Multicast and Broadcast Services Over Mobile Networks: A Survey on Standardized Approaches and Scientific Outcomes

Multicast and broadcast services over mobile networks: a survey on standardized approaches and scientific outcomes

Domenico Striccoli, Giuseppe Piro, *Member, IEEE,* and Gennaro Boggia, *Senior Member, IEEE*

Abstract

Due to the increasing demand for pervasive video and broadcast-like applications, multicast and broadcast communications are expected to assume an important role in upcoming 5G systems. The current research trend is trying to reuse, extend, or adapt reference transmission strategies already designed for the conventional 4G technology. Nevertheless, apart the two reference and standardized methodologies, i.e., Multimedia Broadcast/Multicast Service and Single Cell-Point To Multipoint, many technical extensions and novel solutions were published in the literature so far. Therefore, in order to provide a clear overview on available solutions (already standardized or simply extending conventional approaches), the present work provides a comprehensive survey on network architectures, communication protocols, transmission strategies, and optimization algorithms to improve the performance of multicast communications over mobile radio systems. The core of the conducted study represents a structured taxonomy, able to properly classify scientific contributions based on their reference standard, targeted goal, addressed methodology, considered application domain, and obtained results. Taking into account this taxonomy, more than one hundred of scientific contributions are presented, classified, and reviewed. The study of the state of the art is further enhanced with the discussion on important lessons learned, which clearly highlight the pros and cons of any investigated approach. A focus is also provided on the main issues on Long Term Evolution multicasting that need to be better investigated, and determine the possible future research directions on this subject. The final goal of this work is to support research activities devoted to the identification of promising methodologies, that efficiently support the delivery of real-time and on-demand video contents in a TV-like fashion.

Index Terms

Multicast and broadcast communication schema, MBMS, MBSFN, SC-PTM

D. Striccoli, G. Piro and G. Boggia are with Dept. of Electrical and Information Engineering (DEI), Politecnico di Bari, Via E. Orabona 4, 70125, Bari, Italy. e-mail: name.surname@poliba.it

Multicast and broadcast services over mobile networks: a survey on standardized approaches and scientific outcomes

LIST OF ACRONYMS

3G 3rd generation.

3GPP 3rd Generation Partnership Project.

4G 4th generation.

5G 5th generation.

ADR Aggregate Data Rate.

AL-FEC Application Layer Forward Error Correction.

AMC Adaptive Modulation and Coding.

ARQ Automatic Repeat reQuest.

BER Bit Error rate.

BL Base Layer.

BLER BLock Error Rate.

BM-SC Broadcast Multicast Service Centre.

BS Base Station.

CC Component Carrier.

CH Cluster Head.

CMS Conventional Multicast Scheme.

CQI Channel Quality Index.

CR Cognitive Radio.

CRN Cognitive Radio Network.

CSI Channel State Information.

D2D Device-to-Device.

DL Downlink.

DRX Discontinuous Reception.

EL Enhancement Layer.

eMBMS enhanced Multimedia Broadcast/Multicast Service.

eNB eNodeB.

ESS Exhaustive Search Scheme.

FEC Forward Error Correction.

GA Genetic Algorithm.

HARQ Hybrid Automatic Repeat reQuest.

LTE Long Term Evolution.

LTE-A Long Term Evolution-Advanced.

M2M Machine-to-Machine.

MAC Medium Access Control.

MBMS Multimedia Broadcast/Multicast Service.

MBMS-GW Multimedia Broadcast/Multicast Service GateWay.

MBSFN Multicast/Broadcast Single Frequency Network.

MCE Multi-cell/multicast Coordination Entity.

MCS Modulation and Coding Scheme.

MG Multicast Group.

MIMO Multiple Input Multiple Output.

MME Mobility Management Entity.

MS Mobile Station.

NC Network Coding.

OMS Opportunistic Multicast Scheduling.

P2P Peer-to-Peer.

PDU Packet Data Unit.

PER Packet Error Rate.

PF Proportional Fair.

PLR Packet Loss Ratio.

PSNR Peak Signal to Noise Ratio.

PTM Point To Multipoint.

PtP Point to Point.

QoE Quality of Experience.

QoS Quality of Service.

RB Resource Block.

RLC Radio Link Control.

RLNC Random Linear Network Coding.

RNC Random Network Coding.

RRM Radio Resource Management.

RS Relay Station.

SC-PTM Single Cell-Point To Multipoint.

SE Spectral Efficiency.

SFN Single Frequency Network.

SINR Signal to Interference and Noise Ratio.

SNR Signal to Noise Ratio.

SR Short Range.

SSIM Structural Similarity Index Metric.

SVC Scalable Video Coding.

TTI Transmission Time Interval.

UE User Equipment.

UL Uplink.

VANET Vehicular Ad-hoc NETwork.

I. Introduction

The explosive growth of smart and capable mobile devices continuously requires the deployment of suitable and wireless communication tecnologies, able to distribute data to a massive number of users, while ensuring Quality of Service (QoS), network capacity, spectral efficiency, and service reliability, as well as very low latencies, limited power consumption, and better radio resource utilization [1], [2]. To this end, the current research activities are focusing on several innovations that will constitute the core of the so-called *5th generation (5G) wireless communication systems* [3], [4]. In this context, due to the increasing demand for pervasive video and broadcast-like applications, *one-to-many* communication schema (i.e., *multicast* or *broadcast* communications) are expected to assume an important role in future 5G networks [4]–[6]. Therefore, it is necessary to identify promising methodologies, able to efficiently support the delivery of real-time and on-demand video contents in a TV-like fashion.

The current research trend envisages the possibility to leverage (i.e., reuse, extend, or adapt) multicast or broadcast transmission strategies already designed (or still under development) for the foregoing 4th generation (4G) technology, which include Long Term Evolution (LTE) and Long Term Evolution-Advanced (LTE-A). The two reference methodologies are Multimedia Broadcast/Multicast Service (MBMS) and its evolution, namely enhanced Multimedia Broadcast/Multicast Service (eMBMS) or Multicast/Broadcast Single Frequency Network (MBSFN), and Single Cell-Point To Multipoint (SC-PTM). From one side, MBMS is standardized by the 3rd Generation Partnership Project (3GPP) [7]. It integrates the MBSFN technology which defines a broadcast area spanning over multiple cells, where a multicast flow is transmitted in all of the participating cells, simultaneously and on the same frequency band (i.e., using dedicated, pre-planned, and semi-statically configured time slots) [8]. From another side, SC-PTM is still under standardization and preliminary 3GPP specifications already exist [9]. SC-PTM restricts the broadcasting area to a single cell, allowing a flexible radio resource distribution among unicast and broadcast/multicast flows and supporting more performant physical layer interfaces [10].

At the time of this writing, a lot of scientific contributions already investigated the performances of both MBMS and SC-PTM heterogeneous application domains. See, for instance, the works discussed in [11]–[14] for MBMS, as well as [15], [16] for SC-PTM. In addition, the research community formulated many technical enhancements to the baseline approaches (i.e., those just standardized by 3GPP). As a consequence, the current literature on multicast and broadcast communications over mobile radio systems covers an explosion of possible candidate techniques and solutions for upcoming 5G systems.

Based on these premises, this work wants to provide a comprehensive survey on baseline and novel strategies for multicast and broadcast communications over mobile wireless networks. To this end, the scientific literature is analyzed and the available contributions classified and argued according to a properly suggested structured taxonomy. First of all, two main categories are identified: the first contains all the works focusing on MBMS and MBSFN, and the second embraces all the contributions focusing on SC-PTM. Then, for each of these categories, sub-categories are introduced for efficiently grouping works that leverage similar methodologies and/or reach equivalent goals. In the authors humble opinion, the resulting study offers a clear overview on the state-of-the-art of network architectures, communication protocols, transmission strategies, and optimization algorithms to improve the performance of multicast and broadcast communications over mobile radio systems, and for this reason it could be very useful for researchers working on this topic.

With reference to MBMS, the proposed survey starts with the description of standardized approaches, based on the MBSFN architecture [11]–[14], [17]–[74]. The summary of technical details (including network architectures [17]–[33], cooperation and inter-networking with other communication technologies [34]–[43], error correction techniques [44]–[49], scheduling strategies [50]–[54], physical layer [11], [13], [55]–[63], the implementation of LTE environment through simulation and emulation tools [64], [65]) and the comments on reference results already reported in other contributions [12], [14], [66]–[70] together with the evaluation of performance of LTE multicast [71]–[74], immediately highlight how MBMS is ready to support TV-like services in mobile systems [22]–[24], [68], as well as other multicast applications also in other different contexts like vehicular communications [26]–[28]. Nevertheless, to improve the overall performance of multicast services, MBMS has been extended at different layers of the protocol stack. Indeed, to better differentiate the solutions presented in the literature, the survey introduces four sub-categories, which embrace works proposing some novel techniques for the physical layer (including use multi-antenna transmission schema [75], [76] and strategies aiming to optimize power consumption [77]–[85], coverage [86]–[95], spectrum usage [96]–[100], and data rate [101]–[112]), the Medium Access Control (MAC) layer (like error protection strategies [113]–[119] and novel radio resource allocation mechanisms [120]–[126]), the network layer [127], and the application layer (including multiplexing techniques [128]–[130], cooperative schema for video multicasting [131], error protection strategies [132], [133] and layered video multicasting [134]–[137]). In addition, another sub-category is added to the conceived taxonomy in order to take care about the contributions that propose cross-layer solutions in the fields of error protection [138]–[140], synchronization of the data transmission [141], and cross-layer architectures [142], [143].

A similar study is presented also for the SC-PTM main category. An initial discussion is dedicated to the analysis of standardized aspects characterizing the SC-PTM technology and to the review of works proposing a preliminary performance evaluation [15], [16]. Then, the rest of works available in the current literature and proposing some advancements to the baseline approach are classified in three sub-categories. Specifically, the survey investigates the contributions offering novel strategies for the physical layer (including multi-antenna transmission schema [144], strategies for power optimization and control [145], [146], and algorithms for radio resource allocation [147]) and the MAC layer (like signaling procedures [148], packet scheduling strategies [149], and layered services [150]). Also in this case, the cross-layer sub-category is introduced to properly identify the few proposals that jointly work at different layers of the protocol stack [151], [152].

After having reviewed more than one hundred of works, a list of important lessons learned is presented. The discussion covers several interesting aspects, including key features related to standardized methodologies, and the pros and cons characterizing both standardized methodologies emerging from the state of the art and novel and advanced solutions extending the aforementioned standardized methodologies, in terms of strengths and weaknesses, and complexity issues. Finally, a discussion is carried out on the main challenges and open issues on this very complex theme, and the potential research directions to address them.

The rest of this paper is organized as follows. Section II describes the standardization efforts made for service multicasting in LTE, describing the main features of the multicast transmission schema and the related overall architectures. In section III the main differences are highlighted between this survey and other similar works that classify and describe multicasting strategies in wireless networks. Section IV presents the proposed taxonomy, which classifies the contributions available in the current literature. Section V focuses on the MBMS architecture and presents the main aspects already standardized by the 3GPP, while Section VI further complements the former by discussing all the enhancements available in the current literature. Section VII investigates the Point To Multipoint (PTM) methodology by jointly presenting what has been already standardized by the 3GPP and what are the novel innovations emerging from the current literature. Section VIII provides more detailed insights on the network coding techniques and optimization algorithms, two subjects that are present in several surveyed works. Lessons learned from the conducted study are discussed in Section IX. The main issues and possible future research directions are discussed in Section X. Finally, Section XI reports the conclusions of this work.

II. Standardized multicast/broadcast architectures and functionalities

The goal of this section is to describe the main features of the multicast transmission schema adopted in LTE, and the related overall architectures, to give a comprehensive idea of the research efforts taken by all the works analyzed in this survey. More specifically, Section II-A describes in detail the main functionalities of the MBMS scheme, its goals, the main entities and the phases needed to provide a multicast or a broadcast service to different User Equipments (UEs). Section II-B describes different kinds of architectures that include MBMS, such as the architecture based on the only MBMS, hybrid architectures that utilize MBMS in combination with other technologies, and cooperative architectures where cooperation among the different network entities aims at improving data multicasting and broadcasting. Section II-C describes the point-to-multipoint multicast transmission in a single-cell context, and the main differences with respect to the multicast transmission in a multi-cell scenario. Finally, Section II-D points out the main differences between MBMS and other widely used and promising kinds of wireless networks, for what concerns service multicasting and broadcasting.

A. MBMS and MBSFN architectures and overall functionalities

The main aspects of the MBMS services are covered in [7]. More specifically, MBMS is defined as a service transmitted from a single source to different destinations (i.e., a point-to-multipoint transmission). Two transmission modes are supported in MBMS: multicasting (data are received by many, but not all, destinations) and broadcasting (data are received by all the destinations in the signal range). In both cases, Hybrid Automatic Repeat reQuest (HARQ) feedbacks and Channel State Information (CSI) reports are not supported, because data are multicasted over the same, common channel. The main goal of MBMS is to save as much as possible radio and network resources, by sending data only once on a common channel, regardless of the number of Base Stations (BSs) and UEs wishing to receive it. This very important task is performed by the *MBMS Bearer Service*, i.e., a service provided by the network that multicasts packets to final users. They are first aware of the presence of a MBMS service through the *MBMS Service Announcement* mechanism [7]. Users receive the service provided by the MBMS Bearer Service by means of another service, seen at the receiver side, called *MBMS User Service*. The area containing all the UEs receiving a MBMS session of the MBMS Bearer Service is defined as *MBMS Service Area*, and is usually composed by all the cells transmitting the same MBMS service [153]. According to [7], the provision of a MBMS service is performed through the following phases:

- Subscription: in this phase users agree to receive MBMS services by the provider.
- Service Announcement: through this phase, all the users related to the provider are notified of all the available services.
- Joining: in this phase, users become members of a multicast group, and agree to receive data of a specific MBMS Bearer Service.
- Session Start: this phase indicates the beginning of a multicast session.
- MBMS Notification: the UE is notified on the beginning of data transfer.
- Data Transfer: in this phase, data are effectively delivered to UEs.
- Session Stop: indicates the end of the session since no more data have to be transferred. The bearer resources are released in this phase.
- Leaving: the user leaves the multicast group he/she joined in the Joining phase.

The sequence of phases for the broadcast mode is the same of the multicast mode, except for the Service Subscription, Joining and Leaving phases, which are not present (since they are not necessary for service broadcasting). Fig. 1 illustrates the phases described above, and includes both the multicast and broadcast modes.

The MBMS architecture consists of several functional entities, graphically illustrated in Fig. 2. The

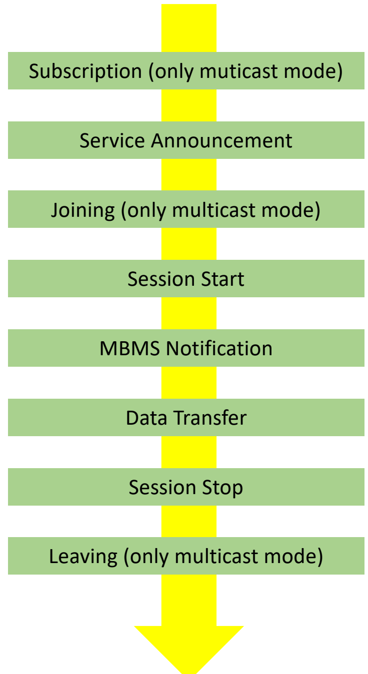

Fig. 1. The sequence of phases needed to provide a MBMS service. Users agree to receive a service (Subscription) and are notified of all the available services (Service Announcement); then, they join a multicast group (Joining) and the multicast session begins (Session Start), notifying UEs on the beginning of data transfer (Notification). Data are then transferred to UEs (Data Transfer) until the session ends (Session Stop) and users leave the multicast group (Leaving).

UE connects to the eNodeB (eNB) that makes part of the Universal Terrestrial Radio Access Network (UTRAN) which logically connects the UE to the core network and provides user and control functionalities to the UE through the related protocols, to deliver MBMS services to MBMS service areas [8]. Specifically, the UTRAN supports the choice of the most suitable radio bearer depending on the number of users in a service area, the initiation and termination of MBMS services, the IP multicast method (that allows to send an IP packet to multiple receivers into a single transmission, without employing a transmission of the packet for each node), and receivers mobility [7]. A UE supports functionalities for activation and deactivation of MBMS bearer service, security, service announcements and storing of MBMS data. The eNB is composed by the antenna providing the physical signal, and the Multi-cell/multicast Coordination Entity (MCE), the logical entity that performs allocation of radio resources in terms of radio bearers and Modulation and Coding Schemes (MCSs), the choice of the multicast transmission mode (i.e., SC-PTM or MBSFN), and the suspension and resumption of multicast services [8], [154]. The MCE is connected to the Evolved Packet Core (EPC) network through the Mobility Management Entity (MME) and Multimedia Broadcast/Multicast Service GateWay (MBMS-GW) entities. MME performs essentially control functionalities to support MBMS services, i.e., session control (by transmitting and filtering different types of control messages to UTRAN nodes), provisioning and modification of the list of MBMS service areas served by the MCEs. It is also interfaced with MBMS-GW for the management of IP multicast addresses. The MBMS-GW provides a Control Plane (CP) and User Plane (UP) interface (see Fig. 2) for entities that use MBMS bearers, for control and used data purposes respectively. It is also responsible for the distribution of IP multicast data, and the allocation of IPv4/IPv6 multicast addresses [7]. The Broadcast Multicast Service Centre (BM-SC) logical entity, placed in the EPC, provides the IP multicast distribution of user data, the allocation of multicasting IP addresses, and an interface for both user and control plane functionalities between the UTRAN and the content provider through the BM-SC, which is the connection point between the content provider and the MBMS-GW (see Fig. 2). It performs functionalities of membership, activation and deactivation of MBMS bearer services, security issues (authentication, authorization and billing), service announcements and initiation, and content synchronization [7], [154].

As explained above, MBMS services are multicasted/broadcasted to different UEs distributed within the MBMS service area, which is the sum of all the local multicast/broadcast areas offering the same MBMS service [153]. Accordingly, if a tight time-synchronization occurs among cells that transmit identical signals, the signal at the terminal will appear exactly as a single signal, transmitted from a single cell but subject to multi-path propagation, which can be implicitly compensated by the OFDM transmission [154]. This time-aligned signal transmission from multiple cells is also identified as MBSFN (or eMBMS), and it has been introduced in Release 9 of LTE [154], [155]. It is worthy to note that thanks to MBSFN the service broadcasting/multicasting in a multi-cell environment is consistently improved, since the received signal strength increases, and the inter-cell interference is strongly reduced. The area where one or more cells transmit the same content is referred to as MBSFN area, which is statically configured and cannot be set-up on-the-fly, nor dynamically change in time [8]. The MBMS Service Area is composed by more MBSFN areas grouped together and providing the same MBMS service. Fig. 3 illustrates what previously described.

The logical architecture for MBSFN is the same of MBMS detailed in Fig. 2. Also the main functionalities of MCE, MBMS-GW, BM-SC and MME remain the same, with the only additional capability of MCE and BM-SC of time synchronization among cells. This operation is performed through the

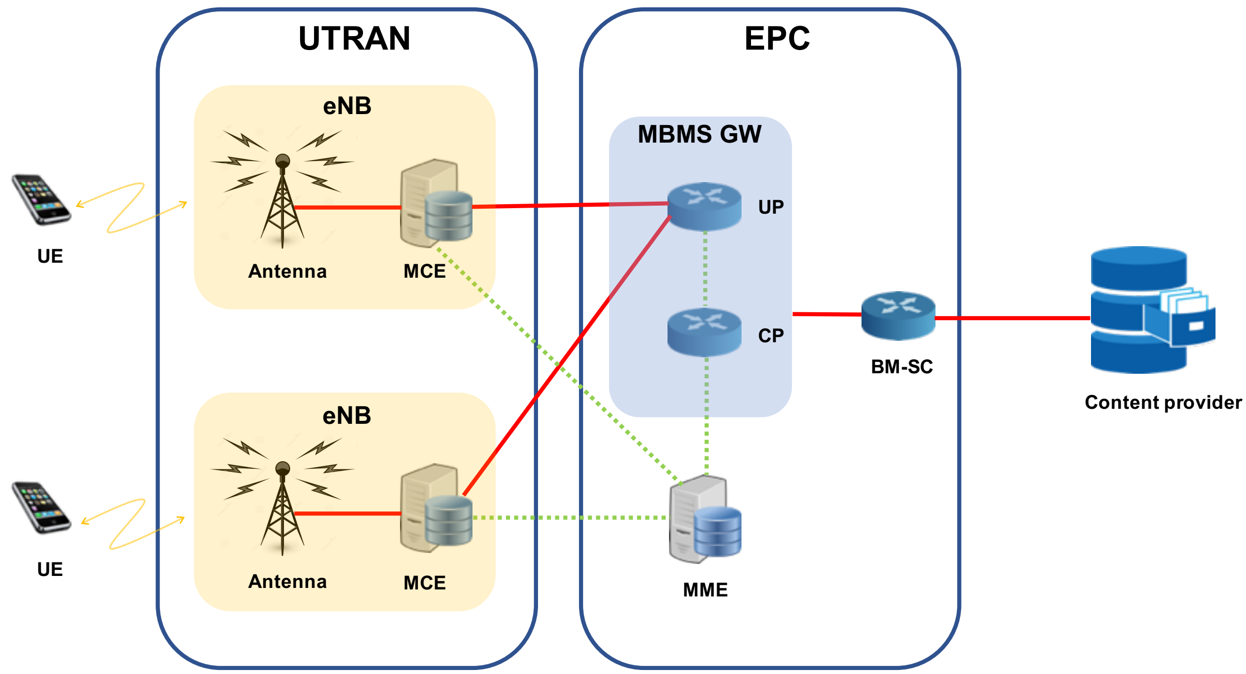

Fig. 2. Schematization of the MBMS architecture. The UTRAN provides connectivity between the end users and the core network (EPC) through the eNBs, which provide the physical signal (through antennas), radio resource allocation and service management to UEs (through the MCE). Connection to the EPC guarantees control functionalities (through the MME) and management of IP addresses (through the MBMS-GW). The BM-SC interfaces the EPC and the content provider, performing functionalities of security, service activation and deactivation, and content synchronization.

transmission of specific synchronization Packet Data Units (PDUs) (SYNC PDUs). Each SYNC PDU contains a time stamp that specifies the temporal beginning of the synchronization sequence. This sequence has a temporal duration, defined as synchronization period, which is configured in the MCE and BM-SC and remains the same for the single MBMS service [8].

B. Further insights on MBMS-based network architectures

Some papers describe MBMS according to specific network architectures. They can be macroscopically classified into MBMS specific architectures, hybrid MBMS and other wireless technologies, and cooperative architectures exploiting MBMS, as described in the following subsections.

1) MBMS specific architectures: The most widely analyzed architecture is the MBMS architecture as described in Section II-A. The MBMS architecture is the main topic of the works [17]–[24], [29], [31]–[33], in different application scenarios. In this context, there are some works that analyze the possibility of exploiting physical resources or enhance functional components of the MBMS architecture to provide TV services through LTE networks [22]–[24], [29].

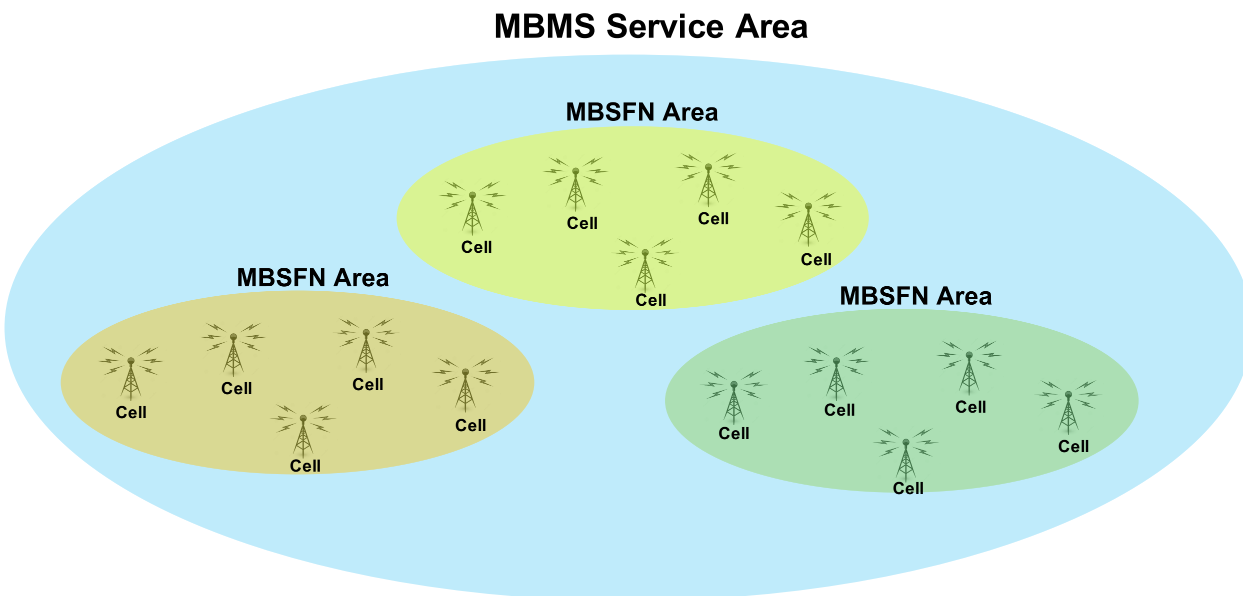

Fig. 3. The hierarchical structure of a MBMS Service Area. A MBSFN area is composed by one or more cells transmitting the same MBMS service with a tight time synchronization. More MBSFN areas providing the same MBMS service are then grouped together into a MBMS Service Area.

2) Hybrid MBMS and other wireless technologies: Other architectures consider the concurrent utilization of different technologies together with MBMS. An example of this configuration can be found in [25], where the backhaul network considers a Passive Optical Network (PON) to connect the BSs with the LTE access network. In this architecture, typically, the BSs providing wireless access to UEs are connected to Optical Network Units (ONUs). An ONU provides services to the radio access network in terms of physical layer protocols, service distribution through the PON and relaying functionalities. Through a network of optical fibers, an ONU is connected to the Optical Line Termination (OLT), that interfaces the nodes providing the service. The OLT manages the PON-related aspects of the transport system [156]. In summary, data are forwarded from a central source node to the OLT, that multicasts them to different ONUs, each one connected to a BS [156], [157]. Such an architecture is advantageous to increase throughput in the backhaul network, but at the cost of increased latencies [25], [157]. Other hybrid architectures consider MBMS as the technology to multicast packets to other ad-hoc networks, employing different technologies and delivering messages to the end users. In this respect, the most widely adopted architecture is composed by the LTE network delivering messages to a wireless vehicular network. Some nodes of the vehicular network can act as relay nodes to disseminate messages to the other vehicles, for different purposes (coverage extension, congestion avoidance, communication robustness and scalability, delay reduction, etc.) [26], [27], [30], [43].

3) Cooperative architectures exploiting MBMS: Other kinds of architectures, identified in this work as cooperative architectures, employ cooperation among different entities to improve multicast performance. In this kind of architecture, some network nodes can act in combination with the classical MBMS architecture for content distribution [34], [37], or act as relay nodes, forwarding data to end users to enhance the multicast performance of LTE [35], [36], [85], [100], [115], [116], [123], [131]. In the latter case, multicasting can also be adopted together with unicast transmission, eventually employing different wireless technologies (IEEE 802.11, WiMAX, etc.). Some LTE nodes can act as relay stations through Device-to-Device (D2D) cooperation [35], [39]–[42], [108]. The main advantages of the cooperative architectures are throughput improvement, coverage extension and energy efficiency.

C. SC-PTM architecture and overall functionalities

Differently from MBSFN, point-to-multipoint multicast transmission can also occur over a single cell. This kind of multicast transmission is referred to as SC-PTM transmission mode [158]. SC-PTM has been officially introduced in release 13 of LTE, to allow the transmission of MBMS services in a single cell, even if some works analyzing PTM transmission in a single cell scenario existed already, as will be described in Section VII-A. Consequently, SC-PTM can be seen as a complementary transmission scheme with respect to MBSFN [159]. It adopts the same overall architecture of MBSFN, together with the same channels, protocols and scheduling techniques to support efficient delivery of bursty traffic [159].

Like MBSFN, HARQ feedback and CSI reports are not used since the Downlink (DL) channel is intended for multicasting, even if studies have been made on the possibility to introduce Uplink (UL) feedback through HARQ retransmissions from all the members of a multicast group to increase Spectral Efficiency (SE) [158]. Each MBMS service is in fact identified by a unique group identifier, which identifies groups of users in the cell, mainly for scheduling purposes. Also control information is included in the DL channel, and scheduled accordingly [159]. The choice of the SC-PTM or MBSFN transmission mode is performed by the MCE.

Some significant studies performed by 3GPP highlight some important differences between MBSFN and SC-PTM [158]. First, there are some scenarios where SC-PTM performs better than MBSFN, and specifically when there is no synchronization among cells (that makes MBSFN unfruitful), when the number of the cells where users receive the same service is very small (up to 3 cells), and when there are some cells in the MBSFN area without users receiving the service [158]. Another advantage of SC-PTM is its higher flexibility in the scheduling of time-frequency resources (that can be done cell-by-cell based of the users distribution), a lower latency, and a higher SE for bursty traffic. Finally, the SC-PTM deployment is simpler because of the lack of synchronization among cells. The main disadvantages of SC-PTM, if compared to MBSFN, lie in a higher interference and a lower SE due to the absence of cell synchronization, and in a lower efficiency for large broadcast areas.

D. Differences with multicasting/broadcasting in other wireless networks

Multicasting and broadcasting of services in different wireless networks scenarios have been widely analyzed and standardized in the past literature. Multicasting and broadcasting of services is basically a methodology to deliver data common to more users exploiting common resources. It has attracted attention especially for the transmission of multimedia services, as testified by the standardization efforts carried out in the recent past, and for different types of wireless networks [7], [160], [161]. Nevertheless, even if the macroscopic definitions of the main elements of multicast services (including the single frequency network definition, the number of BSs involved in simultaneous transmission, synchronization among BSs to reduce signal interference, and the absence of HARQ feedbacks) are common to the different wireless scenarios, necessarily the various details on multicast transmissions, i.e., the amount and location of radio resources, the frame structure, the synchronization procedures, the steps to join/leave a multicast service, the configuration of the BSs participating to the multi-cell area, etc., are peculiar to each wireless technology adopted, and follow precise standard specifications that differ according to the specific wireless network under analysis, and for each layer of the protocol stack [7], [8], [160], [161]. Just to strengthen this thesis, works can be found in literature that perform a detailed survey of strategies and proposals in different wireless networks scenarios, with respect to multimedia multicast (please see Section III for further details). Nevertheless, even if conceptually there are strategies aiming to reach the same goals, they are applied in the specific wireless network scenario [162], [163]. Furthermore, wherever two wireless technologies are concurrently utilized in the same network architecture, wireless interfaces are duplicated accordingly, to support all the wireless architectures implemented [163]. One of the major drawbacks of each wireless data transmission, including service multicasting/broadcasting, is how to exploit at best the limited spectrum resources available at physical layer. In this context, a very interesting technology is Cognitive Radio (CR), a very promising technology in future wireless networks [164], [165]. The spectrum allocation is a problem, due to the presence of a plethora of different wireless services and technologies, anytime and anywhere. Through the CR approach, spectrum resources are dynamically allocated, so that the portion of the spectrum not utilized by some users can be reallocated to some others, according to the so-called “understanding-by-building” methodology [164]. Cognitive Radio Networks (CRNs) are networks based on this approach, where devices detect variations in RF signals and perform changes in some physical layer parameters (transmission power, modulation schema, frequency carrier, etc.) to increase the efficiency in the spectrum utilization and communication reliability. Like the other wireless technologies, also for CRNs the key functionalities at the different layers of the protocol stack, and their cross-correlations, are peculiar of this technology, and are subject to standardization efforts by many committees, including IEEE 802.11 and IEEE 802 families [165]. In this context, few works in the recent past deal with the multicast transmission that take into account the CRNs, especially with respect to scheduling and routing issues [166], [167]. This means that detailed studies on this issue are still a work in progress, but fall outside of the scope of this work, that has instead the focus to provide a detailed review of the literature on multicast transmission adhering to the LTE standardized specifications and recommendations.

III. Related work

There are few works in the recent literature where a descriptive classification is performed of multicasting strategies in wireless networks [162], [163], [168], [169]. Among these, the works [168], [169] do not perform a survey on multicasting; rather, they propose enhanced architectures [168], or discuss the possible challenges and research directions in next-generation wireless networks [169]. Nevertheless, part of these two works is dedicated to the analysis of the most representative literature on the related topic.

The work [168] analyzes multicasting in WiMAX networks. The goal of this work is to analyze in detail Multicast and Broadcast Services (MBS) architecture for WiMAX networks and to propose an enhanced MBS-based architecture that can increase the performance of video broadcasting. MBS for WiMAX is analyzed in detail, in terms of components and functionalities, proposing enhancements to provide video mobile services, but dedicating only a brief discussion on MBMS. If compared to this survey, the goal of [168] is to propose an architecture for MBS to enhance the quality of video broadcasting in WiMAX networks, and not to provide a survey of the state-of-the-art of MBMS and its evolutions. In fact, MBMS is only superficially introduced and described. In addition, [168] deals with WiMAX, which differs from LTE (the central topic of the present survey) in many aspects, since LTE implements significant enhancements to provide high mobility, wide coverage, higher throughputs and capacity, exploiting at best the available spectrum resources, and accordingly presents different specifications in structure and protocols of the stack with respect to WiMAX [161], [170]. Finally, [168] is focused on MBS, which is different from MBSFN in LTE (the latter requires synchronism and stringent timings requirements of transmitted data among adjacent cells).

Also the work [169] aims at identifying the multicast applications suitable for next-generation 5G networks. MBMS is described with respect to its logical components and functionalities, and then the enhancements to MBMS are discussed to meet the requirements of 5G applications. A synthetic review of the literature is also performed, with respect to multicasting in last generation cellular networks. Like [168], the survey of the recent literature on multicasting is not the focus of this work. In fact, only few recent works are considered for discussion. Also the topics cover only part of the set of strategies described in this work; specifically, they comprise group formation techniques, energy efficiency, heterogeneous network architectures, cooperation and D2D, Network Coding (NC) and beamforming. They are all functional to identify and discuss the main challenges of multicasting over next-generation cellular networks. Instead, the goal of the present survey is totally different, since it aims at covering, with a high level of detail and completeness, all the state of the art research on multicasting in LTE/LTE-A systems, including existing strategies and novel proposals, at all levels of the protocol stack.

The papers [162], [163] are instead focused on surveying some aspects of multicast transmission in wireless networks. The goal of the survey [162] is to analyze multicasting in OFDMA-based systems, with respect to the issues of (sub)group formation and Multicast Scheduling and Resource Allocation (MSRA) algorithms. The multicast group formation strategies are described and classified in detail. Fewer works are analyzed and classified to describe the most relevant strategies of scheduling and resource allocation. The other survey [163] analyzes multicasting techniques in different wireless access networks, focusing on packet error rate reduction for multicasting at different layers of the protocol stack, through enhancements to unicast transmission and error correction techniques. Also the resource allocation issue is analyzed for Wireless Local Area Networks (WLANs), IEEE 802.16 (WiMAX) and, secondarily, in 3rd generation (3G) networks. The most part of the work analyzes WLAN-based access networks; topics discussed in this scenario comprise the reception of multicast and unicast transmissions according to IEEE 802.11 standard, data rate adaptation techniques, error correction schema and relaying, with the goal of increasing data rate and coverage, recover from lost packets, or in combination with NC. Hybrid access networks are then discussed, where WLANs act in cooperation with other types of technologies (i.e., 3G networks using MBMS). The only relaying issue is discussed in this part of the work, where WLANs are responsible of the retransmission of lost packets. The last part of the paper analyzes mainly radio access networks based on the IEEE 802.16 standard. The topics analyzed in this context are: resource allocation (through optimal MCS, subcarrier, or power assignment), error correction schema including Forward Error Correction (FEC), HARQ and coding, relaying for packet retransmissions, and energy efficiency for power optimization.

If compared to the present survey, all the works analyzed in [162] are classified in a general multicast scenario, without particularizing them in the specific context of MBMS; in this regard, only a brief description of the main features of MBMS is provided. Furthermore, all the topics treated in [162] belong only to the physical layer of the protocol stack, whereas the present work analyzes more layers of the stack, and all the strategies described are contextualized to to LTE and LTE-A systems, which, as previously said, implement more sophisticated (and more complex) evolutions of MBMS. Finally, the group formation techniques, to which the majority of the work [162] is dedicated, do not consider their most recent evolutions, which instead are considered in the present survey. As regards MSRA algorithms, in [162] they are discussed dedicating a substantial part of the work to an overall description of their features, i.e., performance metrics and tradeoffs, system model and computational complexity, which is a complementary approach to the classification and description of the strategies performed in the present work.

There are also many substantial differences between the survey [163] and this survey. First, the work [163] is primarily focused on IEEE 802.11 networks, whose multicasting protocols adhere to the IEEE 802.11 standard [160], and secondarily to other technologies, especially IEEE 802.16 (WiMAX) [161]. Few works are considered to discuss 3G networks strategies, and anyway all the wireless networks considered in this survey (WLANS, WiMAX, and 3G) differ from LTE, since the latter is a sophisticated evolution of the former ones [170], [154]. The second difference is that, according to what previously explained, MBMS is not the key issue in [163], and instead MBMS and its evolutions are the key issue of the present survey. The third relevant difference lies in the covered topics. The work [163] focuses mainly on error recovery techniques, where lost packets are recovered through retransmissions or NC, and involving only the lower layers of the protocol stack. NC techniques are analyzed, but often in the context of hybrid WLAN-3G networks, where the ad-hoc WLAN network is responsible of recovering packets lost in the 3G network. This is a different background if compared with the relaying techniques specific of LTE systems, that exploit different strategies, like for example D2D, and not only for error recovery issues. As regards the description of the other radio access networks, most of the works cited in [163] deal with resource allocation at physical layer, including scheduling of time-frequency resources, power allocation through optimization algorithms, and MCS adaptation. At MAC layer, error recovery schema are analyzed, also in combination with FEC techniques at application layer. If from one side these topics are also treated in the present work, from the other several other topics are discussed here, including Multiple Input Multiple Output (MIMO) transmission, power control, D2D techniques and analytical models, network layer strategies, multiplexing techniques, cross-layer strategies, more hybrid network architectures, advanced scheduling algorithms, etc., completing the survey with the discussion of the works on performance evaluation and standard analysis of LTE.

Summarizing, the present survey can be seen as complementary to all the abovementioned works, completing the missing topics of each of them in more recent scenarios, and extending, updating and detailing as much as possible the analysis of the state-of-the-art research of multicasting strategies in the context of the last-generation LTE and LTE-A systems.

Table I resumes the topics covered by the surveys described in this section, and compares them with the main themes developed in this survey.

TABLE I
COMPARISON AMONG SURVEY PAPERS ON MULTICAST AND BROADCAST SERVICES.

	[162]	[163]	[168]	[169]	**This survey**
Investigation of no-3GPP architectures (i.e., WLAN and WiMAX)		X	X		
Investigation of 3GPP-oriented solutions based on MBMS and MBSFN				X	X
Investigation of 3GPP-oriented solutions based on SC-PTM					X
Investigation of standardized mechanisms and methodologies	X		X	X	X
Advanced approaches for the physical layer	X	X	X	X	X
Advanced approaches for the MAC layer		X	X		X
Advanced approaches for the network layer		X			X
Advanced approaches for the transport layer		X			
Advanced approaches for the application layer		X			X
Advanced cross-layer approaches		X			X
Network architectures for multicast and broadcast services			X	X	X
Coding schema for multicast and broadcast services		X	X	X	X
Optimization algorithms for multicast and broadcast services	X				X

IV. An overview of the taxonomy of papers on MBMS, MBSFN and SC-PTM

In this section the rationale of the classification of the works on service multicasting in LTE systems is discussed. As already anticipated in Section I, all the surveyed scientific literature is classified through two main categories, i.e., MBMS and SC-PTM. Figure 4 illustrates the proposed paper taxonomy for the main category dedicated to MBMS and its evolved version MBSFN. Specifically, MBMS integrates the following sub-categories:

- **Baseline approaches.** Works comprised in this category evaluate the performance of different aspects of LTE multicast, ranging from MBMS-based architectures to cooperative strategies, error correction and scheduling techniques, physical layer analysis, implementation of the LTE environment by means of simulation/emulation tools, analysis of 3GPP standardization, and performance evaluation. Nevertheless, in all cases, these works do not propose any novel strategy on LTE multicast. Papers analyzing performance of LTE multicast focus on the physical layer performance (in terms of SE, spectrum utilization, resource allocation, distribution of BSs, frequency reuse schema, MIMO techniques, etc.), the MAC layer (scheduling mechanisms, throughput analysis, cooperative multicast, etc.), signalling and control procedures, FEC techniques, coding for error protection, or file recovery at application layer. In some works, MBMS is evaluated in specific application scenarios (i.e., video transmission, intelligent transportation systems, cooperation with other wireless networks, etc.). The main aspects of the 3GPP standard concerning MBMS and eMBMS are analyzed and evaluated in some works. The standard aspects taken into account mainly comprise: frequency channels and bandwidths, MCSs, framing and subframe allocation, spectrum allocation, MIMO implementation, service continuity (handover), scheduling, cell coordination for MBSFN, and Application Layer Forward Error Correction (AL-FEC) techniques.

- **Physical layer strategies.** The group classifies all the works proposing novel strategies at the physical layer of the protocol stack. Some works propose optimizations of the power consumption; this goal is reached through beamforming and antenna selection techniques, MIMO transmission, or power control strategies aiming to optimize power consumption depending on users throughput, mobility, delay, or Quality of Experience (QoE) requirements. Other works propose coverage optimization techniques. To reach this goal, some of them focus on relaying and D2D transmission techniques, some others on beamforming, and others on optimal radio resource allocation and configuration of MBSFN areas. Few works focus on MIMO strategies, mainly dealing with receiver design and Bit Error rate (BER) analysis. Rate optimization strategies are also proposed in several works. The majority of them focuses on subgrouping techniques; but also relaying techniques, MCS selection and schedulers are proposed for rate optimization purposes. Analytical models for traffic analysis can also be found in this group of works. Finally, there are works that propose spectrum optimization strategies, that translate into an improvement of the SE in MBSFN transmission. To this end, Adaptive Modulation and Coding (AMC) and MCS schema are proposed in the majority of the works in this field. Also cooperation techniques and simulation tools make part of this class of strategies.
- **MAC layer strategies.** There are lots of papers focusing on the MAC layer. They can be macroscopically subdivided into two main groups: the first dealing with error protection, and the second proposing resource allocation strategies. Error protection strategies mainly deal with packet repair schema based on synchronization among BSs or the utilization of relay stations, efficient retransmission schema, and NC techniques (even in combination with ARQ or user cooperation). Resource allocation strategies aim to optimize the resources utilized for a reliable data transmission. This is performed through efficient scheduling algorithms to save power or bandwidth or manage user QoS, and through cooperative schema based on efficient grouping of Relay Stations (RSs) and feedbacks, or D2D schema.
- **Network layer strategies.** Only a work focuses on the network layer. It is based on the optimization of the SYNC protocol for content synchronization and on information exchange at network layer between cooperating BSs for scalable multicasting.

- **Application layer strategies.** Innovative strategies are proposed also at the application layer. The most relevant papers dealing with application layer strategies can be grouped into four categories: multiplexing techniques, cooperative strategies, error protection strategies and layered video transmission. The three most relevant papers focusing on multiplexing techniques in LTE propose theoretical models of bandwidth estimation and resource traffic; in all cases, they are used to analyze and eventually optimize the bandwidth allocation for multiplexed services transmission. Cooperative strategies at application layer are found in only one work. It proposes a peer-to-peer recovery protocol to recover damaged or lost multicasted packets. The two works dealing with error protection strategies adopt both FEC schema. Specifically, FEC redundant information is adaptively varied in a packet recovery scenario in one case, and optimized to improve decoding probability and reduce decoding delay in the other. Layered video transmission strategies are mainly focused on Scalable Video Coding (SVC) video transmission through novel adaptive schema, where the allocation of resources is dependent on the specific SVC video layer. Optimization is performed with respect to energy efficiency, user QoE and QoS, and service data rate optimization.
- **Cross-layer strategies.** This category embraces the works proposing a cross-layer interaction between FEC (at application layer) and HARQ (at MAC layer) schema for the optimization of network resource or QoS, synchronization schema for frame alignment at MAC layer through the upper layer SYNC protocol, and strategies jointly involving physical and MAC layers. One paper presents a proposal of a complete cross-layer architecture for video delivery.

Figure 5 illustrates the proposed taxonomy for contributions addressing PTM techniques, where LTE multicast is analyzed in a single-cell scenario. In this case, the following sub-categories are defined:

- **Baseline approaches.** The papers falling into this category compare PTM and MBSFN performance, analyzing the impact of different aspects like MIMO techniques, interference, coverage, data delivery, users mobility, control procedures and signalling.
- **Physical layer strategies.** The physical layer strategies proposed for PTM transmission can be grouped into three different categories: power optimization, spectrum optimization and MIMO strategies. Papers proposing power optimization strategies focus on radio bearer selection for power control and the cooperation among different radio access technologies (i.e., WiFi, UMTS, LTE, etc.) for energy saving purposes. Spectrum optimization strategies act in the scalable video transmission scenario, by adapting the MCSs of the different video layers to meet QoS requirements. Finally, MIMO strategies focus on adaptive MIMO schema, including spatial multiplexing and diversity techniques.

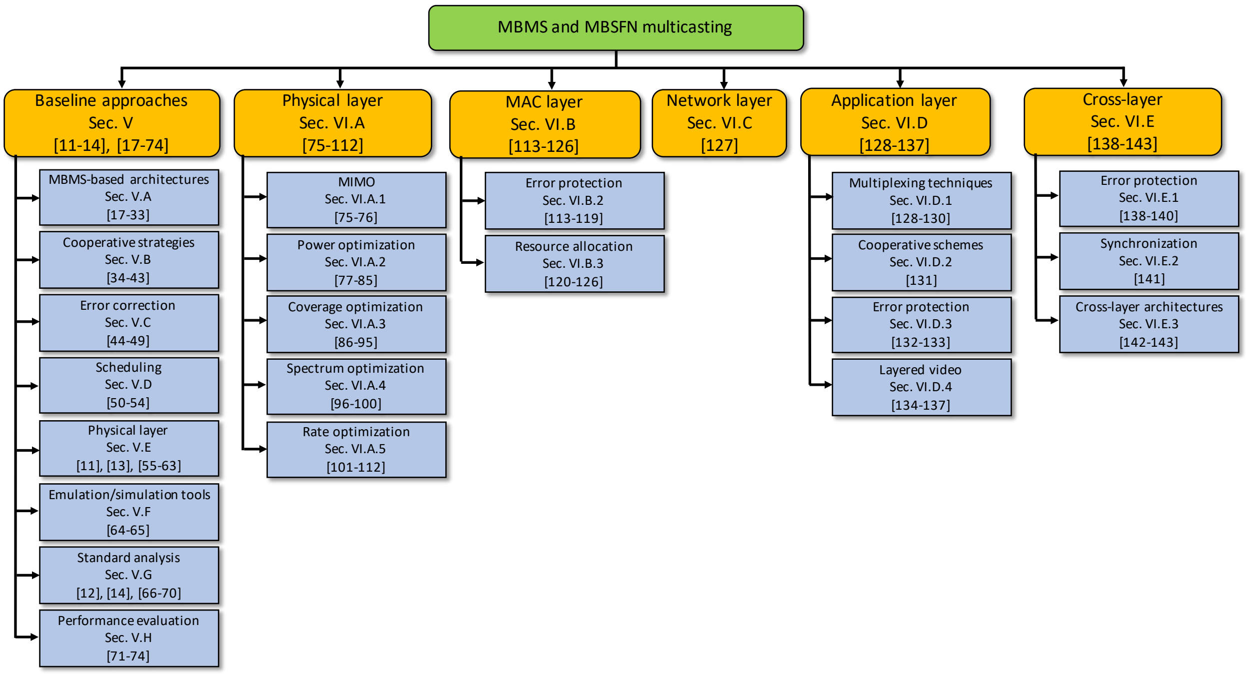

Fig. 4. Paper taxonomy for LTE multicast enabled by MBMS and MBSFN

- **MAC layer strategies.** All the papers dealing with MAC layer strategies for PTM face the issue of the optimization of resource allocation. Their proposals cover users detection strategies to avoid unnecessary transmissions, frequency-based dynamic scheduling algorithms, and the application of Random Linear Network Coding (RLNC) approaches on layered video transmission.
- **Cross-layer approaches.** Cross-layer approaches for PTM include the design and validation of a single-cell transmission scheme (network architecture, signalling, management of radio channels and resource control) and the proposal of analytical models to evaluate the performance of multimedia PTM services on a TV platform.

V. Baseline approaches for MBMS

This section describes the baseline approaches found in the recent literature for service multicasting in LTE and LTE-A. All the papers covering this aspect are classified and described in detail in what follows. The main advantages of MBSFN, more evident at the border between cells involved in the MBSFN transmission, are mainly an increased received signal strength, a reduced interference level, and a higher SE, as highlighted in [11], [13], [55]–[65]. Furthermore, this kind of transmission can be fruitfully adopted in several network architectures and for different purposes (transmission of multimedia and TV

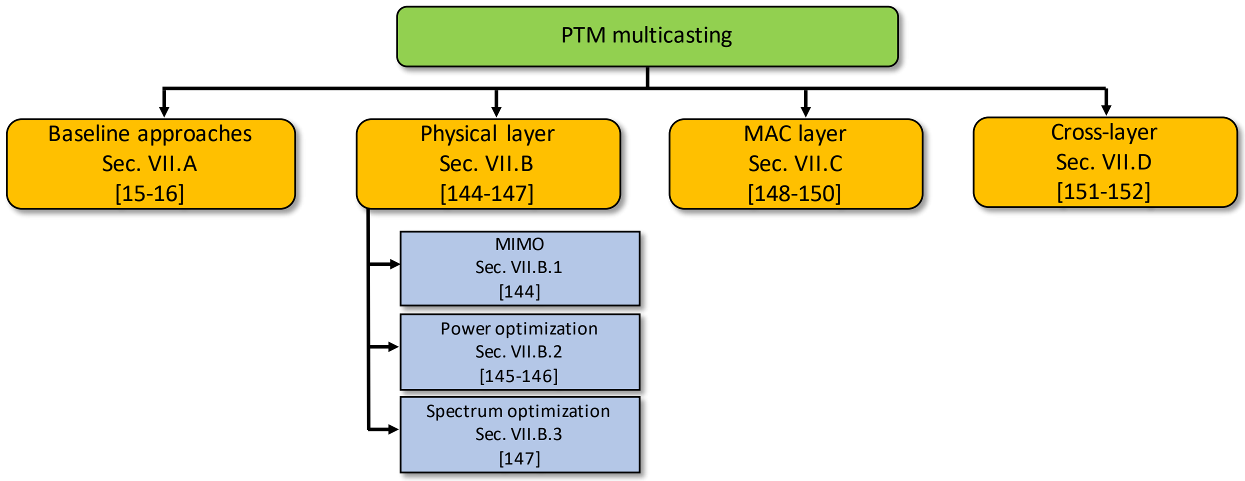

Fig. 5. Paper taxonomy for LTE multicast enabled by PTM

services, public safety, emergency and alert scenarios, vehicular networks communication, cooperative strategies, etc.), as testified by the works [17]–[43].

At physical layer, a device can receive both MBMS and unicast transmissions. This can happen on the same carrier or on different carriers, in the case of a carrier-aggregation capable device. In the latter case, unicast transmissions are received on one Component Carrier (CC), and MBMS transmission on another CC. To guarantee an efficient delivery of MBMS services at this level, error correction and data repair techniques are of great importance as testified by the works [44]–[49]. The physical resource used to transmit MBSFN data is the MBSFN subframe, which consists of a control region, used for control signalling, and a MBSFN region, used to transmit MBMS services and the relative control information. For a correct channel estimation and demodulation, some reference symbols are inserted within the MBSFN subframe. They are placed in the same time and frequency positions, and assume the same values [159].

Scheduling of MBMS services is another very important issue, as testified by the works [50]–[54]. The transmission of MBMS services occurs in bursts, and between two adjacent bursts the device can switch-off its receiver circuitry for battery saving purposes [159]. This mechanism is defined as Discontinuous Reception (DRX) mode. Obviously, this comes at a cost of scheduler restrictions, since the UE can receive data only in the active subframes.

All the most relevant aspects of LTE transmission, i.e., coverage and low device power consumption, are discussed in detail in the LTE standard. The works [12], [14], [66]–[74] analyze and evaluate the MBMS and MBSFN standardization, including performance evaluation of LTE multicast in different scenarios. The issues covered by the papers cited above are now classified and described in more detail in the following subsections.

A. MBMS in network architectures

Analysis of MBMS in different network architectures is performed in the works [17]–[33]. The MBMS-based architecture is analyzed in [17], [18]. The work [17] performs an analysis on the MME, that provides a control-plane support to each UE. The signaling flows that influence the MME processing load are analyzed and quantified analytically for different LTE scenarios. The work [18] focuses instead on the support of broadcast bearers in LTE. The impact of many factors on eMBMS performance and use cases is analyzed, including the influence of radio network engineering for different eMBMS coverage area sizes. The work [19] studies the configuration of the MBMS carrier to send intelligent transportation systems (ITS) messages in LTE networks. In this scenario, ITS applications are described and related to eMBMS architectures. LTE video multicasting is analyzed in [20], [21] in different scenarios. The work [20] analyzes the performance of distributed BSs (DBSs). The system throughput and QoS provisioning are chosen as performance metrics in the video streaming scenario, considering both the UL and DL channels. In the UL, a scheduling algorithm is presented which allocates a Resource Block (RB) to a specific user to maximize a utility function that depends on services and QoS requirements. In the DL, assuming that the same video is sent to all users, a RB is allocated to transmit the same stream to the MBMS users. In [21] a discussion is found on the different options adopted to transmit videos through WiMAX or, alternatively, LTE technologies. After an analysis of the standardization activities, the methodologies that evaluate the system level are presented, together with the related numerical results that quantify and compare WiMAX and LTE capabilities for both multicast and unicast transmission, also considering other data service requirements. In [32] an architecture is proposed, where a single user can obtain multiple contents simultaneously, coming from different BSs within the MBSFN area. Several are the goals of the proposed scheme. First, the increase of the user satisfaction, since more contents are multicasted at a time and users can have alternative choices. Second, the additional revenue of service providers, that exploit the same infrastructure to transmit more content. Third, a more efficient resource utilization that allows more users to be served with high QoS. A scheduling algorithm is performed by the Multi-cell/Multicast Coordination Entity (MCE), that schedules radio resources for each cell based on different attributes (type of user and service, QoS, etc.). A service architecture is analyzed in [33] through the IP Multimedia Subsystem (IMS) framework of LTE system. The IMS is an architectural framework, logically placed into the LTE core network, that allows multimedia services (including audio, video, chat, etc.) to be delivered in a packet switched network [171]. The proposal of this work is an IMS-based architecture for an efficient video conferencing service distribution, aiming to reduce the cost (in terms of utilized resources) and increase the flexibility of the multimedia conference service. To this end, a publicly available prototype is developed, evaluating its performance under realistic working scenarios. This attempt aims at replacing the current IMS-based architecture with a new solution exploiting the capacity of LTE BSs. Finally, three different scenarios are proposed for evaluation purposes. Also the work [31] discusses the emergency issue. An architecture is proposed, that exploits the eMBMS specifications for DL video and file transmission and improve emergency group call networks. Some enhancements are introduced: the centralized Multicast Coordination Entity (MCE) architecture, the group privacy and handling requirements (designed according to 3GPP specifications), the multimedia support feature, QoS capabilities, and the call setup time.

A business-oriented analysis on the possibility of utilization of European TV unused frequencies in LTE is performed in [22]. This work focuses on the business challenges related to the intrinsic characteristics and value of Europe TV white space. An investigation is carried out on the promising potential to successfully exploit white spaces by LTE secondary access; this investigation is supported by a quantitative study performed on eleven different European countries. Some aspects and technical questions are considered in this study, i.e., how much spectrum it is expected to deploy, what deployment technology is best suited for the spectral and regulatory regime (Macro-,micro-, pico-, or femto-cells, TDD, FDD, MBMS-only, etc.), where are the white spaces, and how does all this relate to the population density. TV transmission is taken into account also in [23], where a design of the BM-SC is proposed to support IPTV services in LTE networks. Several functionalities are taken into account in the BM-SC design that supports all the required procedures to receive eMBMS services, according to 3GPP specifications. The protocol stack, the main blocks of BM-SC, and the related interfaces are analyzed for multicast transmission purposes. A computer-based test bed is also set up to implement and test the proposed eMBMS functional design. TV distribution through MBMS is also considered in [24], through the adoption of low-power transmitters exploiting the same frequency. To this end, TV spectrum and the possibility of its reallocation to LTE communication is examined. Finally, a comparison is performed of the power required in the TV and LTE broadcasting scenarios. In [29] an integration of LTE and DVB-H is proposed, with the goal of increasing the efficiency of radio resources and service quality. For this reason, several logical entities are introduced, to provide encoding, description, allocation and handover functionalities. Finally, three different scenarios are analyzed to design the proposed converged architecture.

A new type of architecture, the Virtual Single Cell network, is proposed in [25]. The goals of this proposal are the applicability of the system to wide-ranging mobile networks, to keep both the similarity of the air interface with other interfaces to offload the network, and the system-related economies derived by borrowing technologies and components from the major mobile network. To this end, cells are logically connected by Passive Optical Networks (PONs), that perform multicasting/broadcasting functionalities. The continuity of packet transfer is guaranteed by some small cells placed around the cell that contains the target terminal, that make some information for handover available to the BS before the handover begins, to strongly reduce the service interruption.

Vehicular networks are analyzed in [26]–[28], [30], for congestion control [26], QoE [27], and security issues [28]. For sake of clarity, in Fig. 6 an example is provided of a hybrid architecture involving LTE multicast and ad-hoc VANET communication protocols. The scheme proposed in [26] faces the issue of congestion control in an integrated Vehicular Ad-hoc NETwork (VANET)-cellular environment. The contribution of the work is twofold. First, a heterogeneous networks framework and a beacon rate adaptation scheme are proposed, so that vehicles can deliver emergency messages through the BS in LTE. Second, a simulation platform is implemented to verify the proposed scheme. Another hybrid VANET-LTE architecture is found in [27], aiming to provide uninterrupted multimedia services among groups of vehicles. The contribution of this work is threefold. First, the VANET-LTE heterogeneous network is designed, which incorporates the most relevant QoS components. Second, the essential metrics are identified from the perspectives of both the VANET and LTE from the point of view of the vehicle elected as a Cluster Head (CH), which acts as a gateway between the VANET and LTE. Third, a multicast tree is incorporated, that is a combination of a mesh and a tree, to improve communication robustness and scalability. The design of a Group Key Management (GKM) scheme is proposed in [28], to provide confidentiality in multicast communications for vehicle-based systems. The main goal of the scheme, not explicitly thought for LTE but applicable to it, is to reduce the communication overhead. To this end, an analysis of the communication overhead is carried out, including messages for key distribution and location update, and the storage overhead. The GKM overhead is then further optimized through a heuristic algorithm. Other heterogeneous architectures that include LTE systems are analyzed in [30]. In this work the MBMS architecture, together with the wireless vehicular network, is exploited for alert messages dissemination, with the goal of delivering alert messages to the greatest number of users in an area. The main contribution of this study is an efficient message delivery in different networks scenarios (i.e., LTE and IEEE 802.11p networks). A set of possible solutions relying on multicast is proposed, to jointly decrease the number of messages sent in the core network, and increase the number of users receiving alert messages [30].

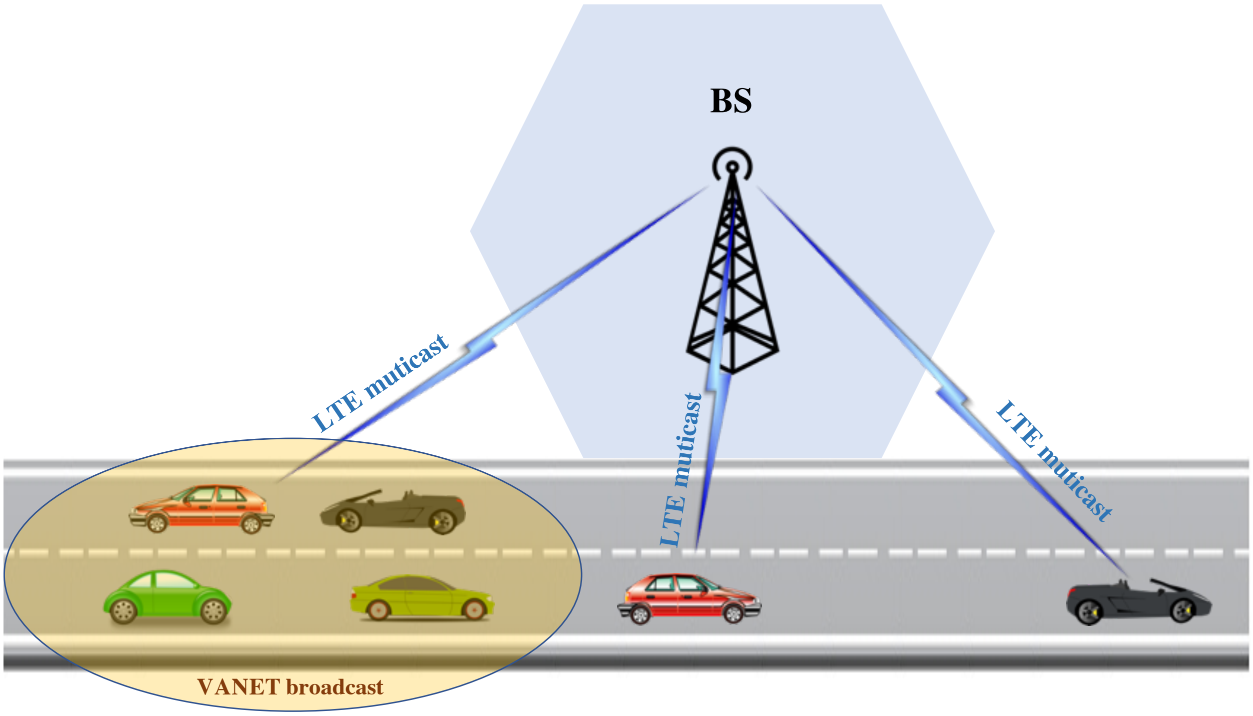

Fig. 6. An example of hybrid LTE-VANET network architecture.

B. Cooperative strategies

The works [34]–[43] discuss cooperative strategies, where additional nodes are exploited to improve coverage and throughput in MBMS transmission, or to relay multicast data to final users. To this category belongs the D2D transmission, where some UEs operate as RSs to forward data to other UEs. A network architecture is studied in [34], that utilizes micro BSs (mBSs) in combination with Macro BSs (MBSs) to enhance the SE of MBSFN transmission. When employed, mBSs multicast packets to cell edge users through specific scheduling algorithms. The methods proposed make use of TDMA scheduling with different reuse levels, so that adjacent BSs transmit in different time slots, and the time slots employed by the BSs determine the number of reuse levels. These schema are combined with the simultaneous multicast transmission of MBSs and mBSs, to increase multicasting performance and the service coverage area.

In [35], [36] multicasting is part of cooperative architectures. Fig. 7 illustrates the general scheme of a multi-hop cooperative architecture, where some nodes (labeled as RS), that can be other UEs or ad-hoc nodes, receive data (from the BS or other RSs) through relay links, and forward them to UEs falling in their coverage area. D2D and Wi-Fi cooperation techniques are compared in [35] in a LTE-A scenario, where all the UEs of the same multicast group broadcast downloaded portions of the received content in a Wi-Fi ad hoc network through D2D-based techniques, where some UEs act as RSs. Two-hop relaying strategies are discussed in [36], where a detailed evaluation of different relaying methodologies in a realistic environment is performed through system level simulations, analyzing different relaying possibilities: the conventional decode-and-forward relaying, a dynamic relaying where the transmission rate in the BS-RS link is varies dynamically and the RSs transmission is not synchronized, and a cooperative relaying between all RSs and BS, where instead all the RSs transmit data synchronously.

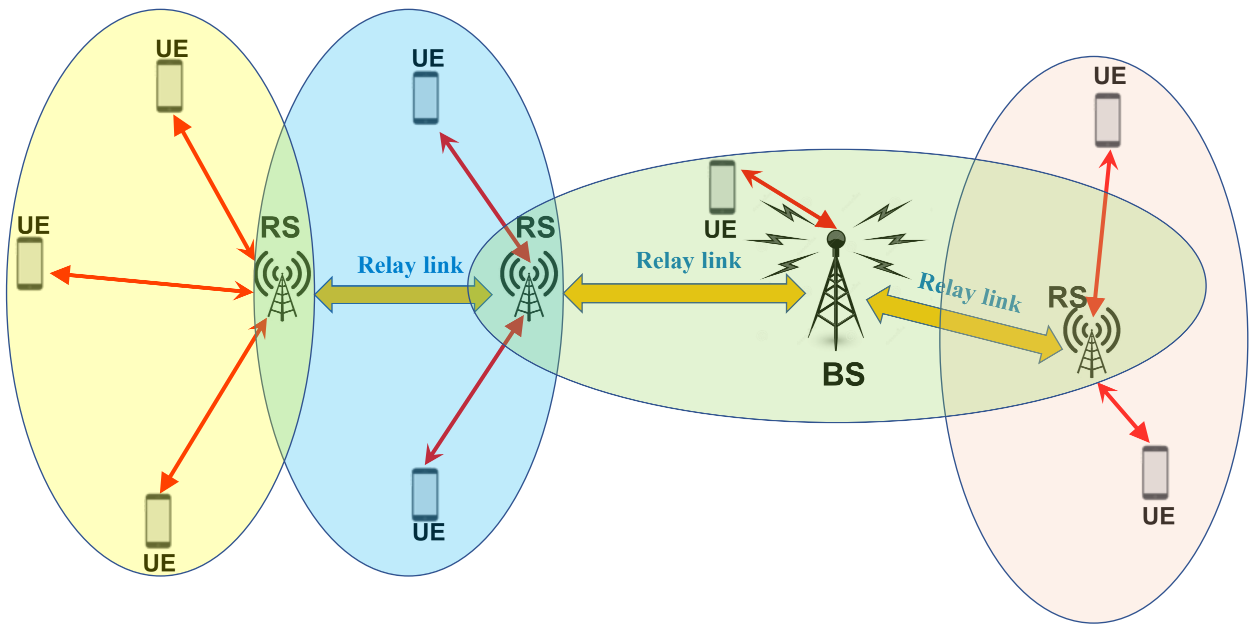

Fig. 7. A general scheme of a multi-hop cooperative architecture.

Cooperative strategies are also analyzed in the works [37]–[43]. Cross-layer frameworks are analyzed in [37], [38]. In [37], the framework is thought for the management and control of large scale emergence across heterogeneous networks. The scenario depicted is the forwarding of emergency messages across all the networks in a defined area. To this end, the IEEE 802.21 standard is considered, that supports different types of networks. It is extended to include multicasting [37]. The cellular network is integrated in this architecture, including MBMS. In [38], a cross-layer cooperation is studied, that combines the IPv6 protocol with some features of the layer 2 of the protocol stack. The goal is to enhance the efficiency of the geographical coverage in a limited geographic area where packets are exchanged at network layer. To this end, at network layer the radio cells are properly selected by the endpoints so that the coverage area is reduced. In the terminal, out-of-scope data are dropped at layer 2 to reduce the workload of the network layer [38]. This study is described using MBMS together with all the multicast/broadcast entities standardized by 3GPP for cellular networks.

The works [39]–[42] deal with D2D communications. Clustering approaches are analyzed in [39]–[41], to save energy [39], [41] and enhance data rate [40]. The method proposed in [39] focuses on the cooperation among UEs for battery saving and energy-efficiency purposes. Initially, all the nodes are directly connected to the BS without any D2D connection, and any cluster is formed by a single terminal. Then, the clusters are iteratively merged into so-called coalitions, to obtain the minimum energy consumption on the link with the BS [39]. In each coalition, only one device can communicate with the BS. This process is performed both in the UL and DL (exploiting multicasting). Machine-to-Machine (M2M) cooperation is investigated in [41] for multicasting of real-time video. Also in this case, mobile devices are grouped into clusters by means of an algorithm for content distribution. For each cluster, the CH receives video data from the BS in multicast or unicast on a long-range link through an appropriate LTE scheduling. Video data are then multicasted to the other members of the same cluster, using a short-range wireless communication. In the clustering approach proposed in [40] a single device is chosen to relay data bidirectionally between the BS and other devices. This approach is mainly thought for the public safety networks scenario. The LTE rate calculations, channel model, and resource allocation strategy are considered in a standard-compliant fashion, and the data rate optimization is performed by increasing the performance of the worst case users. In [42] the D2D potential is investigated with the goal of improving the Conventional Multicast Scheme (CMS) solutions. The proposed scheme is based on the possibility that some UEs can act as relay nodes for all the other UEs experimenting worse channel conditions, that cannot be served in the DL. To this end, for each Channel Quality Index (CQI) level in the DL, the BS verifies that all users can be served, also through D2D links, and the availability of RBs for both the direct DL and D2D links. After the RB allocation is complete, the BS selects the CQI level in the DL so that the system data rate is maximized. Cooperative schema in public safety vehicular networks are proposed in [43] for clusters of moving vehicles. The goal of the proposed schema is to fasten as possible the data delivery through the Short Range (SR) transmission among vehicles. Two schema are proposed. In the first, the BS transmits data to a single vehicle, which in turn forwards them on SR multicast links. In the second, the BS transmits data in multicast mode, and vehicles with better channel conditions that successfully receive them, forward the same data on SR multicast links to vehicles experimenting lower achievable rates [43].

C. Error correction techniques

The works [44]–[49] consider different error correction techniques, i.e., FEC and coding techniques at different layers of the protocol stack, for error correction and data repair purposes. Figure 8 illustrates the possible FEC placements in the protocol stack, and the way a source block can be protected.

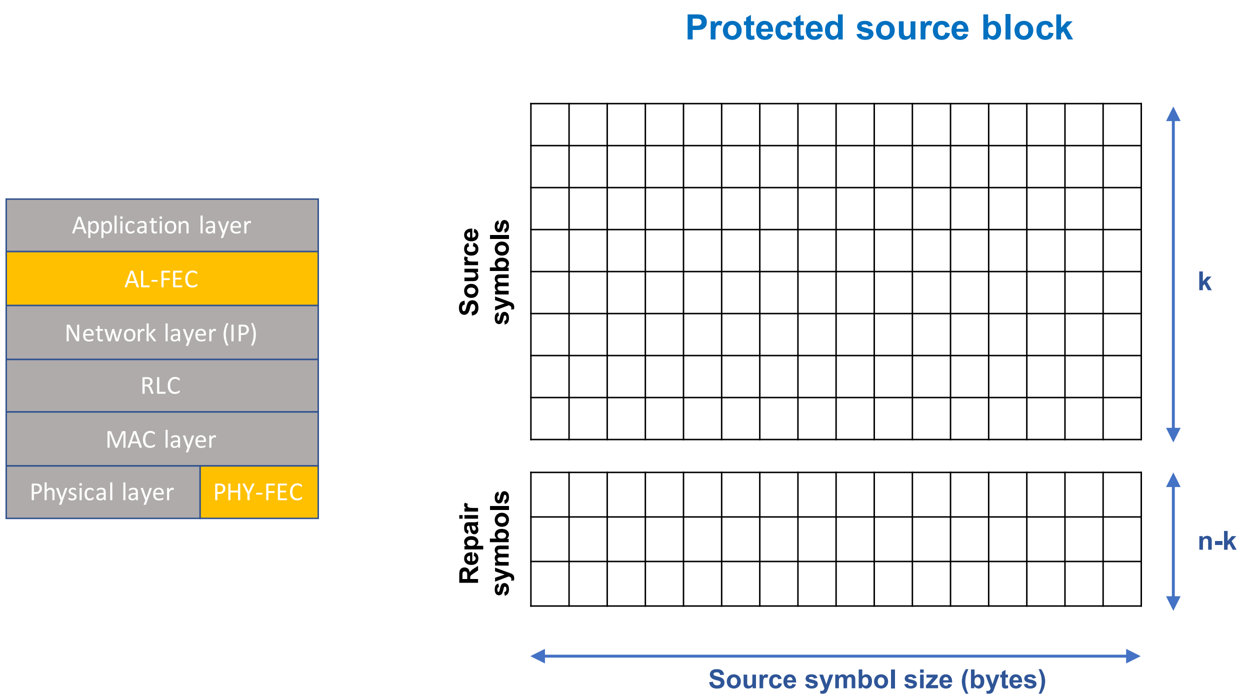

Fig. 8. FEC placements in the protocol stack (on the left) and source block coding (on the right).

FEC implementations at application layer are considered in [44]–[46], for streaming and download of services. The goal is the evaluation of the benefits of AL-FEC for efficient transmission of eMBMS services. Performance is analyzed for different configurations, evaluating the best trade-off between packet overheads and service data rate [44] or cell range and percentage of satisfied users [45]. A comparison with FEC techniques at physical layer is also carried out in [44] to find the best trade-off between the two techniques. AL-FEC is analyzed for QoE purposes in a video streaming framework in [46], where different layers of the protocol stack are considered for the evaluation of end user QoE, represented through different metrics (i.e., delay, rebuffering, and Peak Signal to Noise Ratio (PSNR)). A Markov model can also be found at Radio Link Control (RLC) and MAC layers, that simulates the PDU losses in LTE. A comparison between codes is provided, in terms of PSNR and rebuffering. Performance of error protection codes, i.e., FASTAR codes, is also investigated in [47] and compared to the conventional systematic Raptor codes, in terms of successful decoding probability and average number of transmissions (please refer to Section VIII-A for further details on NC techniques). A system model is found, that takes into account 3GPP specifications. FEC is also considered in [48], and applied to the file repair procedure when multimedia services are multicasted in a DL channel. Different MBSFN scenarios are investigated, with different numbers of both cells surrounding the center of the MBSFN area and transmitting the same

MBSFN service (assisting rings), and cells surrounding the MBSFN area interfering with the transmission of the MBSFN service (interfering rings). Three file repair approaches are implemented, evaluating their performance in terms of the total cost, which takes into account different factors such as the transmission over the air and core interfaces, polling procedures in each BS, and synchronization in the only case of MBSFN [48]. An analysis of RLNC is taken into account in [49] to increase the probability of successful packets delivery and the robustness to network changes or link failures in LTE-A. In addition, this technique is compared with the D2D cooperative in terms of throughput and/or energy consumption at the receivers.

D. Scheduling strategies in LTE multicast

Scheduling is a well investigated issue in LTE, as testified by the works [50]–[54]. Scheduling techniques for MBMS optimally allocate resources based on specific metrics, that depend on the proposed scheduling algorithm, eventually considering the coexistence with unicast transmission, as illustrated in Fig. 9. Scheduling techniques that jointly consider multicast and unicast transmission schema are analyzed in [50]–[52]. In [50] a scheduling mechanism for power saving is proposed. It includes unicast and eMBMS schema. For MBSFN scheduling, MBSFN areas with similar numbers of demanded time slots are grouped together, while for the unicast scheduling, the Proportional Fair (PF) algorithm, as standardized by 3GPP, is adopted [50]. A resource allocation strategy is then proposed for multicast, taking into account interference among overlapping or adjacent MBSFN areas, with the goal of minimizing the number of needed radio resource units to save energy. The joint multicast/unicast scheduling solution proposed in [51] is based on the dynamic optimization of the single LTE frame, to maximize the overall throughput in a MBSFN area. To this end, the proposed technique combines unicast and multicast transmission schema to achieve a target bit rate for all the users receiving a multicast service. The optimal MCS and the optimal number of subframes reserved for multicast transmission are obtained for each LTE frame; furthermore, the unicast scheduling metric is used to allocate the remaining resources and obtain a guaranteed data rate. A fast search algorithm is also developed, to find a suboptimal solution for the multicast transmission parameters, at the same time achieving the dynamic optimization with fewer iterations than an exhaustive search algorithm. Also [52] considers the possibility of a combination of unicast and multicast transmission schema for multimedia delivery, in a scenario where the BS delivers videos with different popularity to UEs. The goal of the proposed approach is an efficient scheduling of video pieces over multicast (or unicast) channels, to maximize the energy of devices and save bandwidth [52]. Like the work [51], algorithms are proposed that are less computationally expensive if compared to the exhaustive search algorithm, even if they provide a suboptimal solution. Another scheduling mechanism for multicast transmission is evaluated in [53], where both TDMA and FDMA scheduling strategies are proposed. The goal is the maximization of the overall throughput rate. To this end, a throughput capacity rate metric is considered, to be maximized in the scheduling algorithms. It quantifies the minimum throughput level per terminal, achieved over all the cells and UEs [53]. Fairness is also taken into account, to guarantee all the mobile terminals a minimum feasible throughput. Multicast scheduling without using mBS is also considered, and a detailed analysis of the throughput capacity rate and failover conditions is conducted, investigating the influence of inter-site distance and the presence and location of mBSs on the multicast SE [53]. The work [54] analyzes a feature of LTE-A, the so-called LTE-Direct. The proposal made in this work consists of a mechanism for broadcasting transmission in LTE networks, with the goal of reaching a certain degree of robustness against BS failure. This feature is implemented in the resource allocation phase, based on an interpretation of the eMBMS standard, where resources are allocated and accessed by the UEs on a large area without connecting with the BS [54]. LTE can allocate part of its resources for direct broadcast transmissions, considered in the proposed mechanism. The performance of resource access schema can be optimized applying principles peculiar of the IEEE 802.16p networks [54]. Both the strategies of resource allocation and resource access scheduling are implemented exploiting the eMBMS features.

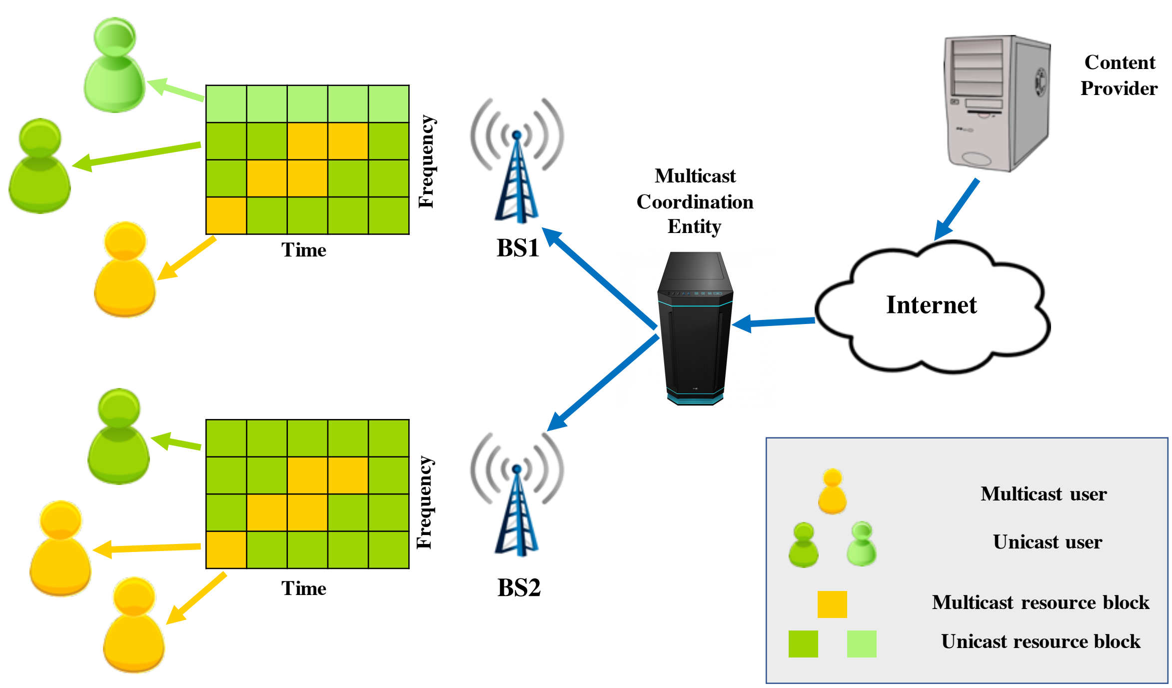

Fig. 9. Scheduling scheme.

E. Physical layer evaluation in MBMS

Several works evaluate the performance of LTE multicast transmission at the physical layer of the protocol stack [11], [13], [55]–[63]. SE analysis and evaluation are performed in [11], [55]–[59]. In [11], Signal to Interference and Noise Ratio (SINR) computation is first performed at a given location; then the SE is derived for any MCS, and an investigation is carried out on the use of different MCSs in LTE, (QPSK, 16-QAM, 64-QAM), through simulation. Relationship between SE and MCS are also considered in [55] to choose the most suitable MCS as a function of the SINR, to achieve a target BLock Error Rate (BLER). Different approaches are proposed that select the most suitable MCS for MBMS data transmission, and optimize the SE and the user throughput [55]. SE related to soft-frequency reuse is analyzed in [56]. In this work, different zones are defined within a cell where different frequency reuse factors are adopted, i.e., at the border of each cell, where only a fraction of the frequency spectrum is used, so that the maximum SE can be achieved in the center of each cell, at the same time obtaining a smaller inter-cell interference at the cell edge [56]. Based on target values of carrier-to-interference (C/I) ratio, the UE knows what type of reuse should be applied (depending on the C/I level) [56]. SINR and throughput are the performance metrics considered in this work. The work [57] proposes a review work of the MBMS services offered by 3GPP in LTE networks. The adoption of MBSFN is analyzed, with reference to the MCSs used and the SE. Achieved data rates for different numbers of UEs are obtained through simulations. The gain in SE is evaluated in [58]. An investigation is carried out on the effect of frequency selectivity on system performance. To this end, the most critical scenarios are defined. A subgrouping algorithm already proposed in literature is considered, to exploit frequency selectivity in subgroup formation and evaluate the gain in SE under different deployment cases [58].

A cost-oriented analysis is presented in [59], [60]. In [59] the MBSFN delivery scheme is evaluated. The goal of this study is to improve the accuracy of results with respect to the evaluation of the only SE in the air interface for different topologies, distributions of MBSFN areas, and users [59]. Based on this analysis, the work [60] performs a quantitative analysis on the cell rings that contribute to the MBSFN area. The goal is to maximize the ratio between the Single Frequency Network (SFN) gain and the telecommunication cost, evaluated as a function of the users distribution, in different scenarios [60].

Other aspects at physical layer are analyzed in [13], [61]. The work [13] focuses on the dynamic creation of MBSFN areas through simulation. The increase of the size of a MBSFN area and its dependence on the cell size is studied, based on SINR computations. Accordingly, a method is proposed that dynamically creates MBSFN areas to optimize the efficiency of multicast transmission of events to large numbers of users in a limited area. Furthermore, a group-based multicasting method is proposed to provide, to all the users in the region, the service with different transmission characteristics (i.e., different MCSs), depending on the radio channel quality of the single UE. Performance of transmission modes and their respective average cell throughputs in the DL channel of LTE-A is analyzed in [61], with respect to MIMO systems.

The work [62] analyzes the multicast delivery of Video on Demand (VoD). The goal is to reduce the blocking probability and increase the throughput, when VoD is transmitted through a limited number of wireless channels. The video is segmented and distributed through different multicast connections, acting in parallel to unicast streams. Based on the value of a pre-defined multicast efficiency factor (which depends on the request rate and video length), a content transmission request is satisfied by means of multicast or unicast connections.

The work in [63] proposes a transmission mode (called MIX) and a switching mechanism that switches transmission among MIX itself, SFN, and Point to Point (PtP) modes. Both these strategies are conceived to alleviate the negative effect of poor signal users in MBMS, since the allocation of wireless resources in MBMS is related to the worst signal user, and this translates into a low SE in presence of poor signal users. Wherever there are users experimenting poor signal conditions, the MIX mode is adopted, which incorporates both SFN and PtP modes, using a fast-select algorithm to switch between them, to optimize the SE and save energy resources. The switching mechanism monitors the state of the target MBSFN area, selects the transmission mode corresponding to the maximum spectral efficiency among MIX, SFN and PtP, compares the candidate transmission mode with the current mode, and decides whether to change mode.

F. Implementation of MBMS and MBSFN in simulation/emulation tools

Some works make an effort to implement some transmission aspects of multicasting in LTE through simulators/emulators, or experimental analysis [64], [65]. The work [64] aims at validating the MB-SFN transmission in LTE and LTE-A environments through the implementation of eMBMS in LTE simulators/emulators (i.e., the proposed OpenAirInterface open-source platform), according to the 3GPP Release 10 specifications, and including some eMBMS aspects like service continuity and MIMO. The physical layer and layer 2 protocols are fully implemented for this purpose. Results obtained through OpenAirInterface are compared with the corresponding theoretical values shown in the 3GPP standard [64]. SVC video transmission on LTE networks is proposed in [65]. An experimental analysis is carried out on the impact of the packet loss on video quality in a LTE system, considering different quality metrics. The SVC transmission through MBMS is then simulated over a wireless link, to investigate the quality of the SVC video delivery over LTE when losses occur due to both the wireless network and the buffer overflow at UE side [65].

G. Analysis of MBMS and MBSFN standards

The works [12], [14], [66]–[70] mainly focus on the analysis of specific aspects of the 3GPP standards involving MBMS and MBSFN. In [14] eMBMS is compared to the unicast transmission scheme, in terms of resource efficiency, expressed by means of the transmission overhead. An estimation of the lower bound of this metric is provided for unicast and broadcast transmission [14]. Mobile Broadband and MBSFN functionalities are investigated in [66]. MBSFN performance is analyzed in terms of SE, which comprises the application of AL-FEC schema, and power consumption in the scenario of DVB-T transmission for mobile users. Considerations are drawn on the possibility of the adoption of MBMS services in broadcasting scenarios, and on the power efficiency in public networks. PTM and MBSFN schema for multicast transmission are compared in [67]. Taking into account the related standards, the choice of the most suitable transmission scheme is discussed, highlighting the related features, i.e., the less waste of resources with respect to time, frequency, call number, user distribution over cells and reception conditions of the UEs [67]. Two mechanisms are proposed to decide which transmission scheme to use: the first based on the user density information as the breakpoint between transmission schema, and the second based on the comparison between single cell and multi-cell transmission, in terms of needed resources [67]. Coordination and mobility issues are also discussed. The relationship between the eMBMS service and next generation TV is discussed in [68]. Different aspects are analyzed according to the respective standards, such as bandwidth deployments for LTE/LTE-A systems, the effective utilization of eMBMS services, spectrum requirements and sharing for the first-generation TV (Digital TV, DTV) as related to population density. Link budget parameters and protocol stacks are compared and qualitatively analyzed. Examples are also discussed, together with the implications of a long-term spectrum planning. Service continuity for eMBMS users is analyzed in [12], presenting a standard analysis that follows the 3GPP specifications, and techniques that ensure service continuity in eMBMS. A method is proposed, thought to be integrated into the 3GPP standard, which aims at guaranteeing the service continuity and reducing the service interruption time during handover. Its goal is to choose the best candidate BS for handover and re-selection, exploiting the information of the neighboring cells. Moreover, the proposed technique allows the UE to favor services provided by neighboring cells transmitted in the same frequency of the current cell, as stated by LTE-A specifications [12].

The analysis of the 3GPP standard for MBMS and eMBMS service delivery is performed in [69], [70]. In [69], the most relevant use cases are described for audiovisual streaming of eMBMS services.

This work is a tutorial on eMBMS, and provides an analysis of the main parts of MBMS, including the radio access and core networks, and the service layer. Different aspects are analyzed: the eMBMS enhancements to unicast applications, the interactions of devices and service layers, the procedures of file download and repair, AL-FEC schema, and QoS issues, together with different aspects of the radio transmission [69]. The possible impact of the deployment of eMBMS services is finally summarized. Potentialities and opportunities of video delivery in LTE networks are analyzed in [70]. In this work, mobile video applications are first classified into four categories: Video on Demand, video multicast, video chat, and video uplink. For each of these categories, the most relevant characteristics (bandwidth, delay, jitter, resiliency, etc.) are then discussed. Issues and opportunities of video delivery in 4G networks are described, providing some considerations on the relationship between the specific application and some features, i.e., packet loss and power efficiency [70]. The most suitable metrics for video quality evaluation are also designed and implemented.

H. Performance evaluation of LTE multicasting

The works [71]–[74] focus on the performance evaluation of multicasting in LTE networks. Dimensioning strategies are studied in [71] for backhaul networks. Two different optical technologies for LTE backhauling are compared in terms of planning costs and network dimensioning, investigating the possibility of their implementation in LTE backhaul networks [71]. The comparison makes use of Linear Programming (LP) techniques, that define a cost function to be optimized. A discussion is carried out on the capability of the two proposed networks to support implementations of multicast. Three different methods are finally proposed to support LTE multicast traffic. The work [72] proposes a framework for wireless cameras. From one side, the proposed system receives video streams from camera stations (CSs) and forwards them to a control center; from the other, it exploits multicasting to transmit aggregated streams to different mobile stations to efficiently monitor video scenes [72]. The goal of this work is the optimization of the mobile stations QoE and the maximization of the efficiency in distributing videos to mobile stations. A policy of resource allocation and an algorithm for video layer adaptation in the case of scalable video are also proposed. In [73] the possibility of MBMS transmission is explored to design a framework in a smart grid scenario. Relating to this aspect, a greedy heuristic approach is proposed to solve the scheduling and Radio Resource Management (RRM) problems peculiar of LTE. Then, the solutions are extended to MBSFN [73]. Performance analysis and simulation results are carried out with respect to latency, packet loss and throughput, comparing the proposed framework with the classical unicast scheme. The work in [74] proposes a scheme to transmit unicast data superimposed to multicast data by means of scheduling. In this scheme, the BS chooses dynamically the transmission rate to guarantee, with a given probability, that UEs can successfully receive data, stopping the transmission when the large majority of the users has received data [74]. The transmitter encodes the multimedia information at source side, generating different layers of streams with different priority levels, mainly identified by the transmission power parameter. This scheme is exploited to mix multicast and unicast transmission, assigning to multicast data the higher priority: multicast transmission depends on the chosen scheduling scheme, while unicast transmission is adopted for the only user with the best channel condition [74].

I. Summary

Several works focus on different network architectures, that include MBMS in LTE networks [17]–[33]. These works follow two different approaches. In the first, the MBMS architecture is analyzed, with particular reference to its logical components, and in different application scenarios (ITS, video streaming applications, TV broadcasting, etc.) [17]–[24], [31]–[33]. In the second approach, the MBMS architecture is integrated with other architectures (like optical networks, DVB, or vehicular networks) [25]–[30]. *Lessons learned:* the analysis of the MBMS is useful to have a deeper insight on its main components. Some works consider the possibility of improving some aspect of MBMS transmission to enhance the user QoS/QoE and the efficiency in service distribution, at the same time saving resources, which are important aspects of multicasting, especially in emergency scenarios. The integration of MBMS with other technologies is surely useful to improve transmission performance. Nevertheless, an additional amount of complexity is needed at nodes side to perform the integration among the different technologies, especially in terms of a higher number of interfaces to support different technologies in the same node, and an increased energy consumption of mobile devices.

Cooperative strategies employ additional nodes to increase the cell coverage, the system data rate and the efficiency in content distribution, at the same time reducing interferences among adjacent cells. Additional nodes can be used to multicast the same service in parallel to BSs [34], or relay data to the final users in multi-hop transmission strategies [35], [36], [39]–[42]. Also the same UEs can be exploited as relay nodes, as happens in D2D cooperation [39]–[42]. Message forwarding can be performed also through different networks [37], [41]. *Lessons learned:* The cited studies testify that cooperative strategies are of great help in increasing spectral and energy efficiency, offloading traffic of BSs, and reducing delay. In such scenarios, scheduling becomes a very important issue, together with the most appropriate choice of the frequency reuse schema to avoid interferences between BSs and RSs. D2D is a promising strategy to relay data, at the same time saving the costs of an additional relay infrastructure. Nevertheless, devices acting as RSs will experiment an increased battery consumption to forward much more data in the UL channel. Cooperative strategies exploiting different technologies and heterogeneous networks suffer the same issues of the heterogeneous architectures: the improved performance comes at a cost of an increased number of interfaces and protocols at RSs side.

The goal of error correction techniques is to recover as much as possible original RBs that arrive corrupted at receiving side, without requesting their retransmission, and increase the robustness of transmission against packet losses in error-prone channels. As illustrated in Section VIII-A, coding techniques are suitable to perform this task. The way the additional data are added to the original source block, and the layer of the protocol stack where error correction is performed, characterize the specific FEC technique adopted in the works on this topic. At application layer, error correction techniques aim at improving the service data rate and users QoE [44]–[46]; more in general, the goal is to increase the robustness of the packet transmission in terms of successful packet delivery [47]–[49]. *Lessons learned:* Coding techniques for file repair and error correction are a key-feature for multicast transmission, where there is not the possibility to acknowledge the successful reception of packets, and increase the system throughput and the transmission robustness in noisy channels. This is counterbalanced by an increased complexity of the transmitters that have to wisely insert redundant data to allow packet recovery at receiving side, and an additional implementation cost in terms of encoding and decoding operations and data overhead (quantified by the code rate) [172].

An efficient allocation of time-frequency resources for data transmission is another very important issue in MBMS, since it can be of great help for energy saving and throughput maximization; in fact, its goal is to share the radio resources among UEs to maximize the efficiency in resource utilization, and minimize the per-user radio resources accordingly. Some works on this subject face this problem by considering the joint transmission of multicast and unicast services [50]–[52]. The basic idea behind these works is that only the UEs experimenting the same conditions (i.e., similar signal strength, assigned MCS, or number of needed time slots/subframes) can receive multicast streams. Other works instead propose scheduling schema to increase throughput [53] and robustness towards the the failure of BSs [54]. *Lessons learned:* Scheduling in both time and frequency domains should be taken into account in MBMS, because it achieves two important results in parallel: optimizes the overall system throughput, and minimizes the per-user energy consumption. Nevertheless, from this point of view, proposing an efficient scheduling scheme is not a simple task: the optimal resource allocation depends on several parameters, like for example the type and number of MBMS services and their average data rate, the channel quality, the number of UEs receiving the same service and their position in the multicast area, the number of available radio resources, etc. Taking into account all these aspects complicates the optimal design of solutions in this direction, as discussed in [50]–[52].

There are several works on LTE multicasting that analyze different aspects at the physical layer of the protocol stack. In this context, one of the most widely analyzed aspect in literature is the evaluation of the SE and the user throughput, related to SINR, MCS, frequency reuse schema, frequency selectivity and cost of transmission [11], [55]–[59], [63]. Other physical layer aspects taken into account are the creation, size and shape of MBSFN areas aiming to improve the users SINR [13]. Physical layer performance are also evaluated with respect to MIMO systems [61]. *Lessons learned:* One of the most important peculiarities of the multicast transmission at physical layer is MBSFN, where signals coming from different cells are considered at UE side as s single signal, with an increased strength. In this way SINR increases and the interference from different cells is eliminated. This explains the interest in SINR analysis and evaluation made by several works on this topic. The drawback lies in an additional complexity of such a scheme, that requires a stringent synchronization of the different signals coming from the cells of the same MBSFN area. Finding the best trade-off between these two opposite aspect is a complicated task [60].

There is a couple of works that consider the implementation of the LTE environment for performance evaluation purposes, following the most recent 3GPP specifications, in the lower or upper layers of the protocol stack [64], [65]. The goal is to evaluate and test some standardized features of multicast transmission [64], even in the most promising application scenarios [65]. *Lessons learned:* Emulating the MBSFN behaviour according to the 3GPP specifications is a very interesting and useful topic, for two different reasons. First, it allows to test the effectiveness of the ongoing standardization efforts on LTE multicasting, before their implementation on a real infrastructure. Second, it highlights the standardization aspects to be improved, and the new possible functionalities to be added. But there is a wide variety of standardization aspects, at different layers of the protocol stack. Considering them all in a complete simulation/emulation tool is a challenging research task.

The analysis of the main functionalities of MBMS and MBSFN according to 3GPP specifications are the goal of several works in literature. The main MBSFN functionalities are investigated, if compared to unicast [14] or multicast transmission in a single cell [67]. In some works, analysis of the eMBMS standardization is related to specific applications, especially video broadcasting [66], [68]–[70]. In almost all the works, multicasting is analyzed according to 3GPP specifications, discussing key-aspects like power consumption [66], efficient resource utilization [14], [67], bandwidth [68], [70], error protection [66], [69] and service continuity [12]. *Lessons learned:* All the works that analyze the 3GPP standardizations related to MBMS and MBSFN are of great help in clarifying the (often complex) descriptions of the multiple-faceted standardization aspects. The studies on this topic highlight that multicasting is particularly suitable for video transmission and digital TV broadcasting. Standardization of multicast transmission in LTE is still an ongoing process, and even if it has reached a good level of stability, there is still a margin for further improvements in this direction [12], [68], [70].

The last class of works discussing baseline approaches on MBMS and MBSFN evaluate the performance of different aspects of LTE multicast. Performance evaluation is carried out with respect to the most relevant metrics like radio resource assignment [72], latency, losses, throughput [73], and even the network dimensioning and cost [71]. *Lessons learned:* The evaluation of the MBMS and MBSFN performance can give a more concrete idea of the gap between standardization, often based on theoretical considerations, and the implementation of multicasting in practical scenarios [72], [73]. Performance evaluation suggests that the dimensioning of the core network and the implementation cost of the additional capabilities of MBSFN, with respect to unicast transmission, are critical issues that do not encourage the real implementation of MBSFN in last-generation cellular networks.

Just to conclude this section, Table II synthesizes the baseline approaches for MBMS and MBSFN transmission schema.

VI. Novel contributions for MBMS

A. Physical layer approaches

The strategies proposed at physical layer of LTE/LTE-A systems for multicasting purposes cover different aspects, such as the adoption of MIMO techniques, and optimization strategies for power, coverage, spectrum and data rate. The MIMO techniques considered in the surveyed literature are used mainly in combination with modulation schema for equalization purposes, or analyzed theoretically to derive analytical expressions of the BER. Coverage optimization is performed through cooperation among entities for coverage extension, optimal radio resource assignments, beamforming, or through analytical approaches. Power optimization is performed by means of power allocation and control techniques, that choose the power levels that optimize specific metrics such as throughput, delay, energy consumption, etc. Beamforming is also used to increase the power efficiency. Rate optimization strategies aim at optimizing the allocation of radio resources to maximize the overall system data rate, depending on several factors like channel conditions, resource availability, bandwidth budget, etc. Optimization algorithms are widely used in this context. Spectrum optimization is performed in the surveyed works through optimal MCS selection, cooperation among devices, and scheduling, even with the support of theoretical analysis. All these aspects are analyzed with more detail in the following subsections.

1) MIMO strategies: The works [75], [76] focus on physical layer proposals based on MIMO techniques. The MIMO configuration is taken into account in [75] to derive a theoretical expression for the BER in LTE and LTE-A. BER performance is evaluated through approaches based on the Moment

TABLE II

SUMMARY OF THE BASELINE APPROACHES FOR MBMS AND MBSFN.

Techniques	**Approach**	**Reference technology**	**Description**	**Reference works**
MBMS in network architectures	Enhancement and integration of MBMS into different network architectures	MBMS and MBSFN	Analysis of architectures based on MBMS	[17]–[33]
Cooperative strategies	Adoption of cooperative strategies (relay nodes and data forwarding)	MBMS and MBSFN	Multicasting is part of multi-hop cooperative architectures, including D2D and M2M cooperation	[34]–[43]
Error correction	Adoption of coding techniques for error correction and data repair	MBMS and MBSFN	FEC implementations at application and MAC layers, including transmission of redundant data, for error correction and data repair purposes	[44]–[49]
Scheduling of MBMS services	Time-multiplexing of MBMS services for battery saving purposes	MBMS and MBSFN	Scheduling techniques to increase power saving, throughput and robustness against BS failures	[50]–[54]
Physical layer evaluation	Performance evaluation of MBMS at physical layer	MBMS and MBSFN	Evaluation of MBMS performance in terms of SE, size and population of a MBSFN area	[11], [13], [55]–[63]
Implementation of LTE environment	Validation of MBSFN transmission through simulation/emulation tools	MBMS and MBSFN	The MBSFN transmission in LTE and LTE-A environments is implemented according to 3GPP standards, through an implementation of eMBMS in LTE simulators and emulators	[64], [65]
Standard analysis and evaluation	Analysis and evaluation of different aspects of the MBMS and MBSFN standardization	MBMS and MBSFN	Evaluation of MBMS main functionalities (SE, error protection, power consumption, transmission schema for service delivery, etc.) according to 3GPP standard	[12], [14], [66]–[70]
LTE performance evaluation	Performance evaluation of LTE multicasting	MBMS and MBSFN	Analysis of MBMS performance, as regards network dimensioning strategies and wireless communication frameworks, in different network scenarios	[71]–[74]

Generating Function (MGF) and Pairwise Error Probability (PEP), to derive analytical expressions of BER over slow fading channels in a closed form [75]. The BER performance is also compared on the basis of Coordinated Multipoint Transmission and Reception (CoMP) and MBMS features, for different modulation techniques in a multi-user environment. Modulation schema in combination with MIMO are analyzed in [76] for LTE systems. The goal of this work is the analysis of a receiver able to perform both MIMO detection and channel estimation, to support broadcast and multicast services in presence of errors in the channel estimation. The proposed receiver can apply different MIMO equalization techniques on-the-run to refine channel estimates [76]. This model is enhanced in such a way to incorporate hierarchical constellations and is tested for different configuration scenarios.

2) Power optimization strategies: Several works propose power optimization strategies [77]–[85]. The large majority of the papers on this issue proposes techniques of power allocation [77], [78] and control [79], [80], also considering QoS and QoE [80], [81]. The optimization approach followed by [77] considers an application scenario where a BS sends different flows to clusters of UEs within an area. Each node of the same group receives the same DL data from the BS. The optimal power level is chosen for each flow to enhance throughput and reduce delay, at the same time respecting the power constraints at BS side [77]. All this, taking into account RLNC schema for DL communication. An implementation of the power adaptation strategy is also carried out, that takes into account UEs fairness in resource utilization. Another algorithm of power assignment for MBSFN services is proposed in [78]. It first decides the power to be assigned to BSs for each OFDM subchannel, with the goal of maximizing the multicast channel gain as a function of the users number, an utility parameter, and the power constraints. The algorithm then determines each subchannel power so that a total utility function is maximized. This function depends on the total power employed and the time slots allocated for transmission [78]. Different utility functions can be modeled, based on system objectives. The work [79] proposes a power control scheme that dynamically selects the most suitable MBMS bearer, given the DL transmission power needed to deliver multicast data. The proposed scheme aims at minimizing the BS transmission power, depending on the parameters of MBMS users and services in each cell [79]. Both PtP and PTM schema are considered in this work, and comparison is performed with other standardized mechanisms. Another power control scheme for MBSFN is proposed in [80]. Based on the MBMS standard, a simulator and optimizer is implemented, that aims at minimizing the power consumption of the network, at the same time keeping the same users SE, through a simultaneous selection of the MBSFN area and power adjustment of the cells. To this end, a Genetic Algorithm (GA) is used, that also takes into account BSs position and UEs mobility thanks to its support to real coordinates for UEs and BSs. Multicasting of SVC video is considered in [81]. The power-efficient solution proposed aims at optimizing the user experience and power consumption in both UE and BS, by analyzing three trade-offs: the maximization of the energy saving versus the delay minimization, the maximization of the sleep time versus the minimization of lost packets, and the maximization of the video quality versus the minimization of unnecessary video transmissions. Users are grouped into multicast groups depending on different factors like multicast sources chosen by a user, video quality, and user location, taking into account cooperation among nodes and availability of LTE femtocells [81]. Bearers and channels are then assigned for multicast transmission to any users group. Video data are finally scheduled and resource blocks allocated for each SVC layer.

Some works propose beamforming and antenna selection techniques to improve the transmission power efficiency [82]–[85]. Multicast beamforming is discussed in [82], where an algorithm is proposed whose goal is the minimization of the total transmission power at the BS through an antenna array, at the same time guaranteeing a minimum Signal to Noise Ratio (SNR) per user. To reach this goal, the BS exploits any user CSI to regulate the antenna weights and selectively guide power towards the subscribers directions, at the same time controlling the amount of interference to other users [82]. Furthermore, the proposed model selects the best antennas subset and finds the corresponding vector of the weights that minimize the transmission power, while matching the SNR constraints per subscriber. In [83], the problem of [82] is extended to many Multicast Groups (MGs). Even if the general problem has a high computational complexity, a convex regularization is possible, which yields to a natural SemiDefinite Programming (SDP) relaxation of the problem with a reduced computational complexity. Sparse beamforming vectors are obtained by selecting antennas in a sparse fashion. In this paper it is shown that the proposed approach can be exploited to find a lower bound of the multicast channel capacity, and that the convex approximation of the general problem can be effectively adopted in real solutions, with the advantage of a reduced computational complexity [83]. Beamforming with the aim of transmitting common information to several users is considered in [84]. Given a set of antennas transmitting to a single group of users, the goal is how to find the optimal values of the beamformer weights that minimize the total transmission power, given the users QoS requirements. The idea proposed in this study is to orthogonalize the channel vectors of the user DL, representing the antenna beams, to satisfy the QoS constraints, i.e. the users SNR, and using one of the decomposition techniques already known by literature [84]. Then, a local search is performed to find the vector of weights that maximizes the worst user SNR. An antenna sharing technique is proposed in [85] for efficient multicast transmission over LTE systems. Antennas of neighbouring BSs are treated as one large array, without a central unit, and shared over a high-speed network connection, with specific algorithms described in literature. Some aspects are discussed, like the inter-cell interference, areas in which joint signal processing is possible, UL and DL differences, and the influence of network bandwidth and computational power. A protocol for information exchange between BSs at network layer is also described. The goal is the dynamic selection of the BSs exploiting multicast transmission to send data to different groups of neighbouring BSs, to reduce the number of packets processed by the BS [85]. The groups of cooperating BSs are dynamically selected according to the interference among them, for the sake of network bandwidth scalability.

3) Coverage optimization strategies: To increase the coverage capabilities of LTE/LTE-A, several approaches are proposed in the recent literature [86]–[95]. The works [86]–[89] propose cooperative techniques for coverage optimization in LTE multicasting. RSs are employed in [86], [87], in the scenario of multicast transmission of critical messages for public safety issues. To this end, scheduling algorithms are considered, whose aim is to increase the coverage area [86]. Considerations are made on the relationship between the reciprocal distance between BSs, the opportunity to employ relaying schema for multicast transmissions from BSs to Mobile Stations (MSs), and the the reuse level of the scheduling schema [87]. Furthermore, the possibility is investigated to jointly choose scheduling and routing schema to prevent from failure events of the BS. An analytical model for D2D transmission is proposed in [88], based on an analysis of the coverage probability at terminal side. The multicast throughput is used as a metric for selecting the optimal transmission rate. Dynamics like terminal mobility or data transmission are investigated and related to SINR, coverage probability, mean number of receivers, and throughput [88]. The work [89] proposes a user cooperation method to increase the coverage capabilities of MBSFN. The retransmission requests by UEs with a lower SINR are redirected by the BS to UEs that successfully received the signal, and falling in the same cooperation radius (a strategy of transmission radius selection is also proposed), which defines the space portion enclosing all the UEs participating in cooperation. In addition, an energy-efficient MBSFN transmission is proposed, that saves the total system power and achieves better performance with respect to the classical MBSFN transmission, thanks to the user cooperation [89].

Some works make use of GAs to face the coverage optimization issue [90], [91]. In [90] a GA is used with the goal of optimally assigning Radio Resource Units (RRUs) in superimposed MBSFN areas, that can be used whenever the resource allocation changes according to the users demand. The set of solutions found by the algorithm allows to correctly allocate the RRU to the overlapping MBSFN area, and represents the optimal resource arrangement even in the case of a dynamic variation of the services, that are not transmitted all the time [90]. In [91], a GA is proposed for an optimal MBSFN area configuration, depending on the users distribution. Taking into account the 3GPP technical specifications for MBMS and MBSFN, the algorithm exploits the MCE employed in the MBMS system architecture, to assign the area with highest UE density as a MBSFN area. After the search is performed, the area coverage is optimized by configuring dynamically the MBSFN areas depending on the MBMS users distribution, and in such a way to avoid area overlapping [91]. The search takes into account the distance among areas and the users percentage that can be supported in an area.

Beamforming is another possible solution to solve coverage optimization problems [92], [93]. As illustrated in Fig. 10, beamforming modifies the shape of the transmitted signal by maximizing the signal strength in the direction of a particular UE, optionally minimizing it in the direction of other UEs managed by other BSs. In [92] a design is performed of the optimal beamformers from a QoS perspective. Starting from a single multicast group scenario, an enhancement of a simple adaptive multicast beamforming algorithm already developed in literature is proposed. This algorithm aims at serving a certain percentage of users, at the same time maximizing the minimum SNR under the hypothesis of a fixed power budget [92]. An approximation of the original problem is also proposed, that improves convergence and performs close to the optimum (as shown by numerical results). Beamforming and admission control for multicast groups are studied in [93]. The goal is the maximization of the number of subscribers reached by the service, at the same time minimizing the required power. To this end, an algorithm is proposed that achieves good solutions, with a relatively low complexity. For the case of a single multicast group, an adaptive algorithm already known by literature is used, identifying its strengths and drawbacks, and proposing some improvements [93]. All the algorithms described are carefully tested by deriving SINR values on indoor and outdoor environments. Other solutions can be found in [94], [95]. In [94] a model is presented to give a realistic evaluation of MBSFN transmission. It takes into account the 3GPP specifications introduced in LTE. First, the SINR is evaluated at a given point of the cell, taking into account delay due to fading and shadowing effects on MBSFN transmission. Then, based on the SINR information within a cell, the probability distribution of the SE is analytically derived [94]. The proposed analytical approach is then extended to MBSFN activated dynamically, if only subscribers are present, and used to model the user traffic. To this end, the probability that at least only one user is active in the cell is used to derive the SE distribution. The work [95] analyzes the performance of synchronization channel for the cell detection in MBMS. The application of the so-called Successive Interference Cancellation (SIC) technique is applied to the cell search procedure, to improve the detection performance of the synchronization channels in LTE, responsible for getting the slot timing, the cell group, the frame boundary and the cell ID [95].

4) Spectrum optimization strategies: Strategies for spectrum optimization can be found in [96]–[100]. Some works propose MCS selection strategies for spectrum optimization [96]–[98]. An AMC and scheduling scheme is proposed in [96] for layered video transmission, with the goal of increasing the spectrum efficiency of the MBSFN channel, at the same time maximizing the users throughput. To this end, based on the measured SINR, MCSs are chosen for Base Layer (BL) and Enhancement Layers

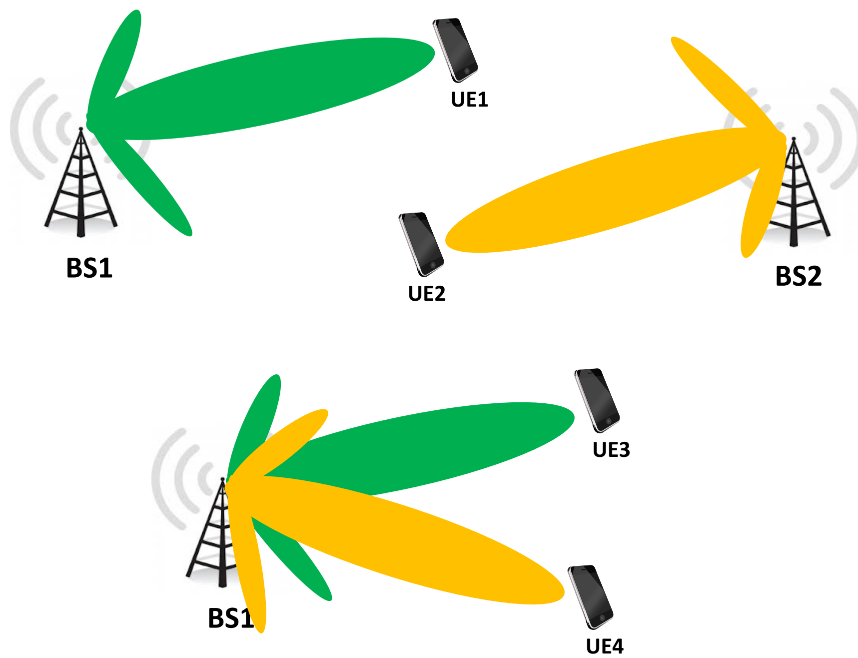

Fig. 10. Examples of beamforming. Each BS contains multiple antennas that regulate the main parameters (magnitude and phase) of the physical signal transmitted by each antenna, to increase the signal strength in privileged directions (coinciding with specific UE locations).

(ELs), depending on the users distribution in three different channel quality regions: the high channel quality region receives all the SVC video layers, the medium quality region the BL and one EL, and the low quality region only the BL [96]. Different approaches for MCS selection in MBSFN transmission are discussed in [97], [98]. In [97] four approaches are proposed. In the first approach, the algorithm selects the MCS corresponding to the minimum SINR, so that all users receive the MBSFN service; in the second, the selected MCS guarantees the maximum throughput and SE; in the third, the lowest MCS is chosen that guarantees a target SE in a MBSFN area, and the fourth approach is a variation of the third one, but it selects the MCS which ensures that the large majority of users reaches the target SE [97]. In each case, MCS values are defined according to the 3GPP specifications. In [98] other three approaches are proposed, with different goals: the provisioning of a guaranteed service, the maximization of the SE, and the achievement of a target SE for various user distributions. MCS selection and SE calculation are performed by first computing the SINR of the UE, then mapping SINR to the available MCSs, and finally calculating the achieved throughput that depends on SINR and MCS values. The SE of each MCS is derived accordingly. The possibility of MIMO techniques is also considered, analyzing the advantages of MIMO utilization.

Other works dealing with SE are [99], [100]. The goal of the work [99] is the optimization of the network energy consumption, at the same time keeping a reasonable SE for the majority of UEs. The SE of a cell is computed as a function of the number of cells surrounding it and transmitting MBSFN services. An analytical method is proposed, based on a robust GA, that accurately computes the SE of cells, also for arbitrary positions of BSs [99]. In [100] a cooperative multicast analysis in LTE-A systems, from the spectrum and energy points of view, is performed. The goal of this work is the design of a cooperative multicast scheme, suitable for high user densities. This scheme is composed by two stages: in the first stage, BS transmits multicast data only to MSs experimenting better channel conditions to successfully receive data; in the second, the selected MSs forward data to the remaining MSs which did not receive them in the first stage [100].

5) Rate optimization strategies: Rate optimization strategies are proposed in [101]–[112]. There are several works that exploit subgroup formation techniques for rate optimization [101]–[107]. Fig. 11 provides a graphical example of a multicasting subgrouping scheme, where multicast receivers are subdivided into different subgroups, generally based on their channel conditions, and resources are assigned accordingly. In [101]–[103] subgrouping strategies based on channel conditions and available data rate are presented. The work [101] proposes enhancements to a subgrouping algorithm already known in literature and designed to maximize the system throughput, represented by the Aggregate Data Rate (ADR). The algorithm iteratively merges two subgroups that, joined together, experiment the highest ADR increase, and so forth until no further ADR increase is achieved, or subgroups cannot be merged anymore [101]. The proposed algorithm is then improved to avoid scalability problems. Performance evaluation is carried out aiming at minimizing the 3GPP guidelines for the physical layer. Another effort to reduce the computational complexity of the procedure of subgroups formation can be found in [102], where a low complexity frequency-based algorithm is proposed. The BS manages the available frequency resources (one or more adjacent subcarriers) within a given scheduling frame, and collects the CSI feedbacks from each member of the multicast group. Accordingly, the best subgroup formation scheme is chosen, based on two different target cost functions to be maximized: the ADR and the Proportional Fairness among users [102]. In [103] a solution is proposed to the problem of subgroup formation, that aims at maximizing the ADR with a reduced computational complexity. This work aims at analytically demonstrating that the solution of the algorithm for subgroup formation can be obtained in a closed form and with a low computational cost, introducing some simplifying hypotheses. The optimal subgroup configuration is found, in terms of user distribution, resource allocation, and MCSs of each subgroup [103]. The proposed strategy is compared with the Exhaustive Search Scheme (ESS) and with other approximated techniques already discussed in literature, according to the 3GPP guidelines for the LTE physical layer. The multicast subgroup allocation scheme proposed in [104] is based on bargaining solutions which exploit game theoretic notions (please, refer to Section VIII-A for further details), to allocate resources in multicast subgroups. Subgroups are formed based on the collected CQI. Then the number of RBs and the transmission parameters are chosen for each subgroup, considering that a subgroup can be formed by UEs with different CQIs [104]. Fairness among users (quantified by the so-called Group Fairness Index, GFI) and throughput are the metrics adopted to test the algorithm performance. The works [105], [106] propose RRM policies to support multicast services in LTE-A, exploiting solutions based on multicast subgrouping. In [105] resource allocation is performed through bargaining solutions, that exploit CCs of LTE-A. MGs are determined based on all the possible combinations of MCSs, then UEs with the same CQI are associated to the same MG, and finally the bargaining solution optimally allocates resources to each MG [105]. Finally, MGs to be activated in each CC are chosen. The system settings are properly tuned to meet fairness and throughput requirements. Based on what developed in [105], in [106] the DL air interface in LTE-A is modeled according to the related standard, including the Carrier Aggregation (CA) scheme that can group together up to five LTE CCs. The proposed framework is compared to other solutions that extend CMSs and Opportunistic Multicast Schemes (OMSs) to LTE-A [106]. The pros and cons of the bargaining solutions adopted are finally evaluated. In the framework of the multicast subgrouping techniques is also the work in [107], that proposes a scheduling technique in the frequency domain for an efficient radio resource management of multicast services. The multicast LTE subscribers are organized into subgroups, based on the experienced channel conditions, and differentiating the subgroups transmission parameters. The scheduler acts as in [104], grouping UEs and assigning the relative MCS and number of RBs depending on the CQI information. The minimization of the Group Dissatisfaction Index (GDI) is taken as QoS parameter [107]. Accordingly, a minimization problem is solved to find the optimal distribution of RBs among the subgroups.

Another RRM policy for rate optimization is found in [108]. The CMS is considered together with D2D communications, to increase the ADR of the eMBMS cell, also preserving fairness among UEs. In the proposed scheme, one or more UEs, defined as Forwarding Devices (FD), are able to retransmit data received from the BS to other members of the same MG. In this case, the CQI feedbacks are collected from all UEs of the same MG with respect to both the BS (CQI collection) and between the UEs themselves (D2D CQI collection) [108]. The BS establishes which UEs are FD, and which UEs are served by a FD through D2D connections, and computes the solution that maximizes the overall system data rate.

Analytical models for rate optimization can be found in [109], [110]. Analytical tools are provided in

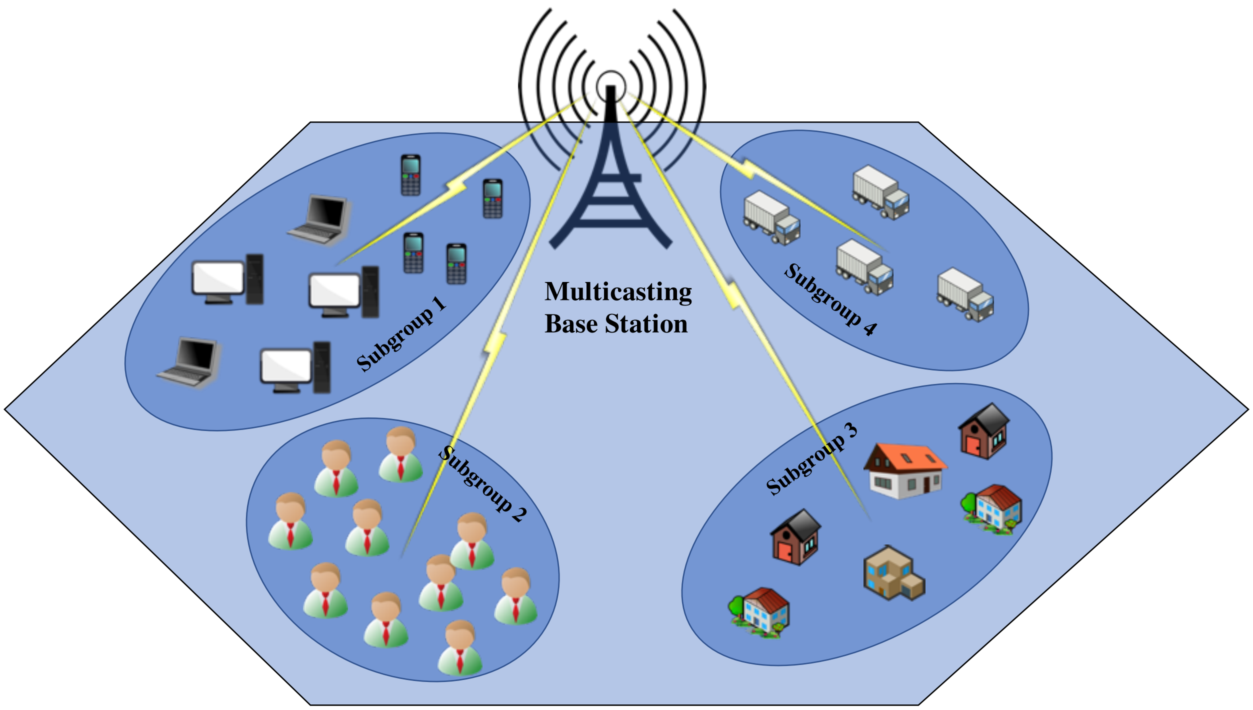

Fig. 11. A multicast subgrouping scheme.

[109] to analyze energy consumption and SE in LTE and LTE-A. A Finite-State Markov Chain (FSMC) model, already adopted in literature, is customized to evaluate average service data rates. Based on the rate analysis, bandwidth and energy consumption are derived, comparing different LTE/LTE-A scenarios [109]. An analytical model is also proposed in [110], with the goal of evaluating the blocking probability in cellular systems transmitting multicast data. The proposed method takes into account all the interfaces connecting the BS with the the Random Network Coding (RNC), to evaluate the effectiveness of the multicast traffic. The proposed model focuses on a multicast group carrying different traffic classes at different rates, and is adopted on a set of links with multicast connections [110]. Finally, analytical and simulation results are compared for different traffic classes (speech, gaming, FTP connection, video conference and web browsing) evaluating blocking probability and throughput.

Rate optimization analysis for video multicasting is performed in [111], [112] in specific application scenarios. In [111] multicasting of layer-encoded IPTV programs over LTE-A with relay functionalities is considered. In a scenario with limited resources, the goal of the proposed scheme is to optimally allocate the number of time slots to different IPTV video layers so that the number of served UEs or the overall users satisfaction is maximized [111].The utility function proposed for the optimization depends on the burst profile of the transmitted data and the maximum number of time slots required to forward a layer of the IPTV program from the BS to the RSs of the multicast group. Multicasting of multi-view 3D video streams is analyzed in [112]. An optimization problem is formulated, aiming to minimize the bandwidth consumption, in terms of number of RBs used to transmit all the selected views of multi-view 3D videos. To this end, a MCS is assigned for each selected view, according to the values assigned in 3GPP specifications [112]. The possibility for each user to synthesize a view from a received left view and right view is also considered, in case that the view is transmitted with a MCS not accepted by the user (because of quality constraints). Different scenarios are analyzed in simulation results.

Table III summarizes the main strategies proposed at physical layer for MBMS and MBSFN.

TABLE III

SUMMARY OF THE MBMS AND MBSFN APPROACHES CONCERNING THE PHYSICAL LAYER OF THE PROTOCOL STACK.

Techniques	**Approach**	**Reference technology**	**Description**	**Reference works**
MIMO techniques	Adoption of MIMO configurations to minimize channel errors	MBMS and MBSFN	MIMO configurations are adopted to analyze the BER and performance and face channel estimation errors	[75], [76]
Power optimization	Strategies of optimal power allocation and control	MBMS and MBSFN	The transmission power level is adjusted in order to improve specific metrics (throughput, delivery, delay, transmission power constraints, MBSFN area, QoS and QoE parameters). Beamforming and antenna selection techniques are employed to minimize transmission power at BS	[77]–[85]
Coverage optimization	Optimal resource allocation to maximize the signal coverage area	MBSFN	The MBSFN coverage area is maximized by optimally allocating resources in terms of D2D cooperation, analytical models, algorithms for optimal radio resource allocation, and beamforming	[86]–[95]
Spectrum optimization	Strategies for the maximization of the spectral efficiency and throughput	MBSFN	The spectrum optimization is achieved through a smart selection of MCS that satisfies a target SE and/or optimizes the power consumption of the network	[96]–[100]
Rate optimization	Strategies of optimal data rate allocation among UEs	MBMS and MBSFN	The aggregate data rate is maximized through subgroup formation techniques in the frameworks of radio resource management, relaying techniques and analytical models	[101]–[112]

B. MAC layer approaches

Multicasting strategies proposed at MAC layer are mainly focused on error protection and resource allocation strategies. Error protection strategies translate into retransmission techniques, that often act in combination with FEC and cooperation schema, with the goal of reducing the packet errors, reduce overhead and save bandwidth. Resource allocation is performed through scheduling strategies that optimize power consumption, delay and fairness among users. Some papers propose cooperation to reduce the UEs feedbacks and increase service continuity and QoS. A detailed discussion of both the abovementioned strategies will be carried out in the following subsections.

1) Error protection strategies: Strategies for error protection can be found in the works [113]–[119]. Retransmission techniques are considered as error protection strategies in [113], [114]. In [113], two kinds of Automatic Repeat reQuest (ARQ) schema with NC are designed. The goal is to set a proper redundancy factor, so that network-coded ARQ can save retransmission resources at the same time ensuring the same Packet Error Rate (PER) if compared to the normal HARQ retransmission scheme. In addition, three kinds of chase combining schema are analyzed, comparing their performance. Retransmission schema for MBSFN are proposed in [114], to decrease the retransmission rate and increase goodput. In the proposed schema, the Physical Random Access CHannel (PRACH) is used to report NACK messages. Two methods are proposed, in dependence of whether two channels can share the same PRACH sequence numbers or not; also retransmission trigger schema are introduced and applied on different feedback channels [114]. The proposed schema are finally compared with the HARQ-based scheme. Simulation parameters of LTE are set taking into account the compatibility with the 3GPP standard.

Another error protection strategy exploiting cooperative transmission is developed in [115]–[117]. NC in combination with cooperation is analyzed in [115], [116]. In [115] the cooperation is provided by RS nodes in a LTE-A scenario, where RSs forward network-coded packets to terminals with different rates. The goal is to find the transmission schedule with minimum delay of the RSs when the multirate transmission is adopted. To this end, a Markov process is employed to model RS retransmissions and derive the optimal retransmission mode and delays [115]. Two heuristics are also proposed as retransmission procedures, to find a trade-off between quality and computational effort. In [116] the user cooperation is exploited together with NC in MBMS, for bandwidth saving and and QoS purposes. User cooperation occurs because of the UE capability to communicate with other UEs on a SR link, and is used to reduce the overhead of repair symbols introduced by NC in the cellular link, saving bandwidth and QoS, without reducing correction capabilities [116]. Moreover, by applying Raptor codes only to the SR link, also delay can be reduced by using smaller blocks. A cooperative retransmission scheme is presented in [117] as a solution to manage errors at receiving side in LTE systems. The goal of this work is to reduce traffic load and energy consumption in the recovery procedure by sending retransmitted data over SR links. To this end, the Cooperative Retransmission Protocol (CRP), already present in literature, is modified to manage multicasting. The energy efficiency is improved by reducing retransmissions, on average, and avoiding as much as possible retransmissions of the same message required by more UEs.

Other error protection strategies can be found in [118], [119]. The possibility of integration of the Fountain codes (please refer to Section VIII-A for further details on Fountain codes) in LTE and LTE-A at MAC layer is analyzed in [118]. The main goal of this work is the implementation of Fountain codes taking into account the specific constraints of LTE. This implementation is then compared with the implementation at application layer of the same codes. The work [119] discusses a method to guarantee a synchronized transmission of multimedia data in MBMS. This strategy is applied when a channel is shared by different variable bit rate streams, where, in case of loss of transport blocks, the BS does not know the distribution of lost packets into transport blocks. Through the use of a deterministic header size, lost data can be resumed without knowing their distribution.

2) Resource allocation strategies: Resource allocation strategies are discussed in [120]–[126]. Scheduling algorithms for optimal resource allocation are discussed in [120]–[123]. QoS-aware scheduling is discussed in [120], [121]. In [120] a hybrid scheduling scheme is proposed to manage the user request in eMBMS. The main goal of the strategy is to reduce the overall waiting time of a requested data, and provide QoS both to the system and to users. The scheme measures and manages the impatient user behaviour, occurring when a user is unable to retrieve the requested data and for this reason leaves the system or performs multiple requests. Accordingly, the proposed scheme schedules the user data and the relative request based on the so-called cutoff point: the scheduler cuts the single request dynamically, also classifying it according to the number of users and the data popularity [120]. Two different queues based on data popularity are considered, and two different scheduling algorithms are applied, one per queue. Scheduling and resource allocation in LTE-A is discussed in [121], with the goal of improving QoS evaluated in terms of delay and fairness among streams. Scheduling is subdivided into two layers, long-term and short-term scheduling. In the long-term scheduling, information on the queue state, like the packet delay and fairness, is used to set the transmission order of the streams. In the short-time scheduling, weights are assigned to the frames to establish the priority in the transmission of the more important packets. A Power saving Scheduling Algorithm (PSA) for MBSFN is proposed in [122]. The goal is the improvement of the terminal energy efficiency, in a scenario of mixed multicast and unicast traffic. To reach this goal, the UE switching times between sleeping and active states in DRX mechanism are reduced, obtaining a better average throughput, both for unicast and multicast traffic, and energy efficiency of the UEs [122]. Also the interference issue in MBSFN areas is considered, by allocating different RBs to overlapping or adjacent areas, depending on the priority assigned to each area. As regards the unicast transmission, the PF scheduling algorithm, as described in the LTE standard, is adopted.

The trade-off between cooperation strategies and flow multiplexing for MBMS transmission is discussed in [123]. The goal of this work is to investigate the real benefits of cooperation, if compared to multiplexing of multicast sessions. To this end, first the sets of active relays are partitioned into subsets with a negligible mutual interference, enabling cooperation between relays that provide the same service. The low interference among subsets is exploited to reuse channels across subsets without cooperation and perform multiplexing, allocating subchannels on relay and access nodes to maximize the system performance [123]. Cooperation is also analyzed in [124], where an integration is discussed between D2D communication and the cellular network for service multicasting. The goal of this work is to increase the QoS and service continuity for D2D users through a control mechanism performed by the network to improve D2D communication, in terms of resource allocation and transmission power. A switching mechanism is also considered from D2D to cellular connection and vice-versa, in case of bad D2D signal reception [124]. The proposed strategy is validated through a simplified implementation of the protocol stacks for user and data planes.

In [125] the goal for resource optimization in MBMS services is reached through a reduction of the set of UE feedbacks, while keeping a high SE. The feedback reduction is obtained by allowing only UEs with worse channel conditions to send feedbacks [125]. Three different algorithms are proposed, based on the parameters that determine the choice of the UEs in the feedback set: the path loss, the SINR, and the SINR together with the short-term BLER (when available). The algorithms are implemented according to the 3GPP guidelines for the LTE physical layer. The application of RNC is adopted in [126] with the goal of improving communications reliability for H.264/SVC video streams. An optimization of the radio resource allocation process in eMBMS is performed through a modification of the MAC layer of LTE, by adding the so-called MAC-RNC sublayer that performs symbol coding. The design of this sublayer is particularized for SVC video streams [126]. Furthermore, a strategy is proposed to differentiate QoS levels depending on MBMS areas, through a joint optimization of block sizes together with the MCS of each area.

Table IV synthesizes the novel approaches involving the MAC layer for MBMS and MBSFN.

C. Network layer approaches

A network layer strategy for LTE multicasting is found in [127]. In this work, the synchronization protocol SYNC adopted in MBMS is discussed. Some alternatives for the setting of the TimeStamp

TABLE IV

SUMMARY OF THE MBMS AND MBSFN APPROACHES CONCERNING THE MAC LAYER OF THE PROTOCOL STACK.

Techniques	**Approach**	**Reference technology**	**Description**	**Reference works**
Error protection	Strategies of improvement of packet error rate	MBMS and MBSFN	Strategies of protection from packet errors: retransmission techniques, cooperative or synchronized transmission of redundant data, integration of error protection codes in the LTE protocol stack	[113]–[119]
Resource allocation	Strategies for optimal transmission of user and control data at MAC layer	MBMS and MBSFN	The optimization of resource allocation is performed through scheduling algorithms, cooperation techniques and service areas differentiation	[120]–[126]

(TS) control field are taken into account and compared to the existing scheme, in which the BM-SC is responsible of the TS setting. Three are the proposals: the TS is configured based on the absolute time each packet is received, or on the absolute time of the first received packet, or on the synchronization sequence number. In the first method, the BM-SC does not need to know the synchronization sequence information and the delay; in the second and third methods, the BM-SC needs to know only the starting time and the length of the synchronization sequence.

D. Application layer approaches

Several strategies for multicasting in LTE are proposed at application layer. They mainly focus on multiplexing techniques, cooperative schema, error protection, and the transmission of layered video. Multiplexing is analyzed through theoretical models for bandwidth estimation and statistical multiplexing. Cooperation at application layer is performed through a protocol in video multicasting scenarios. FEC and NC techniques are adopted as error protection strategies at application layer, while layered video transmission strategies focus on the optimal allocation radio resources to the different layers of a SVC video, so that specific metrics are maximized (like users QoE, QoS or energy efficiency) or minimized (like bandwidth or video distortion). All these categories are described in detail in the following subsections.

1) Multiplexing techniques: Multiplexing techniques are analyzed in the works [128]–[130]. Analytical models are developed in [128], [129]. In [128] a theoretical bandwidth estimation model is proposed. It is a low complexity model, thought for MBSFN transmission of broadcasted video streams. The model considers the factors influencing the bandwidth requirements of a MBSFN transmission, i.e., service blocking, over-provisioning and service multiplexing. The bandwidth is estimated for each group of users receiving a service and for the the aggregated bit rate of the reserved resources, taking into account the three factors mentioned above [128]. The work [129] proposes a Markov-based model for the statistical multiplexing of video streams in MBSFN. The proposed resource allocation scheme is introduced, together with the traffic model that considers the sources as modulated by a semi-Markov process. The probability of buffer overflow and violation of delay constraints is evaluated, exploiting two different approaches known by literature [129]. Finally, the optimal channel rate is derived, based on the QoS constraint (buffer, or delay), by solving an optimization problem. Statistical multiplexing of MBMS services in analyzed in [130]. The goal of this work is the evaluation of the multiplexing gain for MBMS transmission for different services in different overlapping areas that dynamically change their transmission time slots without any inter-cell synchronization (the so-called "dynamic SFN" scenario) [130]. Statistical multiplexing gains are evaluated for H.264 encoded video sources, studying the impact of the number of concurrent services and cells on the multiplexing gain.

2) Cooperative schema: The only cooperative scheme at application layer is proposed in [131], where a cooperative protocol, PULLCAST, is discussed for video multicasting. Videos are subdivided into chunks, and nodes receive them by means of multicast transmission. The goal of the protocol is to exploit a Peer-to-Peer (P2P) overlay network, where unicast transmission is used to transmit lost data among peers. The impact of mobility, together with some other parameters, on the system performance is also discussed for different fractions of clients, adopting as performance metric the minimum number of chunks needed for an acceptable video quality experience [131].

3) Error protection strategies: Error protection strategies at application layer are found in [132], [133]. In [132] the FEC strategy is investigated at application layer. The file recovery scheme proposed in this work analyzes the peculiarities of MBSFN to improve FEC operations in eMBMS, based on the use of Raptor codes. An ACK is sent by the receiver when all the encoding symbols collected allow to completely recover the file, while the sender continues sending redundant encoding symbols until all receivers have sent their ACK [132]. NC applied to eMBMS network architecture is discussed in [133]. This study takes into account the difference between the Packet Loss Ratio (PLR) and the degree of redundancy of the Fountain codes, explained in detail in Section VIII-A. This study highlights that for small differences, an optimal range of the redundancy degree can be found to minimize the probability of decoding failure; for large PLRs instead, an excessive redundancy can be avoided because not useful [133]. The goal of the proposed scheme is twofold: to improve the decoding probability and reduce the decoding delay, so that the optimal compromise can be found between PLR and the degree of code redundancy.

4) Layered video transmission: Fig. 12 illustrates an example of construction of layered video frames. SVC is a standardized video compression technique aiming at generating layered videos, at different resolutions and frame rates. The first video stream (usually defined as BL) is coded at the lowest resolution, requiring a low bandwidth. The other streams (the ELs) are coded at higher resolution, to improve accuracy in details and motion description, at a cost of an increased frame rate and bandwidth. These additional streams complement the information of the BL. Strategies for layered video transmission are developed in [134]–[137]. The works [134], [135] deal with the MCS allocation to each video layer, according to specific metrics. More in detail, in [134] MCS and resource symbol rate are derived for each video layer, exploiting the rate adaptation mechanism provided by the BSs supporting LTE (the AMC mechanism), with the goal of maximizing fairness among users and their QoE. The optimization procedure is performed in terms of the maximization of the user satisfaction in terms of QoE and improvement of fairness among users [134]. Four different resource allocation strategies are analyzed: throughput, PF which takes into account fairness (quantified by a fairness index), QoE (specified by a QoE model), and PF with respect to QoE maximization [134]. The trade-off between QoE and fairness for three different kinds of video distortion is also discussed. The goal of the work [135] is to analyze the advantages of an MCS allocation that depends on the SVC video layer. According to the channel models and available MCS schema peculiar of MBSFN, the BL is transmitted to the whole target area, exploiting only a fraction of the available bandwidth; the rest of the bandwidth is used to transmit the EL to the only UEs with good reception conditions [135].

The contribution of [136] is to present a hybrid transmission approach that exploits multicast transmission schema of both layered and non-layered videos. More specifically, the most suitable transmission scheme for a multicast group is adaptively chosen, based on an energy efficiency analysis performed on both layered and non-layered video transmission schema, and dynamically choosing the more energy efficient scheme [136]. A video adaptation scheme is proposed in [137] for SVC videos. The goal of the scheme is to keep a high QoS for the maximum number of users in a cell area. The proposed scheme can decide for unicast or multicast transmission for Video on Demand (VoD), while considers only multicast transmission for mobile TV. For unicast transmission, different video resolutions can be sent, based on UEs feedbacks, depending on criteria such as channel quality, traffic congestion at BS side, or UE battery and/or processing effort. For multicast transmission, UEs with better and worse channel conditions are subdivided into two different multicast groups, sending the BL of the SVC video to both the groups and the ELs only to the group with better channel conditions [137].

Table V summarizes the main approaches at the application layer for MBMS and MBSFN transmission schema.

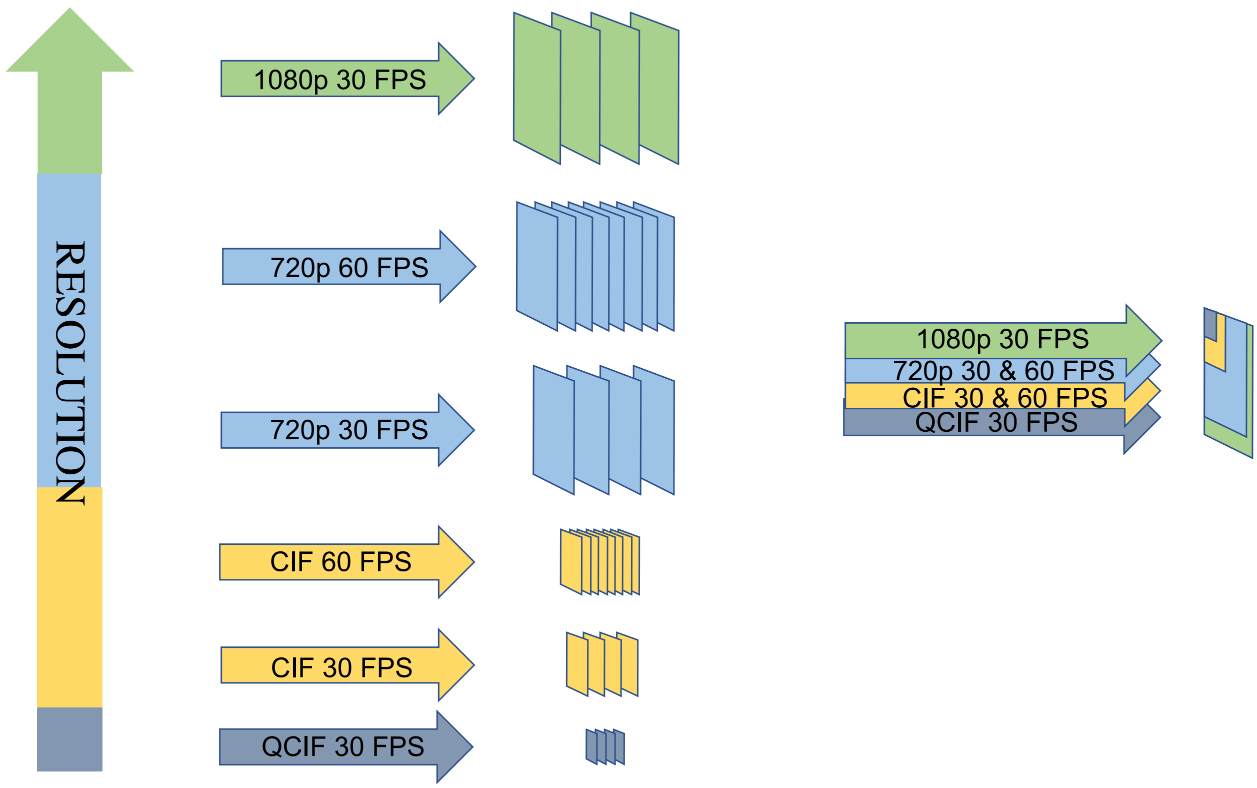

Fig. 12. Construction of a layered video. The same video is compressed with multiple frame resolutions (QCIF, CIF, 720p and 1080p in the figure) and rates (30 and 60 Frames Per Second, or FPS, in the figure). Each video layer is a combination of the abovementioned parameters. The BL has the lowest resolution and frame rate (QCIF resolution and 30 FPS), and is anyway transmitted; the other ELs, with higher resolutions and frame rates (yellow, blue and green colors in the figure) are transmitted in case of higher signal quality and bandwidth capabilities.

E. Cross-layer approaches

Cross-layer strategies are proposed in the works [138]–[143]. The works [138], [139] focus on error protection at different layers, including FEC and HARQ for efficient delivery of eMBMS data. The contribution of [138] is twofold. First, the coexistence of Raptor codes and HARQ is analyzed in LTE networks; specifically, the most appropriate values of Raptor codes and HARQ parameters is found, based on channel conditions. The optimal values of the parameters are used to adapt traffic in the DL channel according to the varying network conditions. The goal of the proposed approach is to keep the signaling procedure simple, save network resources by reducing the number of HARQ retransmissions and the redundancy factor of Raptor code, and satisfy the QoS requirements, evaluated in terms of data reliability [138]. All this, by changing only the system parameters without changing the implementation of the eMBMS protocol stack, which is compliant to the 3GPP standard. FEC schema with Raptor

TABLE V

SUMMARY OF THE MBMS AND MBSFN APPROACHES FOR THE APPLICATION LAYER OF THE PROTOCOL STACK.

Techniques	**Approach**	**Reference technology**	**Description**	**Reference works**
Multiplexing techniques	Data multiplexing approaches for bandwidth optimization purposes	MBMS and MBSFN	Multiplexing of compressed video sources is analyzed in theoretical models for bandwidth estimation purposes	[128]–[130]
Cooperative schema	Cooperative protocol for multicast video streaming	MBMS and MBSFN	A Peer-to-Peer (P2P) overlay network exploits a combined unicast-multicast transmission of video packets	[131]
Error protection	Strategies of error protection performed at application layer	MBMS and MBSFN	FEC and network coding methodologies are employed for file recovery and PLR reduction	[132], [133]
Layered video transmission	Optimization of transmission of layered video streams	MBMS and MBSFN	The different video layers are assigned to different MCSs, multicast groups, or energy levels	[134]–[137]

codes in combination with HARQ mechanisms for eMBMS are also discussed in [139]. Considering the practical system constraints, the goal of the scheme is to simplify the signaling procedure by decreasing the number of HARQ retransmissions, at the same time satisfying the user QoS requirements. To this end, a threshold setting mechanism is proposed that conditionally controls the amount of HARQ feedback, i.e., a NACK packet is sent by the user only when the loss rate is higher than a threshold, set-up by the HARQ entity and depending on the QoS requirements [139]. A transmission scheme based on RLNC is proposed in [140]. It aims at minimizing the energy consumption of the BS per message delivered in a MBMS scenario. RLNC is integrated at MAC layer, instead of application layer, to reduce delay and redundancy of information due to network coding. The proposed Extended-RLNC (E-RLNC) aims at optimizing the number of the copies of each received coded symbol that are needed to correctly deliver a message, based on the information on the UEs channel conditions (through CQI), the length of each information symbol, and the MCS adopted to transmit a MBMS flow [140]. The synchronization issue is considered in [141], that investigates the possibility of a tight synchronization of MBSFN content transmission performed by the BS, through a common relative time reference instead of an absolute time reference used in the past literature. The goal of this work is to provide a temporal SFN alignment among multiple BSs, so that the Reference System Frame Number (R-SFN) is synchronized with a good accuracy at physical level among the BSs in a MBSFN area [141]. This goal is reached through some modifications to the standardized parts of the LTE system concerning the coordination between BS and the MCE and the interworking between RLC and MAC schedulers. In [142], the multicast support in LTE is developed as part of a cross-layer architecture for video delivery. The goal of this work is to illustrate the research on multicast video in next-generation networks. To this end, SVC multicasting is analyzed in the architecture sub-systems, and the end-to-end support is analyzed in a SVC scenario. Solutions to improve the wireless access efficiency are proposed, and eMBMS is also enhanced to improve QoS and the efficiency of the SVC transmission. The work [143] proposes a multicast protocol in LTE. The goal of this work is a cross-layer optimization involving MIMO and cooperative transmission, through a scheduling mechanism that aims at the maximization of the UE throughput. The BS exploits the MIMO spatial multiplexing on a single channel, but can also send other separate sessions to other terminals, through a different channel. Furthermore, UEs can cooperate each other in exchanging data through orthogonal, non interfering channels, so that several independent sessions can be performed simultaneously on the same number of channels, instead of using one session per channel [143].

Table VI summarizes the main cross-layer approaches for MBMS and MBSFN transmission schema.

TABLE VI
SUMMARY OF THE CROSS-LAYER APPROACHES FOR MBMS AND MBSFN.

Techniques	**Approach**	**Reference technology**	**Description**	**Reference works**
Error protection	Strategies of error protection at different layers of the protocol stack	MBMS and MBSFN	Combination of FEC and HARQ techniques, and extension of NC techniques to other layers of the protocol stack	[138]–[140]
Synchronization	Synchronization of MBSFN content transmission	MBMS and MBSFN	The synchronization of MBSFN content transmission is performed by means of a common relative time reference and through a modification to LTE standards	[141]
Cross-layer architectures	Integration of multicast support in a cross-layer fashion	MBMS and MBSFN	cross-layer optimizations of multicasting for efficient delivery of mobile data	[142], [143]

VII. SC-PTM TRANSMISSION SCHEMA

A. Baseline approaches

Complementary to MBMS and eMBMS schema, that require a multi-cell environment for coordinated transmission of services, PTM schema considers service multicasting in a single cell environment. So, it is very useful to compare these two approaches, discussing the pros and cons of each solution. In fact, the comparison between SC-PTM and MBSFN, whose usefulness is testified by the 3GPP studies highlighted in Section II, is discussed in two papers dealing with SC-PTM baseline approaches [15], [16]. In [15] both the MBSFN and SC-PTM scenarios are analyzed, focusing on inter-cell interference coordination methods. In the scenarios analyzed, different factors are considered: fixed RSs, frequency reuse, space diversity and MIMO multiplexing. In the SC-PTM scenario, where no temporal synchronism among BSs is assumed, macro-diversity time synchronization is considered between the two closest BSs [15]. This study is mainly focused on numerical results. The proposed scheme is evaluated through Monte Carlo simulations at link level, carried out according to 3GPP specifications. MBSFN and PTM methods are considered together for the provisioning of MBMS services in [16]. The goal of this work is twofold: a performance evaluation of the combination of MBSFN and PTM, and a more sophisticated evaluation aiming to obtain more accurate results, if compared to the other approaches found in literature. To this end, the total transmission cost is introduced. It includes the cost of polling, the cost to detect cells with UEs that want to receive a specific MBMS service, the cost of the interface for delivering packets to BSs, the cost of synchronization (appearing only for MBSFN transmission scheme) and the PTM cost that includes the air interface and the packet delivery cost at the network nodes [16]. For experimental evaluation, a simulation tool is designed and implemented: it selects the most suitable transmission mode (MBSFN or PTM) that minimizes the total transmission cost, depending on the specific LTE configuration. User mobility in a single cell and variable MBSFN areas are also taken into account.

B. Physical layer strategies

Papers dealing with the PTM physical layer analyze MIMO techniques and power and spectrum optimization strategies. MIMO and spatial multiplexing are adopted to reduce BER. Strategies for power optimization in PTM scenarios aim at minimizing the transmission power by combining together PtP and PTM transmissions, with the aid of cooperation, to save energy of the multicasted services. Spectrum is optimized by aggregating the different carriers in a LTE-A scenario. All the papers exploring these issues are analyzed with more detail in the following subsections.

1) MIMO strategies: The work [144] proposes an adaptive MIMO scheme for for PTM transmission in LTE. The proposed scheme focuses on both spatial multiplexing and diversity techniques. Spatial diversity is adopted to improve the signal strength at cell edge to increase the BER performance, while spatial multiplexing is adopted to increase data rate and throughput of UEs with good channel conditions [144]. Different spatial multiplexing techniques are dynamically chosen, based on UEs distributions in the MBMS area.

2) Power optimization strategies: Two works focus on power optimization strategies for PTM [145], [146]. In [145] a mechanism is proposed, that selects the most suitable radio bearer in eMBMS to optimize the transmission power. The goal of the proposed scheme is the minimization of the BS transmission power through a wise selection of the most suitable combination of PtP and PTM bearers in the DL. This selection relies on an estimation of the optimum coverage in PTM transmission, depending on the UEs distribution in the cell; after that, the algorithm decides the PtP bearers that are used for the remaining parts of the cell [145]. Cooperation between different Radio Access Technologies (RATs) is considered in [146]. The goal is the energy saving of multicasted services. The typical scenario analyzed in this work is a MG that asks for a service, and is placed in an area with different RATs. Given a MG, with each working cell associated to a specific RAT that can select PtP or PTM, the objective is the optimization of the overall energy consumption [146]. To this end, an algorithm is designed, that provides solutions that include the choice of the working cell and the transmission mode (PtP or PTM) according to UEs dynamics, at the same time saving energy efficiently. The optimization problem is formulated by means of Integer Linear Programming techniques.

3) Spectrum optimization strategies: The work [147] focuses on the scenario of scalable video multicast in LTE-A. The goal of this proposal is to exploit SC-PTM and eMBMS to transfer MBMS information from the BS to a group of UEs over shared resources, exploiting the Carrier Aggregation feature of LTE-A. Functionalities such as packet scheduling, link adaptation, adaptive modulation and coding and HARQ are performed with respect to each group of eMBMS subscribers for the transmission of scalable videos, by optimally choosing MCSs for the BL and the EL of the video, for QoS purposes, and proposing independent service objectives for the BL and the EL. To this end, the problem formulation for allocating resources to multicast groups is discussed, proposing a near-optimal greedy approximation for a fair assignment of carriers for the BL and opportunistic assignment of carriers for the EL, to maximize throughput.

C. MAC layer strategies

MAC layer strategies for PTM schema can be found in [148]–[150]. The work [148] proposes a strategy to simplify the signalling procedures in MBMS. The goal of this work is to support streaming and “download-and-play” services to UEs, taking into account radio and network resources. This task is carried out through a strategy which simplifies the signalling procedure that detects the presence of UEs in a cell, in a context in which both PtP and PTM with dedicated feedbacks are supported [148]. The choice of PtP or PTM depends on the number of users and the provided services. The Frequency Domain Packet Scheduling (FDPS) algorithm is extended in [149] for PTM services. RBs are dynamically allocated, based on the information on instantaneous channel conditions, with the goal of optimizing the system throughput with an acceptable loss of coverage. The proposed packet scheduler acts bot in the time domain, by scheduling MBMS (re)transmissions in each Transmission Time Interval (TTI), and in the frequency domain, by scheduling MBMS services on different frequencies [149]. Retransmissions are allowed, but only if all the UEs of the MG do not acknowledge the packet. Different RLNC approaches for PTM layered service delivery are proposed in [150]. The work relies on the Multirate Transmission (MrT) strategies, that allow the delivery of different versions of the same service; so, they are suitable for layered videos consisting of a BL and multiple ELs, like SVC videos. The proposed technique is a resource allocation scheme that aims at optimizing the transmission scheme and minimizing the number of broadcast packets, at the same time guaranteeing a satisfactory QoE of the user [150]. This goal is achieved by jointly optimizing transmission parameters and the RLNC scheme adopted, using packet error probability expressions as performance metrics.

D. Cross-layer strategies

Cross-layer approaches for PTM are discussed in [151], [152]. In [151] SC-PTM is applied to data broadcasting. The single cell scenario is designed, in terms of network architecture, signalling procedures, radio channel mappings and radio resource control. Then, the proposed SC-PTM mechanism is applied to a real LTE-A testbed, consisting of content servers, two BSs and two mobile terminals, while the implemented multicast services are streaming, chatting, and a personal multicast service where a mobile terminal transmits video to other users subscribed to the service [151]. In [152] the issue of mobile TV traffic delivery over OFDMA networks is discussed. Analytical cross-layer models, based on Markovian and fixed-point analysis, are developed first for unicast connections, and then extended to SC-PTM. The proposed cross-layer approach considers different aspects at physical layer (e.g., SINR and throughput), MAC layer (e.g., RRM, priorities and channel assignments), and flow-level (evolution of incoming and outcoming users) [152]. Some analytically tractable solutions are derived, also in closed form wherever possible, and a proposal is made, that varies the MCS according to the users channel conditions.

E. Summary

Some works describe different aspects of SC-PTM [15], [16]. They mainly focus on physical layer aspects like signal interference, frequency reuse schema, BSs locations, MIMO multiplexing and SE [15]. A comparison between PTM and MBSFN multicasting is analyzed for performance evaluation. An effort is also made to increase the accuracy of the evaluation, by taking into account several transmission aspects that contribute to a cost function built for performance evaluation purposes [16]. These works present also accurate numerical results through extensive simulations or ad-hoc tools. *Lessons learned:* Comparison between the two main approaches for multicasting in LTE systems, i.e., PTM and MBSFN, is one of the most interesting issues and expresses the need to quantify the pros and cons of each solution. Accordingly, a very accurate analysis, especially at physical layer, of the two transmission schema is mandatory, which is a very difficult task due to the different standard-related aspects of both PTM and MBSFN.

The physical layer of SC-PTM is analyzed with explicit reference to MIMO and multiplexing techniques, power consumption and optimal selection of transmission parameters at physical layer [144]–[147]. The novel proposals focus on improvements in signal strength and data rate [144], minimization of the power consumption [145], [146] and optimal choice of MCS and CCs in specific application scenarios [147]. *Lessons learned:* Like MBSFN, also for SC-PTM the physical layer aspects are of paramount importance to improve the overall quality of data reception for QoS purposes. In the case of SC-PTM, since multicasting acts within the boundaries of a single cell, the novel proposals should take into account signal interferences with adjacent cells transmitting the same service, especially in a context that does not require synchronization.

The works concerning the MAC layer of SC-PTM multicasting mainly deal with scheduling and error management strategies through feedbacks [148]–[150]. Proposals in this direction optimize the choice of the transmission mode [148], scheduling [149], and resource allocation [150]. The main goal is to increase the robustness of the packet transmission in an error-prone channel to increase data throughput [149] and reduce the packet error probability [150]. *Lessons learned:* The novel strategies proposed for SC-PTM at MAC layer confirm the need to increase the robustness of packet transmission, to save energy and increase throughput. Optimal solutions to these problems are not simple to reach also in single-cell environments, especially when multiple metrics have to be jointly optimized.

A couple of works propose cross-layer optimizations for SC-PTM [151], [152]. In this respect, the novel strategies focus on signalling procedures and resource management and control, in specific application scenarios like broadcasting of data [151] and TV services [152]. *Lessons learned:* Cross-layer strategies are a very interesting and more complete approach, since they jointly take into account different layers of the protocol stack, through theoretical [152] or testbed-based [151] analysis. Nevertheless, this kind of approach is more complex to face, because it should take into account both the different layers of the protocol stack and interfaces among layers.

Table VII summarizes the main approaches for PTM transmission schema.

TABLE VII

SUMMARY OF THE MAIN APPROACHES FOR PTM TRANSMISSION SCHEMA.

Techniques	Approach	Reference technology	Description	Reference works
Baseline approaches	Evaluation of the state-of-the-art of PTM transmission	PTM	Comparison of MBSFN and SC-PTM for inter-cell interference coordination and cost/performance evaluation	[15], [16]
Physical layer	Novel approaches concerning the physical layer of the protocol stack	PTM	MIMO schema, strategies of power optimization and control, and optimal resource allocation	[144]–[147]
MAC layer	Novel approaches concerning the MAC layer of the protocol stack	PTM	Signalling procedures, packet scheduling strategies and RLNC approaches or layered services	[148]–[150]
Cross-layer	Novel cross-layer approaches for PTM transmission	PTM	Proposals of cross-layer architectures and analytical models for SC-PTM transmission	[151], [152]

VIII. FURTHER INSIGHTS ON NETWORK CODING TECHNIQUES AND OPTIMIZATION ALGORITHMS

A. *Network coding techniques*

Network coding is a very interesting technique that aims at increasing throughput by reducing the number of packet transmissions, at the same time increasing the amount of information exchanged between source and destination nodes. With NC there is no need of packet retransmissions and receiver feedbacks, thus saving bandwidth and power, and packet losses are reduced by properly coding packets also at intermediate nodes, rather than only at the source node. Also security is increased, since the transmitted packets are combinations of the original packets. For all these reasons, NC is the right choice in wireless networks scenarios, where the transmission is subject to signal errors and throughput limitations, and securing packets is of great importance. It has been implemented at different layers of the protocol stack, to improve the transmission performance, especially in terms of throughput and robustness toward losses. In the specific LTE scenario, NC is typically applied at MAC layer.

The key-concept of NC is to combine more packets coming from different source nodes into a single packet (the coded packet) [173]. This task can be performed both at source node, by combining different generated packets, and/or at intermediate nodes, by combining together packets coming from different other nodes. In typical packet switched scenarios (like Internet), NC is thought for transmissions over so-called erasure channels, where packets are either received without errors, or are not received. In such scenarios, the most widely used NC techniques are variants of block-based codes for erasure channels, the most representative of which is the Reed-Solomon coding technique. It consists of generating, at source side, a number of coded packets, say N, higher than the original number of packets, say K. Packets are coded in such a way that when the receiver has successfully received any K of the N coded packets, the original K packets can be recovered [172]. A code rate is also defined, as the ratio $K/N \leq 1$. It is obvious that this kind of coding technique is effective only for small values of N and K, and for high code rates. This technique has evolved into another class of erasure correcting codes, the Fountain codes, where packets are coded in such a way that the original K packets can be successfully decoded by receiving a (possibly slightly) higher number of packets, which is not determined a priori, but varies dynamically as the minimum number of coded packets needed to recover the original information [172], [174]. For this reason, Fountain codes are said to be "rateless". In random linear Fountain codes, at each step, a coded packet is obtained on-the-fly, as a weighted sum (performed as a XOR operation) of a number K of original source packets, with K binary weights randomly generated [172]. The first example of such kind of Fountain codes were proposed by Luby, and are called LT-codes [172], [174]. It can be shown that the encoding cost of Fountain codes is not constant if the number of output packets is close to the number of input packets. To overcome this limit, and reduce the computational cost at the encoder and decoder and obtain a constant encoding and decoding cost, another class of LT-codes, the Raptor codes, has been introduced in literature [174]. Basically, Raptor codes keep the same advantages of the Fountain codes, but with a reduced computational complexity at encoder and decoder sides. This goal is reached by appending to the input symbols some redundant symbols, and then use an appropriate LT-code to generate output symbols.

Based on the coding procedure explained above, RLNC linearly combines a number of source packets into a number of coded packets, where the coefficients of the linear combination are chosen randomly in a predefined field. RLNC has been fruitfully adopted in several scenarios with lossy and unreliable channels, like different types of wireless networks, and for different applications like multicast distribution and storage systems [175], [176]. More in detail, block-based RLNC approaches group together the source messages into blocks, and each block, called generation, is composed by a given number of source packets [176]. All the packets in a generation are then coded by means of linear combinations of all the original packets in the generation, and with coefficients extracted randomly from a finite field [177]. The generation of coded packets occurs at each block time. It is worthy to note that Fountain and Raptor codes can be considered as particular cases of RLNC, where the coefficients of the linear combinations are binary. The computational complexity of this kind of approach is tractable both at encoder and decoder sides; nevertheless, the most significant drawback lies in the increased delay to derive the original source packets from the coded ones, since the whole generation must be received before decoding the original packets [176]. To mitigate the impact of this issue, the size of the generation could be reduced, or the generation partitioned in a number of sub-generations with only a coded packet appended to each of them, but in this way the packet loss probability increases because each coded packet protects less source packets [175].

There are some variants of this approach that aim at reducing delay, which is a critical issue in real-time applications. One of them, called systematic block-based RLNC, appends to the source symbols in a generation a number of coded packets, grouped together in a "tail". These coded packets are used only if some of the original packets of the generation are lost [175]. Nevertheless, the delay of this systematic block-based approach keeps low only when packet drops are not so frequent. Vice versa, if only one original packet is lost, it is necessary to wait all the coded packets in the tail to recover it. A reduced delay can be obtained through sliding window RLNC techniques, while keeping a high reliability in data delivery. Sliding window approaches perform the coding of source packets within a window that slides over the sequence of original packets. Nevertheless, a feedback is usually needed to slide the window and include new source symbols [176]. This is a problem in some application scenarios like LTE multicasting, that do not allow feedbacks from the receivers. To avoid feedbacks, some variants of the sliding window approaches have been proposed, that achieve a good compromise between a reduced delay and an increased reliability, and without requiring feedbacks from the receiver. Nevertheless all this comes at a cost of an additional signalling procedure and some modifications of the packet header [176].

B. Optimization algorithms

Several works propose algorithms for the optimization of different aspects of multicast transmission in LTE, ranging from MCS selection and beamforming optimization at physical layer, to scheduling, sub-grouping techniques, power and resource allocation, and layered video transmission. All the optimization algorithms are exploited in general to maximize, or minimize, an objective function with one or more input variables, subject to some inequality constraints. The constrained optimization problems can always be solved through the so-called exhaustive search (or direct search, or brute-force) schema. ESS algorithms find the global solution to the problem, by testing all the values of the input variable(s) in order to find the solution, but they can require prohibitive computational efforts and execution times to find the solution, especially if the number of variables is high. So, they are usually not suitable for real-time computations, and very often approximated searching algorithms are needed, that find a suboptimal solution but with much less computational effort [178]. Different approaches are proposed, that are classified in the following subsections as *Exhaustive search schema*, *Genetic algorithms*, *Standardized programming-based optimization techniques*, *Game theoretic approaches*, *Iterative algorithms*, and *Heuristic strategies*.

1) Exhaustive search schema: As explained previously, ESS approaches guarantee the global optimum solution, but at the expense of a huge computational cost, which very often is not feasible in practical scenarios. In the works [107], [134] examples can be found of the application of this approaches. The work [107] proposes a scheduling algorithm that organizes multicast users into subgroups, while the work [134] proposes different optimization algorithms that allocate resources to the different layers of a layered video for scalable video transmission. An effective technique is to perform the same ESS procedure, but on a reduced search space to reduce the computational cost, as testified by the works [28], [86], [106]. Specifically, in [28], [106] this approach is adopted in a cluster formation algorithm with D2D links, to maximize the overall system data rate. In [86] the algorithm optimizes the location of RSs to maximize the system throughput.

2) Genetic algorithms: Another class of optimization algorithms exploits a genetic approach, that can be synthesized into an evolutionary method based on the "survival of the fitted" concept [178], [179]. In GAs, the starting point is an initial population, usually represented by random strings of fixed or variable length. The individuals of the population are characterized by so-called fitness values, that are computed by means of an objective function, to discriminate between "good" and "bad" individuals. Then, the best fitted individuals, representing the optimal solution at the current step, are chosen for reproduction, and new, mutated individuals are generated from the chosen ones, by crossing them over. Then the process repeats, generating other individuals with a better fitness, and after a number of iterations, the population converges towards the optimum. The drawback of this technique is that there is no guarantee of convergence to the global optimum, and the convergence time can be huge. Examples of the application of GA are found in [80], [99], for power optimization, and in [90], [91] for resource allocation strategies.

3) Standardized programming-based optimization techniques: Some optimization techniques exploit a well defined structure of the objective function, the input variables, and the associated constraints. Depending on this, different kinds of programming problems are introduced. In this subsection, the programming techniques found in the analyzed literature will be explained in detail. The first kind of programming technique is the integer programming, where the input variables can assume only integer values. A particular case of this kind of programming technique is when the variables can assume only binary values (0 or 1). In this case, the optimization problem is called Binary Integer Programming (BIP), or zero-one programming problem [178]. Linear programming techniques identify all that optimization problems where both the objective function and the constraints are linear functions of the input variables [178]. Dynamic programming techniques refer instead to the procedure adopted to solve the problem.

Dynamic programming occurs whenever the main optimization problem is subdivided into a number of subproblems, that are solved sequentially one after the other [178]. The sequential quadratic programming techniques are utilized whenever the objective function has a quadratic behaviour with respect to the input variables, and the constraints are linear. Finally, SemiDefinite Programming (SDP) techniques are used whenever the input variables of a function, that linearly depends on them, can be organized into a matrix which is positive semidefinite [180]. All the programming-based optimization problems mentioned above can be solved through specific procedures [178], [180]. Examples of these kinds of problems can be found in [52], [77], [82]–[84], [92], [93], [111], [115], [123], [143]. The algorithms in [52], [77], [92], [93] face the power optimization issue. In the framework of the physical layer approaches are the algorithms discussed in [82]–[84], while the work [111] develops an optimization algorithm in the framework of resource allocation strategies. Finally, in [115], [123], [143] scheduling algorithms are proposed in the context of cooperative strategies.

4) Game theoretic approaches: Some optimization problems found in the context of resource allocation strategies are solved by means of game theoretic bargaining approaches. Basically, they are based on a set of players, each of them characterized by a set of actions and a function identifying the strategy of the player. Each player is required to have a minimum performance value to enter the game, called as "disagreement point". Examples of this approach are found in [104]–[106], where optimization algorithms are developed for strategies of subgrouping and resource allocation.

5) Iterative algorithms: Some optimization algorithms are solved through iterative procedures that cannot be framed within any of the approaches discussed in the previous subsections. These procedures are formally implemented by means of nested cycles that perform a search over all the input variables, and update the value of the objective function whenever a new optimum value (maximum or minimum) is found. The works [20], [40], [41], [50], [51], [55], [72], [78], [94], [97], [98], [101]–[103], [112], [121], [122] exploit such algorithms. Scheduling algorithms are discussed in [20], [121]. In [40], [41] algorithms are proposed for grouping UEs in clusters. Optimization algorithms are proposed in [50], [51], [122] for combined Unicast/Multicast scheduling. The algorithms presented in [55], [97], [98] aim at finding the best MCS that optimizes a predefined metric through the SE evaluation. The work [72] proposes an algorithm for optimal resource allocation for layered video transmission. A power optimization algorithm for multicasting of video streams is presented in [78]. Optimization algorithms for subgrouping strategies are developed in [101]–[103]. The analysis developed in [94] comprises an algorithm that finds the minimum MCS satisfying a target SE. The optimization algorithm proposed in [112] assigns the best MCS to each view of a multi-view video stream, so that the bandwidth consumption is minimized.

6) Heuristic strategies: Heuristic approaches are often used to find an approximate solution to an optimization problem, sacrificing the precision and correctness of the optimal solution found in favour of a reduced computational complexity and a faster execution time of the algorithm. Heuristic methods are introduced because the optimal solution to the original problem is very hard to find in a reasonable time; the approximations introduced often do not explore all the search space and/or explore only some arbitrary solutions, generally the most likely ones, excluding the others. As a consequence, the approximations introduced do not precisely describe the problem to be optimized and, accordingly, the overall optimal solution will most probably not be found. Rather, a suboptimal solution (the so-called “local optimum") can be found, that is the result of the algorithm in the reduced search space that verifies the constraints. Please note that some classes of optimizations algorithms described in the previous subsections, i.e., genetic algorithms, game theoretic approaches, and iterative algorithms, can be considered as particular classes of heuristic algorithms, with well defined solving procedures. Examples of heuristic-based approximations not falling into the previous categories can be found in [52], [111], [115], [126], [140]. The works [52], [111], [115] propose heuristics to simplify the optimizations of their algorithms already discussed in subsection VIII-B3. In addition, the works [126], [140] adopt heuristic algorithms to solve resource allocation problems.

IX. Lessons learned on multicasting in LTE

A. *Lessons learned on the analysis of MBMS and eMBMS standardization*

The analysis of the eMBMS standard as developed by 3GPP is a key-aspect to provide a deeper insight on the key features of eMBMS, especially related to the application of the standard itself to different scenarios (video delivery services, file delivery and repair, broadcast TV, etc.). Since all these studies rely upon standardized recommendations and specifications, their adherence to the universally accepted and consolidated scenario of LTE multicasting is guaranteed. On the other side, the only degree of novelty introduced by these works consists on the application of eMBMS standards to specific use cases, without proposing any substantial enhancement to the various aspects of the standard itself.

B. *Lessons learned on cooperative strategies*

From the analysis of the papers on this topic, it emerges that cooperative schema are advantageous because they reduce the costs of implementation of the system, have a higher flexibility, can reduce the number of retransmissions for data repair, can be applied to various communication scenarios (the most widely known is the cooperation between several UEs), and can increase the energy efficiency. Nevertheless, all this comes at the cost of a higher complexity, especially for what concerns synchronization among RSs. Furthermore, cooperative multicasting cannot be so beneficial, since data have first to be delivered to relays, during the relaying phase, before being forwarded to UEs. This implies a SE decrease and a delay increase if compared to the non-cooperative multicasting schema that adopt only a single phase for data delivery.

C. Lessons learned on subgrouping strategies

Subgrouping strategies are another very interesting approach proposed by several works in literature, to overcome the limitations of the CMSs, where the various transmission parameters are set-up based on users with the worst channel conditions. Subgrouping improves the overall performance of the multicast session, since each of the subgroups adopts the MCS of the user with the worst channel conditions, and all the subgroups are composed by receivers with similar measured CQI; this mitigates the limiting effect of a low MCS imposed by the user receiving the worst signal. Moreover, subgrouping can be usually managed without employing additional multicast channels. The higher control overhead is counterbalanced by a SE improvement. Nevertheless, fairness among UEs of different subgroups is still an open issue, and in addition, fairness in throughput and the overall system efficiency are two conflicting aspects, so a trade-off has to be found accordingly. Furthermore, the computation complexity needed to allocate subgroups is relatively high, due to the efforts in optimizing the number of subgroups, their relative transmission parameters and resources (i.e., the optimal distribution of RBs among the subgroups). So, an ESS approach is not suitable for real-time implementations. Low-complexity sub-optimal schema are usually proposed to reduce the computation time, but they provide sub-optimal solutions (in terms of SE, user QoE, and fairness).

D. Lessons learned on layered video multicasting

Encoding videos in accordance to the different users radio conditions is surely a straightforward solution to improve resource assignment for video multicasting in LTE. The works facing this issue consider the multicast transmission of SVC encoded videos, by applying different strategies. Video layers are usually mapped to different MGs. Depending on the MG of membership, a receiver can obtain video data from the BL, assigned to a low MCS, up to a certain number of ELs, so that users with better reception conditions can experience a reduced power consumption and an increased video quality. On the contrary, the best assignment of the video layer(s) to MGs implies that UEs can feed back signal propagation or positioning information to the transmitting node, and these features are not implemented in multi-cell MBSFN services. In addition, according to the recent studies on this issue, the resource allocation problem seems not to have been properly addressed by taking into account scheduling and structure of LTE radio frames as stated by 3GPP specifications.

E. Lessons learned on power analysis and optimization strategies

Power analysis and optimization is often related to optimal beamforming and antenna selection in the majority of the works found in literature. The optimal beamforming design aims at minimizing the total transmitted power, given some predefined QoS guarantees, typically measured by the SNR constraints of each multicast group. The maximization of SNR for MGs translates into finding the optimal vector of weights that control the main lobe and side-lobes of the beamformer, and the selection of the antennas contributing to the beam pattern. Unfortunately, this optimization problem under SNR constraints is NP-hard, and requires a large amount of computational capabilities. Approximations have been proposed to reduce the computational load, but they unavoidably bring to suboptimal solutions, which do not guarantee the QoS requirements for the group with the weakest link. The possibility of designing multiple beam patterns to transmit simultaneously to more MGs has also been analyzed in the recent literature, but under the restricting hypothesis of negligible crosstalk interferences. On the other side, some works have considered the case of multiple co-channel MGs, but the proposed solutions are potentially infeasible because of the aforementioned crosstalk limitations. Other works focus on system optimizers that compute the optimal MBSFN area so that the network power consumption is minimized, while keeping the users QoS at the required levels, but again the computational complexity of the solvers is high and periodic re-computations of the optimal solution are required in dynamically varying scenarios. Proposals on MCS selection techniques aim at improving the energy efficiency and SE of the system by choosing dynamically the MCS according to specific criteria, like the users SNR and BLER. This translates into more efficient scheduling and resource allocation mechanisms, more robust towards delay constraints. Nevertheless, the level of integration of these techniques with the MBSFN standards seems to be still a work in progress. As an example, the impact of MIMO and multiple antenna techniques on the overall performance is not properly taken into account. Another drawback, that seems still unsolved, is the negative effect due to the frequent MCS switching, occurring whenever the movement of the worst-SINR user in the MBSFN border area forces the BS to constantly change the MCS.

F. Lessons learned on joint unicast/multicast transmission

The advantages of the simultaneous utilization of unicast and multicast transmission is twofold. First, combining unicast and multicast allows to support many more mobile users. Second, it allows optimizing the power consumption, a very important advantage for battery-powered mobile devices. So, the combined unicast/multicast transmission can be a gain in terms of both network load and energy saving: unicast transmission increases the amount of saved energy, because of the choice of the highest MCS allowed by UEs channel conditions, and multicast transmission avoids multiple transmissions of the same service, reducing the network load. To exploit this technique, efficient scheduling techniques have been studied in literature, but the related resource allocation problem can be solved at best only through solvers that are computationally costly, especially in the case of on-demand services.

G. Lessons learned on performance evaluation of LTE multicasting

There is a consistent amount of works that analyze the performance of LTE systems under different aspects, ranging from the joint transmission of unicast and multicast services, to LTE multicasting for different network topologies, SE, and transmission bearer mechanisms. These studies are useful because they provide a detailed analysis, generally performed through simulations, on the most relevant aspects of LTE and LTE-A multicasting, at the same time taking into account 3GPP standardization. This anyway comes at the cost of an unavoidable simplification of the considered simulation scenarios, for example, considering a limited number of BSs, eMBMS services, and/or UEs, and a consequent simplification of the eMBMS simulation environment. Also the works analyzing different network topologies must face the relevant issue of the switching mechanisms and additional and more complex interfaces needed for the different parts of the network to interact. This is the case, for example, of network topologies based on cooperative resource relaying, or integrated topologies where the LTE system interacts with other networks (like WiMAX, or cooperative vehicular networks).

H. Lessons learned on analytical models

Some works propose analytical models to examine several aspects of MBSFN multicasting, such as bandwidth estimation, coverage and capacity, BER in MIMO systems, blocking probability, D2D transmission, and multicasted video traffic. Modeling can be very useful to understand the behaviour of different multicast metrics, and design LTE systems efficiently. Nevertheless, a model implies always a compromise between the approximations introduced to make the model analytically tractable, and its complexity. In the first case, some aspects of MBMS and MBSFN are not properly taken into account, and this can bring to a (even partially) wrong evaluations of the metrics under analysis. Vice versa, a very accurate model is also analytically very complex and computationally expensive, and cannot be suitable for real-time computations. Finding the best trade-off between this opposite aspects is challenging. Another drawback is the lack of comparison with real or even simulated MBSFN scenarios. In few works a comparison can be found with simulated scenarios or with other models, but this solution can anyway limit the accuracy of the model if applied on real scenarios.

I. Lessons learned on error correction techniques

In the recent literature, the study of broadcast/multicast data correction (through FEC mechanisms at application and/or MAC layers) and retransmission techniques has attracted much attention. The use of coding techniques at application layer is advantageous because erroneous packets can be recovered also for all underlying layers and protocols, adding functionalities of end-to-end error correction without any required change in the standards. The main problem of the FEC techniques is the transmission overhead that brings to a waste of bandwidth and affects also power consumption and delay. Enhancements have been proposed, aiming to minimize the overhead, but they introduce complications in the encoding procedures. The HARQ schema partially mitigate the computation overhead of FEC-encoded packets, but they suffer from the same overhead problems of the FEC techniques. There are also some works considering the joint adoption of coding and HARQ techniques. Nevertheless, they are managed in different entities (BM-SC and BS, respectively), and the lack of coordination among them can strongly degrade system performance, increasing delay.

J. Lessons learned on resource allocation strategies

Resource allocation strategies are present in the recent literature to deliver eMBMS services in LTE and LTE-A systems. An efficient radio resource allocation can be obtained through scheduling, Opportunistic Multicast Scheduling (OMS) and CMS schema, the adoption of RLNC techniques at MAC layer, etc. Each of the solutions proposed anyway presents weak points. An efficient scheduling technique saves resources but is not a trivial task, as testified by literature. CMS approaches are intrinsically fair, since they guarantee the same throughput to all users, but at the transmission conditions of the user with the worst channel quality, with consequent degradation of the system throughput. On the contrary, OMS schema aim at maximizing the total throughput by serving only users with better channel conditions, to the detriment of the others; so the fairness requirement cannot be satisfied in general. Furthermore, resource allocation is intrinsically a very complex issue, given the plurality of aspects to be optimized (the QoS of the served UEs, bandwidth resources, and energy efficiency).

K. Lessons learned on LTE physical layer for multicasting

The approaches proposed at physical layer mainly concern the use of multicast beamforming, MIMO techniques, and content synchronization. The joint optimization of the multicast beamforming and antenna selection is useful to increase the SE and reduce the power consumption. Nevertheless, the solution to this problem requires a very high computational effort. Also MIMO schema are promising methods for power optimization in LTE systems. Anyway, in MIMO systems the accuracy of the channel estimates is a very important issue for a satisfactory system performance. Content synchronization is another important issue faced in some works, since physical symbols should be temporally aligned among all the BSs of a MBSFN area. This, in turn, translates into the same configuration of the RLC/MAC/PHY layers (assuming the same eMBMS service) at BSs side. To this end, enhancements have been proposed to the SYNC protocol, defined in 3GPP specifications and guaranteeing the ordered delivery of MBMS data from the BM-SC to the BS. Nevertheless, the proposed SYNC modifications have not been validated through real or simulated scenarios.

L. Lessons learned on heterogeneous network architectures

Several approaches have been proposed for the cooperation and interworking of heterogeneous networks for multicast transmission. The application fields can be basically identified in videoconferencing, TV and video streaming, emergency messages, MME signaling, and alert messages dissemination. The majority of the works analyze distributed systems for data multicasting. Heterogeneous architectures are a good solution to save resources and increase the transmission performance. But there are some drawbacks. First, the gateway nodes acting as intermediaries between the different parts of the heterogeneous systems are more difficult to manage and set-up. Second, multicasting can introduce a significant overhead, or is not so robust towards signal interferences. Third, changes in network topology due to users mobility heavily affect the system performance. Proposals aiming at an integration of MBMS with digital video distribution systems, like DVB-H or TV broadcasting, in a converged infrastructure can increase the degree of robustness of video delivery, but with an increased communication cost (in terms of synchronization and signaling flow). In the recent literature, converged network architectures have been proposed to avoid service interruption and increase the SE of multicast distribution. These kinds of architectures mainly suffer from scalability problems and, like in the other cases, a higher complexity due to coordination among the various networks entities.

M. Lessons learned on single cell PTM multicasting

PTM transmission is performed in a consistent number of works, and often compared with PtP and MBSFN. As known, PTM mode, performed on a single cell (SC-PTM), allows service multicasting through a common channel, thus minimizing network resources. A group of MBMS users listens to a common channel, with the same MCS, and shares the same time/frequency resources. All the functionalities such as packet scheduling, link adaptation, AMC and HARQ are performed with respect to each group of eMBMS subscribers in the cell range, with a scalability improvement. If compared to MBSFN, SC-PTM is also less costly, since cells can multicast or broadcast the media to all UEs, without requiring a tight synchronization with adjacent cells to improve the signal quality in the overlapping areas. On the contrary, SC-PTM performance, especially in terms of SINR and throughput, is still restricted with respect to MBSFN, mainly because of signal interferences in the overlapping regions and performance degradation of the UE when the distance from the BS increases.

X. Challenges and future research directions

LTE and its evolutions, based on newer releases of the 3GPP standard, is a very complex system, because of its ambitious goals, just to cite some of them: high throughput, mitigation of interference in the wireless channel, spectrum and resource optimization, scheduling and retransmission, power control, energy saving, device cooperation, security, etc [154]. This reflects also on multicasting and broadcasting of services, where there are several issues that need to be addressed, and that determine the possible directions of future research on this subject. The description that follows does not pretend to be exhaustive, but rather to give a comprehensive idea of the most important issues on LTE multicasting that still need to be developed in the opinion of the authors.

A first important issue can be found in the absence of feedbacks and retransmissions in MBMS, that paves the way to the study and implementation of FEC schema and coding techniques. Research efforts could be directed at reducing as much as possible the more critical issues related to the coding procedures, i.e., the bandwidth overhead due to the error protection procedures and the delay needed to recover from lost packets; all this, at the same time guaranteeing an acceptable degree of robustness towards packet loss for QoS/QoE purposes. A first solution to this issue can be a strategy for the dynamic selection of the most suitable coding technique (i.e., the number of redundant packets to be generated to obtain a given packet error rate), depending on the network conditions. Another possible solution is to combine FEC schema with cooperation/relaying techniques and/or other types of network architectures like P2P and mesh networks, to reduce the number of coded packets transmitted at source side at the same time keeping the same number of redundant packets received to save network resources. Also the utilization of MIMO techniques in combination with error correction schema could be useful to improve the system performance.

Cooperative networks are another interesting research theme for LTE multicasting, as testified by the works found in literature. Relay nodes can be of great help to offload the traffic at BSs, increase throughput and coverage, and reduce interference. Future research directions could consider more accurate and sophisticated relaying schema, where BSs can schedule multicast data to relay nodes, to achieve better coordination and reduced interference among D2D nodes. The joint use of MIMO and scheduling techniques is another interesting research topic in the cooperative network scenario, because it can intuitively increase the system performance, even if it complicates the modeling, the analysis and the implementation of the MBMS scheme.

At physical layer, MIMO is a value-added feature in MBMS, because it increases the data rate and reduces interferences through beamforming techniques. Issues related to this topic are the choice of the optimal antenna selection and the shape of beamformers signals. In this context, the optimal choice of the number and position of the antennas is crucial to exploit at best MIMO features. It would be interesting to extend the existing studies to more complicated scenarios, where different services are multicasted to different multicast groups, at the same time keeping a manageable computational complexity of the proposed optimization algorithms.

The MBMS architecture, the main topic of works on LTE multicasting, has been discussed in many works. The analysis of the main logical components of the MBMS architecture has been considered in some works, also evaluating the possibility of exploiting this architecture in different application scenarios, like TV broadcasting, Intelligent Transportation Systems, vehicular networks, etc. Nevertheless, there are some interesting research possibilities in the joint analysis of the mutual interaction among all the components of the MBMS architecture, to improve delivery of multicast services. A detailed analysis of some specific aspects like end-to-end delay and users mobility, that are critical in some emergency scenarios but not widely treated in the recent literature, would be useful in this context. Also theoretical models on these aspects have not been developed but can be of interest to analyze the behaviour of some performance metrics.

The coexistence of different types of networks architectures including MBMS has been discussed in the recent literature, but research on this topic could focus on a more detailed study of coordination capabilities among network architectures, through a cross-layer cooperation among the different layers of the protocol stacks. To this extent, control plane procedures assume a high importance, even if this aspect has been almost neglected in the recent literature. Concurrently, another aspect to take into account in this research field is how to set-up, improve, or modify, the interface specifications of the networks involved in the hybrid architecture, to allow a better adaptation among different protocols of the stacks. Novel hybrid architectures could also be considered, to mitigate some inefficiencies of the wireless transmission. Just to give a possible research direction on this topic, the spectrum scarcity is a factual reality, and hybrid networks that consider the use of CRNs together with MBMS would be of help to this effect. Another interesting research direction is the analysis of the cost of the coexistence among hybrid architectures, in terms of communication metrics (delay, throughput etc.), implementation and interfacing of the different architectures. This aspect has not been properly taken into account in the surveyed literature.

Scheduling is a widely treated topic in the recent literature. The most critical issue in this context are the optimization algorithms adopted, that aim at finding the best solution in a multidimensional search space. Furthermore, the large majority of the proposed solutions perform scheduling in the time domain. Much less efforts have been made to schedulers that act in the frequency domain, which can be instead an interesting research topic due to the intrinsic frequency-selective nature of the channel. An even more interesting effort can be made by developing schedulers that act in both time and frequency domains, to further increase (but complicate) the optimization process of resource allocation.

Coordination among cells to increase SINR and data rate is an open issue for MBSFN. This aspect is important because it can increase throughput at cell edge, the network coverage, and the suppression of interference among adjacent cells in the MBSFN area. All these aspects require further research efforts, translating into strategies and algorithms aiming to reach this goal, at the same time taking into account all the control plane procedures that can guarantee the cell synchronization in the MBSFN area. In this context, research should also move towards an efficient selection of the BSs that can contribute to the creation of a MBSFN area, depending on their mutual position, distance, and transmission power, especially for coverage maximization and interference minimization purposes; to the best of the authors knowledge, this aspect has not been properly taken into account in the works on LTE multicasting.

Secure multicast communication is another research issue in LTE multicast transmission. The main focus of this topic should be the guarantee of confidentiality among authenticated users in multicast group. Few efforts have been made in this direction; so, future studies should point on advanced authorization and encryption mechanisms for multicasting of reserved data. Another interesting research direction could consider anomaly detection algorithms that prevent from Denial of Service (DoS) attacks and unauthorized access to multicast groups.

There is a strong need to use sophisticated simulation/emulation tools that take into account as much as possible all the layers of the protocol stack and their mutual interactions (through cross-layer analysis), to increase the simulation accuracy. A research effort towards this direction could be very appreciated to provide more realistic results in different practical scenarios. To this end, the tools developed should take into account several aspects, i.e., the transmission of different services in a MBSFN area, users mobility and service continuity, and the terminals energy consumption, which have been almost neglected in all the works on this topic. Higher layers parameters should also be considered as simulation outputs, to assess the quality of the received data in application scenarios like video multicasting.

Optimization algorithms have been widely used in LTE multicasting to find the solution to several problems, especially in the context of novel scheduling techniques, SE optimization, and subgroup formation strategies. The main issue is that the problems very often require the optimization of more variables that jointly contribute to the optimization function, and the ESS schema used to find the optimal solution can be computationally very expensive. Research directions in this regard aim to find the approximations in the optimization algorithms that allow the best trade-off between computational complexity and the closeness of the suboptimal solution to the globally optimal one. This task is not trivial because the accuracy of the search algorithm chosen depends on several factors, i.e., the number of variables, the structure of the optimization function, the limits of the search space (if any), etc. and the low complexity of the algorithm is a key-feature to find the solution in real-time, which is a desirable property in practical scenarios.

Transmission of compressed video, including SVC, is the most widely used application scenario for service multicasting in LTE. Novel proposals and algorithms should be evaluated explicitly in this context. Even if some research efforts have already been made in this direction, performance of video multicast in MBMS networks should be evaluated by considering more video quality metrics, like Structural Similarity Index Metric (SSIM), the Video Quality Metric (VQM), or the Perceptual Evaluation of Video Quality (PEVQ), that are more difficult to evaluate and require more computational efforts than the mostly used PSNR, but provide a more exhaustive evaluation of the perceived video quality.

As regards SC-PTM, much less research effort has been spent if compared to MBSFN, but almost all the issues addressed above remain still valid also for single cell multicasting. In addition, a critical issue in this context is an analysis of the inter-cell interference, that worsens the received multicast signal and the overall transmission performance. Novel error protection and retransmission strategies and detailed analytical models of transmission aspects that take into account this further complication are surely of particular interest in this research field.

XI. Conclusions

In this paper, a comprehensive survey is proosed on the multicast and broadcast communications strategies over last-generation mobile networks. A detailed classification of all the scientific contributions on this issue has been performed, grouping them into two main categories: the first focusing on MBMS, and the second on SC-PTM. For each of the two categories, a number of sub-categories has been introduced to efficiently group together works that develop similar methodologies and/or reach equivalent goals. The resulting study offers a clear and detailed overview on the state-of-the-art of network architectures, communication protocols, transmission strategies, and algorithms able to improve the performance of multicast and broadcast communications over mobile radio systems. This study could be fruitfully exploited by all researchers working on this topic, providing useful indications on directions which the new 5G systems should follow.

For what concerns MBMS and MBSFN, the conclusions that can be drawn from the analysis of the proposed contributions highlight several trade-offs. Optimizations performed on spectrum and energy/power saving are counterbalanced by an increased computational complexity, as for subgrouping, beamforming, and combined unicast/multicast transmission scenarios. The increased flexibility, reduced implementation cost and increased robustness of the cooperative strategies are counterbalanced by an increased complexity and overhead due to coordination among different entities/networks, and a higher transmission delay. The same conclusion holds for error correction and retransmission techniques: the increased transmission robustness comes at a cost of an increased overhead, that affects also power consumption and delay. All the contributions focusing on resource allocation strategies highlight the high complexity of this issue, due to different aspects that have to be taken into account (QoS, efficient utilization of bandwidth resources, SE, and energy efficiency). Analytical models are of great help in studying the behaviour of different multicast metrics and optimally design LTE systems, but the unavoidable simplifications introduced to reduce the model complexity provide only an approximated description of the system aspect under analysis. The same holds for all the works simulating different aspects of 3GPP specifications; furthermore, they do not introduce any substantial degree of novelty with respect to the analyzed standards. As regards SC-PTM, the works proposed highlight that SC-PTM can minimize the usage of network resources, improve scalability, and is less costly, but introduces destructive interferences and worsens SINR and throughput.

References

[1] E. Dahlman, G. Mildh, S. Parkvall, J. Peisa, J. Sachs, Y. Selén, and J. Sköld, "5G Wireless Access: Requirements and Realization," *IEEE Communications Magazine*, vol. 52, no. 12, pp. 42–47, 2014.

[2] A. Gupta and R. Jha, "A Survey of 5G Network: Architecture and Emerging Technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.

[3] Z. Kaleem, Y. Li, and K. Chang, "Architecture and Features for 5G Mobile Personal Cell," in *Proc. of International Conference on Information and Communication Technology Convergence (ICTC)*, October 2015, pp. 389–393.

[4] D. Sorbara, M. Caretti, M. Schubert, and A. Georgakopoulos, "Use cases, KPIs and requirements," FANTASTIC-5G project, IR IR2.1, 2015. [Online]. Available: http://fantastic5g.eu/wp-content/uploads/2016/01/FANTASTIC-5G_IR2-1_final.pdf

[5] O. Awobuluyi, J. Nightingale, Q. Wang, and J. Alcaraz-Calero, "Video Quality in 5G Networks: Context-Aware QoE Management in the SDN Control Plane," in *Proc. of IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, October 2015, pp. 1657–1662.

[6] Next Generation Mobile Networks, *NGMN 5G White Paper*, February 2015.

[7] 3GPP, "Multimedia Broadcast/Multicast Service (MBMS); Architecture and functional description," 3rd Generation Partnership Project (3GPP), TS 23.246, Dec. 2017.

[8] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2," 3rd Generation Partnership Project (3GPP), TS 36.300, June 2009.

[9] 3GPP, "Support of single-cell point-to-multipoint transmission in LTE," 3rd Generation Partnership Project (3GPP), RP 151110, Jun. 2015.

[10] X. Zhu, D. Yang, J. Wang, and X. Zhang, "Small area broadcast in LTE," in *Proc. of IEEE International Conference on Wireless Communications & Signal Processing (WCSP)*. IEEE, 2015, pp. 1–5.

[11] F. Wibowo, A. Bangun, A. Kurniawan, and H. Hendrawan, "Multimedia broadcast multicast service over single frequency network (MBSFN) in LTE based femtocell," in *2011 International Conference on Electrical Engineering and Informatics*, Jul 2011, pp. 1–5.

[12] N.-D. Nguyen and C. Bonnet, "Service continuity for eMBMS in LTE/LTE-Advanced network: Standard analysis and supplement," in *Proc. IEEE 11th Consumer Communications and Networking Conference (CCNC)*, Jan 2014, pp. 219–224.

[13] A. de la Fuente Iglesias, R. P. Leal, and A. G. Armada, "Performance analysis of eMBMS in LTE: Dynamic MBSFN areas," in *Proc. OPNETWORK 2013*, Aug 2013, pp. 1–6.

[14] S. Mitrofanov, A. Anisimov, and A. Turlikov, "eMBMS LTE usage to deliver mobile data," in *Proc. 2014 6th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, Oct 2014, pp. 60–65.

[15] P. Gomes, J. Seguro, and A. Correia, "MIMO coordination for E-MBMS transmissions in LTE-A," in *Proc. 2010 6th International Conference on Wireless and Mobile Communications (ICWMC)*, Sep 2010, pp. 291–296.

[16] A. Alexiou, C. Bouras, V. Kokkinos, A. Papazois, and G. Tseliou, "Optimizing the combination of MBSFN and PTM transmissions in LTE systems," in *Proc. 2011 Wireless Telecommunications Symposium (WTS)*, Apr 2011, pp. 1–7.

[17] I. Widjaja, P. Bosch, and H. L. Roche, "Comparison of MME signaling loads for long-term-evolution architectures," in *Proc. 2009 IEEE 70th Vehicular Technology Conference Fall (VTC 2009-Fall)*, Sep 2009, pp. 1–5.

[18] A. Urie, A. Rudrapatna, C. Raman, and J.-M. Hanriot, "Evolved multimedia broadcast multicast service in LTE: An assessment of system performance under realistic radio network engineering conditions," *Bell Labs Technical Journal*, vol. 18, no. 2, pp. 57–76, 2013.

[19] J. Calabuig, J. Monserrat, D. Gozalvez, and O. Klemp, "Safety on the roads: LTE alternatives for sending ITS messages," *IEEE Vehicular Technology Magazine*, vol. 9, no. 4, pp. 61–70, 2014.

[20] E. Yaacoub and O. Kubbar, "On the performance of distributed base stations in LTE public safety networks," in *Proc. 2012 8th International Wireless Communications and Mobile Computing Conference (IWCMC)*, Aug 2012, pp. 927–932.

[21] O. Oyman, J. Foerster, Y.-J. Tcha, and S.-C. Lee, "Toward enhanced mobile video services over WiMAX and LTE," *IEEE Communications Magazine*, vol. 48, no. 8, pp. 68–76, 2010.

[22] J. V. de Beek, J. Riihijarvi, and P. Mahonen, "Intrinsic challenges (and opportunities) to deploy LTE in europe's TV white spaces," in *Proc. 2012 Future Network & Mobile Summit (FutureNetw)*, Jul 2012, pp. 1–8.

[23] O. Bataa, O. Chuluun, T. Orosoo, E. Lamjav, Y.-I. Kim, and K. Gonchigsumlaa, "A functional design of BM-SC to support mobile IPTV in LTE network," in *Proc. 2012 7th International Forum on Strategic Technology (IFOST)*, Sep 2012, pp. 1–5.

[24] J. Huschke, J. Sachs, K. Balachandran, and J. Karlsson, "Spectrum requirements for TV broadcast services using cellular transmitters," in *Proc. 2011 IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN 2011)*, May 2011, pp. 22–31.

[25] T. Yamada and T. Nishimura, "Possibility of a new type of wireless network: Connected small cells with 3G-LTE

signaling primitives and PON for the backhaul," in *Proc. 2014 16th International Telecommunications Network Strategy and Planning Symposium (Networks)*, Sep 2014, pp. 1–6.

[26] Y. Yang, P. Wang, C. Wang, and F. Liu, "An eMBMS based congestion control scheme in cellular-VANET heterogeneous networks," in *Proc. 2014 IEEE 17th International Conference on Intelligent Transportation Systems (ITSC)*, Oct 2014, pp. 1–5.

[27] R. Sivaraj, A. Gopalakrishna, M. Chandra, and P. Balamuralidhar, "QoS-enabled group communication in integrated VANET-LTE heterogeneous wireless networks," in *Proc. 2011 IEEE 7th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2011)*, Oct 2011, pp. 17–24.

[28] M.-H. Park, G.-P. Gwon, S.-W. Seo, and H.-Y. Jeong, "RSU-based distributed key management (RDKM) for secure vehicular multicast communications," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 3, pp. 644–658, 2011.

[29] D. Lee, W. Kim, B. Bae, H. Lim, and J. So, "Converged architecture for broadcast and multicast services in heterogeneous network," in *Proc. 2014 16th International Conference on Advanced Communication Technology (ICACT)*, Feb 2014, pp. 141–145.

[30] D. Camara, C. Bonnet, N. Nikaein, and M. Wetterwald, "Multicast and virtual road side units for multi technology alert messages dissemination," in *Proc. 2011 IEEE 8th International Conference on Mobile Adhoc and Sensor Systems (MASS)*, Jun 2012, pp. 947–952.

[31] J. Song and R. Phung, "Emergency group call over eMBMS," in *Proc. 16th International Conference on Advanced Communication Technology (ICACT)*, Feb 2014, pp. 1017–1022.

[32] B. Damera and P. Babu, "Broadcast the user chosen content over LTE in a simplified approach using evolved multimedia broadcast multicast service (eMBMS) with optimized MCE scheduling algorithm," in *Proc. 2014 International Conference on Intelligent Computing Applications (ICICA)*, Mar 2014, pp. 90–94.

[33] A.-L. Tien, H. Nguyen, and N. Crespi, "IMS-based distributed multimedia conferencing service for LTE," in *Proc. 2012 IEEE Wireless Communications and Networking Conference (WCNC)*, Apr 2012, pp. 3210–3214.

[34] I. Rubin, H.-B. Chang, and R. Cohen, "Robust multicasting in micro base station aided wireless cellular networks," in *Proc. 2012 IEEE Global Communications Conference (GLOBECOM 2012)*, Dec 2012, pp. 5464–5469.

[35] L. Militano, M. Condoluci, G. Araniti, A. Molinaro, A. Iera, and F. Fitzek, "Wi-fi cooperation or D2D-based multicast content distribution in LTE-A: A comparative analysis," in *Proc. 2014 IEEE International Conference on Communications Workshops (ICC)*, Jun 2014, pp. 296–301.

[36] E. Reetz, R. Hockmann, and R. Tonjes, "Performance evaluation of topologies for cooperative and dynamic resource relaying," in *Proc. 2008 14th European Wireless Conference (EW 2008)*, Jun 2008, pp. 1–6.

[37] M. Wetterwald, T. Buburuzan, and G. Carneiro, "Combining MBMS and IEEE 802.21 for on-the-road emergency," in *Proc. 2008 8th International Conference on ITS Telecommunications (ITST 2008)*, Oct 2008, pp. 434–438.

[38] M. Wetterwald, "A case for using MBMS in geographical networking," in *Proc. 2009 9th International Conference on Intelligent Transport Systems Telecommunications,(ITST)*, Oct 2009, pp. 309–313.

[39] E. Yaacoub and O. Kubbar, "Energy-efficient device-to-device communications in LTE public safety networks," in *Proc. 2012 IEEE Globecom Workshops*, Dec 2012, pp. 391–395.

[40] E. Yaacoub, "On the use of device-to-device communications for QoS and data rate enhancement in LTE public safety networks," in *Proc. 2014 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, Apr 2014, pp. 236–241.

[41] E. Yaacoub, Z. Dawy, and A. Abu-Dayya, “On real-time video streaming over LTE networks with mobile-to-mobile cooperation,” in *Proc. 2012 19th International Conference on Telecommunications (ICT)*, Apr 2012, pp. 1–6.

[42] M. Condoluci, L. Militano, G. Araniti, A. Molinaro, and A. Iera, “Multicasting in LTE-A networks enhanced by device-to-device communications,” in *Proc. 2013 IEEE Globecom Workshops (GC Wkshps)*, Dec 2013, pp. 567–572.

[43] R. Atat, E. Yaacoub, M.-S. Alouini, and F. Filali, “Delay efficient cooperation in public safety vehicular networks using LTE and IEEE 802.11p,” in *Proc. 2012 IEEE Consumer Communications and Networking Conference (CCNC)*, Jan 2012, pp. 316–320.

[44] J. Calabuig, J. Monserrat, D. Gozavez, and D. Gomez-Barquero, “AL-FEC for streaming services in LTE E-MBMS,” *2013 EURASIP Journal on Wireless Communications and Networking*, vol. 2013:73, 2013.

[45] C. Bouras, N. Kanakis, V. Kokkinos, and A. Papazois, “AL-FEC for streaming services over LTE systems,” in *Proc. 14th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Oct 2011, pp. 1–5.

[46] U. Kumar and O. Oyman, “QoE evaluation for video streaming over eMBMS,” in *Proc. 2013 International Conference on Computing, Networking and Communications (ICNC 2013)*, Jan 2013, pp. 555–559.

[47] E. Kurniawan, K. Chong, S. Sun, and K. Yen, “Application of FASTAR code in multimedia broadcast multicast service,” in *Proc. 2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, May 2011, pp. 1–5.

[48] A. Alexiou, K. A. C. Bouras, V. Kokkinos, and A. P. G. Tseliou, “Reliable multicasting over LTE: A performance study,” in *Proc. 2011 IEEE Symposium on Computers and Communications (ISCC)*, Jul 2011, pp. 603–608.

[49] N. H. Marcano, J. Heide, D. Lucani, and F. Fitzek, “On the throughput and energy benefits of network coded cooperation,” in *Proc. 2014 IEEE 3rd International Conference on Cloud Networking (CloudNet)*, Oct 2014, pp. 138–142.

[50] X. Sun, C. Wu, and Z. Zeng, “Dynamic power saving mechanism joint delivery of unicast and E-MBMS services in LTE MBSFN networks,” in *Proc. 2013 16th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Dec 2013, pp. 189–194.

[51] A. de la Fuente, A. G. Armada, and R. Leal, “Joint multicast/unicast scheduling with dynamic optimization for LTE multicast service,” in *Proc. 20th European Wireless Conference (EW2014)*, May 2014, pp. 1–6.

[52] M. Rahman, C.-H. Hsu, A. Hasib, and M. Hefeeda, “Hybrid multicast-unicast streaming over mobile networks,” in *Proc. 2014 IFIP Networking Conference*, Jun 2014, pp. 1–9.

[53] I. Rubin, H.-B. Chang, and R. Cohen, “Micro base station aided failover for multicast scheduling in wireless cellular networks,” *IEEE Transactions on Wireless Communications*, vol. 13, no. 9, pp. 5251–5263, 2014.

[54] L. Gallo and J. Harri, “Short paper: A LTE-direct broadcast mechanism for periodic vehicular safety communications,” in *Proc. 2013 IEEE Vehicular Networking Conference (VNC)*, Dec 2013, pp. 166–169.

[55] A. Alexiou, C. Bouras, V. Kokkinos, A. Papazois, and G. Tsichritzis, “Spectral efficiency performance of MBSFN-enabled LTE networks,” in *Proc. 2010 IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Oct 2010, pp. 361–367.

[56] C. Reis, A. Correia, N. Souto, and M. M. da Silva, “On enhanced multimedia broadcast multicast service for 4G,” in *Proc. 21st International Conference on Telecommunications (ICT)*, May 2014, pp. 186–190.

[57] A. Ibrahim and N. Seman, “The investigation of MBMS for delivering multimedia contents in single frequency network,” in *Proc. 2012 IEEE Student Conference on Research and Development (SCOReD 2012)*, Dec 2012, pp. 214–217.

[58] M. Condoluci, G. Araniti, A. Molinaro, A. Iera, and J. Cosmas, “On the impact of frequency selectivity on multicast subgroup formation in 4G networks,” in *Proc. 2013 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Jun 2013, pp. 1–6.

[59] A. Alexiou, C. Bouras, V. Kokkinos, and G. Tsichritzis, “Communication cost analysis of MBSFN in LTE,” in *Proc.*

2010 IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), Sep 2010, pp. 1366–1371.

[60] ——, “Performance evaluation of LTE for MBSFN transmissions,” *Wireless Networks*, vol. 18, no. 3, pp. 227–240, 2012.

[61] S. Lee, K. Cho, K. Kim, D. Yoon, and T. Kim, “Performance analysis of LTE-Advanced system in the downlink spatial channel model,” in *Proc. IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC 2009)*, Nov 2009, pp. 169–171.

[62] O. Bataa, E. Lamjav, S. Batsuuri, U. Purevdorj, C. Naimannaran, K. Young-Il, and K. Gonchigsumlaa, “Service control algorithm of providing efficient video-on-demand service using hybrid mechanism,” in *Proc. 2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet)*, Apr 2012, pp. 3507–3512.

[63] T. Liu, L. Li, J. Yang, and S. Leng, “Work in progress: An optimized transmission mode and switching mechanism towards MBMS business in long term evolution,” in *Proc. 2014 9th International Conference on Communications and Networking in China (CHINACOM)*, Aug 2014, pp. 513–516.

[64] N.-D. Nguyen, R. Knopp, N. Nikaein, and C. Bonnet, “Implementation and validation of multimedia broadcast multicast service for LTE/LTE-Advanced in openairinterface platform,” in *Proc. IEEE 38th Conference on Local Computer Networks Workshops (LCN Workshops)*, Oct 2013, pp. 70–76.

[65] P. McDonagh, C. Vallati, A. Pande, P. Mohapatra, P. Perry, and E. Mingozzi, “Investigation of scalable video delivery using H.264 SVC on an LTE network,” in *Proc. 2011 14th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Oct 2011, pp. 1–5.

[66] J. Huschke, “Facilitating convergence between broadcasting and mobile services using LTE networks,” in *Proc. 2011 Technical Symposium at ITU Telecom World (ITU WT)*, Oct 2011, pp. 123–128.

[67] M. Gruber and D. Zeller, “Multimedia broadcast/multicast service: new transmission schemes and related challenges,” *IEEE Communications Magazine*, vol. 49, no. 2, pp. 176–181, 2011.

[68] G. Walker, J. Wang, C. Lo, X. Zhang, and G. Bao, “Relationship between LTE broadcast/eMBMS and next generation broadcast television,” *IEEE Transactions on Broadcasting*, vol. 60, no. 2, pp. 185–192, 2014.

[69] D. Lecompte and F. Gabin, “Evolved multimedia broadcast/multicast service (eMBMS) in LTE-advanced: overview and rel-11 enhancements,” *IEEE Communications Magazine*, vol. 50, no. 11, pp. 68–74, 2012.

[70] A. Pande, V. Ahuja, R. Sivaraj, E. Baik, and P. Mohapatra, “Video delivery challenges and opportunities in 4G networks,” *IEEE MultiMedia*, vol. 20, no. 3, pp. 57–76, 2013.

[71] Y. Pointurier, B. UscumliÃš, I. Popescu, A. Gravey, Q. Wei, and M. Lott, “A comparison of subwavelength optical switching networks for LTE mobile backhauling,” in *Proc. 16th International Conference onTransparent Optical Networks (ICTON)*, Jul 2014, pp. 1–4.

[72] P.-H. Wu, J.-N. Hwang, J.-Y. Pyun, K.-M. Lan, and J.-R. Chen, “QoE-aware resource allocation for integrated surveillance system over 4G mobile networks,” in *Proc. 2012 IEEE International Symposium on Circuits and Systems (ISCAS 2012)*, May 2012, pp. 1103–1106.

[73] A. Roy, H. Kim, N. Saxena, and R. Kandoori, “LTE multicast communication for demand response in smart grids,” in *Proc. 2014 IEEE International Conference on Advanced Networks and Telecommuncations Systems (ANTS)*, Dec 2014, pp. 1–6.

[74] Q. Li, M. Yao, W. Wang, and X. Zhang, “Performance analysis of superposition coding for mixed unicast and broadcast transmission using opportunistic multicast scheduling scheme in LTE-A system,” in *Proc. IET International Conference on Communication Technology and Application (ICCTA 2011)*, Oct 2011, pp. 197–200.

[75] P. Chawla and B. Gupta, “BER analysis of single/multi-user LTE and LTE-A systems,” in *Proc. 2014 IEEE International Advance Computing Conference (IACC)*, Feb 2014, pp. 262–266.

[76] N. Souto, A. Correia, R. Dinis, J. Silva, and L. Abreu, “Multiresolution MBMS transmissions for MIMO UTRA LTE systems,” in *Proc. 2008 IEEE International Symposium onBroadband Multimedia Systems and Broadcasting*, Mar-Apr 2008, pp. 1–6.

[77] A. Tassi, D. Marabissi, R. Fantacci, D. D. Lorenzo, and M. Maischberger, “A novel convex power adaptation strategy for multicast communications using random linear network coding schemes,” in *Proc. 2012 IEEE International Conference on Communications (ICC)*, Jun 2012, pp. 5270–5274.

[78] D. Kim, T. Fujii, and K. Lee, “A power allocation algorithm for maximizing total utility over an MBSFN,” *IEEE Wireless Communications Letters*, vol. 2, no. 3, pp. 283–286, 2013.

[79] A. Alexiou, C. Bouras, V. Kokkinos, and E. Rekkas, “Optimal MBMS power allocation exploiting MIMO in LTE networks,” in *Proc. IEEE 69th Vehicular Technology Conference (VTC 2009)*, Apr 2009, pp. 1–5.

[80] K. Asimakis, C. Bouras, V. Kokkinos, and A. Papazois, “Mobility-aware power control in MBSFN,” *Telecommunication Systems*, pp. 1–15, 2014.

[81] O. Karimi and J. Liu, “Power efficient high quality multimedia multicast in LTE wireless networks,” in *Proc. 2011 IEEE 8th International Conference on Mobile Adhoc and Sensor Systems (MASS 2011)*, Oct 2011, pp. 161–163.

[82] O. Mehanna, N. Sidiropoulos, and G. Giannakis, “Multicast beamforming with antenna selection,” in *Proc. 2012 IEEE 13th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun 2012, pp. 70–74.

[83] ——, “Joint multicast beamforming and antenna selection,” *IEEE Transactions on Signal Processing*, vol. 61, no. 10, pp. 2660–2674, 2013.

[84] A. Abdelkader, I. Wajid, A. Gershman, and N. Sidiropoulos, “Transmit beamforming for wireless multicasting using channel orthogonalization and local refinement,” in *Proc. 2009 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2009)*, Apr 2009, pp. 2281–2284.

[85] A. Ibing, Y. Hadisusanto, and V. Jungnickel, “Scaleable network multicast for cooperative base stations,” in *Proc. 2008 3rd International Conference on Communication Systems Software and Middleware and Workshops (COMSWARE 2008)*, Jan 2008, pp. 753–756.

[86] I. Rubin, H.-B. Chang, and R. Cohen, “Optimal relay deployment for multihop multicast scheduling in cellular wireless networks,” in *Proc. 2012 IEEE Globecom Workshops*, Dec 2012, pp. 1166–1171.

[87] ——, “Multihop relay-aided multicast scheduling for cellular wireless networks,” in *Proc. 2012 IEEE Global Communications Conference (GLOBECOM)*, Dec 2012, pp. 1824–1829.

[88] X. Lin, R. Ratasuk, A. Ghosh, and J. Andrews, “Modeling, analysis, and optimization of multicast device-to-device transmissions,” *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4346–4359, 2014.

[89] P. Gong and J. Ren, “Energy-efficiency MBMS single frequency networks with user cooperation,” in *Proc. 2013 19th Asia-Pacific Conference on Communications (APCC)*, Aug 2013, pp. 315–320.

[90] M. Azman, N. Tuban, K. Noordin, and M. Ismail, “Genetic algorithm approach for solving radio resource allocation for overlapping MBSFN area,” in *Proc. 2011 International Conference on Information Technology and Multimedia (ICIM)*, Nov 2011, pp. 1–5.

[91] N. Tuban, M. Azman, K. Noordin, and M. Ismail, “Genetic algorithm approach for dynamic configuration of multicast broadcast single frequency network deployment in LTE,” in *Proc. 2011 International Conference on Information Technology and Multimedia (ICIM)*, Nov 2011, pp. 1–5.

[92] E. Matskani, N. Sidiropoulos, and L. Tassiulas, “On multicast beamforming and admission control for UMTS-LTE,” in

Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008), Mar-Apr 2008, pp. 2361–2364.

[93] E. Matskani, N. Sidiropoulos, Z.-Q. Luo, and L. Tassiulas, "Efficient batch and adaptive approximation algorithms for joint multicast beamforming and admission control," *IEEE Transactions on Signal Processing*, vol. 57, no. 12, pp. 4882–4894, 2009.

[94] L. Rong, O. B. Haddada, and S. Elayoubi, "Analytical analysis of the coverage of a MBSFN OFDMA network," in *Proc. IEEE 2008 Global Telecommunications Conference (GLOBECOM 2008)*, Nov-Dec 2008, pp. 1–5.

[95] Y. Ryu and J. Kim, "An improved cell search with SIC detection for 3GPP LTE system," in *Proc. 2011 6th International Conference on Computer Sciences and Convergence Information Technology (ICCIT)*, Nov-Dec 2011, pp. 582–585.

[96] R. Radhakrishnan, B. Tirouvengadam, and A. Nayak, "Channel quality-based AMC and smart scheduling scheme for SVC video transmission in LTE MBSFN networks," in *Proc. 2012 IEEE International Conference on Communications (ICC)*, Jun 2012, pp. 6514–6518.

[97] A. Alexiou, C. Bouras, V. Kokkinos, A. Papazois, and G. Tsichritzis, "Efficient MCS selection for MBSFN transmissions over LTE networks," in *Proc. Wireless Days (WD), 2010 IFIP*, Oct 2010, pp. 1–5.

[98] S. Bochrini, C. Bouras, and V. Kokkinos, "Efficient MCS selection mechanisms for multicasting over LTE networks," in *Proc. 2013 6th Joint IFIP Wireless and Mobile Networking Conference (WMNC)*, Apr 2013, pp. 1–8.

[99] K. Asimakis, C. Bouras, V. Kokkinos, and A. Papazois, "Genetic optimization for spectral efficient multicasting in LTE systems," in *Proc. 2012 5th Joint IFIP Wireless and Mobile Networking Conference (WMNC)*, Sep 2012, pp. 54–59.

[100] Y. Zhou, H. Liu, Z. Pan, L. Tian, and J. Shi, "Spectral- and energy-efficient two-stage cooperative multicast for LTE-advanced and beyond," *IEEE Wireless Communications*, vol. 21, no. 2, pp. 185–192, 2014.

[101] G. Araniti, M. Condoluci, A. Molinaro, A. Iera, and J. Cosmas, "Low complexity subgroup formation in LTE systems," in *Proc. 2013 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Jun 2013, pp. 1–6.

[102] G. Araniti, M. Condoluci, A. Iera, A. Molinaro, J. Cosmas, and M. Behjati, "A low-complexity resource allocation algorithm for multicast service delivery in OFDMA networks," *IEEE Transactions on Broadcasting*, vol. 60, no. 2, pp. 358–369, 2014.

[103] G. Araniti, M. Condoluci, M. Cotronei, A. Iera, and A. Molinaro, "A solution to the multicast subgroup formation problem in LTE systems," *IEEE Wireless Communications Letters*, vol. 4, no. 2, pp. 149–152, 2015.

[104] L. Militano, M. Condoluci, G. Araniti, and A. Iera, "Bargaining solutions for multicast subgroup formation in LTE," in *Proc. 2012 IEEE Vehicular Technology Conference (VTC Fall)*, Sep 2012, pp. 1–5.

[105] ——, "Multicast service delivery solutions in LTE-Advanced systems," in *Proc. 2013 IEEE International Conference on Communications (ICC)*, Jun 2013, pp. 5954–5958.

[106] L. Militano, D. Niyato, M. Condoluci, G. Araniti, A. Iera, and G. M. Bisci, "Radio resource management for group-oriented services in LTE-A," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 8, pp. 3725–3739, 2015.

[107] G. Araniti, V. Scordamaglia, M. Condoluci, A. Molinaro, and A. Iera, "Efficient frequency domain packet scheduler for point-to-multipoint transmissions in LTE networks," in *Proc. 2012 IEEE International Conference on Communications (ICC)*, Jun 2012, pp. 4405–4409.

[108] L. Militano, M. Condoluci, G. Araniti, A. Molinaro, A. Iera, and G.-M. Muntean, "Single frequency-based device-to-device-enhanced video delivery for evolved multimedia broadcast and multicast services," *IEEE Transactions on Broadcasting*, vol. 61, no. 2, pp. 263–278, 2015.

[109] C. Khirallah, D. Vukobratovic, and J. Thompson, "Bandwidth and energy efficiency of video broadcasting services over

LTE/LTE-A," in *Proc. 2013 IEEE Wireless Communications and Networking Conference (WCNC)*, Apr 2013, pp. 1909–1914.

[110] D. Parniewicz, M. Stasiak, and P. Zwierzykowski, "Model of the cellular network servicing multicast connections in iub and WCDMA resources," in *Proc. 2010 14th International Telecommunications Network Strategy and Planning Symposium (NETWORKS)*, Sep 2010, pp. 1–6.

[111] S.-S. Sun, Y.-C. Chen, and W. Liao, "Utility-based resource allocation for layer-encoded iptv multicast service in wireless relay networks," in *Proc. 2011 IEEE International Conference on Communications (ICC 2011)*, Jun 2011, pp. 1–5.

[112] Y.-C. Chen, D.-N. Yang, and W. Liao, "Efficient multi-view 3D video multicast with depth image-based rendering in LTE networks," in *Proc. 2013 IEEE Global Communications Conference (GLOBECOM)*, Dec 2013, pp. 4427–4433.

[113] J. Li, Z. Hu, and Y. Wang, "Compressed multicast retransmission in LTE-A eMBMS," in *Proc. 2010 IEEE 71st Vehicular Technology Conference (VTC 2010-Spring)*, May 2010, pp. 1–5.

[114] C.-C. Lu, W.-C. Chung, C.-J. Chang, and F.-C. Ren, "NACK-based retransmission schemes for MBMS over single frequency network in LTE," in *Proc. 2011 6th International ICST Conference on Communications and Networking in China (CHINACOM 2011)*, Aug 2011, pp. 284–288.

[115] H.-C. Lu and W. Liao, "Cooperative multicasting in network-coding enabled multi-rate wireless relay networks," in *Proc. IEEE INFOCOM 2012*, Mar 2012, pp. 334–342.

[116] Q. Zhang, J. Heide, M. Pedersen, and F. Fitzek, "MBMS with user cooperation and network coding," in *Proc. 2011 IEEE Global Telecommunications Conference (GLOBECOM 2011)*, Dec 2011, pp. 1–6.

[117] A. Laya, K. Wang, L. Alonso, and J. Alonso-Zarate, "Multi-radio cooperative retransmission scheme for reliable machine-to-machine multicast services," in *Proc. 2012 IEEE 23rd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, Sep 2012, pp. 1–6.

[118] H. Hamdoun, P. Loskot, and T. O'Farrell, "Implementation trade-offs of fountain codes in LTE and LTE-A," in *Proc. 2012 7th International ICST Conference on Communications and Networking in China (CHINACOM)*, Aug 2012, pp. 419–424.

[119] M. Gruber and D. Zeller, "Uniform MAC headers for synchronized MBMS transmission," in *Proc. 2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep 2009, pp. 2787–2791.

[120] R. Sachan, N. Saxena, and A. Roy, "An efficient hybrid scheduling scheme for impatience user in embms over LTE," in *Proc. 2013 International Conference on Computer Communication and Informatics (ICCCI)*, Jan 2013, pp. 1–5.

[121] C. Lou and L. Qiu, "QoS-aware scheduling and resource allocation for video streams in e-MBMS towards LTE-A system," in *Proc. 2011 IEEE Vehicular Technology Conference (VTC 2011)*, Sep 2011, pp. 1–5.

[122] C. Wu, X. Sun, and T. Zhang, "A power-saving scheduling algorithm for mixed multicast and unicast traffic in MBSFN," in *Proc. 2012 Computing, Communications and Applications Conference (ComComAp)*, Jan 2012, pp. 170–174.

[123] K. Sundaresan and S. Rangarajan, "Cooperation versus multiplexing: Multicast scheduling algorithms for OFDMA relay networks," *IEEE/ACM Transactions on Networking*, vol. 22, no. 3, pp. 756–769, 2014.

[124] J. Seppala, T. Koskela, T. Chen, and S. Hakola, "Network controlled device-to-device (D2D) and cluster multicast concept for LTE and LTE-A networks," in *Proc. 2011 IEEE Wireless Communications and Networking Conference (WCNC)*, Mar 2011, pp. 986–991.

[125] Y. Cai, S. Lu, L. Zhang, C. Wang, P. Skov, Z. He, and K. Niu, "Reduced feedback schemes for LTE MBMS," in *Proc. IEEE 69th Vehicular Technology Conference (VTC 2009)*, Apr 2009, pp. 1–5.

[126] A. Tassi, C. Khirallah, D. Vukobratovic, F. Chiti, J. Thompson, and R. Fantacci, "Resource allocation strategies for network-coded video broadcasting services over LTE-advanced," *IEEE Transactions on Vehicular Technology*, vol. 64,

no. 5, pp. 2186–2192, 2015.

[127] H. Wang, H. Vandervelde, and S. Kim, "LTE MBMS SYNC protocol for support synchronisation of content," in *Proc. 2009 IEEE International Conference on Communications Technology and Applications*, Oct 2009, pp. 392–395.

[128] Y. Chen, "A bandwidth estimation model for multiplexed E-MBMS services," in *Proc. IEEE 68th Vehicular Technology Conference (VTC 2008-Fall)*, Sep 2008, pp. 1–5.

[129] X. Wang and Y. Zhang, "Optimal video stream multiplexing in MBSFN," in *Proc. 2009 IEEE International Conference on Communications Technology and Applications (ICCTA '09)*, Oct 2009, pp. 360–365.

[130] Y. Chen, "Statistical multiplexing for LTE MBMS in dynamic service deployment," in *Proc. IEEE 2008 Vehicular Technology Conference (VTC 2008)*, May 2008, pp. 2805–2809.

[131] A. Russo and R. L. Cigno, "Pullcast: Peer-assisted video multicasting for wireless mesh networks," in *Proc. 2013 10th Annual Conference on Wireless On-demand Network Systems and Services (WONS)*, Mar 2013, pp. 60–67.

[132] A. Alexiou, C. Bouras, V. Kokkinos, A. Papazois, and G. Tseliou, "Enhancing FEC application in LTE cellular networks," in *Proc. 2010 IFIP Wireless Days (WD)*, Oct 2010, pp. 1–5.

[133] X. Wang, Y. Zhang, and S. Xiao, "How network coding helps eMBMS," in *Proc. IET International Conference on Communication Technology and Application (ICCTA 2011)*, Oct 2011, pp. 511–515.

[134] M. Han and J. Hong, "Balancing QoE and fairness of layered video multicast in LTE networks," in *Proc. 2012 14th International Conference on Advanced Communication Technology (ICACT)*, Feb 2012, pp. 387–392.

[135] C. Hellge, R. Skupin, J. Cho, T. Schierl, and T. Wiegand, "Capacity improvement in EMBMS using SVC and layer-aware bearer allocation," in *Proc. 2011 18th IEEE International Conference on Image Processing (ICIP)*, Sep 2011, pp. 933–936.

[136] M. Park, J. Lee, J. Kim, S. S. Newaz, and J. Choi, "Adaptive layered coding based energy efficient multicast transmission scheme over wireless access networks," in *Proc. 2014 International Conference on Information and Communication Technology Convergence (ICTC)*, Oct 2014, pp. 306–311.

[137] R. Radhakrishnan and A. Nayak, "An efficient video adaptation scheme for SVC transport over LTE networks," in *Proc. 2011 IEEE 17th International Conference on Parallel and Distributed Systems (ICPADS)*, Dec 2011, pp. 127–133.

[138] E. Baik, A. Pande, and P. Mohapatra, "Cross-layer coordination for efficient contents delivery in LTE eMBMS traffic," in *Proc. 2012 IEEE 9th International Conference on Mobile Adhoc and Sensor Systems (MASS)*, Oct 2012, pp. 398–406.

[139] X. Wang and Y. Zhang, "Efficient streaming delivery in eMBMS with HARQ and raptor," in *Proc. 2011 IEEE International Conference on Communications (ICC)*, Jun 2011, pp. 1–5.

[140] A. Tassi, G. Rigazzi, C. Khirallah, D. Vukobratovic, F. Chiti, J. Thompson, and R. Fantacci, "Reliable multicast in LTE-A: An energy efficient cross-layer application of network coding," in *Proc. 2013 20th International Conference on Telecommunications (ICT 2013)*, May 2013, pp. 1–5.

[141] J. Song, I. Singh, R. Phung, H. Kim, J. Lim, and D. Kim, "Relative time based MBSFN content synchronization," in *Proc. 2013 IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2013)*, Oct 2013, pp. 285–290.

[142] S. Figueiredo, M. Wetterwald, T.-T. Nguyen, L. Eznarriaga, N. Amram, and R. Aguiar, "SVC multicast video mobility support in MEDIEVAL project," in *Proc. 2012 Future Network & Mobile Summit (FutureNetw)*, Jul 2012, pp. 1–8.

[143] H. Xu, J. Jin, and B. Li, "YMMV: Multiple session multicast with MIMO," in *Proc. 2011 IEEE Global Telecommunications Conference (GLOBECOM 2011)*, Dec 2011, pp. 1–5.

[144] Y. Kong, Z.-H. Tan, S. Xu, and J. Li, "An adaptive MIMO scheme for E-MBMS in point-to-multipoint transmission

mode," in *Proc. 2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)*, Sep 2010, pp. 1–4.

[145] A. Alexiou, C. Bouras, and V. Kokkinos, "An enhanced MBMS power control mechanism towards long term evolution," in *Proc. 2009 Wireless Telecommunications Symposium (WTS 2009)*, Apr 2009, pp. 1–6.

[146] K. Ying, H. Yu, and H. Luo, "Inter-RAT energy saving for multicast services," *IEEE Communications Letters*, vol. 17, no. 5, pp. 900–903, 2013.

[147] R. Sivaraj, A. Pande, and P. Mohapatra, "Spectrum-aware radio resource management for scalable video multicast in LTE-advanced systems," in *Proc. 2013 IFIP Networking Conference*, May 2013, pp. 1–9.

[148] Y. Sheng, X. Guo, M. Peng, and W. Wang, "Analysis of novel user detection scheme based on polling for E-MBMS networks," in *Proc. IEEE 68th Vehicular Technology Conference, 2008 (VTC 2008-Fall)*, Sep 2008, pp. 1–5.

[149] S. Lu, Y. Cai, L. Zhang, J. Li, P. Skov, C. Wang, and Z. He, "Channel-aware frequency domain packet scheduling for MBMS in LTE," in *Proc. 2009 IEEE 69th Vehicular Technology Conference (VTC Spring 2009)*, Apr 2009, pp. 1–5.

[150] A. Tassi, I. Chatzigeorgiou, and D. Vukobratovic, "Resource-allocation frameworks for network-coded layered multimedia multicast services," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 2, pp. 141–155, 2015.

[151] J.-H. Na, D.-I. Kim, J. Moon, S. Lee, and Y. Shin, "A novel single-cell transmission scheme for multicast services in LTE-Advanced," in *Proc. 2014 International Conference on Information Networking (ICOIN)*, Feb 2014, pp. 186–191.

[152] L. Rong, S. Elayoubi, and O. Haddada, "Performance evaluation of cellular networks offering TV services," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 2, pp. 644–655, 2011.

[153] 3GPP, "Multimedia Broadcast/Multicast Service (MBMS); Stage 1," 3rd Generation Partnership Project (3GPP), TS 22.146, Mar. 2017.

[154] E. Dahlman, S. Parkvall, and J. Skold, *4G LTE/LTE-Advanced for Mobile Broadband*. Academic Press, 2011.

[155] W. Zirwas, "Single frequency network concepts for cellular OFDM radio systems," in *Proc. International OFDM Workshop*, Sep 2000.

[156] ITU-T, "Broadband optical access systems based on Passive Optical Networks (PON)," *Tech. Rep. G.983.1*, 2005.

[157] T. Orphanoudakis, E. Kosmatos, J. Angelopoulos, and A. Stavdas, "Exploiting PONs for mobile backhaul," *IEEE Communications Magazine*, vol. 51, no. 2, pp. S27–S34, 2013.

[158] 3GPP, "Study on single-cell point-to-multipoint transmission for E-UTRA," 3rd Generation Partnership Project (3GPP), TR 36.890, Jul. 2015.

[159] E. Dahlman, S. Parkvall, and J. Skold, *4G, LTE-Advanced Pro and The Road to 5G*. Elsevier, 2016.

[160] ISO/IEC/IEEE, "Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements - Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," ISO/IEC/IEEE, Tech. Rep. IEEE 8802-11, May 2018.

[161] IEEE, "Ieee standard for air interface for broadband wireless access systems," IEEE, Tech. Rep. IEEE Std 802.16-2017 (Revision of IEEE Std 802.16-2012), Mar 2018.

[162] R. O. Afolabi, A. Dadlani, and K. Kim, "Multicast Scheduling and Resource Allocation Algorithms for OFDMA-Based Systems: A Survey," *IEEE Communications Surveys Tutorials*, vol. 15, no. 1, pp. 240–254, 2013.

[163] J. M. Vella and S. Zammit, "A survey of multicasting over wireless access networks," *IEEE Communications Surveys Tutorials*, vol. 15, no. 2, pp. 718–753, 2013.

[164] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, 2005.

[165] Y. C. Liang, K. C. Chen, G. Y. Li, and P. Mahonen, “Cognitive radio networking and communications: an overview,” *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 3386–3407, 2011.

[166] C. Gao, Y. Shi, Y. T. Hou, H. D. Sherali, and H. Zhou, “Multicast communications in multi-hop cognitive radio networks,” *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, pp. 784–793, 2011.

[167] P. Polacek, C.-W. Huang, and J.-W. Chiang, “Unified opportunistic scheduling for layered multicast over cognitive radio networks,” *IEEE Transactions on Mobile Computing*, vol. 15, no. 10, pp. 2627–2639, 2015.

[168] T. Jiang, W. Xiang, H. H. Chen, and Q. Ni, “Multicast Broadcast Services Support in OFDMA-Based WiMAX Systems [Advances in Mobile Multimedia],” *IEEE Communications Magazine*, vol. 45, no. 8, pp. 78–86, 2007.

[169] G. Araniti, M. Condoluci, P. Scopelliti, A. Molinaro, and A. Iera, “Multicasting over Emerging 5G Networks: Challenges and Perspectives,” *IEEE Network*, vol. 31, no. 2, pp. 80–89, 2017.

[170] A. Al-Kandari, M. Al-Nasheet, and A. R. Abdulgafer, “WiMAX vs. LTE: an Analytic Comparison,” in *Proc. of 2014 Fourth International Conference on Digital Information and Communication Technology and it’s Applications (DICTAP)*, May 2014, pp. 164–166.

[171] 3GPP, “Technical Specification Group Services and System Aspects; Network architecture,” 3rd Generation Partnership Project (3GPP), TS 23.002, Mar. 2018.

[172] D. MacKay, “Fountain codes,” in *IEE Proceedings - Communications*, Dec 2005, pp. 1062–1068.

[173] A. Naeem, M. H. Rehmani, Y. Saleem, I. Rashid, and N. Crespi, “Network coding in cognitive radio networks: A comprehensive survey,” *IEEE Communications Surveys and Tutorials*, vol. 19, no. 3, pp. 1945–1973, 2017.

[174] A. Shokrollahi, “Raptor codes,” *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2551–2567, 2006.

[175] S. Pandi, F. Gabriel, J. A. Cabrera, S. Wunderlich, M. Reisslein, and F. H. P. Fitzek, “PACE: Redundancy Engineering in RLNC for Low-Latency Communication,” *IEEE Access*, vol. 5, pp. 20 477–20 493, 2017.

[176] S. Wunderlich, F. Gabriel, S. Pandi, F. H. P. Fitzek, and M. Reisslein, “Caterpillar RLNC (CRLNC): A Practical Finite Sliding Window RLNC Approach,” *IEEE Access*, vol. 5, pp. 20 183–20 197, 2017.

[177] T. Ho, M. Medard, R. Koetter, D. Karger, M. Effros, J. Shi, and B. Leong, “A random linear network coding approach to multicast,” *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4413–4430, 2006.

[178] R. K. Arora, *Optimization: Algorithms and Applications*. CRC Press, 2015.

[179] T. Weise, *Global Optimization Algorithms: Theory and Application*. http://www.it-weise.de, 2009.

[180] B. Gartner and J. Matousek, *Approximation Algorithms and Semidefinite Programming*. Springer, 2012.