the message is forwarded to the associated BS through the low-latency link. Then, the BS broadcasts the message to its served vehicles. Therefore, with the inter-operator relay, a user connected to an operator can receive relevant emergency messages from another operator with low delay. Note that the inter-operator relays do not perform radio transmissions, and hence, they do not interfere with the transmissions sent over other operator's spectrum. Besides, although inter-operator relays involve transmitting more messages, those transmissions happen in rare, event-triggered time-critical conditions which by themselves do not increase communication overhead significantly.

Fig. 6 illustrates the proposed approach. It shows a border area between regional operators with two BSs serving the vehicles in this area. One BS is owned by OpA and the other by OpB. An inter-operator relay associated to the OpB BS, which receives information from OpA BS, is shown in the figure. The distribution of a time-critical message in this scenario according to our approach would be performed in four steps:

1) A user in OpA sends a time-critical message to its serving BS. This transmission is represented by a dotted arrow in the upper part of Fig. 6.
2) OpA ITS local server processes the received message. After detecting that it is a time-critical message, it passes it to the local broadcasting entities of the BS. Then, the BS broadcasts the emergency message through SC-PTM. Broadcasting is represented with a set of gray arches.
3) The inter-operator relay associated to OpB BS receives the message and detects that it is a time-critical message. The relay passes the message to OpB BS to broadcast it, as represented by a curved black arrow in Fig. 6.
4) OpB BS broadcasts the time-critical message through its local broadcasting entities so that vehicles within its service area can receive it. This broadcasting is represented by a set of black arches in Fig. 6.

A. Latency analysis

The deployment of local ITS servers and broadcasting functions in all the BSs to disseminate time-critical messages results in a reduction of the intra-operator communication latency compared to the values reported in Section II-C. The path followed by the time-critical messages between two

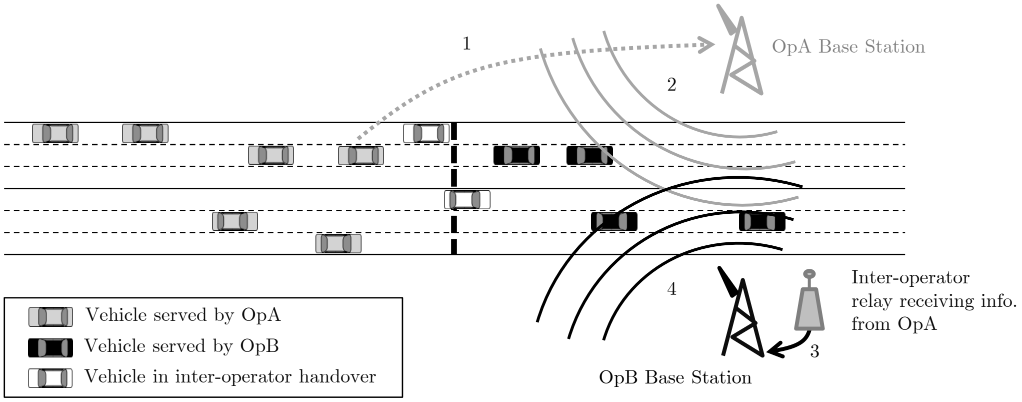

Fig. 6: Illustration of the proposed approach based on the use of an inter-operator relay. Four steps are indicated: transmission of time-critical message from a vehicle to the OpA Base Station (1), broadcasting of the time-critical message by OpA Base Station (2), reception of broadcast message by inter-operator relay and forwarding of the message to OpB Base Station (3), and broadcasting of the time-critical message by OpB Base Station (4).

vehicles served by the same BS in this case would be: *Sender UE→BS→Destination UE*. The total delay can be decomposed as $t_{\text{UE-tx}} + t_{\text{UL}} + t_{\text{BS-rx}} + t_{\text{BS-tx}} + t_{\text{DL}} + t_{\text{UE-rx}}$, which results in a latency of at least 18 ms. This value would be increased a certain amount of time in case the transmitter and receiver vehicles are in different BSs. The additional time would account for the communication between BSs, and as in [15], it could be assumed to be 2 ms. Therefore, for inter-BS intra-operator communications, a latency of at least 20 ms can be expected. If we consider a worst-case for the communication in which both the transmitter and the receivers are involved in intra-operator handovers, the previously calculated values could be increased in 40 ms (two times the intra-operator handover interruption time), thus leading to 58 ms and 60 ms, for intra-BS and inter-BS communication. Note that these values could be increased two milliseconds due to subframe alignments in uplink and downlink transmissions.

In border areas, the E2E communication for a message transmitted from a vehicle in one operator to a vehicle in another operator through an inter-operator relay with localization of ITS server and broadcasting functions in the BS would follow the path: *Sender UE→BS→Inter-operator relay→BS→Destination UE*. Fig. 7 illustrates an example of such kind of communication path and presents its main delay components. The delay components involved in this path have been already presented except the delay of data transmission from the relay to the BS, $t_{\text{Relay}\rightarrow\text{BS}}$. We assume this delay to be around 1 ms, considering that the relay and its associated BS are nearly or totally co-located and have a high-speed connection. With that in mind, the E2E communication latency without inter-operator handovers is at least 25.5 ms ($t_{\text{UE-tx}} + t_{\text{UL}} + t_{\text{BS-rx}} + t_{\text{BS-tx}} + t_{\text{DL}} + t_{\text{UE-rx}} + t_{\text{Relay}\rightarrow\text{BS}} + t_{\text{BS-tx}} + t_{\text{DL}} + t_{\text{UE-rx}}$), while with inter-operator handovers it is at least 65.5 ms. These values could reach 27.5 ms and 68.5 ms, respectively, due to the addition of two times and three times (the number of radio transmissions for each case) the maximum subframe alignment delay.

Table II summarizes the minimum E2E latencies that result from the analysis conducted in previous sections for different situations and the three regional split alternatives considered. The values in the table are the minimum values, therefore, its calculation excludes subframe alignment delays and scheduling delays. For situations with handovers considered, we focus on the worst case in which both the transmitter and receiver experience a handover. As can be observed, the fast inter-operator handover provides a benefit in the situation in which an inter-operator handover occurs making feasible the communication for non-time critical services even in border areas although additional delays may jeopardize the fulfilment of latency requirements in that case. The use of a MEC solution together with inter-operator relays provides a reduction of the latency in all the situations. Specifically, for intra-operator communications, which occur out of border areas, minimum latencies of 20 ms or less are found which is an enabler for time-critical services. Note that with regional split, intra-operator communications are the most frequent case, and hence most of the time-critical packets would experience these low latency values. In border areas, the proposed solution enables latency values lower than 70 ms which is a significant reduction compared to previous solutions. The two solutions presented can be seen as enablers that could be complemented by other improvements such as the reduction of the over-the-air transmission delays (which e.g. 5G standards are providing) or reduction of processing delays to further reduce the E2E delays towards a seamless support of time-critical services.

V. SIMULATION ENVIRONMENT

A. Simulation tool

A proprietary dynamic system-level simulation tool developed in C++ with an implementation of LTE presented in [28] has been used in this assessment. This simulator was used in the framework of the WINNER+ project [29], which was one of the International Mobile Telecommunications-Advanced evaluation groups of the International Telecommunications Union, and more recently in the METIS and METIS-II projects for the evaluation of the 5G system, e.g., in [30].

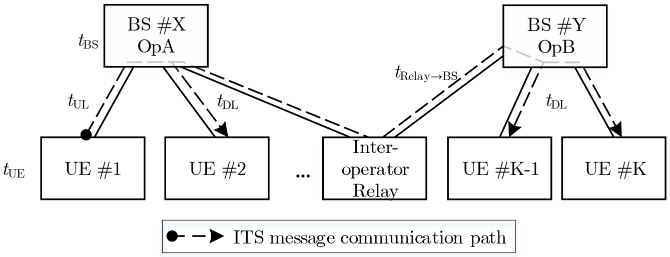

Fig. 7: Example of communication path for inter-operator relaying with the main delay components.

TABLE II: Summary of minimum communication latencies for different communication situations and the studied regional split alternatives.

Type of communication	Normal inter-op handover	Fast inter-op handover	MEC and inter-op relay
Intra-op intra-BS w/o handovers	38 ms		18 ms
Intra-op inter-BS w/o handovers	38 ms		20 ms
Inter-op w/o handovers	58 ms		25.5 ms
Intra-op intra-BS w/ handovers	78 ms		58 ms
Intra-op inter-BS w/ handovers	78 ms		60 ms
Inter-op w/ intra-op handovers	98 ms		65.5 ms
Inter-op w/ inter-op handovers	658 ms	98 ms	65.5 ms

B. Scenario Description

The simulation scenario is a closed-circuit comprising a 5.2 km rectilinear segment of the German A9 highway running in both directions from the Allianz Arena to the city of Garching bei München, an area nearby the city of Munich. The scenario also includes some intersecting fragments of the national road B471 in the northern part, and of Munich's outer ring road A99 in the southern part, together with all on- and off-ramps connecting them, as shown in Fig. 8. Note that there is a border line defining the areas served by the two different operators (OpA and OpB), the reasoning behind the location of the line is explained later in Section V-B3. The data on the geographical area is imported from OpenStreetMap through JOSM editor to SUMO [31], a popular open source tool for road traffic simulation.

1) Channel model: Because of the rectilinear geometry of the highway and the low profile of the terrain, there is line-of-sight or almost line-of-sight in all the scenario. As in [7], the pathloss model proposed for the link between BS and vehicle in a highway scenario is the Okumura-Hata model at 2 GHz:

$$PL(dB) = 128.1 + 37.6 \log(d), \tag{1}$$

where d is the distance in km. Also the shadowing model follows the indications in [7], while, for the sake of simplicity, fast fading has been modelled with a tapped delay line whose power delay profile is the well-known extended vehicular A.

2) Network deployment and system parameters: The network deployment is represented in Fig. 8. The two operators considered had four sectorized BSs (BS1–BS4) each. BS1 of both operators were co-located and their coordinates were taken from the location of an existing BS. For BS2 and BS4, the BSs of both operators were also co-located, whereas BS3 of OpA and OpB were located in opposite sides of the road. BS3 had three sectors, while the others had two sectors. Table III details the exact position of each BS in Universal Transverse Mercator (UTM) coordinates (meters), as well as the azimuth of each of the sectorial antennas in degrees, denoted as Sx, calculated clockwise with respect to the north direction. A central frequency of 2.3 GHz and a bandwidth equal to 10 MHz were considered for the simulations. The antenna pattern was a parabolic pattern, with a horizontal beamwidth of 70°, a vertical beamwidth of 10° and 20 dB of front-back ratio. Other parameters for the BSs and vehicles are summarized in Table IV.

3) Regional split: For the simulations with regional split of operators, the scenario was divided into two regions, each one served by a different operator. The border line between these regions is shown in Fig. 8. The region to the north of the border was served by OpA, while the region to the south was served by OpB. The position of the border line was selected to ensure that a vehicle traversing that line while driving along the highway experiences a good connection to at least one BS of each operator. Specifically, a good connection to the BSs 1 of OpA and OpB was ensured. As a consequence of the border

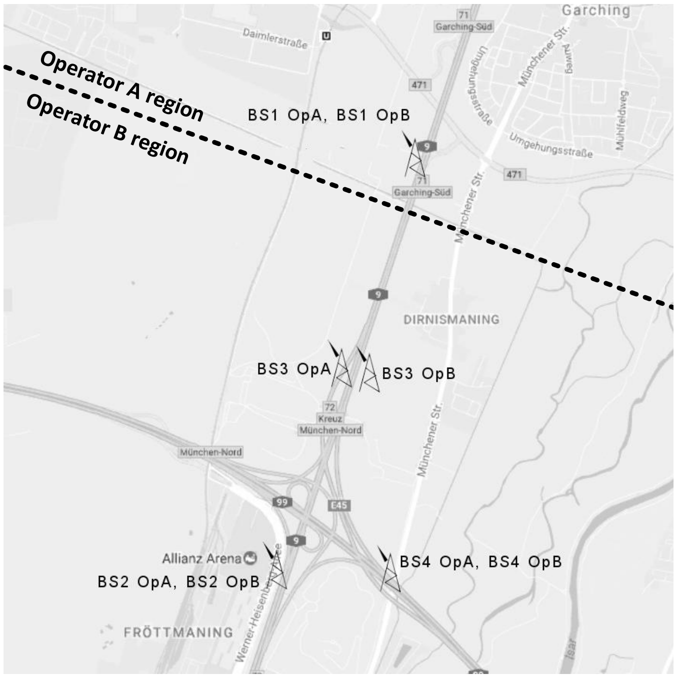

Fig. 8: Map of the simulated region in the surroundings of Munich with the considered base station positions.

TABLE III: BS coordinates and sector azimuth values

BS_name	UTM x (m)	UTM y (m)	S1	S2	S3
BS1 OpA	695870	5346448	23°	193°	-
BS1 OpB	695870	5346448	23°	193°	-
BS2 OpA	695108	5343909	290°	20°	190°
BS2 OpB	695108	5343909	290°	20°	190°
BS3 OpA	695464	5345153	23°	193°	-
BS3 OpB	695635	5345129	13°	203°	-
BS4 OpA	695804	5343915	310°	140°	-
BS4 OpB	695804	5343915	310°	140°	-

TABLE IV: System parameters

Bandwidth	10 MHz
BS TX power	46 dBm
BS number of antennas	2
BS antenna gain	17 dBi
BS antenna downtilt	6 deg
BS antenna height	12 m
BS cable loss	2 dB
BS noise figure	3 dB
Vehicle number of antennas	1 TX - 2 RX
Vehicle antenna gain	2 dBi
Vehicle cable loss	0.2 dB/m (2 m cable)
Vehicle implementation loss	5 dB
Vehicle noise figure	7 dB
UE max TX power	23 dBm
Thermal noise	-174 dBm/Hz

line location, a vehicle traversing the border line and traveling from the north to the south would perform an inter-operator handover from the BS1 of OpA to the BS1 of OpB. Note that in case of regional split, BS2, BS3, and BS4 of OpB would be deactivated, while these BSs would be active in case of considering a single operator or two global operators over the area under study.

C. Mobility model

SUMO was used to generate the mobility traces for 600 passenger cars of length 4.7 m. This amount of cars represents a density that is equivalent to the density considered in [7] for highways. The cars had an initial geographically uniform distribution. The A9 and A99 highway segments have from 3 to 5 lanes per direction (depending on the section) and a speed limit of 120 km/h. The B471 highway has 2 lanes per direction and a speed limit of 70 km/h. The on-ramps and off-ramps have either 1 or 2 lanes. The maximum driving speed was assumed to be 120 km/h and the cars had a probability equal to 0.1 of taking an exit (e.g. joining the A99 through an off-ramp), thus keeping most of the traffic on the A9. The road traffic trace produced by SUMO consisted of 140 s of simulation time, which was collected immediately after the traffic injected into the A9 reached a nearly steady-state geographical distribution. As the roads studied have a closed-loop layout and simulation was carried out after reaching steady state, traffic from on-ramps and off-ramps was well-balanced and in equilibrium.

D. Traffic models

In this work, we considered two types of traffic with different characteristics and requirements:

- Non-critical periodic messages: these messages were 300 bytes long, generated by each individual vehicle with a periodicity of 100 ms, and were relevant to all the vehicles within 320 m range of their originator with a maximum E2E delay of 100 ms, similarly to the CAM messages in [7].
- Critical messages: they had a size of 300 bytes and needed to be received by all the vehicles within 320 m range from the message origin as soon as possible.

E. Latency assumptions

Recall that in this study we assumed the interruption time due to inter-operator handover to be 300 ms in the conventional approach and 20 ms with fast inter-operator handover, as mentioned in Sections II and III, respectively. The communication delay between the inter-operator relay and its associated BS was 1 ms, as presented in Section IV. Inter-BS communication delay is assumed to be 2 ms, as indicated in Section IV-A. For the rest of delay components, the values in Table I were considered.

VI. PERFORMANCE EVALUATION

Non-critical and critical messages coexist in all the simulations, where critical messages are prioritized by the resource scheduler to keep their E2E latency as low as possible. Consequently, critical messages do not experience any scheduling delay, while non-critical messages may experience such delays.

The main aim of this performance evaluation is to illustrate the benefits of the proposed techniques in an exemplary common scenario. Despite this main aim, we think that the performance assessment is providing generally valid insights. For critical messages, given that we prioritize them in the resource allocation and given the use of SC-PTM, we consider that the performance shown is valid for different user densities and user distributions. In addition, we already consider worst-case situations in terms of the latency of messages. For non-critical messages the absolute performance values could depend on the scenario, but we have used a typical deployment (concerning inter-site distances, transmitted power, antenna configurations, etc.) to obtain meaningful values. Furthermore, the impact of our techniques on the performance of non-time-critical messages transmission is not expected to be significantly different in other scenarios.

A. Non-critical messages reception in the border area

In order to study the performance of the proposed techniques, we first focus on the performance statistics related to the reception of non-critical messages. The first relevant performance indicator in this assessment is the E2E latency experienced in the reception of messages by the vehicles within certain relevance distance from the transmitter of the messages. For the statistics of this indicator we only account the messages which are correctly received, and not those that are lost or erroneous. The second performance indicator is the Packet Reception Ratio (PRR) which, as defined in [7], measures the portion of packets successfully received (correctly and with an E2E delay lower than a maximum value) within a distance interval from the transmitter. Specifically, the average PRR in a distance interval $[a, b)$ is calculated as $\sum_i X(i)/\sum_i Y(i)$, where $Y(i)$ is the number of receivers located in the range $[a, b)$ from the transmitter of the i-th packet at the moment the packet was generated, and $X(i)$ is the number of receivers, among those counted in $Y(i)$, that successfully received the packet.

Following the reasoning in Section II-A, we use SC-PTM to distribute non-critical messages in the downlink. To optimize

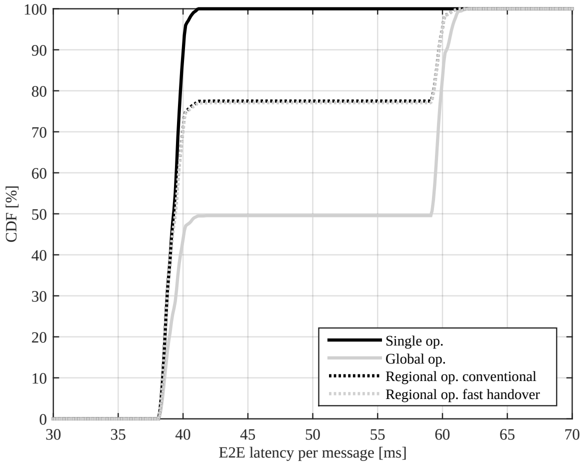

Fig. 9: CDF of the E2E latency of non-critical messages around the regional operators' border area (320 m before and after the border line).

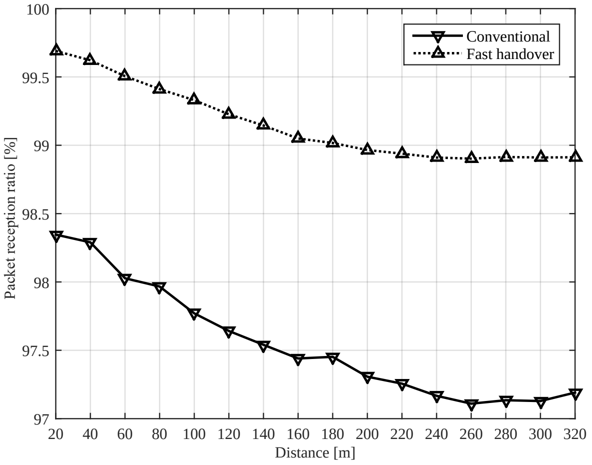

Fig. 10: PRR vs distance for non-critical messages around the regional operators' border area (320 m before and after the border line).

the system performance, we made tests with different modulation and coding schemes, after which we selected the one providing the best results in terms of maximizing the PRR.

With the aim of analyzing the performance of non-critical message transmission, we focus on the border area between the two regional operators, OpA and OpB, considering the messages generated in a segment of 320 m before and after the border line in which the vehicles perform an inter-operator handover. Recall that this inter-operator handover creates a service interruption that delays the transmission and reception of packets. To optimize the system performance, our LTE implementation drops packets whose maximum E2E delay is exceeded. In this case, the service interruption may lead to the loss of some messages.

Fig. 9 presents the Cumulative Distribution Function (CDF) of the E2E latency of non-critical messages generated in the area of the highway under study considering a relevance distance of 320 m for four different configurations. First, we have considered a single operator configuration (Single op.). This configuration provides the best achievable performance in terms of latency because in this configuration there is not any inter-operator delay. This is shown in Fig. 9 as the E2E latency concentrated around the 40 ms value. Second, we show the performance for a configuration with two global operators in which each vehicle stays connected to the same operator during the whole simulation (Global op.). In this case a 50% of message transmissions are intra-operator while the other 50% transmissions are inter-operator. This is the reason to find two steps with a height of 50% in the CDF, which are centered at the latency values expected for intra-operator (40 ms) and inter-operator (60 ms) communications. Finally, we have considered a regional split of operators with the conventional slow inter-operator handover (Regional op. conventional) and with the fast inter-operator handover (Regional op. fast handover). The CDFs for both regional split configurations are almost equal (this is due to the fact that none of the approaches make special provisions for non-critical messages) and similar (in term of general shape) to the global operators configuration (because all three configurations comprise intra-operator as well as inter-operator communications). The CDFs present two steps centered at the same positions as the global operators configuration, but the height of the first step is higher for regional split configurations since the portion of intra-operator communications is higher in this case. In the area under study, this portion is close to the 80% and hence the step height in Fig. 9. In case of showing the results in a broader region, this portion would be higher and the CDF would be close to the CDF of the single operator case. As expected from the explanation in the previous sections, some messages experience higher latencies than those covered by the x-range in Fig. 9 but the percentage of messages that undergo such higher latencies is very low. The latencies for both regional split configurations are similar, despite the higher interruption time of the conventional approach, because the disruption is producing packet losses, and lost packets are not considered in the latency statistics.

In Fig. 10 we show the PRR vs distance for non-critical messages around the regional operators' border, which is a typical representation in the literature. The i-th point of the curves represents the average PRR of all the message receptions within the range of distances $[(i-1)\times 20, i\times 20)$. In this case, for the sake of clarity, we provide the performance for the regional split with conventional inter-operator handover (Conventional) and for the regional split with fast inter-operator handover (Fast handover). It can be observed that the proposed fast handover provides a better PRR thanks to the shorter interruption period that it implies, which is translated in less lost packets. The absolute PRR difference is in this case between 1.25% and 1.8%. This difference would be higher if the statistics were obtained in a narrower area as we will show in the following section.

B. Critical message reception in the border area

Critical messages require a PRR as high as possible. To ensure such high reliability, we use for the transmission of those messages the most robust modulation and coding scheme available in LTE (QPSK modulation and code rate 0.15). This fact implies that the transmission of a critical message of 300 bytes needs two subframes to be completed. Therefore, each radio transmission requires 1 ms more than the values considered in the previous latency analysis, and the actual E2E delays will increase accordingly.

For the performance study of time-critical message transmission, we also focus on the border area between the two regional operators but only on one way of the highway, from Garching bei Munchen to Munich, i.e. from north to south in Fig. 8. Driving along this way, when vehicles reach the regional operators' border, they perform an inter-operator handover from the BS1 of OpA to the BS1 of OpB.

To assess the performance of the proposed techniques we have simulated a specific situation in which a vehicle close to the regional operators' border and served by OpB generates a DENM message. For the sake of studying the worst case scenario, we assume that the vehicle starts an inter-operator handover just before sending the DENM message. We study the PRR, i.e. the percentage of intended receivers that receive a message correctly and before the maximum delay, as a function of the directed distance of the receiver to the border area in the chosen direction. Negative distances mean that the receiver is located to the south of the border, in the OpB region. Positive distances indicate that the receiver is located to the north of the border, in the OpA region. Therefore, in the way of the highway under study, i.e. from north to south, the vehicles move from positive directed distances to negative directed distances.

In Fig. 11 we show the PRR as a function of the directed distance for a maximum E2E delay of 100 ms. As it can be observed, the conventional approach provides a 0% PRR. The reason is that we have assumed that the transmitter is performing an inter-operator handover, which in the conventional case takes 300 ms. This value is longer than the maximum E2E delay considered, 100 ms, and thus prevents the successful reception of packets. With the fast handover approach, the PRR is 100% in the majority of locations. This is due to the fact that vehicles are pre-attached to both operators, and thus the inter-operator handover interruption time is basically shorter and equivalent to the duration of a typical intra-operator handover. However, there is a slight degradation produced around the border between regional operators. The reason is that around the border both the transmitter and the receivers are experiencing an inter-operator handover and then the communication latency is at least 98 ms as shown in Table II. Subframe alignment delays and the above-mentioned additional radio transmission delays may provoke a latency higher than 100 ms and explain the PRR being lower than 100% around the border. The use of MEC and inter-operator relays allows to fulfill the requirement in every position. As shown in Table II, with MEC and inter-operator relays the latency values are small enough to bear additional delays.

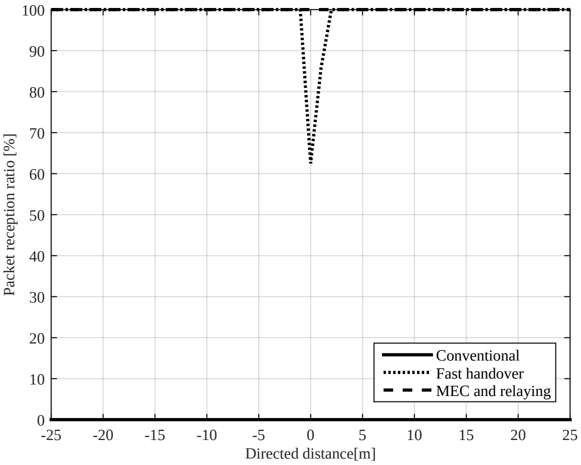

Fig. 11: PRR as a function of the directed distance of a receiver to the border area in the north-to-south way of the highway for a maximum E2E delay of 100 ms with the transmitter performing a handover.

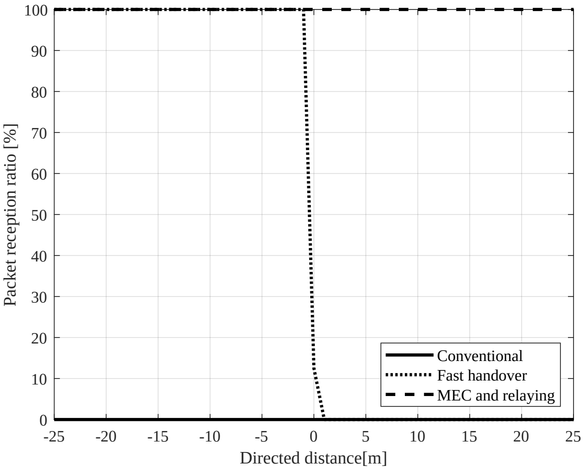

Fig. 12: PRR as a function of the directed distance of a receiver to the border area in the north-to-south way of the highway for a maximum E2E delay of 70 ms with the transmitter performing a handover.

The same results are obtained when the maximum latency is reduced down to 90 ms, approximately. However, as can be seen in Fig. 12, for 80-70 ms the fast handover approach presents a poor performance for the receivers in one side of the highway. Specifically, those served by the operator different to that which serves the transmitter. In that case, the inter-operator handover time of the transmitter plus the inter-operator communications delay result in at least 78 ms latency, which explains the bad performance for maximum delays of 70 ms. The addition of subframe alignment delays and the additional radio transmission delays prevents also the fulfilment of a maximum delay of 80 ms.

For a maximum E2E delay of 50 ms, see Fig. 13, the fast

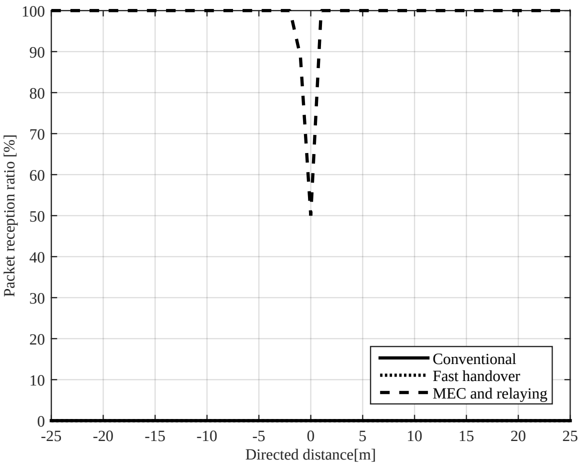

Fig. 13: PRR as a function of the directed distance of a receiver to the border area in the north-to-south way of the highway for a maximum E2E delay of 50 ms with the transmitter performing a handover.

handover approach provides a 0% PRR since, in this case, even the intra-operator communication takes more than 58 ms (the 38 ms reported in Table II for inter-operator communications without handovers plus the addition of the 20 ms needed in this case by the transmitter to perform an inter-operator handover). Only the use of MEC and inter-operator relays is able to provide acceptable PRR values in almost all the scenario except around the border line. At that point is where we find the worst-case and where additional solutions will be needed for the studied maximum E2E delay.

VII. CONCLUSION

Infrastructure-based V2V communications between devices served by a single operator present several benefits compared to multi-operator scenarios such as small latency or reduced complexity. The single-operator scenario also simplifies the deployment of some extra solutions to enforce a localized treatment of V2V traffic. However, it is not easy to imagine that in every country only a V2V cellular operator will exist. Therefore, a good solution would be to allow the existing operators to accommodate a regional split scheme. With a suitable accounting scheme, the operators could even regulate a fairer revenue of money even for those that will serve the less privileged areas. This paper has dealt specifically on how to support such a regional split scheme, by guaranteeing a proper reception of critical and non-critical messages even in the border between the regions covered by two different operators, and with terminals equipped with only one RF unit. Our solution is also valid in the national borders where, even in the case of having national V2V operators, different operators are expected at each side of the border.

The first proposal of this paper has been a solution for fast inter-operator handover based on the UE pre-attachment to all the involved operators. Apart from the description of the new functional entities to be added to the 5G CN, this paper has also dealt with the required signaling of this pre-attachment and its mobility support. The immediate effect of this solution is the improvement in the latency budget and PRR of the transmitted packets, for both non-critical and critical services.

The latency reduction obtained by the fast handover technique is, unfortunately, still insufficient to guarantee the correct operation of critical services. To further decrease the communication delay, this paper has presented a complementary idea where MEC is used to disseminate time-critical messages and is used in conjunction with inter-operator relays in border areas. The new local functionalities of the BSs can reduce the packet latency of critical messages, while the use of relays minimizes the latency among operators.

The numerical analysis conducted has proven that the proposed fast handover approach combined with MEC and inter-operator relaying is useful for the successful delivery of critical messages in the majority of locations given the latency around 20 ms achieved in intra-operator communications. At the same time, the worst-case latency in the border areas with our proposal is lower than 70 ms, which is significantly better than the values obtained with conventional techniques. The performance evaluation has corroborated these findings based on simulations in a realistic scenario.

However, the relaying alternative requires the deployment of relay units in border areas, which entails additional infrastructure costs. Despite this, the relaying alternative is less expensive than a solution based on equipping the cars with two RF units, as the extra cost affects only the infrastructure instead of every terminal.

Future work will include the study of the reduction in processing and transmission delays that could be obtained in forthcoming communication standards together with its impact on the E2E latency. The reduction of those delays is needed to reach latency values lower than 70 ms with high guarantees in border areas as the present work has shown.

ACKNOWLEDGMENT

The work of S. Roger was partially supported by the Spanish Ministry of Science, Innovation and Universities through grant number RYC-2017-22101.

REFERENCES

[1] G. Araniti, C. Campolo, M. Condoluci, A. Iera, and A. Molinaro, “LTE for vehicular networking: a survey,” *IEEE Commun. Mag.*, vol. 51, no. 5, pp. 148–157, May 2013.

[2] J. Calabuig, J. F. Monserrat, D. Gozalvez, and O. Klemp, “Safety on the Roads: LTE Alternatives for Sending ITS Messages,” *IEEE Veh. Tech. Mag.*, vol. 9, no. 4, pp. 61–70, Dec. 2014.

[3] C. Campolo, A. Molinaro, and R. Scopigno, “From today’s VANETs to tomorrow’s planning and the bets for the day after,” *Veh. Commun.*, vol. 2, no. 3, pp. 158–171, Jul. 2015.

[4] 3GPP TSG SA, “Study on LTE support for Vehicle-to-Everything (V2X) services,” 3GPP, TR 22.885 V14.0.0, Dec. 2015.

[5] ——, “Study on enhancement of 3GPP support for 5G V2X services,” 3GPP, TR 22.886 V15.0.0, Mar. 2017.

[6] CAMP Consortium, “Vehicle Safety Communications project task 3 final report: identify intelligent vehicle safety applications enabled by DSRC,” U. S. Department of Transportation (USDOT), Tech. Rep. DOT HS 809 859, Mar. 2005.

[7] 3GPP TSG RAN, “Study on LTE-based V2X services,” 3GPP, TR 36.885 V14.0.0, Jun. 2016.

[8] A. Kaloxylos, "LIVA: an Efficient Local Breakout Scheme for V2V Communications," in *21st Pan-Hellenic Conference on Informatics (PCI 2017)*, Larisa, Greece, Sep. 2017, pp. 1–6.

[9] K. Chatzikokolakis, P. Spapis, A. Kaloxylos, and N. Alonistioti, "Toward spectrum sharing: opportunities and technical enablers," *IEEE Commun. Mag.*, vol. 53, no. 7, pp. 26–33, Jul. 2015.

[10] H. Kamal, M. Coupechoux, and P. Godlewski, "Inter-operator spectrum sharing for cellular networks using game theory," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Tokyo, Japan, Sep. 2009, pp. 425–429.

[11] P. R. Milgrom, J. Levin, and A. Eilat, "The Case for Unlicensed Spectrum," *SSRN*, Oct. 2011. [Online]. Available: http://dx.doi.org/10.2139/ssrn.1948257

[12] A. Kousaridas, C. Zhou, D. Martín-Sacristán, D. Garcia-Roger, J. F. Monserrat, and S. Roger, "Multi-Connectivity Management for 5G V2X Communication," in *IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Instanbul, Turkey, Sep. 2019, pp. 1–7.

[13] S. Roger, D. Martín-Sacristán, D. Garcia-Roger, J. F. Monserrat, P. Spapis, C. Zhou, and A. Kaloxylos, "Forced inter-operator handover for V2X communication in multi-operator environments with regional splitting," in *13th IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Valencia, Spain, Jun. 2018, pp. 1–5.

[14] A. Kousaridas and C. Zhou, "Local end-to-end paths for low latency vehicular communication," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, Porto, Portugal, Jun. 2018, pp. 1–7.

[15] S. Roger, D. Martín-Sacristán, D. Garcia-Roger, J. F. Monserrat, P. Spapis, A. Kousaridas, S. Ayaz, and A. Kaloxylos, "Low-latency layer-2-based multicast scheme for localized V2X communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 8, pp. 2962–2975, Aug. 2019.

[16] F. Giust, G. Verin, K. Antevski, J. Chou, and et. al., "MEC Deployments in 4G and Evolution Towards 5G," ETSI, White Paper No. 24, Feb. 2018.

[17] D. Martín-Sacristán, S. Roger, D. Garcia-Roger, J. F. Monserrat, A. Kousaridas, P. Spapis, S. Ayaz, and C. Zhou, "Evaluation of LTE-Advanced Connectivity Options for the Provisioning of V2X Services," in *IEEE Wireless Communications and Networking Conference (WCNC)*, Barcelona, Spain, Apr. 2018, pp. 1–6.

[18] 3GPP TSG RAN, "Feasibility study for Further Advancements for E-UTRA (LTE-Advanced)," 3GPP, TR 36.912 V14.0.0, Mar. 2017.

[19] K. Lee, J. Kim, Y. Park, H. Wang, and D. Hong, "Latency of Cellular-Based V2X: Perspectives on TTI-Proportional Latency and TTI-Independent Latency," *IEEE Access*, vol. 5, pp. 15 800–15 809, Jul. 2017.

[20] GSMA, "Guidelines for IPX Provider Networks," GSM Association, Official Document IR.34 V13.0, Oct. 2014. [Online]. Available: http://www.gsma.com/newsroom/wp-content/uploads/ IR.34-v13.0-1.pdf

[21] R. Trivisonno, R. Guerzoni, I. Vaishnavi, and D. Soldani, "Towards zero latency software defined 5G networks," in *IEEE International Conference on Communication Workshop (ICCW)*, London, UK, Jun. 2015, pp. 2566–2571.

[22] 3GPP TSG CT, "Non-Access-Stratum (NAS) functions related to Mobile Station (MS) in idle mode," 3GPP, TS 23.122 V13.5.0, Jun. 2016.

[23] 3GPP TSG RAN, "General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access," 3GPP, TS 23.401 V14.0.0, Jun. 2016.

[24] 3GPP TSG SA, "Technical Specification Group Services and System Aspects; Network architecture," 3GPP, TS 23.002 V14.1.0, Mar. 2017.

[25] ——, "Technical Specification Group Services and System Aspects; System Architecture for the 5G System; Stage 2," 3GPP, TS 23.501 V15.0.0, Dec. 2017.

[26] D. Han, S. Shin, H. Cho, J.-M. Chung, D. Ok, and I. Hwang, "Measurement and Stochastic Modeling of Handover Delay and Interruption Time of Smartphone Real-Time Applications on LTE Networks," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 173–181, Mar. 2015.

[27] W. Kim and E.-K. Lee, "LTE Network Enhancement for Vehicular Safety Communication," *Mobile Information Systems*, vol. 2017, no. 12, pp. 1–18, Feb. 2017.

[28] D. Martín-Sacristán, J. F. Monserrat, V. Osa, and J. Cabrejas, "LTE-Advanced System Level Simulation Platform for IMT-Advanced Evaluation," *Waves*, vol. 3, pp. 15–24, 2011.

[29] K. Safjan, V. D'Amico, D. Bultmann, D. Martín-Sacristán, A. Saadani, and H. Schoneich, "Assessing 3GPP LTE-Advanced as IMT-Advanced technology: The WINNER+ evaluation group approach," *IEEE Commun. Mag.*, vol. 3, no. 2, pp. 92–100, Feb. 2011.

[30] D. Martín-Sacristán, C. Herranz, and J. F. Monserrat, "Traffic safety in the METIS-II 5G connected cars use case: Technology enablers and baseline evaluation," in *2017 European Conference on Networks and Communications (EuCNC)*, Oulu, Finland, Jun. 2017, pp. 1–5.

[31] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent Development and Applications of SUMO - Simulation of Urban MObility," *Int. J. on Advances in Systems and Measurements*, vol. 5, no. 3&4, pp. 128–138, Dec. 2012.

David Martín-Sacristán received the M.Sc. and Ph.D. degrees in telecommunications from the Universitat Politècnica de València (UPV), Spain, in 2006 and 2016, respectively. He is currently a Researcher with the iTEAM Research Institute, UPV. He has been involved in European projects such as WINNER+, that was an external evaluator of IMT-Advanced technologies for the ITU, or METIS and METIS-II, involved in the development of 5G. He is currently part of the 5G-CARMEN project focused on 5G for connected and automated road mobility. His research interests are focused on the modeling and simulation of communication networks, resource management, massive MIMO, and vehicular communications.

Sandra Roger (SM'17) received the M.Sc. and Ph.D. degrees in telecommunications from the Universitat Politècnica de València (UPV), Spain, in 2007 and 2012, respectively. During her doctorate studies, she performed two research stays at the Institute of Telecommunications, Vienna University of Technology, Austria. From July 2012 to December 2018, she was a Senior Researcher with the iTEAM Research Institute, UPV, where she worked in the European METIS and METIS-II projects on 5G design. In January 2019, she joined the Computer Science Department of the Universitat de València as a Senior Researcher ("Ramon y Cajal" Fellow). Dr Roger has co-authored around 60 papers in renowned conferences and journals. Her main research interests are in the fields of signal processing for wireless communications and vehicular communications.

David Garcia-Roger received the Ph.D. degree in telecommunications from the Universitat Politècnica de València (UPV), Spain, in 2007. He has participated as a Researcher in European projects ALPHA and METIS-II. He has published five journal and 25 conference papers. His research interests include the evaluation of 5G candidate technologies and vehicular communication systems.

Jose F. Monserrat (SM'14) is full professor at the Universitat Politècnica de València (UPV), Spain. His research focuses on the design of future 5G wireless systems and their performance assessment. He has been involved in several European Projects, like METIS/METIS-II where he led the simulation activities, or currently 5G-CARMEN and 5G-SMART. He co-edited the Wiley book "Mobile and wireless communications for IMT-Advanced and beyond" and the Cambridge book "5G Mobile and Wireless Communications Technology". He has published more than 60 journal papers. Currently his research team consists of 5 Postdoctoral fellows, 8 PhD students and 2 Master students.

Panagiotis Spapis received his Diploma degree in electrical engineering, in 2008, and his Ph.D. degree in telecommunications, in 2015. Between 2009 and 2014 he had been employed as a Researcher in the Department of Informatics and Telecommunications, University of Athens. Since 2014, he has been a Research Engineer with the 5G RAN Group of Huawei German Research Center, Munich. He has been participating in several EU projects and he has published parts of his work in numerous journals and conferences. His research interests include the areas of vehicular communications and situation and context awareness in 5G systems.

Chan Zhou received a Dipl.-Ing. degree in computer science from the Technical University of Berlin in 2003 and a Ph.D. degree in electrical engineering in 2009. From 2003 to 2009 he worked at Fraunhofer Heinrich-Hertz-Institute for telecommunications, where he engaged in research activities with a focus on the realization of high-speed data transmission in cellular networks. In 2010 he joined Huawei German Research Center in Munich. His expertise spans from early 5G research toward prototyping and demonstrations for various 5G use cases. He is currently the head of the 5G radio access network research group at Huawei German Research Center, where he has been deeply involved in mobile and wireless applications for vertical industries.

Alexandros Kaloxylos (B.Sc., M.Phil., Ph.D.) received his Ph.D. degree in Informatics and Telecommunications from the University of Athens in 1999. Since 1994, he has participated, as a researcher, in numerous projects realized in the context of EU Programmes as well as national initiatives. In 2002 he joined the faculty of the University of Peloponnese, where he is currently an Associate Professor in the Department of Informatics and Telecommunications. From 2014 until 2017, he worked as a Principal Researcher in Huawei's European Research Center in Munich, where he led the Radio Access Network team in the design of 5G networks. During 2016, he was the vice-chair of the 5G PPP Architecture WG. Since September 2019, he is the executive director of the 5G Infrastructure Association. He has published over 130 papers in international journals, conferences and book chapters and has filed several patents for topics related to 5G networks.