DOCTORAL SCHOOL
UNIVERSITA' *MEDITERRANEA* DI REGGIO CALABRIA

DIPARTMENTO DI INGEGNERIA DELL'INFORMAZIONE, DELLE INFRASTRUTTURE E DELL'ENERGIA SOSTENIBILE (DIIES)

PHD IN
INFORMATION ENGINEERING

S.S.D. ING-INF/03
XXX CICLO

RRM APPROACHES IN MULTICAST TRANSMISSION FOR SUPPORTING EMERGING 5G SYSTEMS

CANDIDATE
Pasquale Scopelliti

ADVISOR
PROF. GIUSEPPE ARANITI

COORDINATOR
PROF. CLAUDIO DE CAPUA

REGGIO CALABRIA, APRIL 2018

Finito di stampare nel mese di **Aprile 2018**

Edizione Centro Stampa d'Ateneo

Quaderno N. 37
Collana *Quaderni del Dottorato di Ricerca in Ingegneria dell'Informazione*
Curatore *Prof. Claudio De Capua*

ISBN 978-88-99352-29-5

Università degli Studi *Mediterranea* di Reggio Calabria
Salita Melissari, Feo di Vito, 89124 Reggio Calabria

PASQUALE SCOPELLITI

RRM APPROACHES IN MULTICAST TRANSMISSION FOR SUPPORTING EMERGING 5G SYSTEMS

The Teaching Staff of the PhD course in
INFORMATION ENGINEERING
consists of:

Claudio DE CAPUA(coordinator)
Giovanni ANGIULLI
Pier Luigi ANTONUCCI
Giuseppe ARANITI
Francesco BUCCAFURRI
Rosario CARBONE
Riccardo CAROTENUTO
Salvatore COCO
Francesco DELLA CORTE
Fabio FILIANOTI
Sofia GIUFFRE';
Antonio IERA
Tommaso ISERNIA
Gianluca LAX
Aimè LAY EKUAKILLE
Giacomo MESSINA
Antonella MOLINARO
Andrea Francesco MORABITO
Rosario MORELLO
Fortunato PEZZIMENTI
Sandro RAO
Domenico ROSACI
Giuseppe RUGGERI
Francesco RUSSO
Valerio SCORDAMAGLIA
Domenico URSINO

Abstract

This Thesis focused on fifth generation (5G) mobile networks. In particular on the Mobile Video, an application which will reach the 80% of the total amount of mobile traffic within 2019. 5G system aims to manage the ever-growing traffic demand, the huge number of connected devices e to improve the performance of current cellular networks. An important solution to such problems is exploiting group-oriented communications, well known as Multicasting, which exploit point-to-multipoint communications, useful to efficiently manage the available radio resources.

This work focused on the multicast traffic management for the Mobile Video application and, in particular, on the Radio Resource Management (RRM).

The Thesis is divided into four different research contribution. In the first chapter, application scenarios of multicast communications have been described together with the respective enhancement needed for their implementation. RRM plays an important role in multicast communication, and it has been dealt with in the following three chapters.

In the second chapter, resource management issues were addressed in multicast communications in a Dense Heterogeneous network scenario, trying to balance customer services in poor channel conditions compared to those in good channel conditions. In the third chapter, the same issue was extended to Single Frequency Networks (SFNs), while in the fourth, radio resource management was dealt with in another perspective. Specifically, integrating into the LTE cellular system a non-orthogonal access technique (NOMA), typical of other broadcast systems.

Contents

List of Tables

List of Figures

Introduction

The increasing number of enhanced devices (e.g. smartphone, tablet, smart TV, etc.) encourages the demand for high-quality video content over mobile networks. Cisco analysis stated that mobile data traffic would reach 49 Exabyte per month by 2021, with a 7-fold increase from 2016 to 2021.

Mobile video applications (video downloading, video streaming, video conferences, concerts, etc.) became very popular in the last years. For this reason, group-oriented services and on demand services will play a key role in the future wireless networks. With the diffusion of such bandwidth-hungry applications, mobile networks tackle very challenging requirements for the emerging fifth generation (5G) networks, like high data rates and low latency.

Such expected growth in the mobile video demand over the broadband cellular networks is one of the key factors driving the wireless industry to develop fifth generation of network technology.

Indeed, forthcoming 5G networks aim to meet very ambitious target in terms of network performance and capacity. The objective of the 5G systems is to provide even higher data rates (to the order of 10 Gbps), reduced latency (in the order of few milliseconds), mobility support at very high speed (~ 500 Km/h) and augmented capacity.

In order to continue to increase system capacity and improve performance, research on 5G system pushes towards deployment within same coverage area of several cell layers (i.e. macro, micro, pico, and femto), diverse Radio Access Technologies RAT) (e.g. GSM, UMTS, LTE, WiFi) and multiple Point-to-Point (PtP) user links (e.g. Device-to-Device communication, mmWave).

The deployment of several small cells within a macrocell area served by a Base Station (BS) can thus provide an improved coverage (either indoor or in the coverage holes, for example) and an increase in the system capacity through the offloading of some of the macrocell's traffic. Furthermore, edge-cell users connected to a femtocell should benefit from a higher data rate, low latency, and improved levels of Quality of Service (QoS) and Quality of Experience (QoE). Nevertheless, this huge growth of different cell deployment leads to high densification of networks and creation of the so-called Dense Heterogeneous Network (DenseNet). The existence of DenseNets means that a very high number of users are inside the same coverage area and could connect to several different individual networks via diverse Access Points (AP) or BS. Nevertheless, the increase of network complexity introduces challenges about interference coordination, radio resource management (RRM), and user mobility. The former is exacerbated in such kind of networks, because many BSs with low power and low coverage are deployed within the same area, hence users pass through the coverage of different BSs/APs very frequently.

Furthermore, current mobile networks based on point-to-point (PtP) transmission limit the network capacity especially when many users access to the same video content. This scenario is fuelling the need for group-oriented services (i.e., multicast and broadcast) in order to efficiently manage the radio resources, and consequently, grant different groups of users simultaneous access to the same multimedia content with differentiated quality of service (QoS). Indeed, to overcome this scalability issue, the Third-Generation Partnership Project (3GPP) standardized the evolved Multimedia Broadcast/Multicast Service (MBMS), with the aim to introduce the point-to-multipoint (PtM) transmissions. Thus, the same service can be delivered to many users through a PtM link by exploiting the same radio resources. In this scenario, the usage of and multicast services over current LTE and future 5G systems have been identified as possible enabling approaches to efficiently manage the traffic load and provide a better Quality of Experience (QoE) to end-users. In particular, multicasting will allow a large number of users to be simultaneously served with relatively low latency and high throughput.

The employment of multicast transmissions in this scenario has potential to address the problems. On one hand, the large number of smart user mobile devices and user expectations for high-quality rich media services has determined a growing demand for network resources; in DenseNets, mobile users have to make the choice in terms of the network to connect to, in order to balance energy saving and delivery performance. On the other hand, the proliferation of user accesses to the existing and future network infrastructure will bring along with it the operators need for optimizing the radio resource usage.

Nevertheless, although multicasting aims to offer several enhancements in content delivery towards large groups of users, several open issues are still under consideration. The most challenging issue is the multi-user diversity and different channel quality levels experienced by users. Least channel gain users affect the performance of the whole multicast group as they can only support a transmission with low Modulation and Coding Scheme (MCS) level, thus achieving transmissions with bad spectral efficiency. On the contrary, serving multicast users experiencing high channel quality levels improves the system spectral efficiency at the expense of users that have bad channel conditions.

This Thesis focus on the multicast traffic management over 5G mobile networks for the Mobile Video application and, in particular, on the Radio Resource Management (RRM) over Multicast transmissions. This work is divided into four different research contribution.

In the first chapter, application scenarios of multicast communications have been described together with the respective enhancement needed for their implementation. The major challenges come from the fact that multicast traffic does not only target groups of end-user devices, but it also involves machine-type communications (MTC) for the Internet of Things (IoT). The increase in the MTC load, predicted for 5G, calls into question the effectiveness of the current Multimedia Broadcast Multicast Service (MBMS). Aim of this chapter is to provide a survey of 5G challenges in the view of an effective management of multicast applications, and to identify how to enhance the mobile network architecture to enable multicast applications in future 5G scenarios. By accounting for the presence of both human and machine-related traffic, strengths and weaknesses of the state-of–the-art achievements in multicasting are critically analyzed to provide guidelines for future research on 5G networks and more conscious design choices. As emerged RRM will play a key role in multicast communication. Indeed, it has been dealt with in the following three chapters.

In the second chapter, resource management issues were addressed in multicast communications in a Dense Heterogeneous network scenario, trying to balance customer services in poor channel conditions compared to those in good channel conditions. Starting from a mobility management analysis, a hybrid unicast-multicast algorithm have been proposed in order to smartly manage multicast-group users within a 5G dense environment. The proposed HUMANS algorithm offers the additional option of selecting multicast transmissions in the network selection process during video delivery. By serving users with good channel conditions via unicast transmissions and users with poor channel quality conditions via multicast, HUMANS allows outperforming other solutions in terms of outage percentage and average quality of transmission, in both low- and high-density scenarios. Most importantly, at the same time it guarantees operators a more efficient resource utilization.

In the third chapter, the RRM issue have been extended to Single Frequency Networks (SFNs). SFNs offer new video delivery solutions to Telco operators. Multimedia Broadcast/Multicast Service SFN (MBSFN) support the transmission of identical content simultaneously in multiple cells using the same radio resources, with improved network scalability and spectral efficiency. In this chapter, a Dynamic MBSFN Area Formation (DMAF) algorithm have proposed. It selects the adjacent cells included in a MBSFN area following the principle to increase the system Aggregate Data Rate (ADR) with no outage. The proposed algorithm dynamically creates MBSFN Areas by exploiting the multicast subgrouping approach. Thus, all the cells are grouped into different MBSFN Areas. Each cell could be part of more overlapping MBSFN Areas, each of them delivering a video with a different quality level. The Base Layer is delivered to all users, whereas users with better channel conditions can receive also Enhancement Layers. Furthermore, Dynamic Radio Resource Allocation is performed by efficiently assigning resources to each MBSFN Area.

Finally, in Chapter 4, the fourth, radio resource management was dealt with by a different perspective. Specifically, integrating into the LTE cellular system a Non-Orthogonal Multiplexing Access technique (NOMA), typical of other broadcast systems. Indeed, NOMA techniques are being also considered as a driver to increase the efficient use of the spectrum in multi-user environments with asymmetric data delivery. Layer Division Multiplexing (LDM) is a NOMA approach that exploits the hierarchical spectrum re-use to send two or more different flows over the same frequency resources during the same time. In this chapter, LDM is jointly applied with multicast subgrouping, thus including LDM as an additional resource allocation mechanism in the Radio Resource Management (RRM).

Conclusions and future recommendations for further developments are finally drawn in final section.

1 Multicast Over Emerging 5G Networks

The demand for multicast applications over cellular systems is in continued rapid growth [1]. As a consequence, multicasting is expected to play a key role in the emerging 5G networks, as outlined in white papers (e.g., from the NetWorld2020 technology platform), research projects, and standard documents from 3rd Generation Partnership Project (3GPP) [2]. Indeed, multicasting represents a viable and effective solution to simultaneously convey data to a group of terminals through point-to-multipoint (PtM) communication, with positive consequences on the capacity and the spectrum efficiency of cellular systems. Both features are crucial for the deployment of 5G networks [3], as witnessed, for instance, in the recent METIS and 5GNOW European research projects.

Presently, video communication is considered as the "killer" *human-oriented multicast application.* Cisco stated that the video traffic carried by mobile networks will reach 15 Exabyte by 2019 (13 times larger than 2014). As a matter of fact, enhanced video services, e.g., Ultra High Definition (UHD), 4K and 3D videos, are becoming popular thanks to the Quality of Service (QoS) capabilities of Long Term Evolution (LTE) and beyond systems [1]. These services, together with the myriads of entertainments, interactive and real-time applications filling our daily lives, pave the way to future 5G human-oriented multicasting.

At the same time, the wide diffusion of low-power devices is leading to a fast network densification aimed at supporting the deployment of Internet of Things (IoT) [3]. In this field, *machine-type communications (MTC)* [3] pushes towards the design of effective solutions to deliver small amount of data simultaneously to a very large (and unpredictable) number of MTC/IoT devices.

The 5G multicast scenario becomes more complex when considering that small-cells, underlying the macro coverage, will be used to (i) enhance the strength level of received signals associated to human services, and to (ii) increase the capacity in MTC scenarios, where devices are typically located in challenging positions (e.g., indoor, basement) [1] [3]. It becomes evident that the 3GPP Multimedia Broadcast Multicast Service (MBMS) [2] needs novel architectural and procedural definitions to meet the multifaceted constraints of the expected 5G multicast services.

For the above-mentioned reasons, multicasting is attracting the attention of a wide research community. Notwithstanding, by surveying the relevant literature, it emerges that up to now the primary efforts focused on approaches (based on short-range links usage [4] [5], beamforming [6] [7], network coding [8] [9], etc.) to boost the data rate performance in the view of an improved quality experienced by subscribers. Not enough attention has been given to the design of architectural and procedural solutions to meet the new challenges of multicasting in 5G networks, in which the same dignity is granted to both human-type and machine-type group services.

The objective of this chapter is to provide a critical evaluation of the enhancements required by MBMS to meet the constraints of 5G human- and machine-oriented multicast applications, among which the most significant ones are outlined together with their relevant key features. The recent literature on multicasting, enhanced by the use of 5G technologies, is analyzed. The main open challenges on the 5G MBMS architecture design are highlighted and hints for the enhancement of the standard procedure to support group-oriented machine-type communications are given. Finally, the primary future research trends are discussed.

1.1 Applications Scenarios

The mobile market scenario for future 5G multicast applications is expected to be characterized by two types of services [1] [3]: (i) the evolution of 4G applications tailored for human users; (ii) the definition of novel machine-based services. It is thus of primary importance to analyze in depth these new services in order to highlight the requirements they dictate on the 5G network. A summary of these applications is given in Fig. 1.1 and will be discussed in the remainder of the Section.

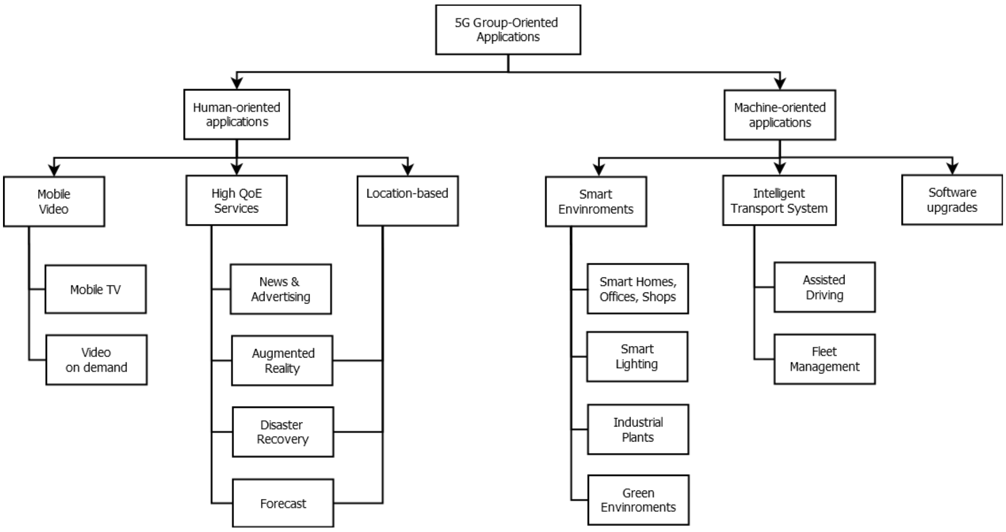


Fig. 1.1 - Application Scenarios For Multicasting Over 5G Systems

1.1.1 Human-oriented applications

Human-oriented applications represent the current trend of multicast services [1]. By considering the current trends and their possible evolutions, the future mobile market of human-oriented multicast applications can be envisioned as composed of the following service categories:

- Mobile video services;
- High-QoE and location-based services.

Mobile Video. Video downloading, video streaming, video conferences, sporting events, concerts and operas became very popular in the last years and, consequently, group-oriented *mobile TV* and *video on-demand* services [2] are expected to play a key role also in 5G systems. These classes of applications will be transmitted over 5G systems at Ultra High Definition (UHD) quality and, at the same time, enhanced 3D video capabilities.

The commercial success of such bandwidth-hungry applications is strongly tied to the development of effective solutions for resource allocation and management of the co-existence with other (unicast and broadcast) services.

This kind of applications requires high data-rate, low-jitter and connectivity everywhere, anytime, especially in mobility. Timing synchronization within the multicast group needs to be taken into account as well.

High-QoE services. The improvement of users' QoE is a hot topic in the current literature [1] and it will remain so in several multicast scenarios relevant to 5G [3]. For instance, *news and advertising* applications can be enhanced by allowing a group of users to receive customized information according to their profiles (e.g., interests, hobbies, preferences). Besides, a special class of high-QoE applications are those implying a service reception based on the user location, namely **location-based applications**. This implies the 5G concept of considering users being fully connected with the surrounding environment (as referred in a NetWorld2020's white paper [10] and in METIS and 5GNOW EU projects). A first example of location-based applications are *augmented reality* multicast applications, especially conceived for commercial or touristic services. They allow users to receive additional information from the surrounding environment (for instance, visitors in a city/museum form a multicast group and receive interactive content related to the art work/rooms they are observing in that moment). Another target scenario for high-QoE location-based multicast applications is public safety for *disaster recovery.* In case of disasters (e.g., earthquake, fire, explosion), a group of users (both victims and rescuers) receive information of common utility to properly re-act to the emergency.

The above-discussed classes of applications pose further challenges related to the management of users' position and profiles: this requires fine-grained user tracking mechanisms as well as effective procedures of group formation/joining and service announcement. In order to accommodate these new features, the current MBMS standard architecture needs to be upgraded. Furthermore, enhanced-QoE multicast applications ask for very-low latency data transmission, high-reliability, and extended coverage to guarantee an adequate quality also to the terminals located in disadvantaged positions.

1.1.2 Machine-oriented applications

The benefits achieved by group-based, instead of unicast, communications towards machines are clear when considering the unprecedentedly huge number of connected sensors/machines expected for 5G systems [3]. A disruptive novelty introduced by group-oriented MTC is that the owner of the devices (i.e., a customer paying for cellular connectivity of all its equipment/appliances) can decide which device or group of devices involve in the communication.

The capability of managing simultaneous multiple MTC devices supports the delivery of different multicast machine-oriented applications such as those in the following subsections.

Smarts Environments. The deployment of smart environments is a key goal of 5G/MTC platforms aiming at supporting stakeholders and users to take decisions based on real-time information in the view of significantly reducing costs, improving the quality of life, optimizing industrial processes, etc. [3]. In this context, multicast MTC applications can be beneficial for *smart homes/offices/shops*, when for example users out of their own homes send messages towards a group of actuators to switch on/off electronics appliances (e.g., the heating/cooling system). Another case is the *smart lighting* application for homes, offices or streets. For example, to save energy, lights on a mountain street can be turned on only when a car is getting close; thus, a group of lights can be switched on/off according to the movement of the cars. In a similar way, lights at home or office can be monitored and managed according to users' needs. Furthermore, multicasting is mandatory for *smart industrial plants* to enable the efficient transmission of security control, warning or management messages. For instance, in case of a problem in the industry chain, the whole group of devices belonging to the assembly line could be stopped or re-organized to react to the critical event.

Above considered classes of applications can be also suitable to multicast *green environments*, where group-based management of sensors/actuators can reduce energy consumption (one of the most critical issues of 5G networks), thus increasing the battery lifetime.

In conclusion, multicast applications related to smart environments ask for low-latency data transmission, low-energy communication, and location- and customer-based group creation procedures. These requirements introduce tough challenges in the efficient and wise group formation (several classes of location-based applications) and in the reduction of the overhead for multicast transmission towards involved devices.

Intelligent transport systems. Vehicular applications developed in the last few years are claimed to be efficiently supported over 5G networks. Roads and vehicles will be equipped with sensors and tags to receive/transmit control/data messages [1] [3]. An example is the *assisted driving*, where terminals involved in the same services (e.g., traffic management) or within the same area (e.g., cars close to the position of an accident) can be grouped to better disseminate data (e.g., traffic measurement, positions, speed) among interested vehicles. Similarly, *fleet management* applications can benefit from multicast transmissions.

Multicast transmissions for intelligent transport systems pose challenges in terms of design of low-latency group formation/re-formation (also location-based) procedures, which are made more complex by the high speed of involved devices.

Software/firmware Upgrade. Sensors, smartphones and, more in general, all smart devices need software/firmware updates, periodically or at the occurrence of specific events/dates. Smart devices could receive software/firmware upgrades either in case a new version is available or to fix bugs and add/change functionalities. In addition, software upgrades can be also location-based when, for instance, sensors installed on a given area receive upgrades to enhance/update their sensing capabilities (e.g., novel route directions on a street). In this case, the key challenge is related to the group formation, which can be driven by the owner of sensors (who, for example, could be interested in sending data *only* to its own devices according to their location or functionalities/tasks). This means that not the network provider, but the sensors' owner should manage group formation; therefore, the definition of effective customer-based group formation procedures becomes an issue.

The key requirements discussed in this section with reference to the different kinds of applications are summarized in Table 1. In the following, we will discuss how to meet such requirements by considering both architectural and data transmission point of views. To this aim, we first analyze the pros and cons of current group-oriented architecture and its possible evolution to meet 5G group-oriented requirements, and then we focus on the contributions in literature mainly covering data transmission aspects. We finally discuss the further enhancements needed to handle effective group-oriented applications in 5G systems.

Group-oriented applications		**High data rates**	**Jitter**	**Low-latency data transmission**	**Location-based group**	**Low-latency group creation**	**Customer-based group**	**Low-energy communicatio**
Human-oriented services	Mobile video	✓	✓	✓				
	High-QoE	✓			✓	✓		
Machine-oriented services	Smart environments			✓	✓		✓	✓
	Intelligent transport system			✓	✓	✓	✓	✓
	Software upgrades				✓		✓	✓

Table 1 – Key requirements of 5G group-oriented applications

1.2 5G Multicast/Broadcast Network Architecture

The focus of this section is on the network architecture and procedures needed to handle multicast services. First, we present the current (4G) MBMS architecture, its design drivers and operation, and then we identify the major architectural and procedural changes needed to handle 5G group-oriented services in order to fit the requirements of Table 1 [11].

The scenario of future 5G multicasting is expected to be composed by heterogeneous environments with different communication ranges as depicted in Fig. 1.2, where 5G MBMS manages different radio access and transmission technologies. In brief, wide area coverage is offered through GEO satellites, managed by the satellite-eNodeB (S-eNB) located in the ground component, while macro base stations (i.e,. eNodeBs) provide group services in urban/suburban areas; finally, small-cells (e.g., femto-cells a.k.a. Home-eNodeBs) and short-range (either 3GPP or non-3GPP) links enhance the connectivity for indoor home/industrial services and extends the coverage of traditional macro-cells, respectively.

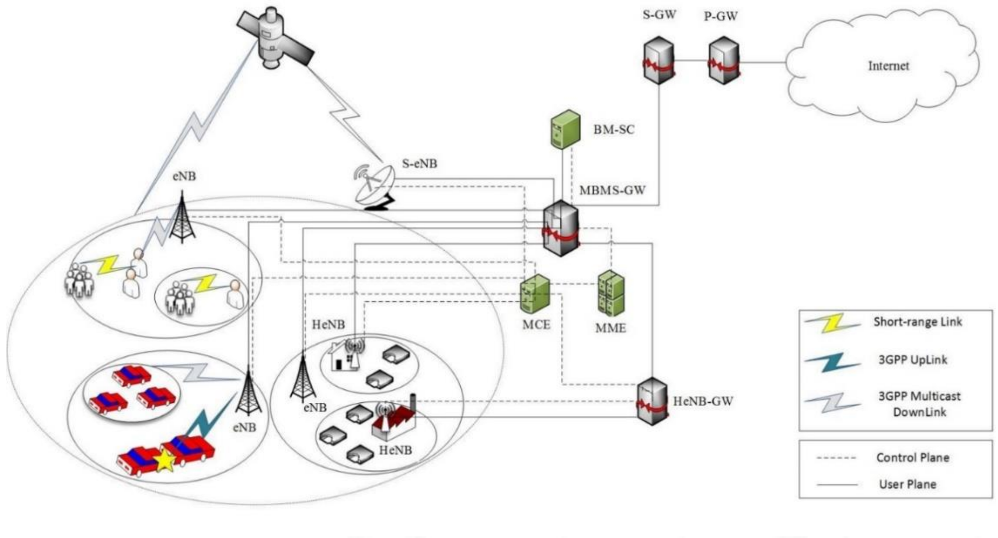


Fig. 1.2 - 5G multicast environments with related 5G MBMS architecture

MBMS represents the reference standard architecture for multicast and broadcast service delivery in cellular systems [1] [2]. It specifies the network entities and the related interfaces as well as the procedures for supporting multicast services over 3GPP networks. The MBMS architecture includes: *(i) Broadcast Multicast-Service Center (BM-SC)*, i.e., the source of multicast content which authorizes and initiates the MBMS bearer services and delivers MBMS data; *(ii) MBMS-Gateway (MBMS-GW),* which accomplishes data content forwarding to the eNode-Bs (eNBs) involved in the MBMS session; *(iii)*

MultiCell/Multicast Coordination Entity (MCE), in charge of the session control signalling towards the involved eNBs.

More in detail, the BM-SC is in charge of providing *Membership*, *Session and Transmission*, *Service Announcement*, and *Security* functions that manage authorizations for MBMS subscribers. The BM-SC *Session and Transmission* function schedules MBMS session transmissions and retransmission and sends MBMS data. Through the *Service Announcement* function, the BM-SC is able to provide the user equipment (UE) with media descriptions, specifying the media to be delivered as part of an MBMS user service. The MBMS *Security* function is used for distributing MBMS keys (Key Distribution Function) to authorized UEs.

The MBMS-GW is located between the BMSC and eNBs and its principal functions is the sending/broadcasting of MBMS packets to each eNB transmitting the service. It allocates the IP Multicast address to the eNBs involved in the delivery of MBMS traffic, and implements MBMS Session Control Signaling (session start/update/stop) towards the E-UTRAN via the Mobility Management Entity (MME).

The MCE is a logical entity with tasks of admission control and radio resource allocation to all eNBs, to decide suspension of MBMS session(s), to decide not to establish the radio bearer(s) of the new MBMS service(s) if the radio resources are not sufficient for the corresponding MBMS service(s). The MCE is involved in MBMS Session Control Signaling. Moreover, an eNB is served by a single MCE

A summary of the functionalities of MBMS entities, its limitations and future enhancements can be found in Table 2.

	Node function and/or procedure	**Limitations and/or future enhancement**
BM-SC	• Service announcement Function	• Extension for MTC missing
	• Membership Function	• Extension for MTC missing
	• Security Function	• Low processing capabilities of machines

MBMS-GW	• Data forwarding	• Heterogeneous eNB management
	• Session control signaling	• Signaling messages overloading • Signaling overhead
MCE	• Admission Control and Radio Resource Allocation	• Huge number of devices • MTC traffic with low energy requirements
	• Session control signaling	• Signaling messages overloading • MTC signaling

Table 2 - Entities, related functionalities and limitations of 4G MBMS

MBMS services are delivered through a set of procedures defined in the TS 23.246 [2] (see Fig. 1.3).

Subscription establishes the relationship between the user and the service provider, which allows the user to receive the related MBMS multicast service. Service Subscription is the agreement of a user to receive service(s) offered by the operator. Subscription information is recorded in the BM-SC. Subscription information and other BM-SC functionality may be on separate entities, which is enabled by proxy capability of the G_{mb} interface.

Service announcement: MBMS user service announcement/discovery mechanisms shall allow users to request or be informed about the range of MBMS user services available. This includes operator specific MBMS user services as well as services from content providers outside of the PLMN. Service announcement is used to distribute to users information about the service, parameters required for service activation (e.g. IP multicast address(es)) and possibly other service related parameters (e.g. service start time).

Joining (i.e. MBMS multicast activation by the user) is the process by which a subscriber joins (becomes a member of) a multicast group, i.e. the user indicates to the network that he wants to receive Multicast mode data of a specific MBMS bearer service. An MBMS user service may also be carried by more than one MBMS bearer service. In that case the MBMS user service part in the UE initiates the relevant MBMS bearer services to receive the service).

Session Start is the point at which the BM-SC is ready to send data. This can be identified with the start of a "Multicast session". Session Start occurs independently of activation of the service by the user, i.e. a given user may activate the service before or after Session Start. Session Start is the trigger for bearer resource establishment for MBMS data transfer. If an MBMS user service is carried by more than one MBMS bearer service, a Session Start message is sent for each MBMS bearer service. In that case, the UE may need to initiate the reception of multiple relevant MBMS bearer services to receive the MBMS user service.

MBMS notification informs the UEs about forthcoming (and potentially about ongoing) MBMS multicast data transfer.

Data transfer is the phase when MBMS data are transferred to the UEs.

Session Stop is the point at which the BM-SC determines that there will be no more data to send for some period of time this period being long enough to justify removal of bearer resources associated with the session. At Session Stop, the bearer resources are released.

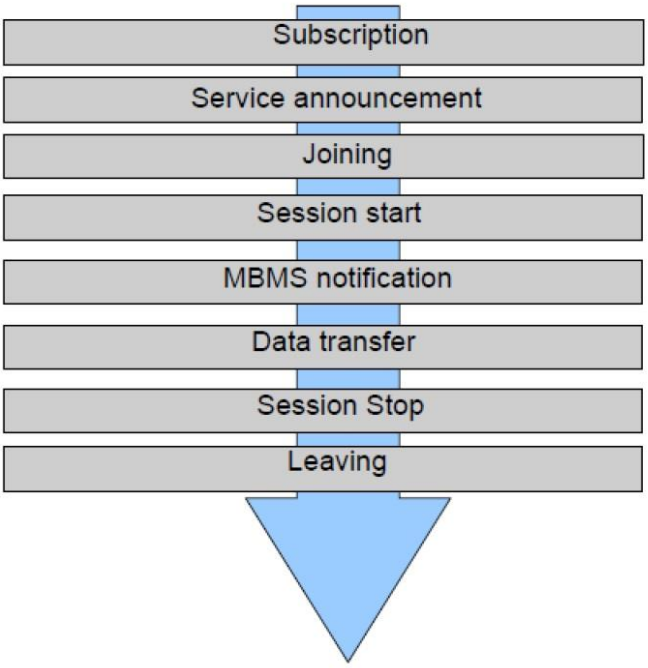


Fig. 1.3 – MBMS Procedures in Multicast Mode

1.2.1 *Enhancing MBMS to 5G multicast services*

The future scenario of 5G multicast services will be composed of heterogeneous environments, characterized by dissimilar communication ranges and capabilities, wherein an evolved MBMS manages different radio access and transmission technologies. The considered heterogeneous scenario dictates a wide set of enhancements in the functionalities of the MBMS entities in 5G systems. The main architectural changes to the MBMS system are needed in order to manage multicast MTC.

Fig. 1.4 shows an extended network architecture that leverages some enhancements to MBMS in order to support group-oriented MTC via small cells. The BM-SC must be enhanced to offer "customer-based" machine-oriented group services besides the traditional (human) services initiated by network providers. In the MTC case, in fact, only a set of MTC/IoT devices belonging to the same service and controlled through their own Service Capability Server (SCS), could be interested in receiving data.

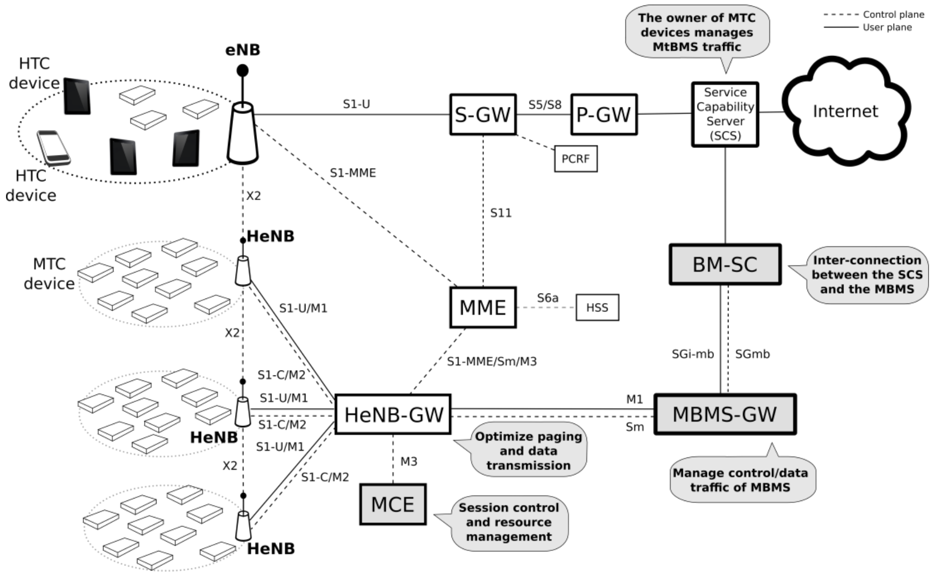
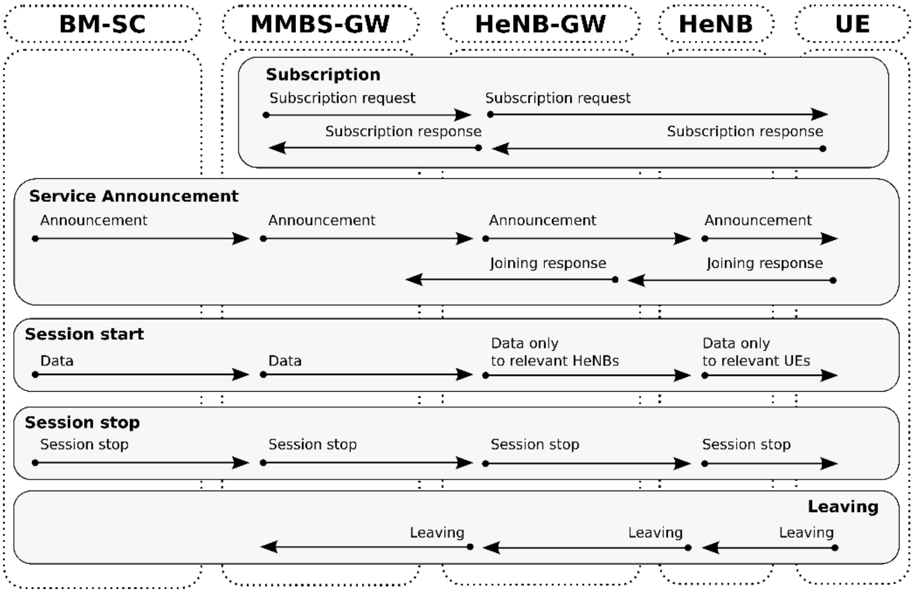


Fig. 1.4 - Enhanced MBMS architecture to support 5G group-oriented machine-type communications.

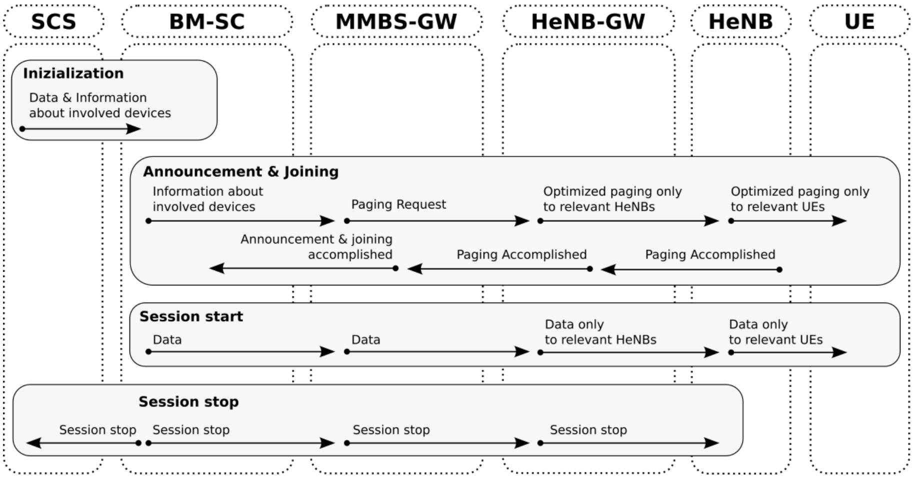
This implies that service announcement and membership functionalities should be tailored to deal with a pre-defined list of devices provided by the SCS. The reason is that machines cannot autonomously decide to join a multicast group; thus, membership has to be *customer-driven,* instead of provider-driven, being only the devices' owner aware of what terminals belong to the MTC/MBMS service. The BM-SC should also be in charge of informing the devices involved in the group-oriented machine-type traffic, by triggering a *paging procedure* to enable MTC/IoT devices in *idle* mode to receive data. This aspect is highlighted in Fig. 1.5 where the differences between the legacy MBMS procedures and the enhanced MBMS procedures for MTC are illustrated [11].

The MBMS-GW will have to manage heterogeneous types of eNBs involved in the MBMS session, i.e., eNB(s)/HeNB(s)/S-eNBs. The issues relevant to this feature are exacerbated in 5G environments, where the number of involved base stations might be large, especially when considering small-cells. *Signaling* towards the mobile core network could be overloading, due to the huge number of involved eNBs.

At the same time, the MCE needs to implement more complex control signaling procedures to account for both human and machine-related group services, wherein the former are served with legacy MBMS control procedures while the latter ask for improved control procedures to support membership and paging functions. The current MBMS architecture does not provide any mechanism for *paging coordination* among different types of base stations, which is an issue raised by heterogeneous 5G networks.



(a) Legacy MBMS procedures



(b) Enhanced MBMS procedures for machine-type communications

Fig. 1.5 - Enhanced MBMS vs. Legacy MBMS Procedures to Support Group-Oriented Services

1.3 Toward 5G Group-oriented Data Transmission

The technology enhancements expected in 5G systems offer unprecedented potentialities for the delivery of multicast applications. In this section, recent achievements on data transmission for multicasting over cellular systems are surveyed.

1.3.1 Short-range enhanced communications

Most of the literature on mobile multicasting has traditionally focused on boosting the data rates experienced by mobile users through the synergistic use of macro-cellular and short-range (both 3GPP and non-3GPP) links [5]: this approach can also assure a low data transmission delay. As for machines, advantages of short-range communications also include prolonged terminal battery lifetime due to the reduced transmission power. In the remainder, we will analyze the different improvements short-range communications bring to 5G multicasting.

The literature on this topic mainly focuses on the use of device-to-device (D2D) [12] [13] cellular links to *increase data rates* of multicast users. Researches in [4]- [14], for instance, proposed to serve only the portion of users with good channel conditions directly from the eNodeB, and use them as relay-entities to forward data towards the remaining receivers over short-range D2D links. The main issue in this scenario is cluster formation, i.e., the selection of the right amount of relaying devices to serve via base station.

In this research area, the literature mainly addresses the impact of different cluster formation techniques on the multicast data rate. For instance, the main idea in [4] is to use intra-cluster D2D retransmissions to minimize the amount of needed resources and this allows achieving about 40% gain in resource utilization when compared to legacy multicasting in cellular systems. This result witnesses the spectral efficiency improvement guaranteed by the joint exploitation of cellular and D2D links.

Energy-efficiency issues are considered in [15], where the authors focus on a two-stage cooperative multicast scheme to minimize the total transmission power without affecting the service coverage. Obtained results demonstrate that in a macro-cell about 80% of the power consumption can be saved thanks to the wise use of D2D communications. Recently, the authors of [16] proposed a novel approach for D2D cluster formation, which encompasses the previous techniques. By assuming macrocell-driven synchronization of the transmitting relay nodes and the interest of all receiving devices in the same multicast content, they propose to allocate resources on the same frequency to all relay nodes for multicast data forwarding over D2D links. This reduces inter-cluster interference. Besides, they design a clustering scheme which minimizes the number of relays, thus *reducing the overall power consumption*.

A different family of works addresses the use of non-3GPP (e.g., Wi-Fi, Wi-Fi Direct) links to enhance the performance in macro-cells. The authors in [14] proposed a clustering algorithm that exploits macro-cellular and Wi-Fi Direct links to improve the session quality experienced by multicast users by *reducing the content delivery delay and the energy consumption*. They also demonstrate that Wi-Fi links bring more beneficial effects than D2D links when the available spectrum is scarce.

A further concept of interest for multicasting enhancement over 5G networks is the *social network group (SNG)*, where group members not only share their personal interests and keep in touch, but also share contents. In this field, [17] introduced a cooperative multicast algorithm where a portion of SNG members download a content of common interest through a cellular link and disseminate it among other group members through short-range ad-hoc wireless interfaces. The node(s) responsible for content downloading is dynamically chosen based on cellular link quality, residual battery level, and amount of the content already downloaded by the members.

Literature review discussed above fits requirements presented in the Table 1. For example, from the point of view of Mobile-video scenario, it is already demonstrated that short-range communication is suitable for providing high-data rate, low-latency and low-energy communication.

1.3.2 Macro/small cell cooperation

Cooperation among macro and small cells is another key aspect that characterizes 5G systems; hence the interest in evaluating cooperation also in multicast scenarios. For instance, [18] introduces an enhanced architecture for multicast transmissions over heterogeneous cellular networks aiming at improving the energy efficiency in multicast delivery. In this scenario, macro-cells provide wide-area coverage while small base stations (e.g., femto-cells) provide *high data rates* locally *in hotspot regions*. By analyzing the user behavior, the idea is to multicast the data stream in the macro-cell region. Subsequent arriving users that (within a certain time window) request the same stream will immediately join the multicast group and the missing fraction of the stream will be served through small cells. It is worth noting that this solution is quite different w.r.t. the exploitation of D2D links, as no buffering is required on the user side. The main benefit of macro/small cell cooperation is that cooperative schemes outperform conventional unicast and multicast scheme in terms of *energy efficiency*. The cooperation between heterogeneous cells satisfies some of the requirements listed in Table 1, such as low-latency and low-energy communications. Although many works in literature deal with this topic, 5G multicasting requirements of low-latency and low-energy need to be further investigated.

1.3.3 Network coding

Network coding (where the packet decoded by a receiver is obtained as the combination of different information with different coding characteristics) has proven to be a further effective solution to enhance throughput and robustness of data transmission.

In [8], the authors proposed the random linear network coding integrated with an unequal error protection (UEP) technique to improve the reliability of a layered multicast service. A layered service (e.g., H.264/SVC video) consists of a base layer, for basic service quality, and multiple enhancement layers, to improve the quality; with UEP, layers that are more important require a higher protection level. They also design a resource allocation framework that minimizes the number of coded packet transmissions needed to deliver service layers. Analyses on H.264/SVC video delivery over LTE-Advanced show that this approach *improves the service quality* (i.e., the number of consecutive layers recovered by a user) w.r.t. legacy PtM transmission of, at least, a factor of 1.35.

The use of network coding has been also considered in short-range D2D scenarios, as in [9].Here, network coding is used to enable the transmission of different information coded into a single multicast packet destined to different receivers. A user-specific bit mapping algorithm applies different coding rates to different information before network coding is performed. Furthermore, a user-specific link adaptation scheme chooses an optimal modulation and coding scheme (MCS) for D2D multicast so that each user obtains its information from the same packet with different MCS according to the respective channel quality. This solution achieves a *throughput gain* ranging from 13% to 45% compared to conventional multicast transmissions.

Other preliminary works show that network-coded transmissions can improve the spectrum re-use. In [19], for example, a new transmission system based on hierarchical spectrum re-use, named *Cloud Transmission (CloudTxn)*, is proposed for terrestrial broadcasting (DVB-T) and can be properly adapted for PtM multimedia services. The idea is that two flows (*A* and *B*) share the same resources; specifically, a (clock synchronized) data stream B is injected onto the same resources used by stream A, and transmitted with a reduced power compared to stream A. This mechanism permits, for instance, to assign a more robust MCS to stream A (e.g., for mobile services in very noisy environments) and more efficient MCS to stream B (e.g., to provide high quality HD and UHD services); it also *enhances spectrum efficiency and re-use*, which are key benefits expected in 5G systems.

Works presented above demonstrate that network coding techniques allow a high throughput performance; hence these may be suitable for high data rate human-oriented applications. Since Network coding provides high data-rate performance, it is helpful to fit requirements of both mobile-video and High QoE applications.

1.3.4 Beamforming

Beamforming, i.e., the possibility of using directive antennas to achieve spatial diversity and throughput improvements, is also a technology that can boost the performance of multicasting over future 5G systems. Indeed, beamforming *improves the data rate* of users with the lowest channel gain without increasing the transmission power level, as for instance addressed in [6]. In detail, the authors present a low-complexity method that reduces the complexity of selecting the beamforming vector (i.e., the information needed by the antennas to make a directive transmission) and increases the throughput performance up to 52% w.r.t. conventional transmissions.

Focus of [7] is the joint transmit beamforming and antenna selection in scenarios with multiple co-channel multicast groups. The algorithm presented in [7] finds, for each multicast group, the beamforming vector that minimizes the overall transmission power. A key goal achieved by this approach is the reduction in the number of antennas required to meet the quality constraints of multicast services.

Finally, adaptive beamforming for scalable video coding (SVC) scenarios is considered in [20], where the authors propose to schedule different SVC layers with different beams and MCSs while guaranteeing the respect of the QoE constraints of all multicast users. The increase in the throughput performance, favored by beamforming, satisfies *high data rate* requirements of human-oriented applications, whereas directive antennas reach devices/machines with low channel quality, thus *extending the coverage*. Consequently, also machine-oriented applications can gain from the implementation of beamforming techniques. As discussed for network coding, beamforming is useful to fit high-data rate communications. Hence, mobile video and high-QoE applications could benefit of this technique.

1.4 Expected benefits and open issues

The technology enabler discussed above mainly focused on the radio-related aspects of multicast communications in order to improve the efficiency of data transmission. Indeed, the exploitation of short-range links, macro-cell and small-cell cooperation, network coding and beamforming brings key benefits to the data plane. More in detail, by considering the requirements listed in Table 1, we can now summarize the benefit provided by such techniques.

Short-range communications support *enhanced data rates* thanks to the better channel conditions experienced by devices that are closer to each other w.r.t. the base station. In a similar way, small cells can be exploited to increase data rates in hotspot areas such as stadiums in case of events. In addition, data rates can be boosted through network coding techniques, which improve the robustness of data transmission. Finally, beamforming can boost data rates by exploiting directive antennas, thus improving the channel gains due to the increase in the received power by devices.

Short-range communications, macro/small cell cooperation, network coding and beamforming *reduce latency* in data delivery as they improve the data rates of data transmission from the base station to the devices. Nevertheless, the exploitation of short-range links introduces delays due to relaying operations, while network coding introduces delays to decode the received packets. These aspects did not receive enough attention in order to understand in which scenarios the final latency is effectively cut w.r.t. that experienced with legacy transmissions from the base station.

By exploiting the intrinsic feature of local communications, short-range links or small cells could be exploited to perform a *location-based group creation*. In this, only the devices in a limited area will

receive the control information to join the multicast transmission, in contrast with the case when the joining procedure is performed through the macro-cell. Location-based services could be potentially offered by exploiting beamforming with a proper re-direction of beams. Nevertheless, the above-mentioned aspects have not yet been investigated. Finally, from a session management point of view, location-based group creation could be enabled by extending the functionalities of the BM-SC, with the joint exploitation of SCS in case of machine-type services.

Short-range communications as well as the exploitation of small cells could *reduce the energy consumption* of devices when receiving data thanks to the proximity nature of these communications.

From this brief synthesis, it is evident that a missing topic for the effective provisioning of 5G group-oriented applications is the control traffic needed to handle multicast services. Indeed, some of the requirements listed in Table 1 have not yet been addressed in the current literature, i.e., *low-latency, location-based, and customer-based group formation*. For this reason, in the remainder of this Section we will focus our attention on future research directions dealing with control traffic management.

1.4.1 Open Issues

The support of machine-based services especially raises several challenges to be solved.

A first issue is related to low-latency group creation. MTC devices need to be switched on (from *idle* to *connected mode*) in order to receive multicast data. Consequently, an issue for the customer is to identify the machines in the multicast group and to page them before data transmission begins. Group paging procedures in 5G systems are expected to be designed with the aim of simultaneously switching a large number of devices into connected mode, with further positive effects on the *control overhead* and the *latency* of group-based machine applications. Indeed, MTC applications may require transmissions of a few bytes each. Thus, minimizing the amount of control information exchanged between the network and the devices becomes an issue of utmost importance, not only from a latency point of view. Besides, due to the machines' low processing capabilities, MTC/IoT communications suffer from *security issues*, since it is not easy to implement security functions, e.g. key exchange procedures. These aspects need to be considered as they affect the group formation delay. As emphasized in [4] [12]- [14], a large number of related works in the literature exploiting 3GPP (i.e., D2D) and non-3GPP (e.g., Wi-Fi) short-range links assume that the unit responsible for resource allocation (generally, the base station) is "somehow" aware of the channel conditions of short-range links among group members. In such scenarios particular interest in future research should be given to *control messaging* (for the purpose of nearby device discovery, cluster formation/updating, etc.). In practical systems, the way the base station becomes aware must be specified. It is straightforward to think that each terminal measures (or estimates) the link quality with nearby devices and transmits this information to the resource management unit. Nonetheless, effective solutions in the case of large multicast groups still need to be properly investigated and designed to avoid overloading or bottleneck effects at the base station in 5G scenarios. In addition, control messaging among interested devices needs to be properly defined, and their performance assessed by considering their impact on MTC traffic and devices in terms of latency and energy-consumption. The above mentioned issues are exacerbated in the case of non-3GPP short-range links, which require the effective solutions to manage control and billing messaging between, e.g., LTE and trusted/non-trusted Wi-Fi interfaces.

The highlighted problems do not yet find answers in the literature. Thus, there is an urgency to focus greater attention on them to enable effective short-range solutions for enhanced multicasting.

A further research need in terms of low energy communications deals with the procedure of *relay selection*. Indeed, while current works mainly focus on channel conditions, additional information about buffer state, residual battery levels or position of nodes could be relevant to enhance 5G relay selection/updating procedures. In particular, relay selection based on device position could potentially enable location-based multicast services. Hence, a constantly increasing number of devices with constrained power and memory capabilities challenges multicast IoT services.

The higher the number of involved devices, the higher the *complexity* of the algorithms, and this affects latency in terms of both group creation and data delivery. This is still linked to the above discussed limited computational power of machine-devices. More powerful devices, instead, face the problem of increasing energy consumption.

In general, the complexity represents the most challenging issue to be solved in 5G systems. Although the use of network coding has emerged as a promising approach to improve the reliability of multicast transmissions [9] without additional resource requirements, the complexity problem is exacerbated by the expected large amount of MTC devices with reduced computational capabilities. In addition, the computational cost for beam form calculation at the base station increases data delivery latency. This calls for low-complexity near-optimal solutions valid in a wide range of environments and users' locations and propagation conditions. From a more general point of view, an effective low-latency (as well as a customer-based) group creation can be achieved through a re-design of joining procedures achieved by extending the functionalities of BM-SC and MBMS-GW entities (Table 3) [11].

Enabling features	**Expected benefits**	**Open Issues**
Short-range direct links	• High data rates • High spectrum efficiency (spectrum re-use in case of 3GPP D2D links) • Offloading of cellular spectrum (in case of non-3GPP links) • Coverage extension • Reduced energy consumption	• D2D cluster formation (device discovery, control messaging among group members) • non-3GPP cluster formation (device discovery, control messaging among group members) • New criteria for relay nodes selection (i.e.,

		buffer status, battery level) • Interference induced by D2D links on non-multicast cellular users • Cluster re-formation procedures (e.g., triggered by device mobility, or running out-of-energy)
Network coding	• High data rates • High robustness • High spectrum efficiency	• Complexity at the transmitter and receiver sides • Latency • Robustness to packet losses • Control overhead
Beamforming	• High data rates • High spectrum efficiency (spatial diversity)	• Complexity • Control overhead • Fast re-computation in case of mobile users • Jointly use of directive antennas and MIMO
Machine-type communication	• MTC group-services • New business opportunities for Telco • Simplification of group management procedures (e.g., announcement, join)	• Extension of legacy MBMS architecture • Definition of machine-oriented MBMS procedures • Definition of multicast paging procedures • Control overhead • Improvements of resource efficiency and system capacity

Table 3 – 5G Multicast Enabling Features: Expected Benefits and Open Issues

1.4 Expected benefits and open issues

2 RRM over 5G multicast – HetNets/ DenseNets management

2.1 Dense Heterogeneous Networks

Today's heterogeneous networks comprised of mostly macrocells and indoor small cells will not be able to meet the upcoming traffic demands. Indeed, it is forecasted that at least a 100-fold network capacity increase will be required to meet the traffic demands in 2020.

In view of such significant future traffic demands, the mobile industry has set its targets high, and has decided to improve the capacity of today's networks by a factor of 100× or more over the next 20 years [21].

In order to achieve this goal, vendors and operators are currently looking at using every tool they have to improve network capacity. Mainly three paradigms are noteworthy, i.e. network densification, the use of higher frequency bands and spectral efficiency enhancement techniques [22]. Network densification is accomplished through the deployment of Heterogeneous Networks (HetNets) and small cells [23] - [24]; larger bandwidths, exploit higher spectrum frequencies, both in licensed and unlicensed spectrum [25]– [26]; spectral efficiency is enhanced through multi-antenna transmissions [27], cooperative communications [28], dynamic TDD techniques [29]– [30], etc. Nevertheless, such paradigms introduce some issues. Network densification complicates network deployment as well as backhauling and mobility management, while higher carrier frequencies suffer from larger path losses, and usually require more expensive equipment. Most spectral efficiency enhancement technologies depend on a tight synchronisation as well as relatively complex signal processing capabilities, and may be compromised due to inaccuracies in Channel State Information (CSI).

In order to meet the exponentially increasing traffic demands [31], mobile operators are already evolving their networks from the traditional macrocell-only networks to HetNets [32], [33], in which small cells reuse the spectrum locally and provide most of the capacity while macrocells provide an umbrella coverage for mobile UEs. Currently, small cells are deployed in large numbers. Indeed, according to recent surveys, the number of small cell BSs was already larger than that of macrocell BSs. These small cell deployments are mainly in the form of home small cells, known as femtocells [34], [35], [36], but many operators have also already started to deploy outdoor small cell solutions to complement their macrocell coverage [37].

However, in a co-channel deployment of small cells with the macrocell tier, due to the large difference in transmission power between both types of BSs, being attached to the cell that provides the strongest pilot Reference Signal Strength (RSS) may not always be the best strategy. UEs will tend to connect to macrocells rather than to small cells, even if they are at a shortest path loss distance. This effect is aggravated as the distance between small cell and macrocell BSs becomes smaller. Indeed, the closer the small cell BS is to the macrocell BS, the smaller is the resulting small cell coverage due to macrocell BS power dominance. This leads to a poor macrocell off-load [32], [33]. Moreover, the transmissions of UEs connected to macrocells will also severely interfere with all small cells located in their vicinity in the Uplink (UL). Note that due to the lower path loss, if a macrocell UE would connect to the small cell with the smallest path loss, this UE would transmit with a much lower UL power. This would allow load balancing as well as UL interference mitigation, thus improving network performance. In order to address these problems arising from the significant power difference between co-channel BSs in HetNets, new cell selection methods that allow UE association with cells that do not necessarily provide the strongest pilot RSS are necessary.

Network densification has the potential to significantly increase the capacity of the network with the number of deployed cells through spatial spectrum reuse, and is considered to be the key enabler to provide most of the capacity gains in future networks.

In order to better understand the implications of network densification on network capacity, let us define network capacity based on the framework developed by Claude Shannon [38] as

$$C[bps] = \sum_{m}^{M} \sum_{u}^{U_m} B_{m,u}\,[Hz] \log_2\left(1 + \gamma_{m,u}\right) \quad (1)$$

where $\{1, \ldots, m, \ldots M\}$ is the set of BSs deployed in the network, $\{1, \ldots, u, \ldots, U_m\}$ is the set of UEs connected to BS m, B[Hz] is the total available bandwidth, and $B_{m,u}$[Hz] and $\gamma_{m,u}$ are the bandwidth granted to and the SINR experienced by UE u when connected to BS m. This model assumes Gaussian interference. At the network level, network densification increases the number of geographically separated BSs M that can simultaneously reuse the available bandwidth B, thus linearly improving spatial reuse and increasing network capacity with M. At the cell level, a consequence of network densification is cell size reduction, which directly translates into a lower number of UEs U_m connected to BS m and thus a larger bandwidth $B_{m,u}$ available per UE. In this way, network capacity linearly increases with the number of offloaded UEs. Moreover, at the cell level too, the average distance between a UE and its serving BS reduces, while the distance to its interfering BSs does not necessarily reduce at the same pace assuming idle mode capabilities. This leads to an increased UE signal quality $\gamma_{m,u}$, and thus the network capacity logarithmically increases with $\gamma_{m,u}$. As can be derived from the above discussion, network densification increases M and in turn improves both $B_{m,u}$ and $\gamma_{m,u}$, resulting in an increase of the network capacity.

Next some of the main challenges faced on the way by dense small networks are highlighted [22].

Mobility management. Future network architecture comprise different small cell tiers with different types of small cell BSs, target at different types of environments and traffic, where dedicated channel mid-frequency small cell deployments with the macrocell tier may be ultra-dense to enhance network capacity. Within this architecture, mobile UEs should be kept in the macrocell tier, while static UEs should be handed over to the ultra-dense small cell tier. In order to realise this, a new mobility management approach is needed, in which

UEs only take measurements and access the cells of the appropriate network tier according to their velocity.

Modulation and Coding Schemes: Deploying higher order modulation and coding schemes is critical to take advantage of the high SINRs resulting from ultra-dense small cell deployments. Even higher modulation schemes than currently used in LTE and Wireless Fidelity (WiFi), i.e., 256-QAM, may be required, e.g., 1024-QAM. However, this brings about the need for accurate channel state information for coherent de-modulation. However, the implementation feasibility of 1024 or higher QAMs is still unclear due to the Error Vector Magnitude (EVM) issues at transmitters [39]. In addition, the Peak-to-Average Power Ration (PAPR) problem should also be re-considered for 1024 or higher QAMs.

Radio Resource Management: In terms of radio resource management, current scheduling and other network procedures have to be revisited since due to the lower number of UEs per cell, the current approaches used in macrocell may not be optimum anymore. For example, proportional fair scheduling may not be the most efficient solution for very small cells, since there are not many UEs to be fairly served and channel fluctuation may be low due to LOS channel conditions.

2.1.1 Mobility management issues in DenseNets

Mobility Management (MM), in the presence of femtocells, is one of the most challenging issues, owing to the dense network layout, the short cell radii and the potentially unplanned deployment. The key challenges of MM support for femtocells are posed during the phases of a) cell identification, b) access control, c) cell search, d) cell selection/reselection, e) HO decision, and f) HO execution.

Cell selection/reselection is a critical issue in large-scale deployments of femtocells, where the tracking area size has a major impact on the user equipment (UE) battery lifetime and the network signaling load. More sophisticated HO decision algorithms are also required, in the presence of femtocells, to mitigate the negative impact of user mobility and cross-tier interference on the Quality of Experience (QoE) and Signal to Interference plus

Noise Ratio (SINR) performance at the UEs. Attaining a low service interruption probability for medium to high speed users is another challenging issue for the HO decision phase. Certain network architectural and procedural enhancements are also required to lower the delay and signaling overhead of the HO execution to/from femtocells.

This section discusses the open issues for MM support in the presence of femtocells and overviews the key aspects of MM in the LTE-A system.

Support of femtocells necessitates the deployment of certain network architectural and procedural enhancements in the cellular system. In the context of LTE-A, a macrocell station is referred to as E-UTRAN Node B (eNB) and a femtocell station as Home eNB (HeNB). Two of the evolved packet core (EPC) network entities are involved in the support of HeNBs: the Mobility Management Entity (MME) and the Serving Gateway (S-GW) [40].

In the presence of femtocells, the E-UTRAN architecture consists of eNBs, HeNBs, and HeNB gateways (HeNB GW). The eNBs provide user and control plane protocol terminations towards the UE, while they support the functions of radio resource management, admission control, scheduling and transmission of paging/broadcast messages, measurement configuration for mobility and scheduling, as well as routing of user plane data towards the S-GW. The functions supported by the HeNBs are the same as those supported by the eNBs, while the same implies for the procedures run between the HeNBs and the EPC.

Support of femtocells necessitates the deployment of more complicated MM procedures. The dense network layout and the short cell radii augment the negative impact of user mobility, enlarge the number of candidate cells during the HO decision phase and increase the HO probability even for low speed users.

The decision part of a cell HO is referred to as the HO decision phase, while the signaling part as the HO execution phase. In prominent cellular networks, the HO decision phase is performed at the serving cell and is based on signal quality measurements provided by the UE, i.e., UE-assisted network-controlled HO [41].

The impact of the HO decision phase is even more prominent in the presence of femtocells, owing to the short-range nature of communications, the denser network layout and the fast varying radio environment.

The HO decision can also be used to improve the energy-efficiency of the network nodes and handle the interference in a macroscopic level, i.e., without using power control, radio resource or interference management. Current literature includes various HO decision algorithms for the two-tier macrocell-femtocell network [42]- [43]. The vast majority of existing algorithms prioritize femtocell over macrocell access based on signal strength [44] [45], [46], UE speed [47] [48], or traffic-type criteria [49], [50]. In most of the cases, the impact of the HO algorithms on the energy consumption, interference, system capacity and network signalling is not investigated.

Attaining a good performance trade-off between exploiting the femtocell utilization opportunities and sustaining a low HO probability is another critical issue. The joint optimization of the interference and energy consumption performance at the network nodes should be integrated within the HO decision phase as well.

Current literature includes various HO decision criteria and parameters for the two-tier macrocell-femtocell network [51]. Below, we describe the most widely used.

Received Signal Strength (RSS) refers to the received power on the reference or pilot signals transmitted by a specific cell [42] [52].

Received interference power (RIP) refers to the total received power from cells or users in proximity. When performed at the UE, the RIP measurement is usually referred to as the Received Signal Strength Indicator (RSSI) [43].

Received Signal Quality (RSQ) refers to the ratio of the RSS from a target cell to the total RIP at the UE [53] – [46].

UE speed is a widely used parameter for enhancing inbound mobility to femtocells and reducing the number of unnecessary HOs for medium to high speed users [47] [50].

Some of the key *energy-efficiency* parameters in current literature are the UE battery power [54], the mean UE transmit power [52], and the UE power consumption [43].

RS transmit power corresponds to the cell transmit power on the RS. Existing algorithms use this parameter to assess the path loss between the UE and the target cell [52].

Traffic-type: Existing classifications of the UE traffic-type mainly include: a) real time or non-real time traffic [49], [50], [55], [54], and b) voice/video or data traffic [56].

Available bandwidth is a measure of the resource availability in the target cell. This parameter is used to minimize the HO failure probability due to admission control. Other bandwidth-related parameters include the cell load [55] and the cell capacity [48].

UE residence time within the cell refers to the duration that a tagged UE is expected to remain within the coverage of a cell. This parameter is used in combination with other speed-related parameters to minimize the number of unnecessary HOs [57].

Finally, a classification of HO algorithm is carried out according to the following five groups [51]: a) received signal strength based, b) speed based, c) cost-function based, d) interference-aware, and e) energy-efficient.

2.2 Mobility-aware Energy-Quality Trade-off for Video Delivery in Dense Heterogeneous Networks

The deployment of small-range base stations offers increased coverage and user capacity as well as higher throughput and lower transmission power for the users. This is as the mobile devices will be closer to the base stations. This helps reducing power consumption, one of the key challenges in the next generation mobile multimedia networks. On network side, DenseNet allows to offload traffic from the macro cell and solve any potential coverage problems, supporting also better capacity in terms of amount of traffic.

In the context of a DenseNet, the network selection gives great importance to improving balance between quality of experience (QoE) of the video service offered to the user and energy saving. In [58] authors proposed a hybrid multimedia delivery solution, which balances the benefits of multimedia content adaptation and of network selection in order to decrease power consumption in a heterogeneous wireless network environment, composed of UMTS and WLAN. Trade-off between energy and quality has been considered via a utility-based function. A similar approach has been used in [59], where the authors employ network reputation in a utility-based network selection mechanism. The proposed solution uses user preferences and service requirements to define a network reputation factor, which reflects user satisfaction with the network service quality provided to the mobile user. A DenseNet scenario introduces additional issues. Since the femtocell and WLAN coverage is small, mobile users could experience several unnecessary handovers with consequent reductions in terms of user QoE and system capacity. Hence, user mobility pattern and speed needs to be taken into account [48] [60]. The authors of [48] proposed a handover algorithm based on the user speed and quality of service (QoS). They considered a femtocell with a small coverage where a user with high speed crosses the femtocell in a short time. In these conditions, the authors suggested that users with high speed do not need to make handover, especially when non-real-time services are taken into account. Nevertheless, they did not consider the energy saving issue. An energy efficient handover algorithm is proposed in [60] with the aim to reduce power consumption and frequent and unnecessary handovers. Users' speed is taken into consideration in order to allow handover only to slow users. On the other hand, power saving is accomplished by decreasing the femtocell power transmission in particular conditions. However, the energy management proposed in [60] is only network-side, and does not consider mobile device power consumption, very important for users.

This chapter introduces an innovative user mobility-aware utility-based network selection, which balances energy consumption and quality for video deliveries in dense heterogeneous network environments. The proposed solution considers the estimated energy consumption of the mobile device when running real-time video applications, estimated network conditions and speed of users.

2.2.1 Scenario and System Model

Fig. 2.1 illustrates a DenseNet scenario, where a user moving from home to office passes through different network coverages. Points A, B, C, D, represent different situations in which the mobile user needs to select and connect to the most appropriate network in order to avail from the video services and best balance energy and quality in the given conditions.

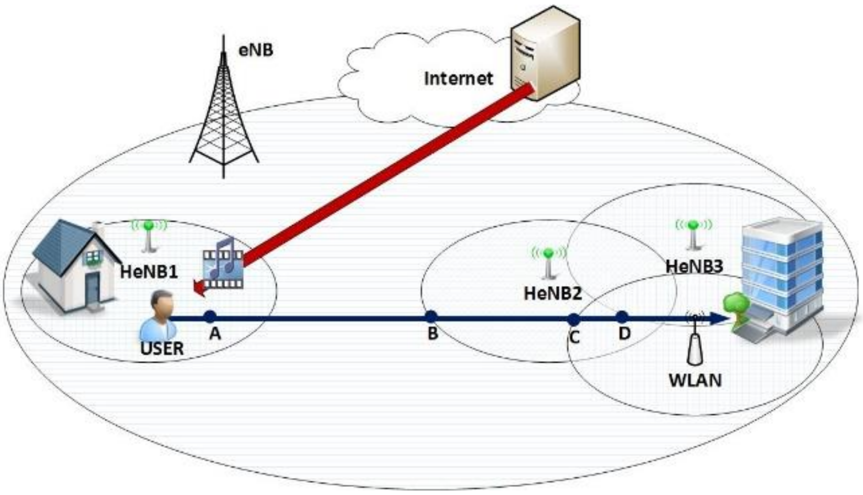
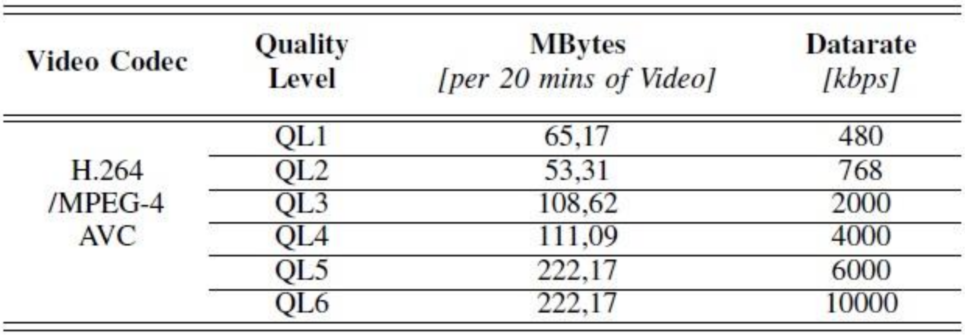


Fig. 2.1 - Example of a mobile user in a DenseNet environment

We consider a wireless network scenario where different types of small networks (the term cell is also used in this paper) (e.g., femtocell, WiFi hotspot, etc.) are deployed in an uncoordinated manner within a macro cellular coverage, as shown in Fig. 2.1. We denote with C_M and C_m the set of C macrocells and c small cells within the considered scenario, respectively. In particular, let us denote with $c \in C = C_M \cup C_m$ the index of a generic cell and with $C_a = C_1, \dots, C_n$ the set of available cells for the user i. Since the handover decision measurements are performed in the downlink direction, we focus on the transmission from the BS of the generic cell c to a generic UE i.

The useful received power by the user i at a generic distance d from the BS c can be expressed as $P_{R_{c,i}}(d) = P_{T_{c,i}} \cdot h_{c,i}(d)$, where $P_{T_{c,i}}$ is the transmitted power from BS c and $h_{c,i}$ the channel gain from BS c to user i located at distance d. In the channel gain coefficient are included all the losses due to the path loss attenuation, shadowing and other factors such as fading and multipath. A dual approach is considered involving network selection and adaptation. Network selection is accomplished through computing of a Utility function U (eq. (1)) that takes into account energy consumption of the mobile device when running real-time video applications, estimated network conditions and speed of users. Adaptation is performed based on a Datarate Quality Mapping Table, which includes the bandwidth/datarates b_{r_l} required to receive the video content at *l*-th Quality Level (QL). Table 4 is an illustration of such a table which has six quality levels.



Video Codec	Quality Level	MBytes *[per 20 mins of Video]*	Datarate *[kbps]*
H.264 /MPEG-4 AVC	QL1	65,17	480
	QL2	53,31	768
	QL3	108,62	2000
	QL4	111,09	4000
	QL5	222,17	6000
	QL6	222,17	10000

Table 4 - Datarate Quality Mapping Table

Network selection is executed according to the utility function defined for each Radio Access Network (RAN) *i* by the following equation:

$$U_i = u_{e_i}^{\omega_e} \times u_{q_i}^{\omega_q} \times u_{s_i}^{\omega_s} \qquad (1)$$

In equation (1) U_i is the overall score function for RAN i and u_{ei}, u_{qi}, and u_{si} are the utility functions defined for device energy consumption, video service quality, and user speed, respectively. w_e, w_q, and w_s are weights for the considered criteria, representing the importance of a parameter in the decision algorithm, where $w_e + w_q + w_s = 1$.

The estimated energy consumption for a real-time application is computed using eq. (2) as defined in [61]]:

$$E = t\,(r_t + Th_{req} r_d) \qquad (2)$$

where t represents the transaction time, which can be estimated from the duration of the video stream; r_t is the mobile device's energy consumption per unit of time (W), Th_{req} is the required throughput (kbps), r_d is energy consumption rate for data/received stream (J/Kbyte), and E is the total energy consumed (J). The parameters r_t and r_d can be determined by running different measurements for various amounts of data and defining an energy consumption pattern for each interface (LTE, WiFi). They were determined by running different simulations for various amounts of multimedia data (i.e., quality levels) while measuring the corresponding energy levels and then used to define the energy consumption pattern for each interface/scenario [62]. Based on the estimated energy consumption E, the utility for the energy criteria u_e is computed using Eq. (3) [62]:

$$u_e\,(E) = \begin{cases} 1 & , E < E_{min} \\ \frac{E_{max} - E}{E_{max} - E_{min}} & , E_{min} \leq E < E_{max} \\ 0 & , otherwise \end{cases} \tag{3}$$

A zone-based quality sigmoid utility function [59] is used to map the throughput to user satisfaction.

$$u_q\,(Th) = \begin{cases} 0 & , Th < Th_{min} \\ 1 - e^{\frac{-\alpha * Th^2}{\beta + Th}} & , Th_{min} \leq Th < Th_{max} \\ 1 & , otherwise \end{cases} \tag{4}$$

The minimum throughput (Th_{min}) is a threshold to maintain the multimedia service at a minimum acceptable quality level, values below this threshold result in unacceptable quality levels. Whereas values above the maximum throughput (Th_{max}) threshold will not add any noticeable improvements in the user perceived quality.

The mathematical definition of the speed utility is given in eq. (5):

$$u_s(S) = \begin{cases} 1 & , S < S_{min} \\ \frac{S_{max} - S}{S_{max} - S_{min}} & , S_{min} \leq S < S_{max} \\ 0 & , otherwise \end{cases} \tag{5}$$

subject to $u_s = 1$, if $i \in C_M$

where S_{min} is the pedestrian speed, i.e. 3 km/h and S_{max}, the urban vehicular speed limit (in Italy 50 km/h). This utility considers urban dense networks only. Equation 5 does not affect the overall utility function if the target cell is a Macro cell. This is due to the fact that the cell range is large and UEs with high mobility should not perceive differences in their transmissions.

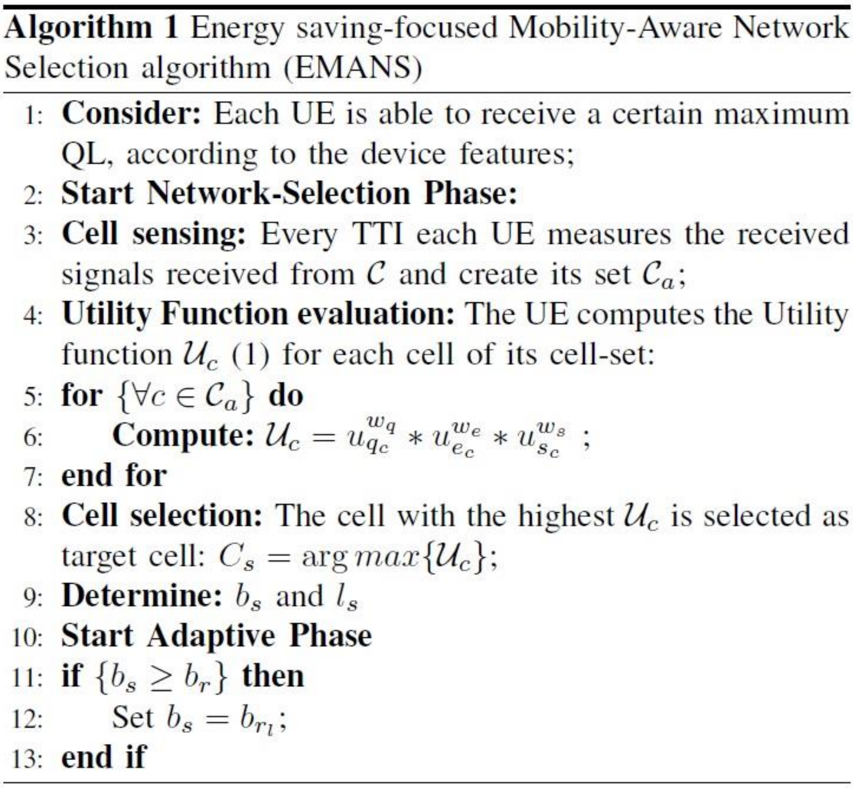
The utility function from eq. (1) is computed for each of the candidate networks and the network with the highest score is selected as the target network.

2.2.2 EMANS algorithm

The proposed Energy saving-focused Mobility-Aware Network Selection algorithm (EMANS) [63] consists of two phases: *(i) network-selection phase and (ii) datarate adaptation phase.* A more in-depth description of the algorithm in provided in Algorithm 1 pseudo-code.

During the network-selection phase, at every TTI the UE creates its cell-set according to the data received from the cells (line 3). Lines 3-8 describe the network-selection phase of the algorithm. First the UE computes the utility function for each cell, according to the utility equation 1. Then, the cell providing the highest utility U_c is set as the selected cell C_s. Next the algorithm sets the datarate provided by the selected cell C_s as serving datarate b_s. Then, b_s is used to find the received quality level ls according to the Datarate Quality Mapping Table (i.e. Table 4) (line 9).

Since the algorithm takes into account the features of the user's device (line 1), once determined b_s the adaptive phase takes place (lines 10-12). If the datarate received b_s is greater than the minimum datarate b_{r_l} that allows to receive the QL required by the UE, the algorithm sets the new datarate as b_{r_l}. This potential decrease of the datarate does not cause a significant decrease of QoE received by the user, enabling also to save energy.



Algorithm 1 Energy saving-focused Mobility-Aware Network Selection algorithm (EMANS)

```
Consider: Each UE is able to receive a certain maximum QL, according to the device features;
Start Network-Selection Phase:
Cell sensing: Every TTI each UE measures the received signals received from $\mathcal{C}$ and create its set $\mathcal{C}_a$;
Utility Function evaluation: The UE computes the Utility function $\mathcal{U}_c$ (1) for each cell of its cell-set:
for $\{\forall c \in \mathcal{C}_a\}$ do
    Compute: $\mathcal{U}_c = u_{q_c}^{w_q} * u_{e_c}^{w_e} * u_{s_c}^{w_s}$ ;
end for
Cell selection: The cell with the highest $\mathcal{U}_c$ is selected as target cell: $C_s = \arg max\{\mathcal{U}_c\}$;
Determine: $b_s$ and $l_s$
Start Adaptive Phase
if $\{b_s \geq b_r\}$ then
    Set $b_s = b_{r_l}$;
end if
```

In order to investigate the benefit of the proposed EMANS algorithm, Matlab simulations have considered a reference scenario where several small cells (LTE femtocells, WiFi Hotspots) are deployed within the coverage of a LTE macrocell.

A dense urban scenario was considered where users are free to move with different speeds from 3 km/h to 60 km/h. The simulations are carried out in a time interval of 20 minutes, with users downloading a real-time video. Video features are described in Table 4. According to the datarate of the selected cell, a corresponding QL for the video delivery is used.

Algorithm performance was compared with that of E-PoFANS [64] and with a classic algorithm transmitting at constant bit rate (CBR), labelled “Conservative”. The “Conservative” approach serves all users with the lowest quality level. A conservative approach delivering the highest quality level at CBR was not tested because loss would severely affect the overall quality.

Different user requirements are also considered and therefore, different user devices are taken into consideration in order to better show the outperforming behaviour of the proposed algorithm. In particular, two cases are considered: (i) users requiring a QL2 service and (ii) users requiring a QL3 service.

Furthermore an analysis with different speeds is also carried out, in order to show the benefit of the proposed solution, especially in terms of unnecessary handovers reduction.

Simulation results in low-speed condition (i.e. 3 km/h) are presented as following.

Fig. 2.2Fig. 2.2 shows the datarate received by users in the both considered cases (QL2 and QL3 respectively). In both cases E-PoFANS has higher bitrate, as it neither takes into consideration the mobility of users, nor the effective user quality requirements. At the same time, EMANS bitrate never exceeds the value associated to the target quality level (QL).

Note, the lower datarate achieved by our proposed algorithm does not significantly affect the QoE as perceived by users. Fig. 2.4 shows that EMANS provide to users the same QLs as E-PoFANS with QL3 as target level (Fig. 2.3). The better quality provided by E-PoFANS in Fig. 2.2 is not very useful since it is greater than the maximum quality supported by the user's device.

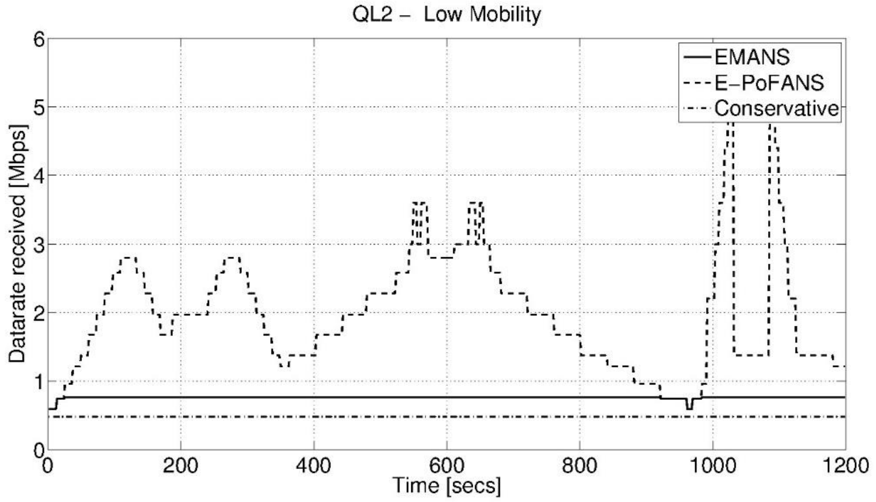


Fig. 2.2 - Datarate Received in Low-Mobility Scenario (QL2)

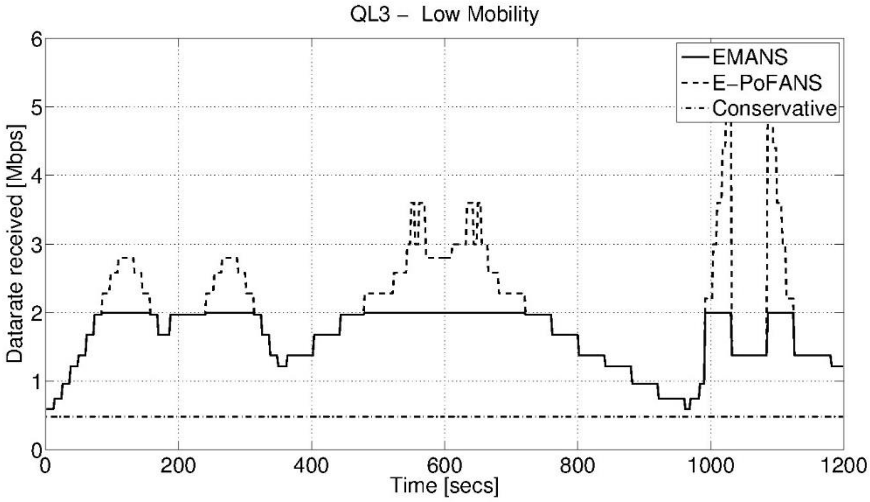


Fig. 2.3 - Datarate Received in Low-Mobility Scenario (QL 3)

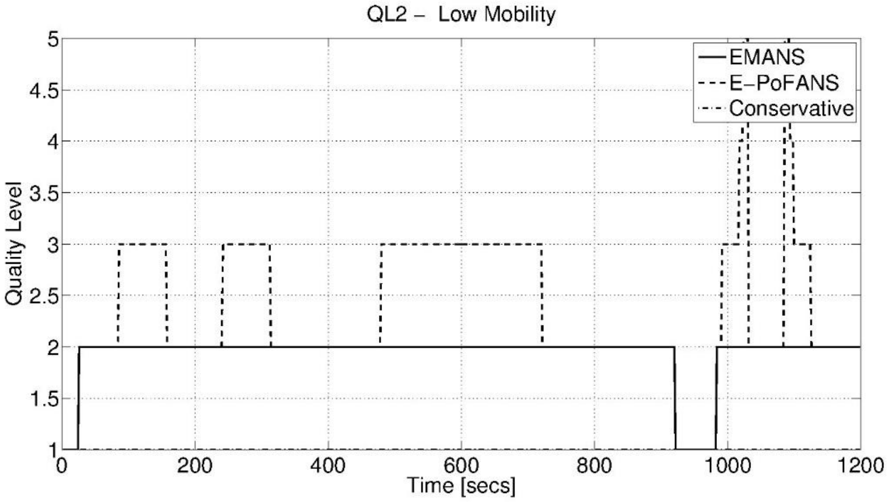


Fig. 2.4 - Quality Level Perceived in Low-Mobility Scenario (QL 2)

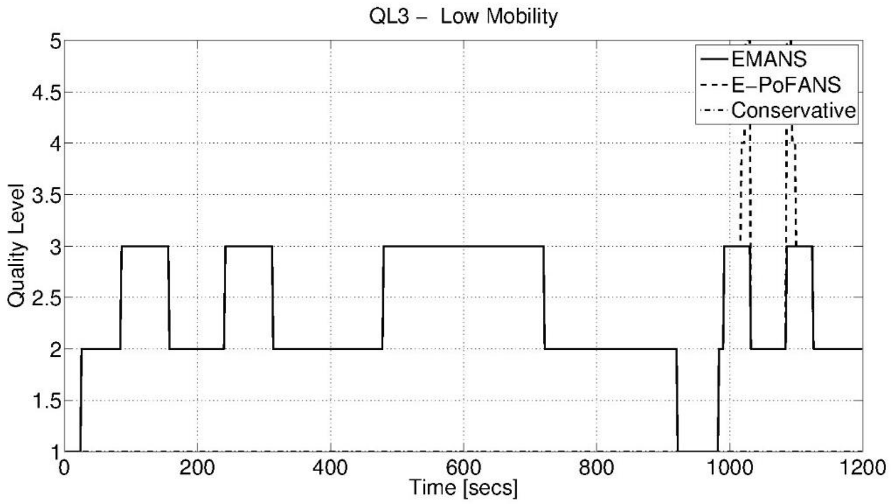


Fig. 2.5 - Quality Level Perceived in Low-Mobility Scenario (QL 3)

The above considerations show how EMANS achieves better performance in terms of network resources (see Fig. 2.6 - Fig. 2.7). In the considered cases, the proposed algorithm never overcomes the red line that represents the optimal delivered datarate, meaning that resources are not wasted because no more than the required bitrate is transmitted. However, the reduction of the datarate is also associated with significant energy saving achieved by our proposed EMANS (Fig. 2.8 - Fig. 2.9), as we intended to demonstrate.

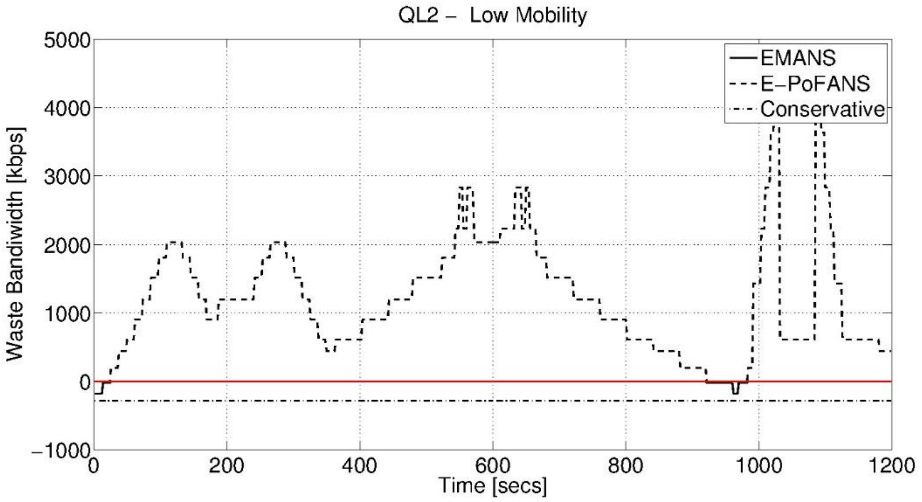


Fig. 2.6 - Bandwidth Utilization in Low-Mobility Scenario (QL 2)

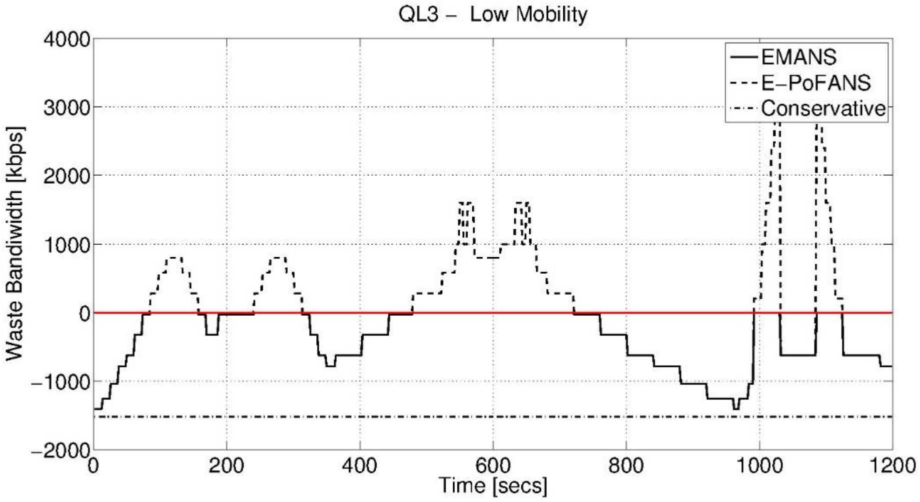


Fig. 2.7 - Bandwidth Utilization in Low-Mobility Scenario (QL 3)

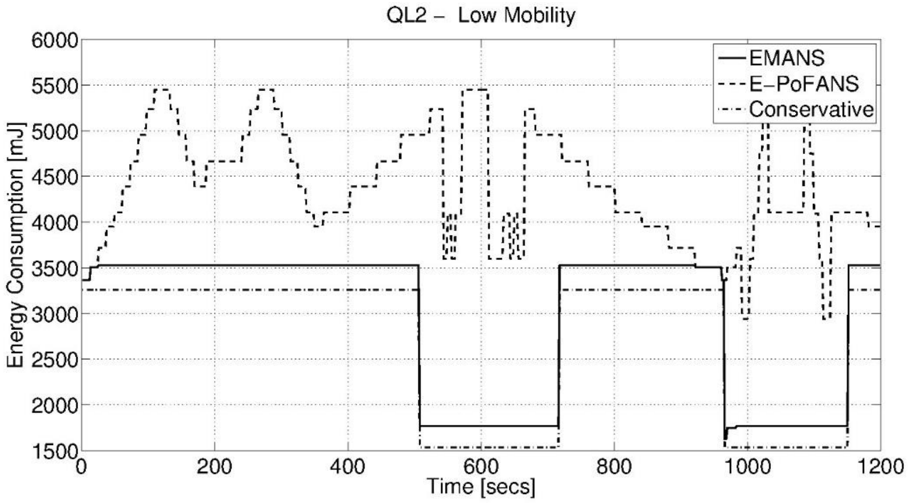


Fig. 2.8 - Energy Consumption in Low-Mobility Scenario (QL 2)

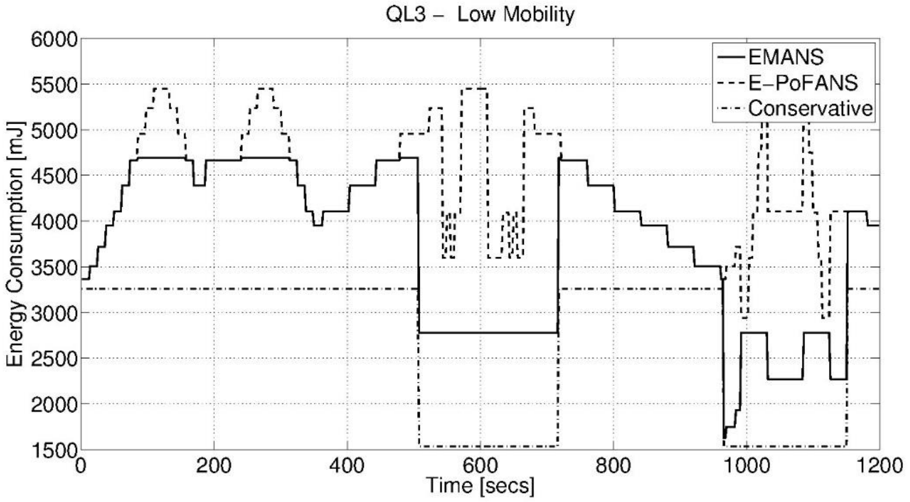


Fig. 2.9 - Energy Consumption in Low-Mobility Scenario (QL 3)

Last considerations refer to the behaviour of the EMANS algorithm in a high-mobility scenario. Datarate, energy consumption and bandwidth utilisation maintain the same trend as in the low-mobility case. There is an expected little loss in terms of Quality levels (Fig. 2.10) but there is a considerable gain in terms of number of handovers: EMANS has 36 % of the number of handovers performed by E-PoFANS as showed in Table 7. Other numerical results are summarized in the Table 5 and Table 6.

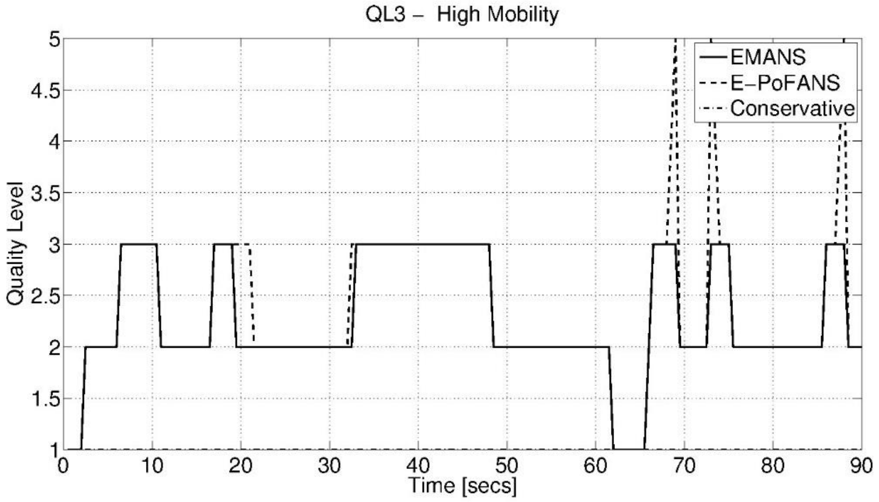


Fig. 2.10 - Quality Level Perceived In High-Mobility Scenario (QL 3)

Target quality level: QL2 *[768 kbps]*			
	EMANS	E-PoFANS	Conservative
Datarate [Mbps]	0.764	1.974	0.480
Quality Level	1.93	2.35	1
Energy Consumption [mJ]	2945.70	4385.77	2689.73
Difference to the Target [%]	-0.55 %	+157%	-37%
No. Handovers	22	34	13

Table 5 - Summary Results (QL2)

Target quality level: QL3 *[2000 kbps]*			
	EMANS	E-PoFANS	Conservative
Datarate [Mbps]	1.65	1.974	0.480
Quality Level	2.31	2.35	1
Energy Consumption [mJ]	3740.55	4385.77	2689.73
Difference to the Target [%]	-17 %	-0.01 %	-76 %
No. Handovers	22	34	13

Table 6 - Summary Results (QL3)

Speed [km/h]	*3*	*7*	*15*	*30*	*45*	*60*
EMANS	-36 %	-53 %	-64 %	-60 %	-64 %	-45 %
E-PoFANS	34	40	112	200	330	397
Conservative	-62 %	-45 %	-42 %	-35 %	-41 %	-35 %

Table 7 - Summary Results: Number Of Handovers

Finally, testing the EMANS algorithm in low- and high-mobility scenarios has demonstrated that when employing similar user perceived quality levels are achieved in comparison with other solutions, while EMANS results in higher energy and bandwidth saving. Furthermore, testing has also showed that EMANS achieves better performance in terms of the number of handovers.

The proposed EMANS demonstrated that a proper network selection in a Dense Heterogeneous scenario is helpful in order to enhance the system performance. Nevertheless, the management of radio resources is still one of the main issues faced by multicasting in forthcoming 5G systems.

Therefore, in next section, we describe a network selection approach that takes into consideration both multicast groups and radio resource management.

2.3 Hybrid unicast-multicast utility-based network selection algorithm

In the view of an increased capacity and improved performance of the system, recent researches push towards the deployment, within the same area, of several coverage layers (associated to macro, micro, pico, and femto cells), diverse Radio Access Technologies (RAT) (e.g. GSM, UMTS, LTE, WiFi), and multiple Point-to-Point (PtP) user links (e.g. Device-to-Device communication [65], mmWave). This massive growth of dissimilar cell deployments is leading to a high densification of networks and to the creation of the so-called Dense Heterogeneous Network (DenseNet) [22] paradigm.

Moreover, radio resources management (RRM) is stressed by the huge number of smart devices requiring video services. In this scenario, device-to-device (D2D) communications [66] and multicast services over current LTE and future 5G systems [11] have been considered as possible enabling approaches to efficiently manage the traffic load and provide a better Quality of Experience (QoE) to end-users. In particular, multicasting allows a large number of users to be simultaneously served with relatively low latency and high throughput. To support such services, the Third Generation Partnership Project (3GPP) offers basic support to the standardization of multicast services over LTE under the name of enhanced Multimedia Broadcast Multicast Services (eMBMS) [67].

One of the most important issues for multicast transmission is the management of multi-user diversity. In fact, each user within a multicast group experiences a different channel quality level. Least channel gain users affect the performance of the whole multicast group as they can only support a transmission with low Modulation and Coding Scheme (MCS) level, thus achieving transmissions with bad spectral efficiency.

On the contrary, serving multicast users that experience high channel quality levels improves the system spectral efficiency, at the expense of users under bad channel conditions. This introduces challenging issues for the RRM in multicast transmissions.

Our research focuses on a DenseNet deployment scenario characterized by overlapping of an LTE-A macro cell and LTE-A small cells (i.e., femtocells), in the presence of multicast groups in each cell. In this scenario, mobile users want to access video content at high user QoE levels and with a low energy consumption. Indeed, energy/power management as well as user mobility management are key challenges in the next generation mobile multimedia networks [68]. Innovative RAT selection [69] solutions help in managing energy issues for smart users in mobility. However, mobile users have to face the issue of the wise selection of the access network to connect to, which is made more challenging by the highly dynamic network environment [70]. In particular, in a DenseNet context, network selection should place great importance to improving the balance between QoE of the video service offered to the user and energy saving [64]. Furthermore, a proper management of network selection is needed, in order to avoid issues such as frequent and unnecessary handovers (i.e. the so-called ping-pong effect).

Access network selection schemes can be user-centric or network-driven. In the user-centric approaches, the focus is on maximizing user QoE levels. However, this presents several limitations as users are only aware of their link quality, and have no information about the network load, which could clearly affect user perceived quality and induce instability due to frequent handovers. In the network-driven solutions, the objective is to maximize the network operator revenues and maintain high overall user satisfaction by avoiding network congestion and by selecting the optimum interface for each user.

In this context, there is a need for a resource allocation mechanism to provide the highest available performance to the largest number of users possible. The presence of a multicast transmissions helps to obtain such a requirement, but at the cost of a compromise in terms of data-rate achieved by users within the multicast group. Generally, the methodology of resource allocation is to model it as an optimization problem whose objective function and constraints are determined by user requirements and network specifications. The objective function is usually referred to as utility function, which characterizes a user satisfaction when allocated given resources [71].

In this chapter the Hybrid Unicast-Multicast utility based Network Selection algorithm (HUMANS) is proposed, a network selection approach that exploits the benefits of multicast. In such an approach both bandwidth utilization (an operator priority) and the trade-off between quality and energy consumption (a user priority) are considered in deciding how to deliver video content in a DenseNet. HUMANS, in taking access-related choices, considers the estimated energy consumption of the mobile device running a real-time video application, the estimated achievable data-rate, the utilized resources and the expected user satisfaction level.

A major contribution is, thus, a mechanism allowing for a wise network selection choice, also considering a multicast group joining option, which at the same time meets users exigencies and enables a smart bandwidth management.

2.3.1 Literature review

On the one hand, the 5G DenseNet environment will provide increasing coverage and system capacity with respect to the current cellular networks. On the other hand, the DenseNet's associated higher complexity exacerbates problems of interference coordination, power consumption, RRM and mobility management. In such a DenseNet scenario, there is a need for proper Network selection and resource allocation in order to meet both 5G requirements and user and market expectations in terms of, high QoE levels, increased power saving, reduced cost, etc. State-of-the art related to our research is discussed next, from the perspectives of network selection solutions and RRM algorithms in multicast transmissions.

A hybrid multimedia delivery solution which balances the benefits of multimedia content adaptation and network selection in order to decrease power consumption in a heterogeneous wireless network environment, composed of UMTS and WLAN networks, was proposed in [64]. The trade-off between energy and quality has been considered via a utility function. Similar approaches have been introduced in [72] and [73]. In [72] authors propose an adaptive real-time Multi-user access network selection load balancing algorithm, taking into account not only the real-time global traffic load on each network, but also considering the different classes of traffic.

Whereas, the solution proposed in [73] combines several inputs such as power of the received signal, throughput, packet delay, cost-per-user, the requested type of traffic, and type of device.

In [74], authors propose a network selection solution based on a novel algorithm, which relies on the concept of Fittingness Factor (FF). The novel solution maximizes a function that reflects the suitability of the available spectrum resources to the application requirements. The selection is carried out by taking into account specific parameters and QoS metrics. The suitability of a network is determined by using the data bit rate required by the new flow and the bit rate that the network can support.

Furthermore, in a DenseNet scenario with several small cells deployed, users are moving near the small cells and enter and exit in/from their coverage area with high frequency. This introduces additional issues such as unnecessary handovers with consequent reductions in terms of user QoE and system capacity. Authors in [75] propose a RAT selection algorithm that efficiently manages the RAT handover procedure by (i) choosing the most suitable RAT that guarantees high system and user performance, and (ii) reducing unnecessary handover events. They introduce a parameter named Reference Base Station Efficiency that considers the BS transmitted power, BS traffic load and user spectral efficiency.

A different approach to avoid unnecessary handovers is a user mobility-aware technique that takes into account users' speed [48] [60]. The authors of [48] proposed a handover algorithm based on the user speed and QoS. The authors suggested that users with high speed do not need to handover, as they cross the coverage area fast and especially when avail from non-real-time services, as this is inefficient. Nevertheless, they did not consider any energy saving issue. An energy efficient handover algorithm is proposed in [60] with the aim to reduce power consumption and frequent and unnecessary handovers. Users' speed is accounted for in order to allow only slow users performing handover. On the other hand, power saving is accomplished by decreasing the femtocell power transmission in particular conditions. However, the energy management proposed in [60] is network-side only, and does not consider mobile device power consumption, a key aspect for users.

EMANS [63], instead, proposes an energy-saving network selection algorithm, which provides a good trade-off between energy consumption and perceived quality when delivering video content. EMANS includes a method to adapt the delivered video stream bitrate according to the available network resources such as maintaining good user perceived quality levels. Furthermore, it also reduces the number of handovers in comparison with other state-of-the-art approaches.

All the works presented above deals with unicast transmission. Nevertheless, a dense 5G scenario should take into account also group-oriented transmissions. In such a solution, the selection of the most proper MCS with which serve all users is a challenging issue. A typical solution is represented

by the conventional multicast scheme (CMS) where all the users within a multicast group are served with the lowest level MCS, representing users with worst channel condition [76].

The opportunistic multicasting [77] has been proposed in literature as a possible solution to overcome the typical limitations of the conservative approach and to efficiently exploit multi-user diversity, thus providing a more effective selection of the MCS based on the users channel information. CMS

and OMS are both single-rate transmission modes, where the BS transmits to all users in each multicast group at the same rate. In Multi-rate, instead, the BS transmits to each user at different rates exploiting users frequency diversity, according to the heterogeneity of wireless channel.

The work presented in [78] optimally forms multicast groups, based on the users data rate. Whereas, the authors of [79] propose an approach for Single-Frequency Networks aiming to increase the aggregate datarate of the multicast group by pushing out of the transmission bad channel users, which are served through unicast transmissions. Nevertheless, differently from our work, this approach does not account for resource utilization and, like some other innovative works, may cause waste of resources. In a 5G scenario, where several users require high quality services, a big issue is the limited availability of radio resources. Multicast transmissions have become a solution for both increasing network capacity and improving spectral efficiency. Hybrid unicast-multicast approaches [80] can provide an efficient radio resource exploitation. Differently from previous works, this paper introduces a utility-based network selection algorithm, which takes into consideration hybrid unicast-multicast transmissions and balances energy consumption and quality for video deliveries in DenseNets.

Besides taking into account the trade-off between throughput and estimated energy consumption of the mobile device, the selection of the network is also affected by the radio resources required by the users, in order to achieve an efficient usage of radio spectrum. In particular, the approach proposed considers that users with good channel conditions, which consequently need less resources, could be served via unicast, whereas users with bad channels can be served via multicast.

2.3.2 Scenario and System Model

The reference scenario consists of a DenseNet scenario, represented by a LTE base station (eNB) and several small-range LTE femtocells (HeNB) under the same coverage area (Fig. 2.11).

Multicast flows are activated within each cell belonging to the reference area. Users within this area access multimedia video content and pass through different cell coverages. In each overlapping point users need to select the most appropriate network to connect to.

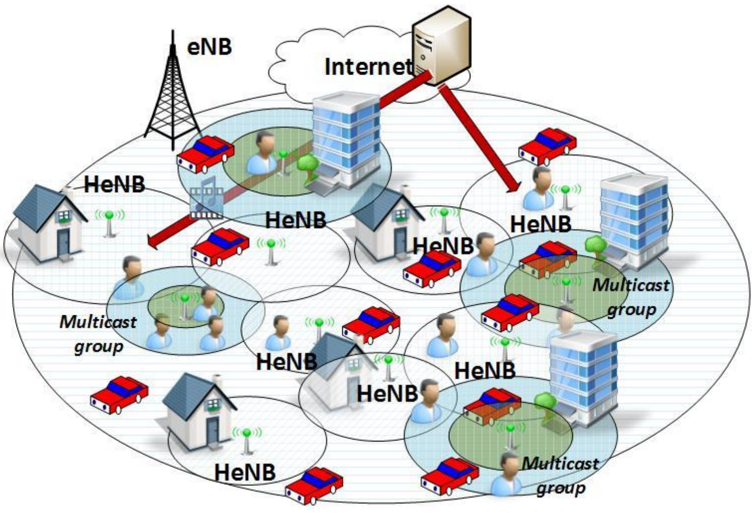


Fig. 2.11 - DenseNet environment with the presence of multicast groups

In LTE systems [40], Orthogonal Frequency Multiple Access (OFDMA) and single carrier frequency division multiple access (SC-FDMA) are used to access the downlink and the uplink, respectively. The available radio spectrum is split into several *Resource Blocks* (RBs) and, in the *frequency domain*, each RB corresponds to 12 consecutive and equally spaced sub-carriers. One RB is the smallest frequency resource that can be assigned to a user equipment (UE). The overall number of available RBs depends on the system bandwidth and can vary from 6 (1.4 MHz channel bandwidth) to 100 (20 Mhz). The eNodeB (eNB), which is the node that communicates with UEs, is in charge to assign the adequate number of RBs to each user. The packet scheduler properly manages the transmission parameters and the allocation of the $\boldsymbol{B}$ RBs according to the Channel Quality Indicator (CQI) feedbacks received from the users. Based on the CQI received by each user, the transmission from the BS to the user is set with a given MCS. For each MCS level, a certain spectral efficiency is achieved by the transmission. The greater is the spectral efficiency, the lower is the number of RBs required to achieve a given datarate. It worth specifying that depending on the spectral efficiency guaranteed by the MCS assigned to that transmission, the frequency scheduler has to decide how many RBs should be assigned to the user.

In case of multicast service, it is typically the UE that experiences the worst CQI that drives the MCS selection for the multicast transmission. It means that the multicast flow is delivered with very low spectral efficiency. On the other hand, during a multicast session all the bandwidth dedicated for the MBMS service could be assigned to the multicast transmission. Furthermore, it is worth noting that, according to the eMBMS standard [67], at least 40% of whole available bandwidth has to be dedicated to unicast transmissions.

We consider a wireless network scenario where different types of small networks, e.g. femtocells, are deployed in an uncoordinated manner within a macro cellular coverage, as shown in Fig. 2.11.

Let $U = \{u_i | i = 1, \dots, n\}$ the set of Users and $C = \left\{C_j | j = 1, \dots, c\right\}$ is the set of all cells of the scenario, and each cell C_j can be either a eNodeB (i.e. a macrocell) or a HeNb (i.e. a small cell). Since the handover decision measurements are performed in the downlink direction, we focus on the transmission from a the generic cell C_j to a generic UE u_i.

τ is the time interval (TTI) in between regular system updates. Every τ each *i*-th UE u_i collects measurements from all cells which it is able to sense.

Network selection is then accomplished through computing of a utility function *U* (eq. 6) that takes into account the energy consumption of the mobile device when running real-time video applications, estimated network conditions, utilized resources and estimated user's satisfaction level.

The proposed Hybrid Unicast-Multicast utility-based Network Selection algorithm (HUMANS) is designed for DenseNet scenarios and is based on appropriate network selection carried out by users. Fig. 2.12 presents a step-wise description of the algorithm phases. Each user first senses the neighbour cells and send the CQI of the respective downlink channel to all of them. According to the received CQI, each cell selects the most appropriate MCS level for the user, and announces the multicast service (eMBMS Service announcement).

In such message is also included the MCS level of the multicast group. According to such information, the user performs the network selection as follow.

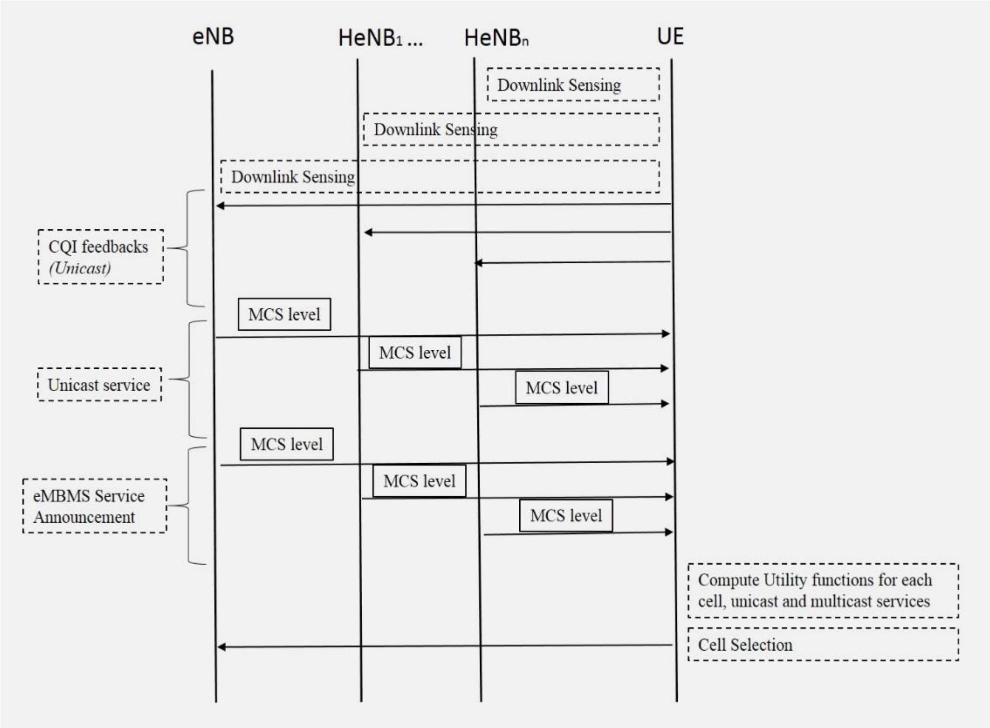


Fig. 2.12 - "HUMANS" Procedures

Network selection is executed according to the utility function defined for each Radio Access Network (RAN) *j* by the following equation.

$$U_i = u_{e_i}^{\omega_e} \times u_{q_i}^{\omega_q} \times u_{s_i}^{\omega_s} \times u_{b_i}^{\omega_b} \qquad (6)$$

Where U_i is the overall score function for RAN *j* and u_{e_i}, u_{q_i}, u_{s_i}, u_{b_i} are the utility functions defined for video service quality, device energy consumption, user satisfaction and radio resource usage, respectively. $w_e, w_q, w_s, w_b,$ are weights for the considered criteria, representing the importance of the associated parameter in the decision algorithm, where $w_e + w_q + w_s + w_b = 1$.

The equation (1) have been properly modified in order to exploit the benefit introduced by group-oriented communications.

The novelty of the proposed approach is that it takes into account the multicast transmission as an additional option during the RAN selection. It means that, when sensing a new cell, each user exploits the opportunity to select either a unicast or a multicast transmission. In such a case if a user decides to join a multicast group following the evaluation of eq. (6), then it could suffer from a lower performance in terms of throughput. This is due to the level determined by the least channel gain user in the multicast group, because it is assumed that the scheduler implements the CMS scheme. On the other hand, higher radio resource savings will be achieved since the resources for the multicast group have been already reserved. Therefore, the user joining a multicast group does not introduce additional resource waste. In such a way, the system has more resources available, i.e. more users could be served.

A zone-based quality sigmoid utility function [81] is used to map the throughput to user satisfaction.

$$u_q\,(Th) = \begin{cases} 0 & , Th < Th_{min} \\ 1 - e^{\frac{-\alpha * Th^2}{\beta + Th}} & , Th_{min} \leq Th < Th_{max} \\ 1 & , otherwise \end{cases} \qquad (7)$$

The minimum throughput (Th_{min}) is a threshold to maintain the multimedia service at a minimum acceptable quality level, values below this threshold result in unacceptable quality levels. Whereas values above the maximum throughput (*Th_{max}*) threshold will not add any noticeable improvements in the user perceived quality. The quality utility has values in the [0,1] interval and no unit. In order to determine the exact shape of the utility function the values of and need to be calculated. Knowing that: (i) for Th_{max}=3500 kbps the utility has its maximum value; (ii) Th_{req}=250 kpbs; and are determined by performing some mathematical computations of [81] and their values are 1.64 and 0.86, respectively.

The estimated energy consumption for a real-time application is computed using equation (8), as defined in [61]:

$$E = t\,(r_t + Th_{rec} * r_d) \qquad (8)$$

where *t* represents the transaction time, which can be estimated from the duration of the video stream; r_t is the mobile device's energy consumption per unit of time (W), Th_{rec} is the received throughput (kbps), r_d is energy consumption rate for data/received stream (J/Kbyte), and *E* is the total energy consumed (J). The parameters r_t and r_d can be determined by running different measurements for various amounts of data and defining an energy consumption pattern for each interface (LTE, WiFi) [82]. They were determined by running different simulations for various amounts of multimedia data (i.e., quality levels) while measuring the corresponding energy levels and then used to define the energy consumption pattern for each interface/scenario [81]. Based on the estimated energy consumption *E*, the utility for the energy criteria u_e is computed using Eq. (9) [62]:

$$u_e\,(E) = \begin{cases} 1 & ,\ E < E_{min} \\ \frac{E_{max} - E}{E_{max} - E_{min}} & ,\ E_{min} \leq E < E_{max} \\ 0 & ,\ otherwise \end{cases} \qquad (9)$$

Where E_{min} and E_{max} are computed considering Th_{min} and Th_{max}, respectively.

The user satisfaction utility function us is defined as the ratio between the datarate received and the datarate required by the user.

$$u_s = \frac{Th_{rec}}{Th_{req}} \qquad (10)$$

Obviously, the satisfaction achieved by users connected via unicast is, on average, closer to the value of 1 since the eNB tries to assign users all RBs they need. Oppositely, the satisfaction of users connected via multicast is affected by users with worse channel gain. Finally, the bandwidth utilization utility reflects the amount of resources used by the user in the context of the total amount of available resources.

The utility is calculated as the ratio between the new RBs used by the user and the number of available RBs in the cell for the corresponding type of transmission (i.e. unicast or multicast).

$$u_b = 1 - \frac{RB_{used}}{RB_{avail}} \qquad (11)$$

In case of a multicast transmission such a percentage is equal to zero. Indeed, if a user joins a multicast group, no more resources are used by the cell. In such a way, radio resources could be saved and, therefore, also the capacity of the system could be increased. The greater is u_b, the higher is the efficiency of the bandwidth utilization.

Eq. (10) and eq. (11) together represent the two factors that differentiate the selection between unicast and multicast transmission within a cell.

2.4 Performance Evaluation

An extensive numerical evaluation is conducted by using Matlab. The performance analysis is performed following the guidelines for the LTE system model in [83]. The main simulation parameters are listed in Table I. The parameters for the LTE system are set according to [40].

Simulations have considered a reference scenario where several LTE femtocells are deployed within the coverage of a LTE macrocell. The coverage area of the Macrocell is 500x500 m. The number of the small cells within the macrocell varies from 10 to 60. A dense urban scenario was considered where users are free to move according to Random Waypoint Mobility model [31]. Users' speed values are uniformly distributed within the interval from 3 km/h to 60 km/h. The simulations are carried out in a time interval of 3 minutes, with users downloading a real-time video.

HUMANS algorithm performance is compared with that of E-PoFANS [10] and EMANS [18]. Furthermore, in the presented simulation campaign, the weights of all four utility functions are considered the same (i.e., equal to 0,25).

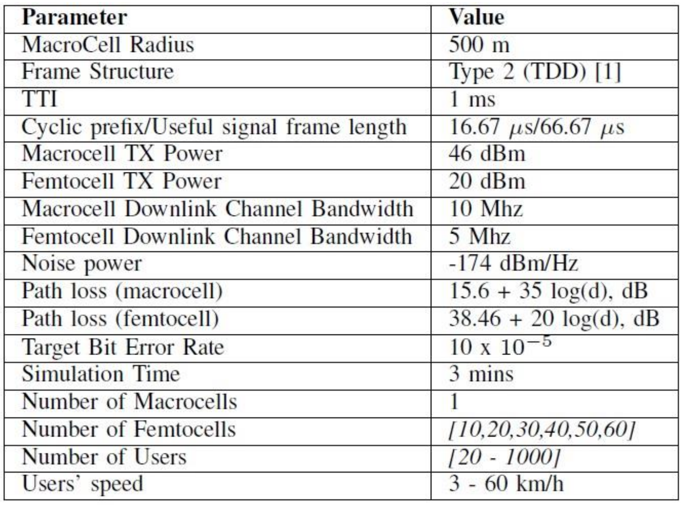
To compare the three algorithms and to simulate the dense scenario of the emerging 5G systems, simulation campaigns have been carried out in different network load conditions.

The algorithms performance has been computed every TTI, i.e., the throughput received Th_{rec} by users and the relative energy consumption have been recorded in each TTI of the simulation.

Simulation results, indeed, have been evaluated in *High density* or *Low density* conditions (i.e. both users density and femtocells density). The number of users has been varied from 20 to 1000. The following simulation metrics have been considered:

- *Average throughput*: the average quality transmission accomplished to users;
- *Aggregate Data rate (ADR)*: the sum of the throughput of the users among overall system;
- Estimated *Energy Consumption*: the estimated energy consumption of the devices when downloading a video flow;
- *User Satisfaction*: the satisfaction perceived by users in terms of the ratio between the datarate received and the datarate required by each user;
- *Percentage of served users*: the measure of how efficiently the algorithms work in terms of system capacity;
- *Percentage of resource usage*: the measure of the efficiency of the algorithms in order to save resources. The lower is this metrics, the higher the performance;
- *CQI variation*: the distribution of users with different CQIs among multicast and unicast transmissions.

It is worth noting that the energy consumption of each user has been calculated according to eq. (8) at every TTI. Th_{rec} is the throughput received by users in the given TTI and *t* is the duration of the TTI.



Parameter	Value
MacroCell Radius	500 m
Frame Structure	Type 2 (TDD) [1]
TTI	1 ms
Cyclic prefix/Useful signal frame length	16.67 μs/66.67 μs
Macrocell TX Power	46 dBm
Femtocell TX Power	20 dBm
Macrocell Downlink Channel Bandwidth	10 Mhz
Femtocell Downlink Channel Bandwidth	5 Mhz
Noise power	-174 dBm/Hz
Path loss (macrocell)	15.6 + 35 log(d), dB
Path loss (femtocell)	38.46 + 20 log(d), dB
Target Bit Error Rate	10×10^{-5}
Simulation Time	3 mins
Number of Macrocells	1
Number of Femtocells	*[10,20,30,40,50,60]*
Number of Users	*[20 - 1000]*
Users' speed	3 - 60 km/h

Table 8 - Main Simulation Parameters

Low Density conditions. In this Section the performance of the three algorithms in low density conditions are presented. Simulation results are shown in the case of 10 small cells within the macrocell. The analysis has been carried out with users moving at different speeds from pedestrian (i.e. 3 kmph) to low vehicular speed (i.e., from 30 to 60 kmph) in a dense urban scenario.

Fig. 2.13 shows the average throughput received by users. The proposed HUMANS algorithm outperforms the other ones, guaranteeing a relatively constant trend even when increasing the number of users within the reference area.

This is due to the presence of the multicast groups, whose users are always served with the same number of RBs and with the minimum CQI experienced by group members. At the same time both E-PoFANS and EMANS experience a decrease in their performance with an increasing number of users, as these two algorithms use only unicast transmissions. This is expected because the availability of resources decreases when increasing the network load. The system ADR achieved by the algorithms is shown in Fig. 2.14. Exploiting multicast communications allows HUMANS algorithm to increase the ADR of the system when increasing the number of users.

Indeed, each new user contribute to add rate to the system ADR. Whereas the other two algorithms saturate after around 200 users within the system.

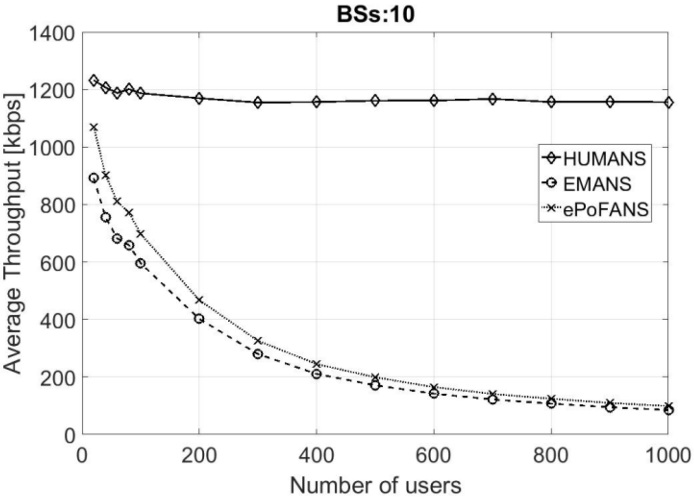


Fig. 2.13 - Average Throughput – Low Density

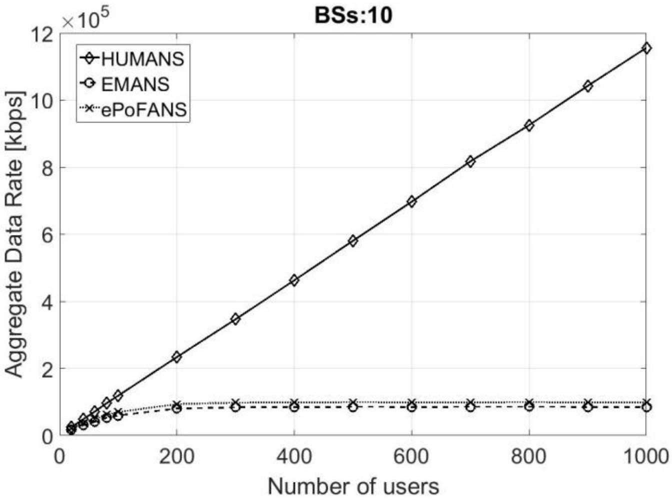


Fig. 2.14 – ADR – Low Density

According to eq. (6), also the estimated device energy consumption has to be taken into account (Fig. 2.15). For all algorithms, the highest energy consumption is met with few users in the system. That is because there are enough resources to serve users requiring higher datarate and greater resources, consequently consuming more energy. Compared to other algorithms, HUMANS achieves a gain ranging from 4% to 9% with respect to ePoFANS, whereas it gains up to 4% against EMANS with a few users in the system.

Following the average throughput trend, the user satisfaction (Fig. 2.16) achieved by HUMANS is always high (i.e., around 80%), whereas the users satisfaction decreases with the number of users when adopting the two other algorithms. This is due to the limited amount of resources for unicast transmissions considered in both EMANS and ePoFANS.

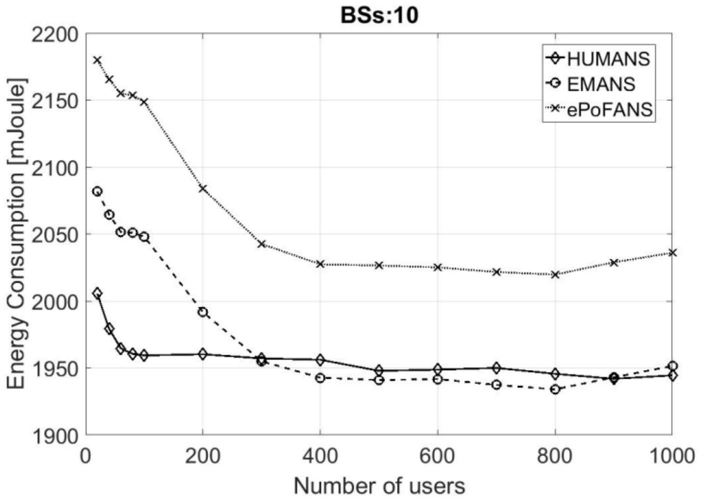


Fig. 2.15 - Devices Energy Consumption – Low Density

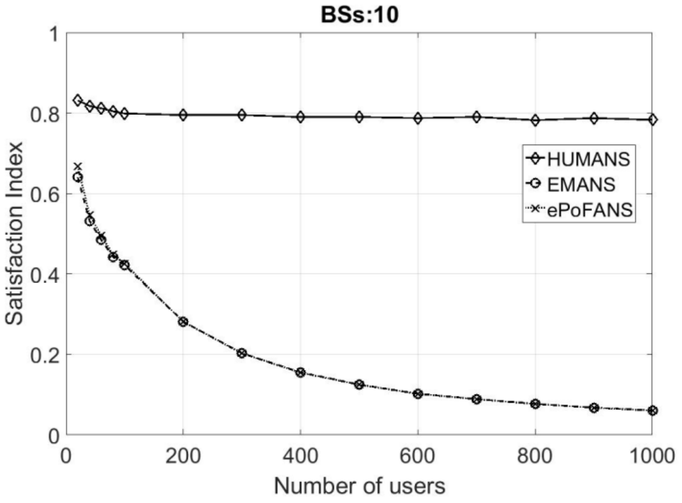


Fig. 2.16 - Users Satisfaction – Low Density

However, the strength of HUMANS is illustrated in both Fig. 2.17 and Fig. 2.18. The former shows how the proposed solution is able to serve all users requiring access to the video flow. This is due to the intrinsic behaviour of the multicast approach, which can serve all users. Whereas, the two other algorithms have a limited capacity as they serve users via unicast only. At the same time, HUMANS also achieves resource utilization savings between 25% and 15% compared to both EMANS and ePoFANS when increasing the density of the network in terms of number of users.

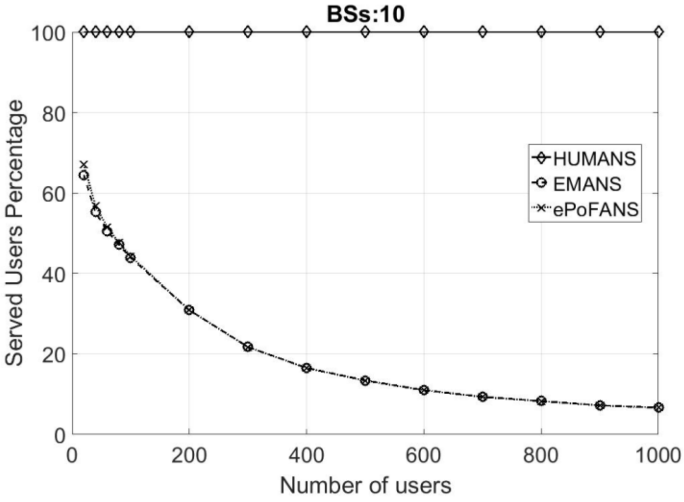


Fig. 2.17 - Percentage Of Served Users – Low Density

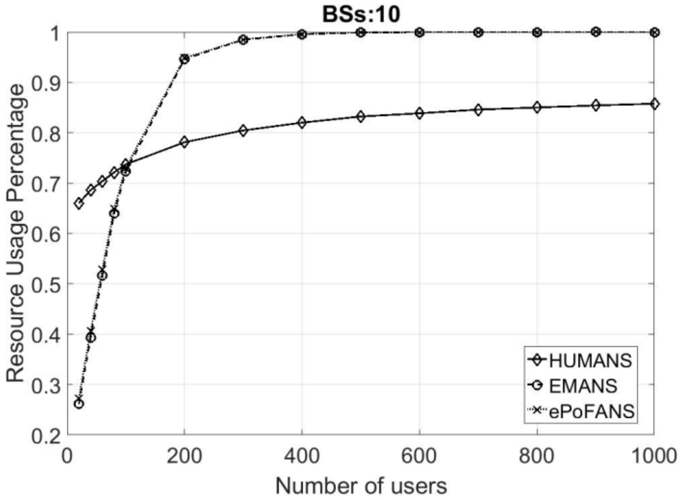


Fig. 2.18 - Percentage Of Resource Usage – Low Density

High-density conditions. The performance of the proposed algorithm has been evaluated also in a high density scenario, when the number of small cells within the macrocell increases to 60.

Fig. 2.19 shows the average throughput received by users in this high density scenario. Similar to the low density case, HUMANS outperforms the other solutions it is compared against. Nevertheless, with few users E-PoFANS has still a very good performance as, in these conditions, the high number of cells deployed provides enough resources to satisfy all user requests. However, E-PoFANS and EMANS decrease their performance with the increasing number of users. This happens as these algorithms employ unicast transmissions only and an increase in the offered load adversely affects their performance.

Fig. 2.20 shows the ADR of the whole system for users in High density scenarios. At a certain point (500 users) EMANS and E-PoFANS do not bring any additional improvements, whereas HUMANS continues to follow the growing ADR. This is because multicast transmissions allow all users requiring the service to receive it with no additional resource requirements. Indeed, as shown in Fig. 4(e) HUMANS provides overall coverage to all users in the system. On the contrary, EMANS and E-PoFANS suffer from high user outage when increasing the overall number of users.

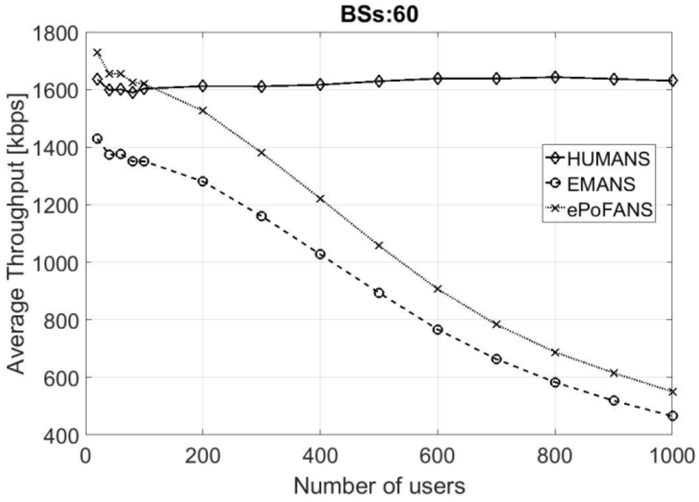


Fig. 2.19 - Average Throughput – High Density

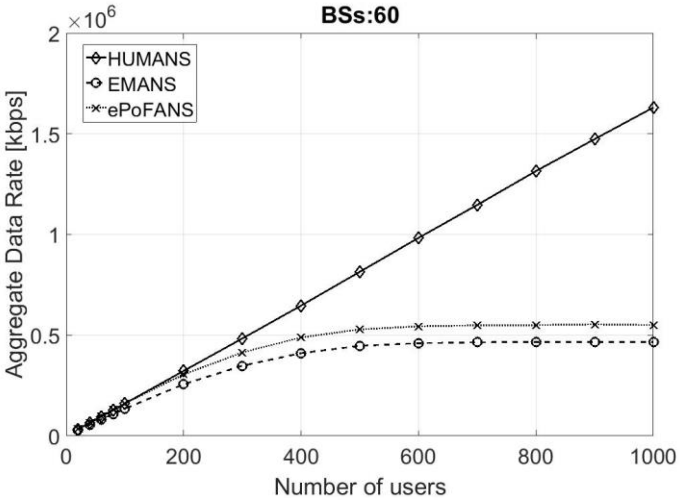


Fig. 2.20 - ADR – High Density

In the High density scenario, the performance of HUMANS in terms of energy consumption (Fig. 2.21) shows a degradation with respect to the other algorithms. Since there are many more resource available for users, both EMANS and e-PoFANS are able to serve more users with lower datarate, and consequently low energy consumption, with respect to the low density scenario. On the other hand, in HUMANS case, the unicast component is more prominent just because more resources are available, thus consuming more energy.

Whereas, as for users satisfaction (Fig. 2.22), HUMANS maintains the same trend of the low Density scenario, as expected, with higher achievable values (i.e., around and 95%). Increasing the available resources in the system allows the other two algorithms to achieve a performance closer to HUMANS, but only with a few users in the system. When increasing the number of users, simulation results show that HUMANS gains up to 65% (with 1000 users).

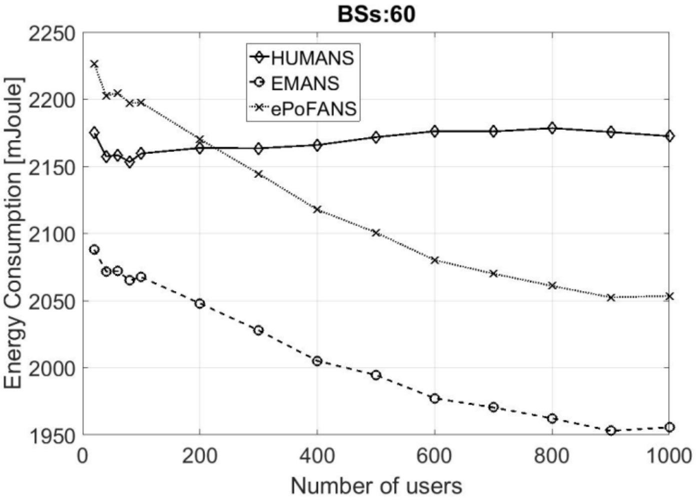


Fig. 2.21 - Devices Energy Consumption – High Density

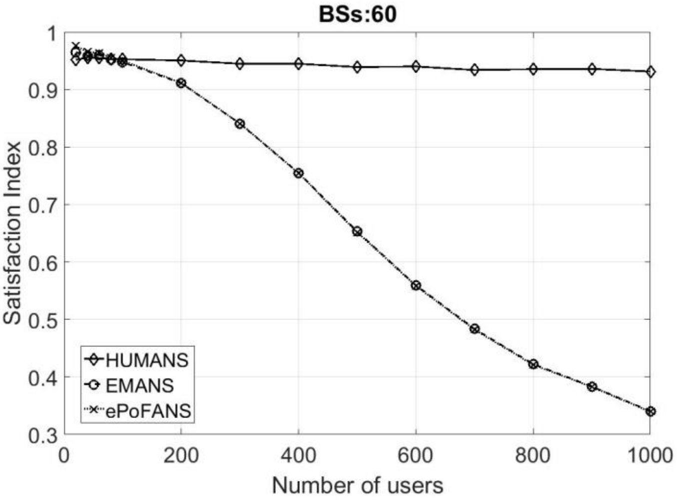


Fig. 2.22 - Users Satisfaction – High Density

All the above considerations are the result of a different behaviour of the algorithms in terms of resource usage (Fig. 2.23). Since multicast transmission consumes many RBs, resource utilization is better for EMANS and E-PoFANS algorithms with a few users in the system. On the other hand, when increasing the number of users, these two algorithms use all the available resources, thus reaching saturation, which leads to the high outage percentage illustrated in Fig. 2.24. Furthermore, as expected, in a high density scenario the increasing number of cells leads to a consequent overall improvement in the performance of all algorithms thanks to the greater number of resources available.

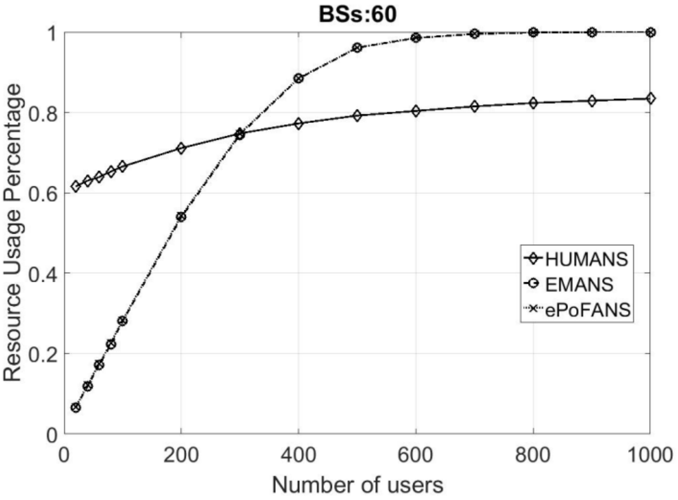


Fig. 2.23 - Percentage Of Resource Usage – High Density

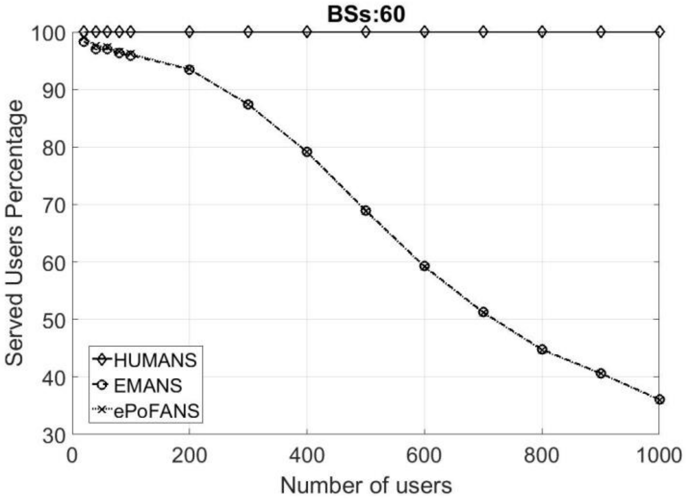


Fig. 2.24 - Percentage Of Served Users – High Density

CQI distribution. The final discussion is about distribution of user CQI levels between multicast and unicast transmissions. Results are presented in Fig. 2.25, where the percentage of users served for each value of CQI is shown, for each kind of transmission (multicast and unicast). In all cases unicast transmission is activated to users with high CQI levels (i.e., in good channel condition) only. This is because users in good channel condition require a few RBs, whereas users experiencing bad channel conditions need more resources to obtain the required datarate. In HUMANS, the users with lower CQI levels are served via multicast transmissions and this has a double advantage: *(i)* they do not waste additional resources and *(ii)* make use of multicast flows (i.e. they receive all the RBs dedicated to the multicast group). Therefore, thanks to this approach, users requiring many resources that cannot be served if an only-unicast oriented algorithm is implemented, can always receive the video service, especially when the system is in high load conditions.

Fig. 2.25 best depicts the objective of the proposed HUMANS, which is to guarantee an increasing user capacity either by using the multicast or by saving resources. It is demonstrated HUMANS provide such benefit independently by neither the number of users nor the femtocells density within the system (e.g.,

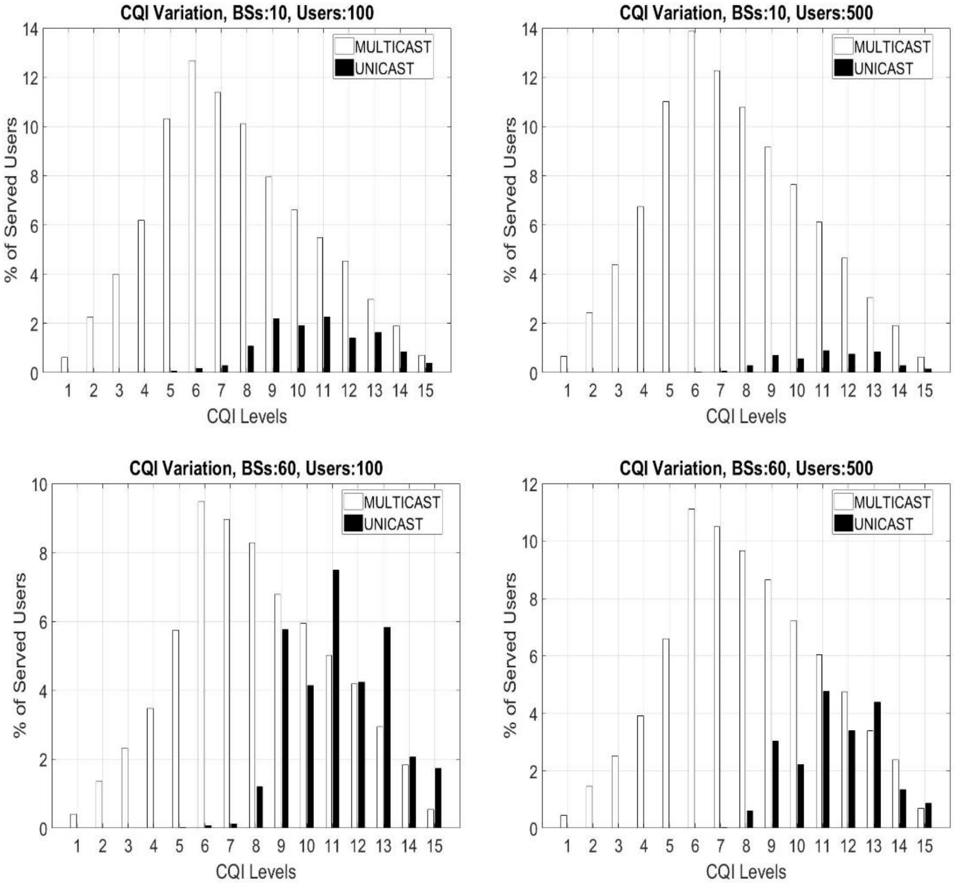


Fig. 2.25 - Users' CQI Distribution - High Mobility

the plot are depicted with both high- and low-amount of either users or BSs).

HUMANS considers bandwidth utilization and the trade-off between quality and energy consumption when delivering video in DenseNet scenarios. A major contribution of HUMANS is the consideration of joining a multicast group as a possible option in the network selection process, thus allowing for smart bandwidth management. HUMANS serves users with good channel conditions via unicast transmissions and the remaining users via multicast.

Performance evaluation carried out in low- and high-density scenarios, demonstrate how the proposed hybrid unicast-multicast approach provides a significant improvement in terms of capacity and radio resource utilization in comparison with other unicast-only solutions. The performance gain is much higher when user density increases within a system, thus providing an interesting solution for the emerging dense 5G systems.

3 RRM over 5G multicast – Single Frequency Networks

With the advent of Long Term Evolution (LTE), evolved-MBMS (eMBMS) has been introduced in the 3GPP Release 9 [84]. According to the current literature, it is foreseen that PtM transmissions will play a key role over 5G Wireless Networks [11], therefore the emerging 5G systems seem to be the most promising systems to better satisfy the eMBMS requirements [67]. In eMBMS, the Modulation and Coding Scheme (MCS) is selected on a per-group basis; in this way, in order to serve all users within the group, the choice of multicast/broadcast transmission parameters is affected by users with the poorest channel conditions. As cell-edge users require a transmission with a robust MCS, and thus poor spectral efficiency, the increase of system efficiency becomes a challenging issue. Several approaches have been investigated to handle the increasing number of smart devices and to manage the radio resources. The conservative approach [76] is the basic solution for Radio Resource Management (RRM). It consists of delivering the content to all users with a robust MCS. The conventional transmission guarantees fairness, but it suffers from poor spectral efficiency, low throughput and high users' dissatisfaction. On the other hand, the opportunistic approach [77] tries to overcome the throughput limitation by serving the set of users with better channel conditions. At the

same time, this approach is affected by short term fairness. The trade-off between throughput and fairness is achieved with the subgrouping approach [78], where users are grouped in different subgroups served with different MCSs.

In order to improve the efficiency of eMBMS, MBMS Single Frequency Networks (MBSFN) has been introduced by 3GPP in Release 7 [85]. In the MBSFN the same content is transmitted at the same time within several cells, by exploiting the same radio resources. Such cells are coordinated to achieve a MBSFN Transmission and are grouped forming a MBSFN Area [40]. Furthermore, the base stations (named as eNBs) of a given MBSFN Area need to be synchronized both in the time and in the frequency domain in order to make cell-edge receivers perceive the transmissions from different eNBs as broadcasted by a single base station. Then, the multipath phenomena is exploited to generate constructive interference at the receiver. Therefore, cell-edge users experience an improved quality by combining the video flows replies.

The process for the selection of the cells belonging to the same MBSFN Area is known as MBSFN Area Formation. When performing the MBSFN Area Formation process, several issues have to be faced, such as the choice of cells to group, the selection of contents to broadcast and users to include, the management of radio resources, and so on. These aspects make MBSFN Area Formation be a tough issue to solve.

This chapter provides a new contribution about RRM over 5G multicast dealing with MBSFN area formation issue. In the next sections, we first presents a brief description of MBSFN and then present a dynamic MBSFN area formation algorithm.

3.1 MBSFN

In a MBMS transmission, cell-edge users suffer from both signal attenuations due to path loss and interference from transmissions of neighbour cells. Thus, it would be necessary to largely increase the amount of radio resources required for each service [1]. To overcome these limitations, in Release 7 [85] the 3GPP introduces MBSFN, that represents the best exploitation of the Single Frequency Network concept for MBMS. MBSFN stands for MBMS Single Frequency Network and consists of procedures, which allow the transmission of the same signal at the same time on the same radio resources.

All cells, instead of avoiding interference between their respective transmissions by using non-overlapping frequencies, use the same frequency for broadcasting services. Signals, coming from multiple sources, are viewed as several components of the same signal reaching the receiver from different paths within the CP (Cyclic Prefix) to avoid Inter-Symbol Interference (ISI). The CP is introduced because of the different time delay due to the distance of the UE from eNBs. Therefore, eNBs synchronisation is the main requirement to enable the receiver to process signals, which appear as multiple signal versions of multipath propagation. Thus, components constructively interfere and let the signal combine and be enhanced at the receiver. In this way, MBSFN transmission is seen as a single one by a UE.

MBSFN provides the following benefits [86]:

- Better cell-edge performance by increasing the received signal level;
- More consistent user experience: signals, received from synchronised neighbouring cells, do not appear as interference but are received as constructive signals, letting the whole network behave as a \single cell";
- Higher overall efficiency thanks to the better cell-edge performance;
- Increased spectral efficiency because the enhanced SINR compared to the unicast transmission.

In Release 9 TS 36.300 [40] the 3GPP introduces the following definitions dealing with MBSFN:

MBSFN Synchronization Area: an area of the network where all eNBs can be synchronized and perform MBSFN transmissions. MBSFN Synchronization Areas are capable of supporting one or more MBSFN Areas. On a given frequency layer, a eNB can only belong to one MBSFN Synchronization Area. MBSFN Synchronization Areas are independent from the definition of MBMS Service Area.

MBSFN Transmission or a transmission in MBSFN mode: a simulcast transmission technique realised by transmission of identical waveforms at the same time from multiple cells. An MBSFN Transmission from multiple cells within the MBSFN Area is seen as a single transmission by a UE.

MBSFN Area: an MBSFN Area consists of a group of cells within an MBSFN Synchronization Area of a network, which are co-ordinated to achieve an MBSFN Transmission. Except for the MBSFN Area Reserved Cells, all cells within an MBSFN Area contribute to the MBSFN Transmission and advertise its availability. A cell within an MBSFN Synchronization Area belongs to only one MBSFN Area for all MBMS services.

MBSFN Area Reserved Cell: A cell within a MBSFN Area, which does not contribute to the MBSFN Transmission. The cell may be allowed to transmit for other services but at restricted power on the resource allocated for the MBSFN transmission e.g. PTP for users at the centre of the cell.

So, according to the standard, an MBSFN Area consists of a group of cells, synchronized in time to achieve an MBSFN transmission and is generally compact without any hole, but it may exclude from the MBSFN transmission some cells, known as MBSFN Area Reserved Cells [87]. One or more MBSFN Areas, overlapping in space, form a larger MBSFN Synchronisation Area [88], which is defined as the region where all eNBs are synchronised in time to perform MBSFN transmissions and its size can vary greatly, from a few cells serving a stadium to many cells delivering content to an entire country. An MBSFN area is included in one MBSFN Synchronisation Area. An eNB can only belong to one MBSFN Synchronization Area. As MBSFN synchronization areas can be larger than MBSFN area, a cell can then belong to several MBSFN areas (at a time a single cell can be a part of 8 MBSFN Areas). MBMS Service Area is defined as the region within which it is transmitted the same content [2]. It comprises of one or more cells.

E-MBMS introduces efficiency, scalability and flexibility and is favourably looked by operators for the following reasons:

- Reusable LTE infrastructure;
- High-SNR achieved by combining at user the level of signals from eNBs belonging to the same MBSFN;
- Hundreds and thousands of reachable users simultaneously by employing the same radio resources as for a single user.

However, in the literature several challenges exist. Among them, there are issues related to physical-layer (i.e. the best approach for the selection of the MCS utilised to improve the spectral efficiency in the transmission of MBSFN data, etc.) and radio resource management (i.e. choose the best algorithm for radio resource allocation to optimise system throughput or the advantages of broadcast with the respect to unicast, etc.).

Furthermore, there is another work [89] dealing with the effect of the MBSFN area on the coverage. In particular, the MBSFN operation is designed to increase the range of coverage. Thanks to MBSFN, the mean SINR is increased and interference is decreased. The result is a better coverage level, which depends on the MBSFN area size, which is the number of synchronised cell providing the same data on the same radio resources. Therefore, users have a better coverage level when receiving the same signal from a greater number of cells.

However, by increasing MBSFN area size there is a lower flexibility to adapt contents to different geographical areas. Also [90] demonstrates that an improvement in the performance of UEs is achieved by increasing the MBSFN area. This study provides how to dynamically cluster MBSFN areas with a method based on the CQI received from UEs, modifying the size of areas according to the dynamic state of users radio channel.

Furthermore, adapting the size of the MBSFN area (adding or removing cells to or from it) allows the system to optimise radio resources in order to satisfy users QoS requirements.

A fundamental aspect scarcely treated in the literature is MBSFN area formation. [91] presents a model for broadcasting in LTE networks. The approach introduced consists in two main tasks:

- Assigning the cells to the areas;
- Deciding the content to broadcast in each area.

In general, different cells have different demand for different content, but this solution is infeasible because the maximum number of areas that can be created is 256 and undesirable because of the inter-area interference among neighbouring areas. In MBSFN, there is no interference between cells belonging to the same area.

The idea is to put neighbouring cell with similar content popularity in the same area and serve only the content that is popular in both cells.

Thus, the merge and the grow approaches are followed. The former starts from assigning a cell per area and merging neighbouring areas in order to maximise the performance improvement. The algorithm stops when reaching the maximum number of areas and merging more pairs of areas cannot increase the performance. The latter is more complex than the previous one, but it introduces a higher level of flexibility. It is selected a cell and created a new area containing only this cell. Then, the new area is grown by adding a new cell. Among the neighbouring cells, the one that is more profitable to add is chosen. If there is not profit, there are no more cells to add to the area and a new area is created.

In the next section, a dynamic MBSFN area formation is presented, with the aim to enhance the performance of the system in terms of overall system data rate.

3.2 Dynamic MBSFN Area Formation (DMAF) Algorithm for Multicast Service Delivery in 5G Wireless Networks

In this chapter, we propose a Dynamic MBSFN Area Formation (hereinafter referred to as "DMAF") algorithm that dynamically creates MBSFN Areas with the aim to increase the system Aggregate Data Rate (ADR), under the constraint that all users must be served through the MBSFN transmission.

Differently from SCF algorithm where users with poor channel conditions are served via unicast link, the main idea behind DMAF is to cluster MBSFN Areas by treating them like multicast subgroups. Indeed, the subgrouping approach is applied to DMAF algorithm in order to form either disjoint or overlapping MBSFN Areas. In the former case a single video flow is delivered within the Area. Whereas, in the latter case, a Base Layer is delivered to all users and only users able to support higher MCS receive Enhancement Layers. This allows improving the video quality perceived by the best users. DMAF further enhances the system performance by: *(i)* exploiting the multi-rate approach of subgrouping; *(ii)* choosing the best MBSFN Area configuration for increasing the ADR; *(iii)* performing dynamic radio resource allocation for an efficient spectrum utilization; *(iv)* guaranteeing total coverage with 100% served users. Results provided through simulations demonstrate the effectiveness of the proposed algorithm, which improves the MBSFN Area Formation approaches already existing in the literature.

3.2.1 Literature review

Multicast/Broadcast transmissions feature a wide range of applications by playing an important role in 5G wireless systems. In the eMBMS architecture , the periodic users' Channel Quality Indicator (CQI) feedback determines the MCS level for multicast traffic delivery by the eNB towards the user. Therefore, the evaluation of users' channel conditions is needed to select the adequate transmission parameters for the content broadcasting. In the literature, different RRM per-group-based schemes are proposed and classified according to the followed approaches: conservative, opportunistic and subgrouping. The Conventional Multicast Scheme (CMS) [76] broadcasts the content to all the multicast users with the same data rate and, so it is a very fair policy. At the same time, the conventional schemes are characterized by low spectral efficiency, affected by the user with poor channel condition.

On the other hand, the Opportunistic Multicast Scheme (OMS) [77] improves the spectral efficiency by privileging the service of the users with the highest CQI. This technique is affected by short-term fairness, although long-term fairness can be achieved. Between these two opposite approaches, multicast subgroup-based approaches have been introduced to overcome limitations on fairness and throughput. The Multicast Subgrouping scheme [78] consists of splitting users in different subgroups and serving all of them with the appropriate MCS. Therefore, the negative effects due to users with worst channel conditions are minimized. As Multicast Subgrouping scheme attains a good trade-off between fairness and throughput, the proposed work is based on the subgrouping policy by exploiting the multi-rate approach.

MBSFN introduces other issues to eMBMS RRM like the MBSFN Area Formation [40]. However, as the best of our knowledge, few works consider how to group cells within MBSFN areas and which content should be broadcasted. Indeed, both [91] and [92] do not consider MBSFN Area formation, but focus on the best MBSFNs configuration in terms of performance. In [91], results show that performance improve with the increasing of the number of areas. Nevertheless, the algorithm only maximizes coverage and neglects the system throughput. Whereas, the work proposed in

[92] provide better performance with the increasing of MBSFN area size when the minimum separation among eNBs gets higher. A legacy approach defined by the 3GPP [93] considers that all cells involved in the MBSFN transmission statically form a single MBSFN Area. Nevertheless, such an approach does not consider neither any users' diversity nor users' content interests.

Differently from the legacy approach, Single Cell Point-to-Multipoint (SC-PtM) [94] supports eMBMS over single cell and dynamically adjusts the MBSFN Area cell by cell. With SC-PtM the number of MBSFN Areas in a Synchronization Area is the same as the number of cell. However, this approach does not fully exploit SFN benefits. In this optic, a dynamic approach, named Single-Content Fusion (SCF) [79] have been proposed for both MBSFN Area Formation and for determining the content to be broadcasted, considering users' interests. The MBSFN Area Formation is dynamically carried out by taking into account the content to broadcast, in order to maximise the overall throughput by using both multicast and unicast transmissions. In particular, SCF increases the minimum MCS level of the MBSFN Transmission and serve users with a lower CQI through unicast link.

By doing so, SCF handles a mixed unicast and multicast traffic by dedicating up to 60 percent of the available Radio Resources to MBMS transmission and the remaining 40 percent to unicast ones, following the 3GPPP standard constraint [67]. Nevertheless, this constraint has been removed in Release 14 [95], thus allowing to dedicate up to 100 percent of resources to eMBMS service, in order to support both static and dynamic resource allocation balancing between unicast and broadcast/multicast transmissions. SCF follows an approach that considers a fixed threshold when splitting the resources between unicast and multicast services. Different works deal with such a kind of balancing. The approach proposed in [96] takes into account Signal-to-Noise Ratio (SNR) threshold values only, not strongly correlated with system capacity, and no minimum performance guarantees are assured to unicast or multicast services. A comparison between unicast and multicast transmission is presented also in [97], where the authors have determined switching thresholds, as a function of the number of users per cell, to switch between unicast and multicast modality for downloading or streaming services. Differently from there works, authors in [80] present a dynamic approach that makes use of a channel-aware subgrouping principle to provide fair throughput to both unicast and multicast users. The idea is to assign unicast subscribers to a virtual group, thus allowing them to compete for network resources on an equal footing with multicast users.

The balancing between unicast and multicast is not the core of our proposal, nevertheless, in order to compare DMAF with other works, we considered fixed threshold between the two services.

One of the most recent work dealing with dynamic area formation is [98]. It maximises users' QoE, firstly considering UEs' display capabilities. Nevertheless, this work does not consider either dynamic radio resource allocation for eMBMS or the overall system throughput.

The contribution provided by this proposal is an algorithm that tries to overcome the limitations of both the 3GPP legacy approach [93] and Single Cell Point-to-Multipoint (SC-PtM) scheme [94] by dynamically creating MBSFN Areas. DMAF exploits the multi-rate approach [78], thus providing the double advantage to minimize cell-edge users' negative effects and to improve the perceived video quality of users with best channel condition by delivering Enhancement Layers. Furthermore, Dynamic Radio Resource Allocation is performed for reducing resource waste and for further enhancing the system ADR.

3.2.2 System Model

In this work, we consider a MBSFN scenario, in which a set C of cells, deployed in a coordinated manner within a Synchronization Area, is interested to the same content. A Synchronization Area [40] may include one or more MBSFN Areas coordinated to achieve an MBSFN Transmission. Let us denote by M the set of all MBSFN Areas activated over the Synchronization Area. Each MBSFN Area $m \in M$ consists in adjacent cells, which broadcast the same content at the same time over the same set of radio resources. The available radio resources are managed in terms of Resource Blocks (RBs) where each RB corresponds to 12 consecutive and equally spaced sub-carriers. One RB is the smallest frequency resource that can be assigned to a UE. The overall number of available RBs depends on the system bandwidth configuration and can vary between 6 and 100, for a bandwidth of 1.4 MHz and 20 MHz, respectively.

The proposed DMAF algorithm exploits the multicast subgrouping technique [78] jointly with Scalable Video Coding (SVC) [99] that produces more video substreams of different quality, namely, a Base Layer (BL) at standard quality, and an Enhancement Layer (EL) at enhanced quality (i.e., with MCS level or frame rate). According to [100], the optimal subgroup configuration that maximizes the system ADR has always to be searched between a single or two-subgroups configuration. Therefore, each substream is delivered to a different MBSFN Area, respectively called as BL MBSFN Area (including all cells broadcasting the Base Layer) and EL MBSFN Area (including all cells broadcasting the Enhancement Layer). In order to avoid interference, a different set of RBs is used in each overlapping Area. If the two areas overlap (i.e., one or more cells belong to both areas), the available bandwidth is split between BL and EL MBSFN Areas. SVC will allow users belonging to EL MBSFN Area to also decode the content provided by BL MBSFN Area, thus increasing their datarate.

Let *RB* be the available amount of radio resources within a cell. We denote RB_{BL} and RB_{EL} the amount of RBs assigned to the Base Layer and to the Enhancement Layer, respectively.

Let *U* be the set of users interested to the broadcasted content. Each UE transmits its Channel Quality Indicator (CQI) feedback to the eNodeB. The CQI is associated to the maximum supported Modulation and Coding Scheme (MCS) [40] according to LTE-A standard, as reported in Table 9.

In case of eMBMS, the user with the worst CQI determines the selection of the MCS for the broadcast/multicast transmission. In order to create MBSFN Areas, to properly allocate Radio Resources and to successfully exploit SVC, the proposed algorithm must meet a number of constraints, briefly discussed as follows.

(i) MBSFN Area Constraints:

- We introduce the binary variable $y_{c,m}$ with $c = \{1, \dots, C\}$ and $m = \{1, \dots, M\}$, such that

$$y_{c,m} = \begin{cases} 1, & \textit{if the } c\!-\!th \textit{ cell belongs to the } m\!-\!th \textit{ MBSFN area} \\ 0, & \textit{otherwise} \end{cases} \quad (12)$$

- A cell shall belong to at most 8 MBSFN Areas:

$$\textstyle\sum_{m \in M} y_{c,m} \leq 8 \qquad \forall\, c \in C \qquad (13)$$

- Within a Synchronization Area, there shall be at most 256 MBSFN Areas:

$$|M| \leq 256 \qquad (14)$$

(ii) Resource Constraints

- The RBs allocated shall not exceed the number of those available:

$$\left(\left|RB_{m,BL}\right| + \left|RB_{m,EL}\right|\right) \leq |RB_m|\ \forall\, m \in M \qquad (15)$$

(iii) Layer Constraints

- The Base Layer shall be delivered to all the users of a given MBSFN Area

$$U_{m,BL} = U_m, \forall\, m \in M \quad (16)$$

Where $U_{m,BL}$ is the set of users in the m-th MBSFN Area receiving the Base Layer.

- Finally, all users who support the MCS assigned to the area for the transmission of the Enhancement Layer shall be selected to receive such a layer. They shall be already scheduled for the reception of the Base Layer:

$$U_{m,EL} \subseteq U_{m,BL}, \forall\, m \in M \qquad (17)$$

Where $U_{m,EL}$ is the set of users in the m-th MBSFN Area receiving the Enhancement Layer.

The proposal idea is to create MBSFN Areas that meet the above constraints in order to maximize the system Aggregate Data Rate (ADR) by dynamically performing resource allocation.

Given a set M of MBSFN Areas, the ADR per frame is given by:

$$ADR = \sum_{m \in M} \left(\left(ADR_{m,L} + ADR_{m,EL} \right) + ADR_m \right) \qquad (18)$$

Where

$ADR_{m,BL} = \sum_{u_{bl} \in U_{m,BL}} R_{minCQI(U_{m,BL})} \times \left| RB_{m,BL} \right| \;\; \forall\, m \in M$ (19), is the aggregate datarate of users receiving the BL,

$ADR_{m,EL} = \sum_{u_{el} \in U_{m,EL}} R_{minCQI(U_{m,EL})} \times \left| RB_{m,EL} \right| \;\; \forall\, m \in M$ (20), is the sum of data rates achieved by the EL delivery and

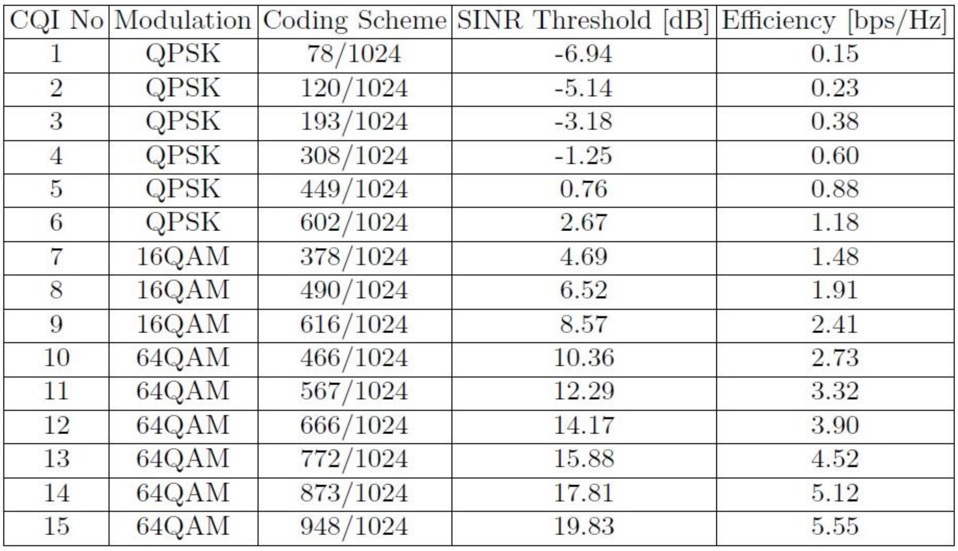
$ADR_m = \sum_{u \in U_m} R_{minCQI(U_m)} \times |RB| \;\; \forall\, m \in M$ (21), is the aggregate data rate of users within each disjoint *m* MBSN area. It is worth remarking that in case of overlapping Areas, $ADR_m = 0$. Whereas, in case of disjoint Areas, $ADR_{m,BL} = ADR_{m,EL} = 0$.

Therefore, the aim of the proposed algorithm is to solve the following problem:

$$arg \max_{RB} ADR$$

subject to (12) – (18)

when serving 100% of users interested in the broadcasted content.



CQI No	Modulation	Coding Scheme	SINR Threshold [dB]	Efficiency [bps/Hz]
1	QPSK	78/1024	-6.94	0.15
2	QPSK	120/1024	-5.14	0.23
3	QPSK	193/1024	-3.18	0.38
4	QPSK	308/1024	-1.25	0.60
5	QPSK	449/1024	0.76	0.88
6	QPSK	602/1024	2.67	1.18
7	16QAM	378/1024	4.69	1.48
8	16QAM	490/1024	6.52	1.91
9	16QAM	616/1024	8.57	2.41
10	64QAM	466/1024	10.36	2.73
11	64QAM	567/1024	12.29	3.32
12	64QAM	666/1024	14.17	3.90
13	64QAM	772/1024	15.88	4.52
14	64QAM	873/1024	17.81	5.12
15	64QAM	948/1024	19.83	5.55

Table 9 - CQI-MCS Mapping

3.2.3 DMAF algorithm

The proposed DMAF algorithm is designed to tackle the MBSFN Area Formation problem by dynamically creating MBSFN Areas in order to improve the overall system performance.

In particular, DMAF aims to achieve higher ADR with respect to already existing MBSFN Area Formation procedures. Generally, in a MBSFN Transmission the user with the worst channel condition (i.e., the worst CQI) drives the system performance. Indeed, all the users within a MBSFN Area receive the same broadcast content with the same MCS that determines the datarate (see Table 9). Therefore, if a user supporting only low MCS level belongs to the MBSFN Area, all users within that Area will suffer poor spectral efficiency. In general, the higher the MCS level of the worst user in the Area, the better the ADR performance. The proposed algorithm aims to increase the ADR by enhancing the MCS level of the MBSFN Transmission.

In order to avoid the outage of users with a CQI less than the CQI required for decoding the transmitted MCS, DMAF exploits the multi-rate approach for multicast scheduling [78] by providing different data rates to two subsets of users with good and bad channel quality conditions, respectively. Each subgroup represents a different MBSFN Area, where the Base Layer and the Enhancement Layer are delivered. If the MBSFN Areas overlap, the Base Layer is guaranteed to all users whereas the Enhanced Layer is decoded by only users with higher channel qualities [101]. DMAF algorithm also takes advantage from Scalable Video Coding (SVC) [99] allowing users with higher CQI to decode both flows, thus improving the perceived video quality. If MBSFN Areas are disjoint the whole bandwidth (i.e., all available RBs) is assigned to such Areas, otherwise, the RBs are split between the Areas. It worth noting that two MBSFN Areas are disjoint if no cells belonging to one are also belongs to the other. Therefore, DMAF dynamically allocates to avoid resource waste.

DMAF Algorithm consists of four phases, which could be considered as four different sub-algorithms. The main phase is *MCS-level Increase*, which includes the other three phases, *Area Formation*, *SVC Improvement* and *Cell Re-Clustering*.

Fig. 3.1 shows an example of a MBSFN Area configuration for a scenario with 57 cells. Three MBSFN Areas have been created. The MBSFN Area 1 delivers the Base Layer; thanks to SVC Improvement phase, the MBSFN Area 2 is formed and broadcasts the Enhancement Layer by enhancing the perceived video quality for users with higher channel conditions. Finally, the disjoint MBSFN Area 3 is created in Cell Re-Clustering phase in order to further enhance the system ADR.

The four algorithms are described in detail in the following subsections.

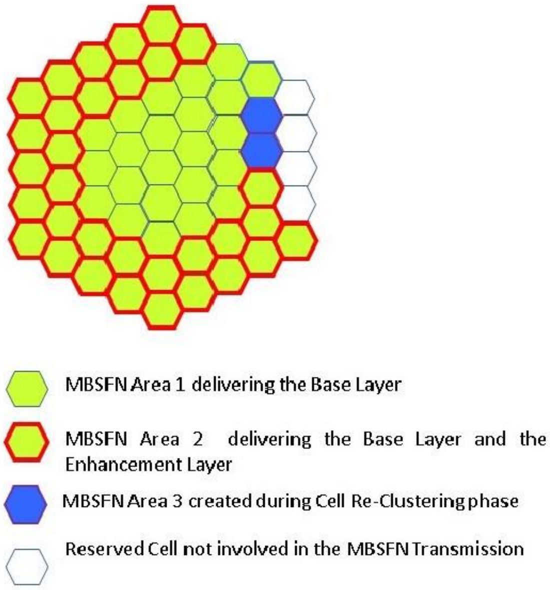
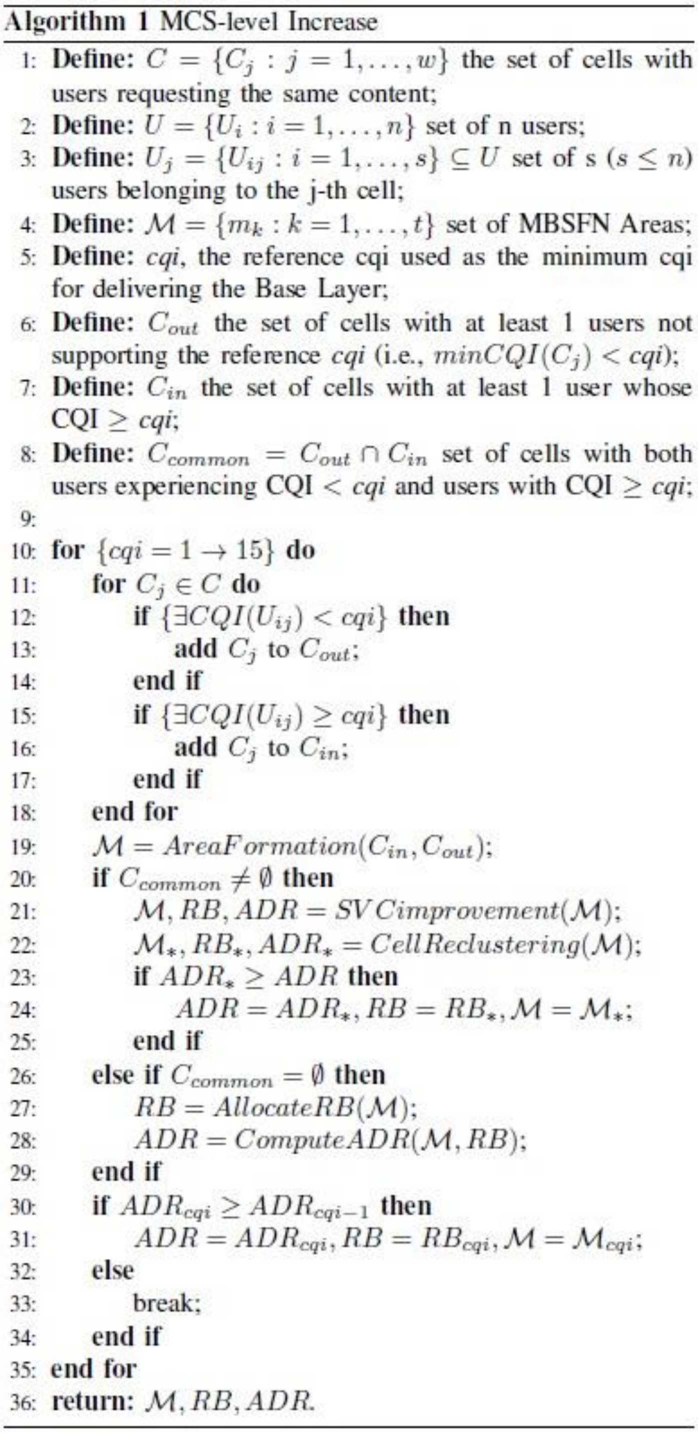


Fig. 3.1 - Example of MBSFN Area Configuration with a 57-cells Scenario

MCS-level Increase. MCS-level Increase is the main task of the DMAF algorithm. It includes the other successive sub-phases, Area Formation, SVC Improvement and Cell Re-Clustering phases (see Algorithm 1 MCS-level Increase). The objective t of this step is to define the best set of MBSFN Areas M and the proper allocation of RBs in order to maximise the system ADR. The algorithm iteratively increases the reference cqi from the minimum CQI index recorded in the Area to the maximum value achievable by the system (i.e. 15, [102]). Then, cells are grouped according to the CQIs experienced by their users. Two cell groups are created, one (i.e.C_{in}) with all cells with at least one user experiencing a CQI greater than the reference cqi, and another (C_{out}) with cells where there is at least one user not supporting the reference cqi (lines 11-18).

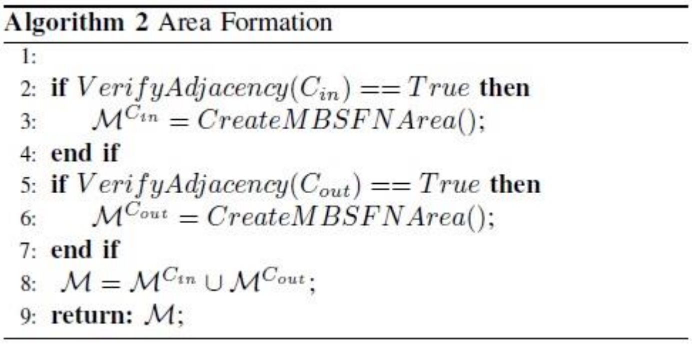
This means that the cells belonging only to C_{in} can support the MCS level increase decided by the current reference *cqi*, whereas in the remaining cells the minimum *cqi* determines the supported MCS. After this splitting, the MBMS Areas are created according to Algorithm 2 – *Area Formation* by including adjacent cells. Then, if the two Areas overlap, Algorithm 3 - *SVC Improvement* and Algorithm 4 - *Cell Re-Clustering* algorithms are carried out (lines 20-25); the same content is broadcasted at two different MCS levels and the available RBs are split between the two Areas. On the contrary if the Areas are disjoint, all available RBs are used for the content delivery with the selected MCS level to each Area (lines 26-29). The algorithm iteratively executes the operations above until the new system ADR is not lower than the *ADR* of the previous iteration (lines 30-34). The Algorithm terminates (line 36) and provides the best set M of MBSFN Areas, the number of RB to be assigned to each MBSFN Area, and the maximum ADR.

Area Formation. The Area Formation is described in Algorithm 2. It takes as outputs the two sets of cells, C_{in} and C_{out}, and checks whether there is an adjacency among the cells belonging to two sets. Each group of adjacent cells in the two sets form a MBSFN Area. The output of this phase are the two sets of MBSFN Areas $M^{C_{out}}$and $M^{C_{in}}$, where video bitstreams can be broadcasted at standard and high quality, respectively.



Algorithm 1 MCS-level Increase

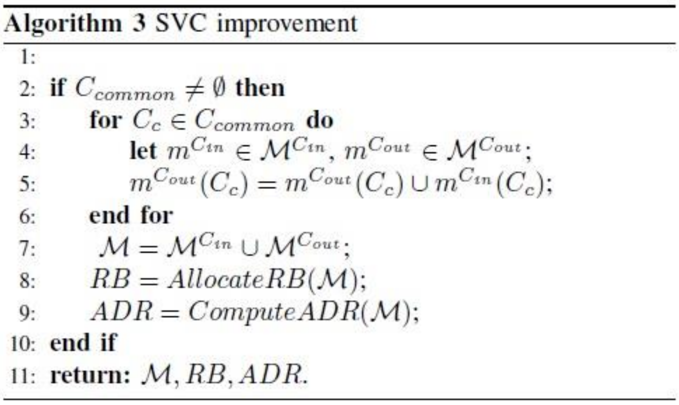
```
Define: $C = \{C_j : j = 1, \dots, w\}$ the set of cells with users requesting the same content;
Define: $U = \{U_i : i = 1, \dots, n\}$ set of n users;
Define: $U_j = \{U_{ij} : i = 1, \dots, s\} \subseteq U$ set of s ($s \leq n$) users belonging to the j-th cell;
Define: $\mathcal{M} = \{m_k : k = 1, \dots, t\}$ set of MBSFN Areas;
Define: $cqi$, the reference cqi used as the minimum cqi for delivering the Base Layer;
Define: $C_{out}$ the set of cells with at least 1 users not supporting the reference $cqi$ (i.e., $minCQI(C_j) < cqi$);
Define: $C_{in}$ the set of cells with at least 1 user whose CQI $\geq cqi$;
Define: $C_{common} = C_{out} \cap C_{in}$ set of cells with both users experiencing CQI $< cqi$ and users with CQI $\geq cqi$;

for $\{cqi = 1 \rightarrow 15\}$ do
    for $C_j \in C$ do
        if $\{\exists CQI(U_{ij}) < cqi\}$ then
            add $C_j$ to $C_{out}$;
        end if
        if $\{\exists CQI(U_{ij}) \geq cqi\}$ then
            add $C_j$ to $C_{in}$;
        end if
    end for
    $\mathcal{M} = AreaFormation(C_{in}, C_{out})$;
    if $C_{common} \neq \emptyset$ then
        $\mathcal{M}, RB, ADR = SVCimprovement(\mathcal{M})$;
        $\mathcal{M}_*, RB_*, ADR_* = CellReclustering(\mathcal{M})$;
        if $ADR_* \geq ADR$ then
            $ADR = ADR_*, RB = RB_*, \mathcal{M} = \mathcal{M}_*$;
        end if
    else if $C_{common} = \emptyset$ then
        $RB = AllocateRB(\mathcal{M})$;
        $ADR = ComputeADR(\mathcal{M}, RB)$;
    end if
    if $ADR_{cqi} \geq ADR_{cqi-1}$ then
        $ADR = ADR_{cqi}, RB = RB_{cqi}, \mathcal{M} = \mathcal{M}_{cqi}$;
    else
        break;
    end if
end for
return: $\mathcal{M}, RB, ADR$.
```



Algorithm 2 Area Formation

```
if $VerifyAdjacency(C_{in}) == True$ then
    $\mathcal{M}^{C_{in}} = CreateMBSFNArea();$
end if
if $VerifyAdjacency(C_{out}) == True$ then
    $\mathcal{M}^{C_{out}} = CreateMBSFNArea();$
end if
$\mathcal{M} = \mathcal{M}^{C_{in}} \cup \mathcal{M}^{C_{out}};$
return: $\mathcal{M};$
```

SVC Improvement. The Algorithm 3 shows the pseudocode of SVC Improvement phase. If the Areas formed in the previous phase overlap (line 3), i.e., they have at least one cell in common, then the RBs in each common cell are split to avoid interference between the Areas. This means that a subset of RBs is dedicated for low-rate (Base Layer) video delivery and another subset of RBs is allocated to high-rate (Enhanced Layer) video delivery. This RB splitting is implemented in whole Area the overlapped cells belong to (line 5). In the overlapped Areas, users with good channel conditions receive both bitstreams, so they get increased data rates and exploit all the available bandwidth.



Algorithm 3 SVC improvement

```
 1:
 2: if $C_{common} \neq \emptyset$ then
 3:     for $C_c \in C_{common}$ do
 4:         let $m^{C_{in}} \in \mathcal{M}^{C_{in}}, m^{C_{out}} \in \mathcal{M}^{C_{out}};$
 5:         $m^{C_{out}}(C_c) = m^{C_{out}}(C_c) \cup m^{C_{in}}(C_c);$
 6:     end for
 7:     $\mathcal{M} = \mathcal{M}^{C_{in}} \cup \mathcal{M}^{C_{out}};$
 8:     $RB = AllocateRB(\mathcal{M});$
 9:     $ADR = ComputeADR(\mathcal{M});$
10: end if
11: return: $\mathcal{M}, RB, ADR.$
```

Cell Re-Clustering. Algorithm 4 presents Cell Re-Clustering procedure. Its aim is to find other possible MBSFN Area configurations in order to further increase the system *ADR* in case of overlapping Areas.

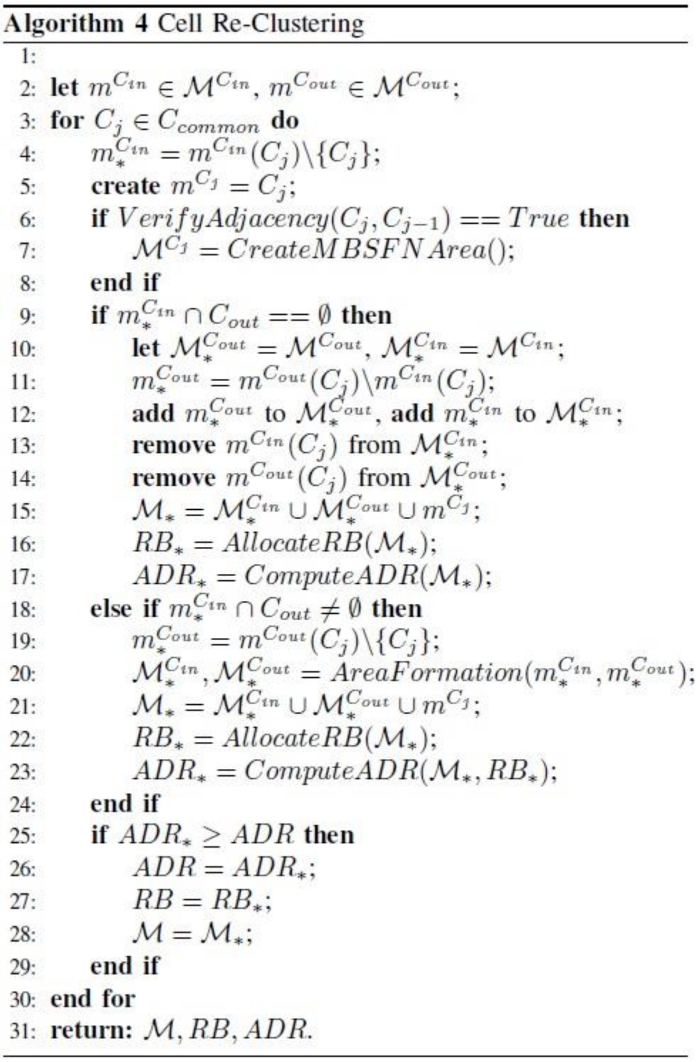
The Cell Re-Clustering phase attempts to increase the overall *ADR* by taking away from the MBSFN Areas those cells negatively affecting the system performance (line 4). Specifically, the algorithm proceeds by removing the cells in the overlapping region. This is due to the fact that in the previous SVC Improvement phase, all cells within the MBSFN area with higher MCS suffer a datarate reduction in some RBs, when the two flows (Base and Enhancement Layer) are considered. Whenever a cell is removed, Cell Re-Clustering checks if the remaining cells still form an MBSFN Area by verifying the adjacency (lines 6 - 8).

This phase of DMAF algorithm distinguishes two cases:

- Case 1: MBSFN Areas overlap over only once cell (lines 9-20);
- Case 2: MBSFN Areas overlap over more cells (lines 21-26).

In Case 1, the removed overlapped cell (i.e., C_j) from the possible MBSFN area $m^{C_{in}}$ constitutes a new area (i.e, m^{C_j}) where all RBs are allocated, and two other possible areas: one area broadcasting video at high data rate, $m_*^{C_{in}}$, and a second area broadcasting video at lower data rate, $m_*^{C_{out}}$ (lines 10-14). By doing so, the ADR of the considered cell (C_j) will decrease because only the Base Layer will be delivered; whereas, the ADR of the cells within $m_*^{C_{in}}$ will increase because all bandwidth will be dedicated to the Enhancement Layer.

In Case 2, Cell Re-Clustering includes Area Formation (line 22) and verifies whether all the cells left out are adjacent to each other in order to form a single MBSFN Area. Then, RBs are efficiently allocated (line 24) in order to further enhance the system ADR, that is recomputed in line 25. Therefore, in case of overlapping MBSFN Areas, Cell Re-Clustering checks whether it is better to keep cells inside the Enhanced MBSFN Area or to leave them out, providing as outputs the new set M of MBSFN Areas and the proper set RB of Radio Resources in order to further improve the system performance in terms of ADR.

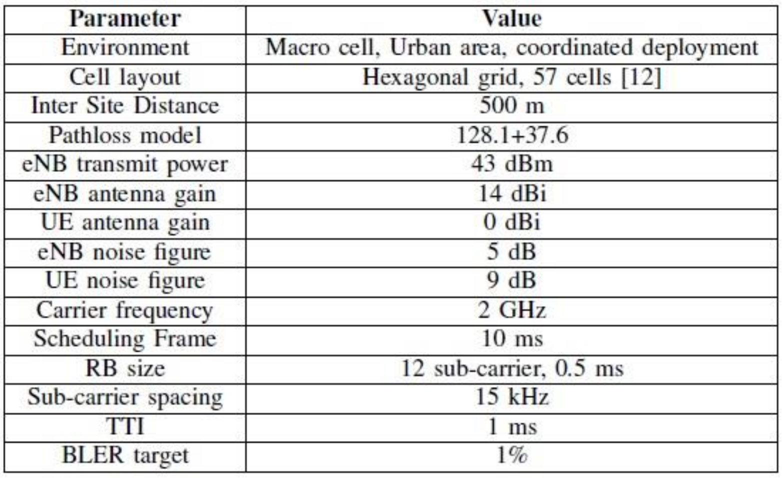


Algorithm 4 Cell Re-Clustering

```

let $m^{C_{in}} \in \mathcal{M}^{C_{in}}$, $m^{C_{out}} \in \mathcal{M}^{C_{out}}$;
for $C_j \in C_{common}$ do
    $m_*^{C_{in}} = m^{C_{in}}(C_j) \backslash \{C_j\}$;
    create $m^{C_j} = C_j$;
    if $VerifyAdjacency(C_j, C_{j-1}) == True$ then
        $\mathcal{M}^{C_j} = CreateMBSFNArea()$;
    end if
    if $m_*^{C_{in}} \cap C_{out} == \emptyset$ then
        let $\mathcal{M}_*^{C_{out}} = \mathcal{M}^{C_{out}}$, $\mathcal{M}_*^{C_{in}} = \mathcal{M}^{C_{in}}$;
        $m_*^{C_{out}} = m^{C_{out}}(C_j) \backslash m^{C_{in}}(C_j)$;
        add $m_*^{C_{out}}$ to $\mathcal{M}_*^{C_{out}}$, add $m_*^{C_{in}}$ to $\mathcal{M}_*^{C_{in}}$;
        remove $m^{C_{in}}(C_j)$ from $\mathcal{M}_*^{C_{in}}$;
        remove $m^{C_{out}}(C_j)$ from $\mathcal{M}_*^{C_{out}}$;
        $\mathcal{M}_* = \mathcal{M}_*^{C_{in}} \cup \mathcal{M}_*^{C_{out}} \cup m^{C_j}$;
        $RB_* = AllocateRB(\mathcal{M}_*)$;
        $ADR_* = ComputeADR(\mathcal{M}_*)$;
    else if $m_*^{C_{in}} \cap C_{out} \neq \emptyset$ then
        $m_*^{C_{out}} = m^{C_{out}}(C_j) \backslash \{C_j\}$;
        $\mathcal{M}_*^{C_{in}}, \mathcal{M}_*^{C_{out}} = AreaFormation(m_*^{C_{in}}, m_*^{C_{out}})$;
        $\mathcal{M}_* = \mathcal{M}_*^{C_{in}} \cup \mathcal{M}_*^{C_{out}} \cup m^{C_j}$;
        $RB_* = AllocateRB(\mathcal{M}_*)$;
        $ADR_* = ComputeADR(\mathcal{M}_*, RB_*)$;
    end if
    if $ADR_* \geq ADR$ then
        $ADR = ADR_*$;
        $RB = RB_*$;
        $\mathcal{M} = \mathcal{M}_*$;
    end if
end for
return: $\mathcal{M}, RB, ADR$.
```

3.3 Performance Evaluation

To demonstrate the effectiveness of the proposed DMAF algorithm, simulations are performed in MATLAB according to the guidelines for the coordinated multi-cell system model defined in [103]. We consider a 57-cells scenario, typically used by 3GPP for LTE network evaluation [79]. The coverage radius of each eNB is 250 m. The eNB transmit power is 43 dBm and its antenna gain is 14 dBi. For the UE, the antenna gain is 0 dBi. Table 10 lists more setting details.



Parameter	Value
Environment	Macro cell, Urban area, coordinated deployment
Cell layout	Hexagonal grid, 57 cells [12]
Inter Site Distance	500 m
Pathloss model	128.1+37.6
eNB transmit power	43 dBm
eNB antenna gain	14 dBi
UE antenna gain	0 dBi
eNB noise figure	5 dB
UE noise figure	9 dB
Carrier frequency	2 GHz
Scheduling Frame	10 ms
RB size	12 sub-carrier, 0.5 ms
Sub-carrier spacing	15 kHz
TTI	1 ms
BLER target	1%

Table 10 - Main Simulation Assumptions

Users are randomly distributed within the Synchronization Area and their position is assumed to be constant. We consider that cells belonging to the same MBSFN Area constructively interfere, while cells belonging to different MBSFN Areas act as interference sources. Furthermore, only one multicast video session is activated by different MBSFN Areas in the Synchronization Area. 'News' is the video flow for video streaming we considered in our analyses and its minimum data rate for the Base Layer is 121 kbps [101]. According to [100], where the optimal subgrouping configuration is achieved with two subgroups, we consider only one Enhancement Layer, despite the algorithm allows to create more Enhancement Layers. We compared the performance of DMAF algorithm with SCF algorithm [79] and the static MBSFN Area configuration specified by the 3GPP standard [93]. Furthermore, DMAF is also compared with the Single Cell Point to Multipoint (SC-PtM) [94] scheme. SC-PtM supports broadcast/multicast services over single cell, and the broadcast/multicast area can be dynamically adjusted cell by cell according to users distribution. We considered three simulation scenarios:

- Scenario A, where we varied the bandwidth (3, 5, 10, 15, 20 MHz), the number of cells is 57 and the number of users per cell is fixed to 60.

- Scenario B, where the number of users per cell varied from 60 to 100, the number of cells is fixed to 57 and bandwidth of 10 MHz (i.e., 50 RBs[1]).
- Scenario C, where we considered a variable number of cells (19, 26, 36, 46, 57) within the Synchronization Area, a variable number of users per cell (from 60 to 100) and fixed bandwidth of 10 MHz.

The described algorithms have been evaluated in terms of the following performance metrics:

- *Mean Throughput:* the average data rate experienced by users; the greater the throughput the higher the service quality and the "satisfaction" level of the multicast users.
- *Aggregate Data Rate (ADR) per cell:* the sum of the data rates experienced by multicast members in each cell.
- *Spectral Efficiency:* the ratio between the number of bits received by multicast users and the channel bandwidth exploited for the multicast transmission; this metrics indicated how efficiently the system resources are exploited during the multicast service provisioning.
- *User Outage:* the percentage of users excluded from the MBSFN transmission.
- *ADR Gain:* the percentage of improvement in terms of ADR introduced by DMAF with respect to literature-existing works.

Scenario A. The performance of Scenario A are evaluated by analysing the mean throughput achieved by multicast users, the ADR per cell and the spectral efficiency, when varying the bandwidth. As expected, both mean throughput and ADR per cell increase when increasing the bandwidth. Fig. 3.2 depicts the mean throughput experienced by multicast members. DMAF outperforms both legacy and SC-PtM approaches, achieving an 8-fold improvement with respect to them, when the bandwidth is 20 MHz. In lower bandwidth (3-5 MHz) cases DMAF shows a little loss with respect to SCF. This trend is due to the dependence among the created MBSFN Areas and the

---

[1] In the simulations we considered that at most 60% of RBs could be available for eMBMS, i.e., 30 RBs (of 50 RBS)

available RBs. Indeed, if few RBs are available DMAF algorithm creates only disjoint MBSFN Areas without benefiting from SVC advantages. In this case, the mean throughput suffers of cell-edge users' negative effects. Whereas, when increasing the bandwidth (from 10 to 20 MHz), DMAF mean throughput ranges from 7.11 Mbps to 15.8 Mbps providing an increasing gain, compared to SCF.

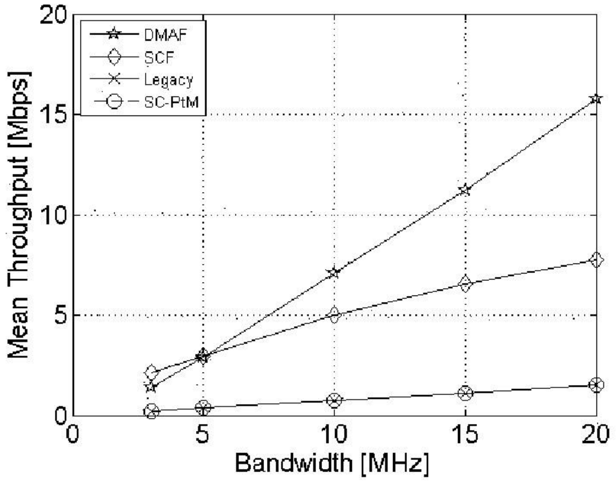


Fig. 3.2 - Mean Throughput – Scenario A

When focusing on ADR per cell (Fig. 3.3), we notice that DMAF algorithm provides better performance than both SC-PtM and the legacy. This is thanks to the exploitation of SVC technique that splits the video stream into a Base Layer and an Enhancement Layer. The former is delivered in the BL MBSFN Area and the latter in the EL MBSFN Area. In this way, users with good channel conditions are able to decode both layers, thus increasing their data rates that leads to such ADR improvement. Hence, DMAF enhances the legacy and SC-PtM by about 270% and 267%, respectively. The ADR achieved by SCF is higher than that of DMAF algorithm because the ADR maximization targeted by SCF is an opportunistic approach. This means that users who could achieve highest datarate are favoured with respect to bad channel conditions users. When increasing the bandwidth, the percentage of ADR loss of DMAF algorithm respect to SCF decreases to 9%, for a 20 MHz bandwidth. Nevertheless, SCF suffers of a high percentage of outage, just because of the intrinsic behaviour of such opportunistic approach (see Table 11). Whereas, DMAF serves all users interested in the eMBMS content.

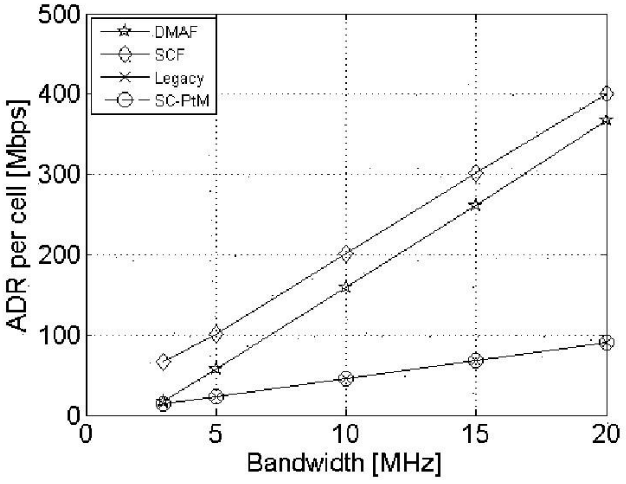
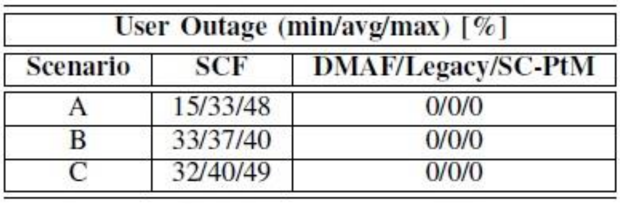


Fig. 3.3 - ADR per Cell - Scenario A



User Outage (min/avg/max) [%]		
Scenario	**SCF**	**DMAF/Legacy/SC-PtM**
A	15/33/48	0/0/0
B	33/37/40	0/0/0
C	32/40/49	0/0/0

Table 11 - User Outage Analysis

Results of spectral efficiency are illustrated in Fig. 3.4. It clearly emerges that both the legacy and SC-PtM suffer of poor spectral efficiency, which is 0.1523 bps/Hz and 0.1541 bps/Hz, respectively. Furthermore, their trends keep constant by varying the bandwidth. Finally, for a 3 MHz bandwidth the proposed DMAF has a 17% mismatch in spectral efficiency with respect to the SCF one. As for the means throughput, this is because with narrow bandwidth DMAF is not able to exploit SVC advantages. On the other hand, with a 20 MHz bandwidth, DMAF achieves the highest spectral efficiency of 1.6674 bps/Hz, whereas SCF obtains a spectral efficiency of 1.4624 bps/Hz. Hence, DMAF is the most performing algorithm also in terms of spectral efficiency.

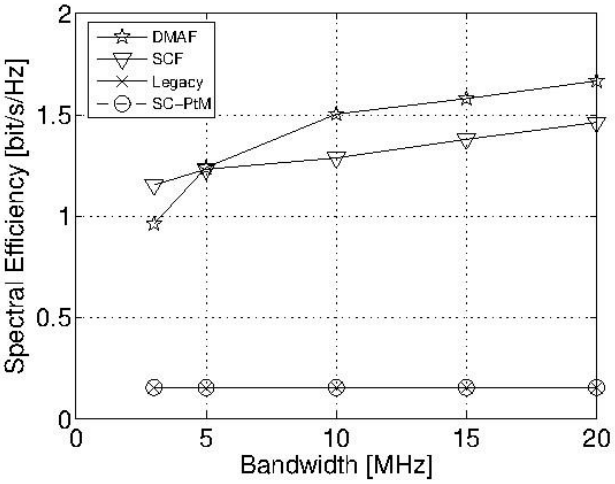


Fig. 3.4 - Spectral Efficiency - Scenario A

Scenario B. Scenario B provides performance results when varying the number of users per cell. Users' mean throughput is plotted in Fig. 3.5. As expected, both legacy and SC-PtM approach are the worst performing policy due to cell-edge users who experience the lowest CQI. Indeed, when increasing the average number of users per cell, the mean user throughput of these techniques keep constant to 0.74 Mbps. In the case of the DMAF algorithm, the mean user throughput decreases when increasing of the number of users in the system because of the greater probability of the presence of users with poor channel conditions. Whereas, in SCF case the mean throughput of MBSFN users increases thanks to the opportunistic approach of the ADR maximization, because the MCS level of the MBSFN transmission gets increased. It worth noting that the mean throughput is obtained only by users served without considering users left out of the MBSFN area, because of the limited unicast resources. Indeed, SCF shows a high users outage (Table 11). When focusing on Fig. 3.6, we notice that the greater the number of users per cell, the higher the ADR per cell.

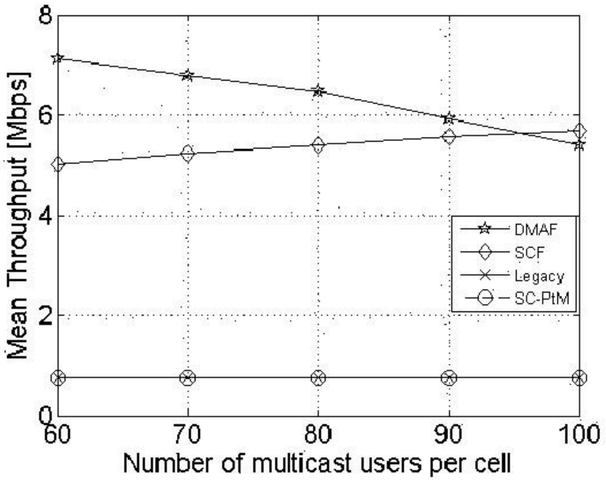


Fig. 3.5 - Mean Throughput – Scenario B

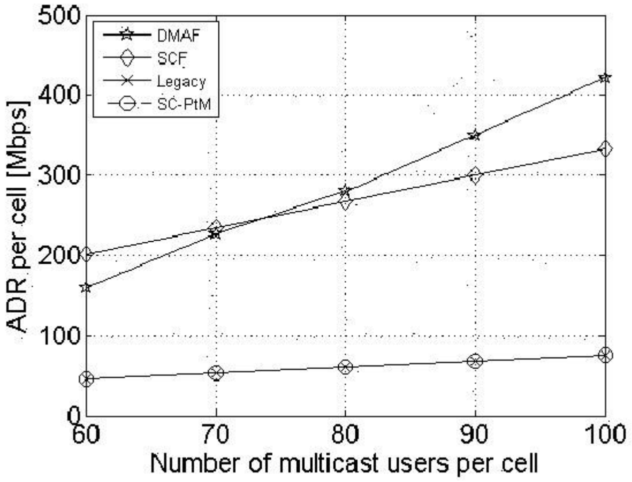


Fig. 3.6 - ADR Per Cell – Scenario B

Final comments are on spectral efficiency, plotted in Fig. 3.7. Also in this scenario, both the legacy and SC-PtM approaches suffer of poor spectral efficiency of 0.1526 bps/Hz and 0.1536 bps/Hz, respectively. Their trends do not change when varying the number of users per cell. Although ADR and mean throughput metrics are sometimes better in SCF case, DMAF achieves better spectral efficiency, thus demonstrating that the proposed approach better exploits the total bandwidth.

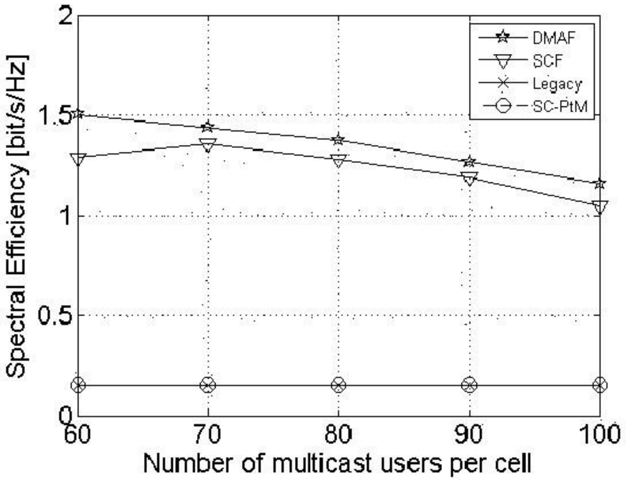


Fig. 3.7 - Spectral Efficiency – Scenario B

Scenario C. The objective in this scenario is to assess the performance of DMAF algorithm when varying both the number of cells over the Synchronization Area and the UEs distributions within each cell involved in the MBSFN Transmission. The Synchronization Area is progressively extended from a 19-cells scenario until a 57-cells one and users are randomly distributed. According to Fig. 3.8, we notice that the ADR per cell increases with the decrease of the number of users per cell because the probability of the presence of users with bad channel conditions gets lower. Furthermore, the greater the number of cells in the Synchronization Area, the more DMAF algorithm performs. Indeed, the best results in terms of ADR per cell is about 421.25 Mbps when considering a 57-cells scenario with 60 users per cell. To further investigate the effectiveness of DMAF algorithm, we used SCF as a benchmark. Therefore, in Fig. 3.9 we present the ADR gain in video streaming analysis for DMAF versus SCF when fixing the bandwidth to 10 MHz. With this additional analysis, we notice that DMAF introduce a 70% improvement in ADR performance with respect to SCF when increasing the number of cells with a fixed number of users per cell. Furthermore, the ADR per cell of DMAF decreases, when increasing the number of users per cell due to the high presence of users with poorest channel conditions. Therefore, we could sometimes observe a loss in the ADR gain compared to SCF. It is worth noting that SCF increases its ADR by delivering the content to the best users via broadcast; whereas users with poor channel gain are served via unicast, but not all of them can establish a Point-to-Point (PtP) communication due to limited RBs. In such a case, SCF achieves up 49% percentage of user outage (see Table 11) whereas DMAF serves 100% of users.

Finally, Table 12 summarizes the performance results of DMAF in Scenarios A and B. Whereas, Table 13 refers to Scenario C and shows Mean Throughput gain, ADR per cell gain and Spectral Efficiency gain introduced by the proposed DMAF w.r.t. the existing SCF algorithm.

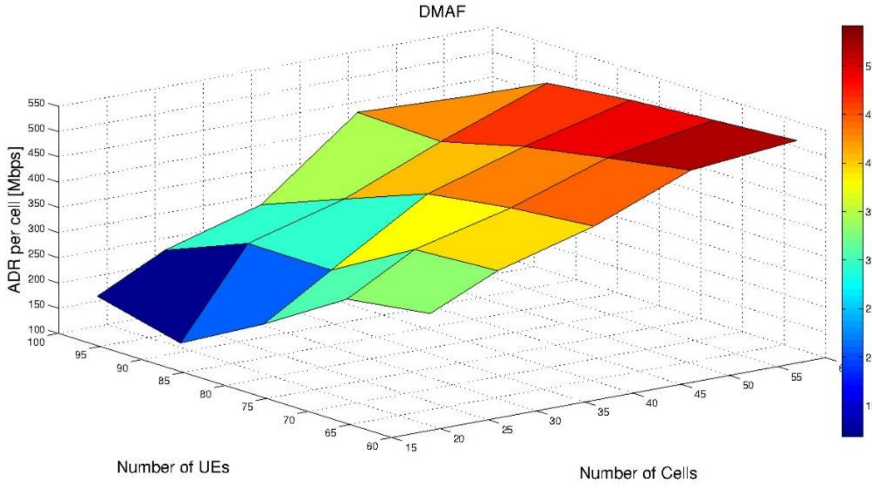


Fig. 3.8 – ADR – Scenario C

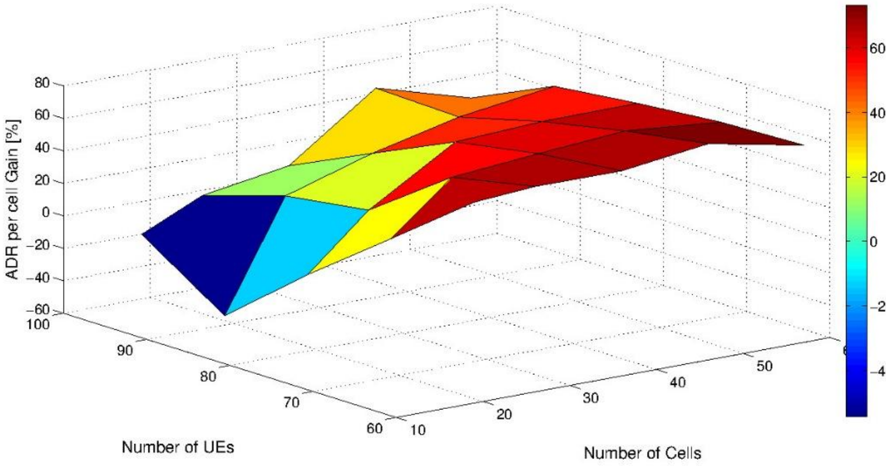
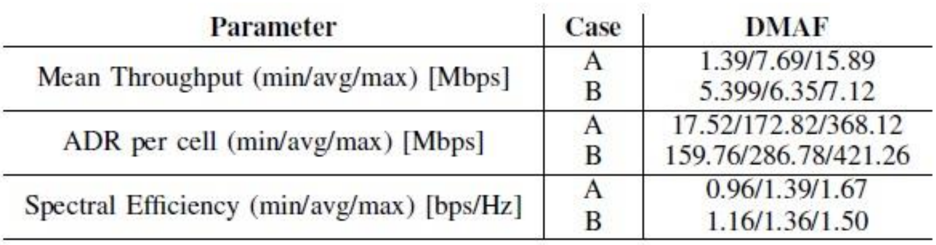
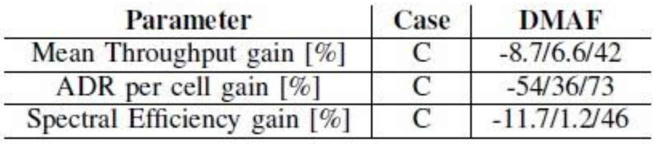


Fig. 3.9 - ADR Gain of DMAF w.r.t. SCF



Parameter	Case	DMAF
Mean Throughput (min/avg/max) [Mbps]	A B	1.39/7.69/15.89 5.399/6.35/7.12
ADR per cell (min/avg/max) [Mbps]	A B	17.52/172.82/368.12 159.76/286.78/421.26
Spectral Efficiency (min/avg/max) [bps/Hz]	A B	0.96/1.39/1.67 1.16/1.36/1.50

Table 12 - Summary Results Of DMAF Algorithm



Parameter	Case	DMAF
Mean Throughput gain [%]	C	-8.7/6.6/42
ADR per cell gain [%]	C	-54/36/73
Spectral Efficiency gain [%]	C	-11.7/1.2/46

Table 13 - Comparison DMAF-SCF – Scenario C

The proposed DMAF algorithm dynamically creates MBSFN Areas by exploiting the multicast subgrouping approach. Since each cell could be part of more overlapping MBSFN Areas, each of them delivering a video with a different quality level, the SVC technique have been also exploited. Therefore, the Base Layer is delivered to all users, whereas users with better channel conditions can receive also Enhancement Layers. The obtained results confirm that DMAF algorithm: *(i)* enhances the overall performance of all users, thanks to the multi-rate approach of subgrouping, *(ii)* improves the perceived video quality for users with higher CQIs, thanks to SVC technique; *(iii)* increases the ADR by choosing the best MBSFN Area configuration; *(iv)* reduces resource waste, thanks to the dynamic radio resource allocation and (v) also guarantees total coverage by serving 100% of users.

4 Multimedia Multicast Services in 5G Networks: Subgrouping and Non-Orthogonal Multiple Access Techniques

Following the current requirements and expectations from both users and operators, one of the main challenges of future 5G wireless networks is to include the efficient provision of mass mobile multimedia services through one or several broadcast transmission modes [104]. The demand for video services in mobile networks (i.e., streaming, downloading, conferences, live social broadcasting etc.) is rapidly increasing and it is expected that video will account for more than 80% of mobile data traffic by 2019[2]. Consequently, the 5G ecosystem will integrate seamlessly different network technologies [11] - including unicast, multicast and broadcast- in order to satisfy the increasing Media and Entertainment (M&E) content distribution[3]. In 5G systems, conventional application scenarios such as mobile pedestrian, vehicular to vehicular (V2V) and vehicular to infrastructure (V2I) communications, generally referred as to V2X communications, will gain higher interest with specific focus on dense urban scenario within the framework of Smart Cities [105].

Meanwhile, the mobile and broadcast industries have both developed independently several point to multipoint technologies to support large-scale consumption of mass multimedia services on mobile devices. Broadcasters have proposed Layered Division Multiplexing (LDM) to enable mobile TV on top of conventional terrestrial digital TV services [106]. LDM, which has been included in the ATSC 3.0 standard [107], is a non-orthogonal multiplexing (NOM)

---

[2] Cisco, White paper: "Cisco visual networking index: Forecast and Methodology", 2016-2021.

[3] New European Media (New), White paper, "5G and Media & Entertainment", January, 2016.

technique that has proven to be an efficient resource allocation scheme in terms of spectrum exploitation. A similar concept was previously presented for cellular environments in [108].

Regarding the broadband industries, 3GPP introduced eMBMS within the Long Term Evolution (LTE) systems and beyond to provide data transmissions from a single source to multiple devices [1]. Efficient resource allocation in case of multimedia delivery is a key issue for eMBMS. As to this, the implementation of subgrouping techniques has shown to be a promising solution [109].

This chapter presents a solution to jointly apply two of the aforementioned techniques. In particular, it is proposed to include LDM as an additional resource allocation mechanism, and what is more, consider it in the subgrouping decision process.

The remainder of the chapter is organized as follows. In the next section, the transmission based on LDM is descripted. Then, the proposed joint resource allocation scheme is presented, and following that, the solution is applied to a typical 5G scenario, where high quality video services are delivered to a group of users with different receiving conditions. Performance evaluation is finally discussed in the last section.

4.1 Cloud Transmission based on Layered Division Multiplexing (LDM)

Cloud Transmission (Cloud Txn) is a robust transmission system for terrestrial broadcasting or point-to-multipoint multimedia services [19]. This technology is robust to co-channel interference, immune to multipath distortion and spectrum re-use friendly. The system has the robustness required to provide mobile, pedestrian and indoor reception. Cloud Txn is considered the candidate technology to implement features of the next-generation broadcast systems, which are listed below:

- Robust performance against noise and co-channel interferences;
- Able to use all RF channels;
- More robust to interference from unlicensed devices and systems;
- Be robust against multipath distortion;

- Be able to provide reception for fixed, mobile, and indoor services while using low power (green) transmission;
- Allow for transmission towers to be located anywhere in the designated coverage area without worrying about co-channel interference and Single Frequency Network (SFN) multipath distortion delay spread;
- Be able to broadcast different programs from different SFN transmission towers for local programming or advertisement;
- Be good for both the traditional high-power high-tower approach and the small-cell approach;
- Provide low cost of implementation, maintenance, and operation.

The key issue for the Cloud Transmission System in exhibiting the above listed characteristics is its high degree of robustness against co-channel interference. In a digital system, the impact of co-channel interference is equivalent to that of Additive White Gaussian Noise (AWGN) [110]. Therefore, a desirable parameter for the proposed system is that it should have an AWGN SNR threshold of -2 to -3 dB. The negative SNR threshold value indicates that the system can withstand combined noise, co-channel interference and multipath distortion powers that are higher than the desired signal power. This means that with this system, the transmitters can be placed anywhere in the designated service area and transmit different programs. In a Cloud Txn SFN, transmitters can emit identical signals, different signals, or a combination of both cases. Different transmitters only require RF frequency lock. It is not necessary to lock the signal phase since a Cloud Txn receiver should be able to synchronize to, and decode the strongest signal.

Based on the information theory, there are two ways to improve the spectrum efficiency– increasing the data throughput, or increasing the reception robustness. Most of the current researches are focused on increasing the data throughput, or moving along the Shannon Limit to the right. [19] is going opposite direction– moving to the left. Although this will reduce the data throughput per RF channel, it will significantly increase the reception robustness and, at the same time, greatly increase the spectrum spatial re-use capability and transmission tower location flexibility. When a system has a low receiving threshold, its data rate must also be very low, representing a typical case of trading robustness for data rate.

As discussed above, although the Cloud Txn system is very robust, its data rate is very low. However, a simple approach to increase the system data throughput is to co-locate two Cloud Txn stations i.e., transmitting two Cloud Txn signals from the same tower and antenna. This is possible because the system can withstand strong co-channel interference. This effectively doubles the system data throughput.

Since the Cloud Txn system is very robust, it allows the use of **hierarchical spectrum re-use**, as shown in Fig. 4.1, to further increase its data throughput. With this approach, it is possible to insert a second digital stream (Stream B), e.g., a DVB-T2 signal, which has the same RF channel bandwidth, is frequency locked, and clock synchronized with the Cloud Txn signal (Stream A). The injection level, for example, can be 5 dB below the Stream A (see Fig. 4.1).

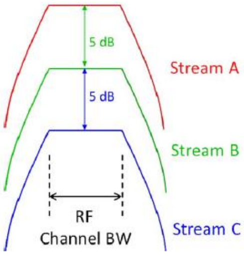


Fig. 4.1 - Hierarchical Spectrum Re-Use

In the next Figure is shown how, at the receiving end, the Stream A, is first decoded to get rid of transmission errors.

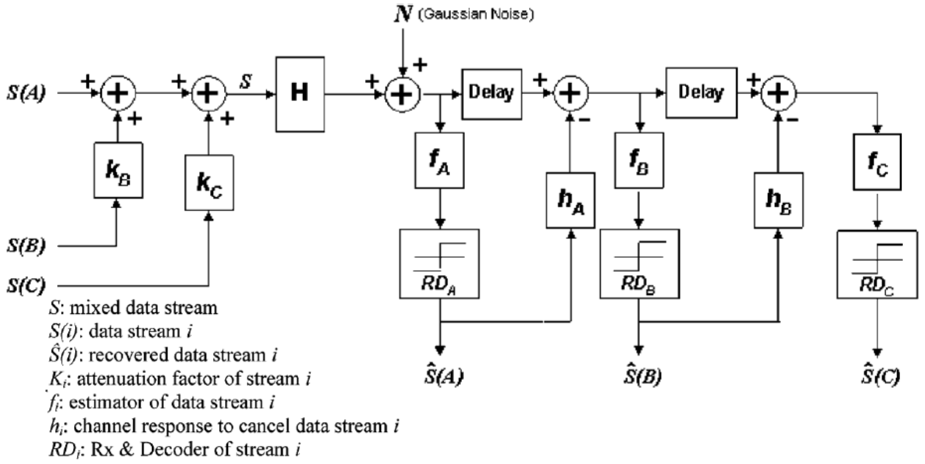


Fig. 4.2 - Cloud TXN Systems Diagram With Hierarchical Spectrum Re-Use

The decoded signal is then fed back and subtracted from the received combined signal (Stream A+B). The resulting output signal is then used to decode the Stream B (e.g., DVB-T2). To successfully decode the Stream B, assuming it is injected at 5 dB below the Stream A, a cancellation gain of about 10 dB on the Stream A signal needs to be obtained. This can be easily achieved, as presented in [111], [112].

The baseband system model shown in Fig. 4.3 provides an insight into the proposed transmitter. In a first approach, the FFT size and cyclic prefix (CP) length are equal for both layers, so that some modules are shared for both layers and the overall system performs cost-effectively. For example, the synchronization and carrier recovery system can work for both layers simultaneously. This configuration will reduce the computational complexity and simplify the design of a frequency domain cancellation algorithm. In order to allow the frequency domain cancellation at the receiver side and for hardware simplicity, both layers should share the same FFT size, in-band pilots and Guard Interval value. The additional computation power requirements for the second layer are OFDM demapping and subtraction.

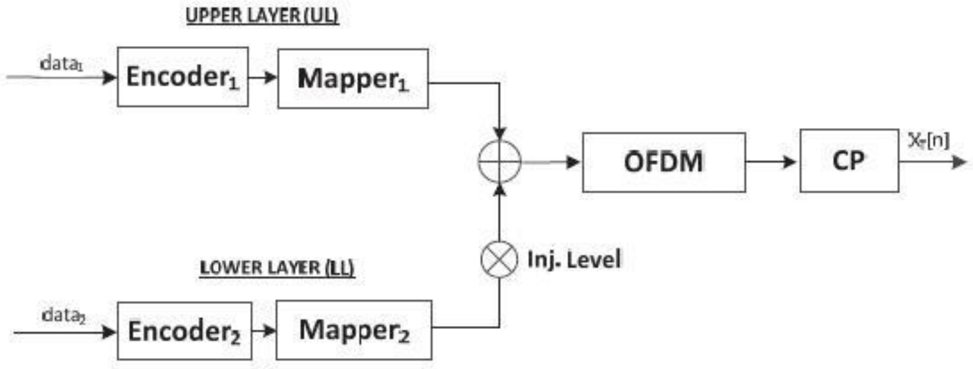


Fig. 4.3 - Cloud Tx transmitter scheme

Nevertheless, the modulation and encoding parameters defined for each data stream might be completely different depending on service providers' preference. The purpose of this design is to grant each layer its own robustness and data rate for different services (mobile TV, HDTV, UHDTV).

For instance, a low code rate and a low modulation order may be assigned to the upper layer (UL), targeting mobile services under very noisy environments; whereas the lower layer (LL) may be configured with high bitrates to provide high quality services (HD, 4K-HD) for fixed channels.

In a first approach, the upper layer binary bits are coded using an LDPC encoder prior to being mapped to a QPSK constellation, whereas the lower layer coded bits are mapped to a 16-QAM constellation. Another difference can be found on the code rate, i.e., the upper layer is to support robust mobile services, and thus a low code rate (R = 1/4) is used. However, the lower layer code rate should be higher (R = 1/2) to satisfy the increasing bitrates demand of the fixed services.

Once both signals are modulated, the UL is considered as the primary signal at the mappers output, superimposed to the LL stream using the frequency hierarchical modulation (See Fig. 4.3). In this architecture, the injection range (Δ) is the key parameter indicating how deep the LL is embedded, i.e., the superimposed signal is constrained by an injection level high enough, so that LL does not interfere with UL. It is important to note that both layers' decoding thresholds are affected by this value, a critical feature in the Cloud Transmission overall design.

LOWER LAYER	UPPER LAYER
(U)HDTV Mobile TV	Mobile TV
Fixed receivers with high antenna/high locations	Mobile receivers
Lower Tx power	Higher Tx power
Weaker channel coding	Robust coding
16/64/256-QAM	QPSK

Table 14 - Lower layer (LL) vs Upper Layer (UL)

One advantage of the Cloud-Txn system with LDM is that the lower-layer can use any existing good technology suitable for delivering high data rate digital broadcasting services, such as DVB-T2/NGH. The lower layer (LDM-LL) is used to deliver (U)HDTV service or multiple HDTV services to fixed receivers, where the operational SNR is usually high due to the large and possibly directional receiver antenna.

The use of hierarchical structure for delivering multiple streams is not new in broadcasting and has been proposed previously. Nevertheless, none of the existing proposals allows all streams (layers) to *transmit using 100% of the time and 100% of the RF channel bandwidth*. In comparison with Time Division Multiplex (TDM) system (ATSC mobile), frequency division multiplex (FDM) system (ISDB-T), or combined TDM and FDM system (DVB-T2), which either transmit data in part of the time or part of the RF channel bandwidth, the Cloud Txn system has the advantage on the *total aggregated data rate and better time-frequency diversity*.

The spectrum efficiency of the Cloud Txn broadcasting system depends to a great extent on the degree of robustness against co-channel interference and noise, especially for the top layer signal. It needs to perform well at very low SNR conditions, even in the negative SNR range. Cloud Txn employs a flexible ultra-robust coding and modulation scheme based on LDPC codes [113] [114] which provide extremely robust detection performance and enable successful demodulation and decoding even with negative SNRs.

In addition to the error correction capability, the other challenge that a new generation system must face is the quality of reception for mobile receivers.

In particular, a sizable interleaver is required to deal with different speeds, which may range from low mobility scenarios, corresponding to pedestrian users, to very fast time-varying scenarios, such as highway reception.

When the system is working in a multilayer hierarchical transmission, with two or more layers transmitted within the same RF channel, inter-layer interferences will appear. The lower layer signal will act as interference to the upper layer, which will reduce its noise tolerance capacity. Meanwhile, assuming a fixed total transmission power, adding the lower layer signal will also reduce the transmission power of the higher layer. Therefore, there is a two-fold impact from the lower layer signal to the upper layer signal: reducing the transmission power and acting as noise interference.

Channel estimation is critical for signal detection of the Cloud Txn system, firstly to decode the upper layer under very challenging conditions, and afterwards to perform accurate signal cancellation.

Cloud Txn Challenges	
Self-error correction	LDPC codes
Quality of reception for mobile receivers	Sizable interleaver
Performance of cancellation algorithm	Channel estimation
Inter-layer Interfercence	Impact of injection function

Table 15 - Cloud Txn Challenges

[115] presented a frequency domain cancellation algorithm that can efficiently decode Cloud Transmission multi-layered signal. As it can be seen in Fig. 4.4, the two layers are received simultaneously, and therefore, the first signal processing blocks are common to both layers: first, the cyclic prefix (CP) is removed, and then, the FFT is applied.

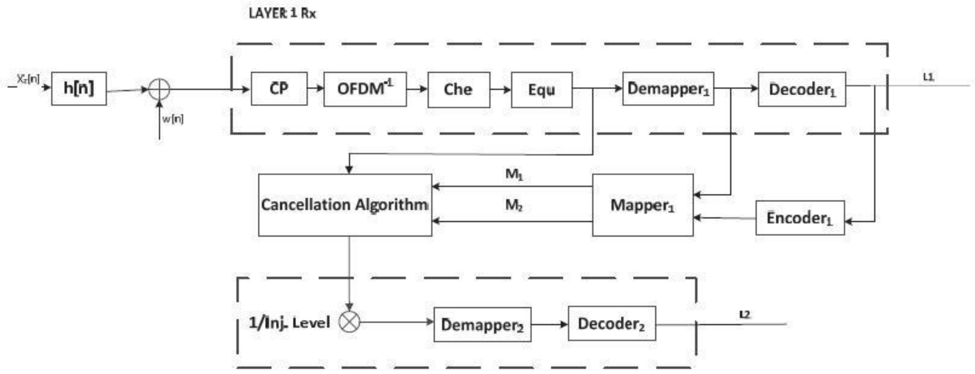


Fig. 4.4 - Frequency Domain cancellation receiver structure

At this point, it is important to note that the error due to non-ideal channel estimation is going to be amplified when the LL equalization is performed after the UL cancellation. That is the reason why the channel estimation algorithm is so important within this process. At this point, the first layer can be decoded without further processing, as the second layer is treated as regular white noise. Once the first layer is decoded, the next step is to subtract it. The cancellation algorithm is simple and is based on removing the first layer from the received signal at the equalization output. Afterwards, the second layer equalization is based on the application of a correction factor, which is the inverse of the injection range.

Nevertheless, as it can be seen in Fig. 4.4, there are two points where the first layer can be extracted, and each of these points leads into a different cancellation algorithm. The first one, M1, considers the equalization output after mapping for signal cancellation, whilst the second one, M2, rebuilds the transmitted signal from the decoded signal. The former method main advantage is that the complexity is notably reduced, whereas in M2 the whole transmitter chain should be created. Nevertheless, it is also true that the second method offers an error free feedback vector, which considerably reduces the cancellation error. In this case, the additional coding/interleaving delay is another consideration, and thus, signal structures that allow low latency should be used.

In [116] authors investigate the possibility of using a long FFT size (32k) for both signals is a very interesting option, since longer FFTs allow higher data rates. However, longer FFTs suffer from higher Doppler degradation in mobile scenarios, so a compromise between capacity and robustness should be considered before deciding the FFT size. Nine different Cloud Transmission signals will be broadcasted in an experimental network, using three FFT sizes (8k, 16k and 32k) combined with three different code rates (3/15, 4/15 and 5/15). For each configuration scheme, the received power will analyzed together with the percentage of correctly decoded frames in order to analyze the system performance.

In some cases the coverage percentage obtained with the largest FFT size (32K) is larger than the coverage obtained with the smallest FFT size (8k). One possible reason for the previous outcome is that the use of robust LPDC codes allows signal reception with negative C/N and, therefore, the multipath could be considered as the main distorting factor in the received signal as the ICI noise is masked by AWGN.

Using larger sized FFT might have other advantages such as a reduced guard interval percentage and higher data throughput. Also, for the same pilot and data carrier ratio, larger FFT means that the distance between adjacent pilots is smaller in Hz. This will improve the channel estimation accuracy, which will lead to better signal cancellation and better system performance.

The following table summarize the difference between the Cloud Txn, designed for broadband services and the LTE-based multicast subgrouping. In the following sections, a joint subgrouping-LDM approach have been proposed and evaluated.

	MULTICAST subgrouping	Cloud Txn
RF resources	50 RBs (49 group B, 1 RB group A)	100% BW in 100% time
BS total Tx Power	20 W	1 – 10 KW ERP
Pw per RF resources	Uniform distribution per RB	≈ 76% UL, ≈ 24% LL (5 dB injection function)
Receiver antenna	Omni directional	High directional antennas (Fixed receiver for LL)
Channel Coding	Turbo Code	LDPC (Low Density Parity Check)

Modulation	QPSK; 16/64 QAM	QPSK; 16/64/256 QAM

Table 16 - Multicast Subgrouping vs Cloud Txn

4.2 Subgrouping with NOMA Resource Allocation Strategies

In a conventional cellular network the devices physical connection with the network architecture is handled by the radio access network (RAN). This is the entity in charge of providing resources through different channel access multiplexing techniques to the user equipments (UEs) involved in the communication process. The wide range of solutions present in the literature can be gathered in two categories: the orthogonal multiple access (OMA) techniques and the non-orthogonal multiple access techniques (NOMA) [117]. In the first case, the available frequency/time resources in the network are orthogonally assigned to the different UEs. Thus, at the receiver site, under perfect conditions, the desired data can be unequivocally separated from the rest of the information. In the second case, the available resources, both frequency and time, are completely shared among different users, and therefore, when decoding the desired content the rest of the signals are considered an additional source of noise (See Fig. 4.5a).

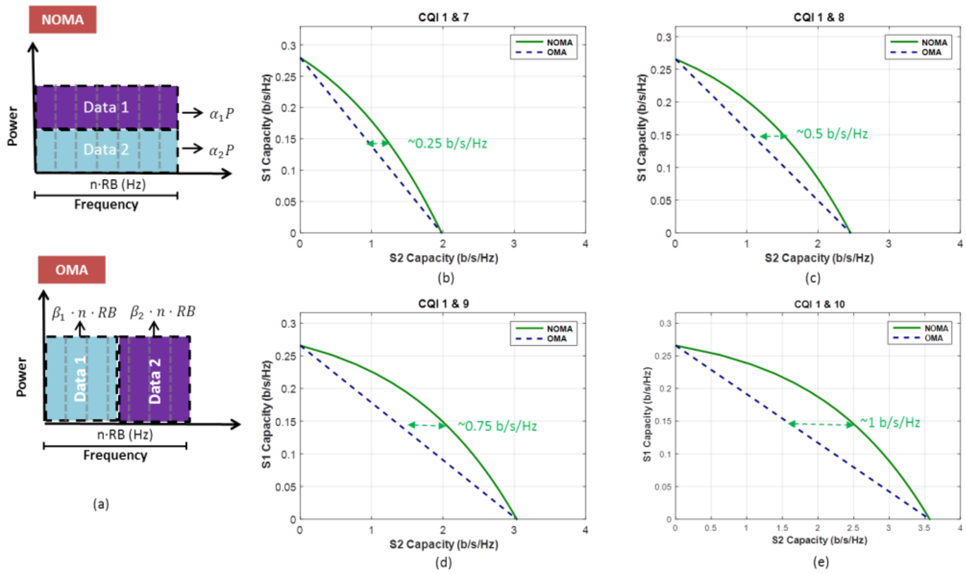


Fig. 4.5 – (a) The implementation of two different services in both NOME and OMA techniques. (b), (c), (d), (e). The theoretical capacity for different receiving conditions.

In the literature, the spectral efficiency of both proposals has been widely studied from an information theoretic point of view. In principle, NOMA based techniques had shown a higher efficiency, especially when the throughput rate among different users is asymmetric [117]. In Fig. 4.5 there have been depicted the different achievable theoretical spectrum efficiencies when two different services are multiplexed, either with NOMA (green solid line) or OMA (blue dotted line). The reception thresholds have been obtained from the set of the signal-to-interference-plus-noise ratio (SINR) values associated to the channel quality indicators (CQI) included at the current LTE release. As shown, the NOMA always performs better than OMA, especially for asymmetric scenarios, where the gain can be up to 1.0 b/s/Hz for the high SINR service. When both services CQI values are closer (i.e., CQI 1&7), the reception threshold difference is smaller, and consequently, the maximum gain is reduced down to 0.25 b/s/Hz.

3GPP has divided the 5G normative work into two phases, first of all, Phase-1 (Rel-15) will address the more urgent subset for commercial deployments, whereas Phase-2 (Rel-16) will address all the identified use cases and requirements. According to the preliminary technical reports 5G RAN will keep the LTE principles, with the same Orthogonal Frequency Division Multiple Access (OFDMA) technology for the downlink for both cases, either New Radio (NR) or Rel-16 [118]. What is more, this radio access should cover mainly the enhanced mobile broadband (eMBB) using mmWave frequencies, where the multicast and point-to-multipoint services will be a key for empowering vertical industries. Consequently, the subgrouping and NOMA techniques, which independently proved to be valuable for previous releases and standards, are jointly proposed as an innovative joint technology for facing the challenge of satisfying the mass media video consumption in the point-to-multipoint scenarios.

It is expected that in NR-5G, the Resource Block (RB), which corresponds to twelve consecutive subcarriers, will be also the smallest frequency resource, which can be assigned to a terminal. The set of available RBs is managed by the packet scheduler to efficiently handle resource allocation to mobile users in both frequency and time domain. Our proposal mainly focus on the frequency domain packet scheduler (FDPS), which is in charge to execute the fast link adaptation procedures by selecting the most appropriate MCS level and number of RBs for each multicast service. The NR-5G Base Station carries out such selection every Transmission Time Interval (TTI), which lasts 1 ms, by considering the CQI feedbacks received by all multicast users. According to the CQI each user will support a given MCS. In principle, a higher MCS level does not guarantee a better quality transmission for the user, whereas a lower MCS causes a poor spectrum exploitation.

Multicast transmissions are affected by users with the worst channel quality condition. Therefore, subgrouping techniques [109] are then presented as a valid solution splitting multicast users in different subgroups according to their capabilities. The main challenge is then the formation of the most appropriate subgroup configuration: the number of subgroups to create, the set of users, the MCS levels and the number of RBs assigned to each subgroup. Such configuration is tackled as an optimization problem, aiming at maximizing (or minimizing) a given objective function. The potential of using subgrouping techniques relies on the independency of the optimization from the objective function considered (e.g. throughput maximization, fairness optimization, minimum dissatisfaction index, etc. [109]).

Considering the conventional time and frequency resource allocation techniques, the best subgroup analytical configuration was found in [100], if Maximum Throughput (MT) is set as cost function. Authors demonstrated that the optimal solution can be found within 1 TTI, the optimal number of subgroup is no greater than two and the number of RBs to assign to the lowest subgroup is 1 or 2, whereas remaining RBs are assigned to the highest level subgroup. Starting from this result, this proposal aims to exploit the LDM technique as an alternative to optimize the resource allocation in the subgrouping process. In such a way, the FDPS also considers the LDM multiplexing in order to assign the different resources to the groups. In particular, in the case of LDM every subgroup will access to the whole frequency band the 100% of the time. The FDPS will assign a different weighted power to each service, performing a layered delivery, which is supposed to be more efficient than the classical multiplexing schemes.

Our proposal is to exploit the LDM concept of splitting the total available power for multicast subgrouping. In such an approach, the different layer defined through the LDM technique are matched with the subgroups derived by the multicast subgrouping approach.

4.3 Performance Evaluation

The performance of the proposed technique have been evaluated according to three different objective function: Maximum Throughput (MT), Proportional Fairness (PF) and Minimum Dissatisfaction Index (MDI) [100]. Maximum Throughput is based on the maximization of a cost function defined as the Aggregate Dara Rate (ADR), which is the sum of the data rate obtained by all the multicast members. The Proportional Fairness scheduling can be accomplished by maximizing the sum of the logarithm of the data rate. Finally, MDI is defined as the weighted difference between the data rate achieved by the

UE and the maximum possible value of data rate achieved by a UE when all RBs are assigned to the user.

A dense urban scenario is the typical environment for characterizing 5G communications in a Smart City context, with different types of users (i.e., pedestrian, vehicular, mixed, and fixed). Several mobility models can be used to describe the activity pattern of users changing position, speed, and similar location related characteristics. In this work, in order to compare the performance between the classical multicast subgrouping and the multicast LDM subgrouping methods, and taking into account the use cases presented in previous section, the Random Way Point model has been implemented. In this model, the users are uniformly distributed over the environment at the initial stage. At every iteration they move along line segments toward a random destination position, with a fixed speed. When the user reaches the target position, waits for a predetermined time interval and then a new random target position is assigned. In our simulations, users are allowed to change direction every second. This allowed to realistic reproduce pedestrians and slow vehicles moving behavior in dense urban environments, where pedestrians walk in crowd spaces (e.g., sidewalk, square, mall, etc.), whereas vehicles often circulate in severe traffic condition (e.g., traffic jam, peak hours).

Afterwards, the channel conditions for each UE are evaluated in terms of the experienced SINR, which results on an effective CQI level assuring a Block Error Rate less than 10%. Then, these results are fed to a simulation model, where the LTE scheduler RRM procedures are simulated, and the resources are shared among the different created subgroups with ideal MCSs. Eventually, the final output parameters for evaluation are the Aggregated Data Rate (ADR) and Throughput per User (TU), which are obtained in order to evaluate and compare the proposed methodologies.

The mobility model simulations follow 3GPP standards. The simulation duration is 180 seconds. The RandomWayPoint mobility scenario metrics have been obtained for an assumed UE speed of 3km/h. In this case, among all the use cases explained previously, we are focusing our research on low speed mobility scenarios. Nevertheless, the results can be extended to other mobility models and different speed values. Performance analysis has been carried out comparing the proposed solution with the LTE-based multicast subgroup technique [78] (i.e., labeled with T/FDM in the next figures).

The first evaluated parameter is the aggregated data rate of the cell, which depends on the assigned optimization cost function of the subgrouping process (See Fig. 4.6). The first important outcome is that the maximum throughput metric does not provide meaningful difference between the different multiplexing methods (Fig. 4.6(a)). That is because, for both approaches, the best solution is obtained offering the maximum number of possible resources to the subgroup with best reception conditions, while the users dealing with more challenging conditions are poorly satisfied. In the case that PF or MDI are used (Fig. 4.6(b) and (c)), it can be clearly noticed how LDM offers a significate gain (about ~10 Mbps in the best case) with respect to the OMA subgrouping technique (i.e., T/FDM in the figures). This means that LDM achieves higher performance in terms of ADR, when exploiting such "fair" metrics.

Finally, the maximum throughput metric is modified, adding a minimum bit rate constraint of 0.5 Mbps (Fig. 4.6 (d), i.e. Constrained MT (CON.MT)) to one of the services. In this case, it is easily shown that the gain is much higher for NOMA subgrouping, because in OMA subgrouping approach, according to this constraint, many RBs must be assigned to the low-level subgroups. On the contrary, with LDM both subgroups exploit the whole bandwidth. This assumption represents those use cases, which need to guarantee a minimum data rate to all users

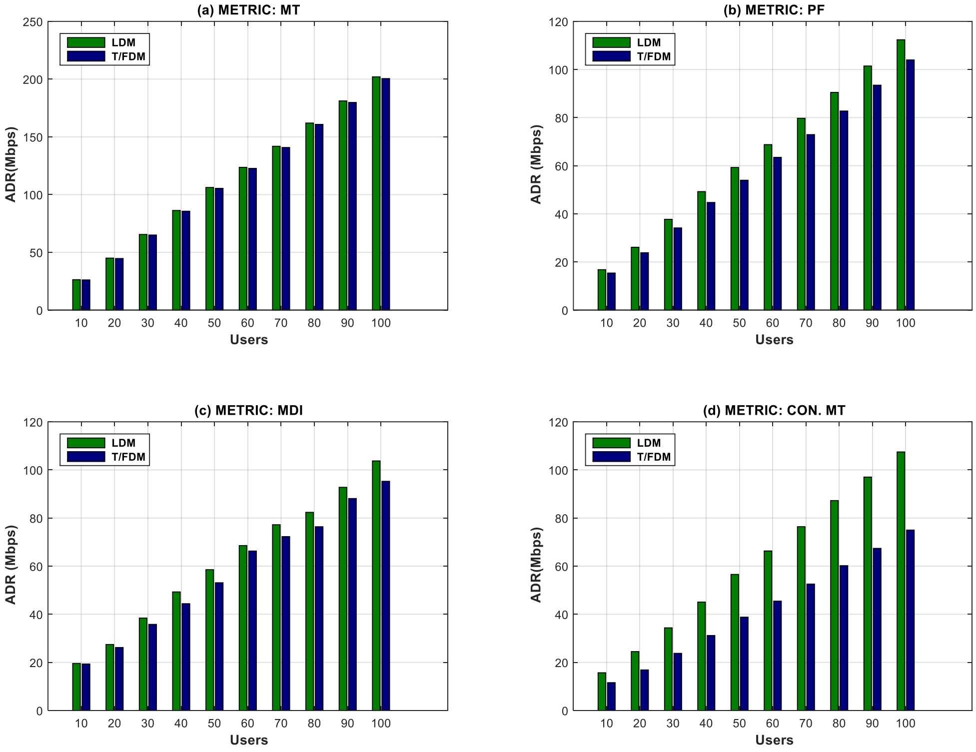


Fig. 4.6 - ADR – Low Mobility Case, with Different Objective Metrics.

.The reason for this behavior can be found in Fig. 4.7. In this case, each subplot indicates the throughput per user, differentiating the average throughput of each of the two subgroups. In this case, the good SNR condition service is marked as HTH (High Throughput) and the poor SNR condition group is tagged as LTH (Low Throughput). As expected, according to the theoretical facts explained before, the higher the required bitrate assigned to most challenging group the bigger is the gain offered by LDM. It can be noticed that in the MT and MDI cases the LTH subgroup receives a very poor rate (Fig. 4.7(a), (c)). Fig. 4.7(b) with PF and in Fig. 4.7(d) when the MT is modified, the gain can be up to 2 Mbps per user for the high capacity subgroup. It is expected that this trend will be maintained in the newly designed 5G environments, where the available bandwidths will be bigger and the required minimum bitrates will be higher due to the fact of the user expectations.

NOMA is considered a strong candidate to be included in the 5G ecosystem. Consequently, the challenges associated with its efficient integration in the RAN architecture have drawn a lot of attention. Some of them, will also be useful for the technical solution proposed in this paper. For instance, there are studies analyzing the impact of the path loss in the NOMA performance or the feasibility to combine non-orthogonal multiplexing techniques with cooperative communications. In addition, NOMA is also considered a very powerful tool for massive MIMO technologies. What is more, recently there has been also proposed in one of the most promising communication paradigms: the visible light communications (VLC) 0. Finally, regarding the LDM implementation, the error cancelation impact on the overlaid layers and the low complexity implementations should also be studied.

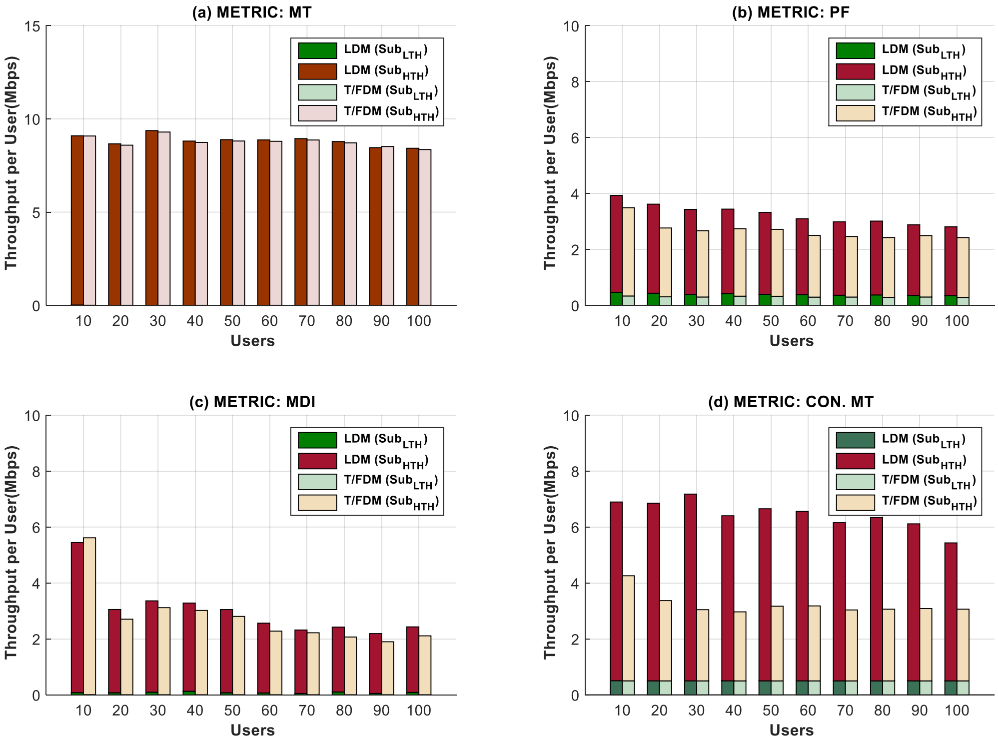


Fig. 4.7 - Throughput – Low Mobility Case, with Different Objective Metrics.

NOMA is considered a strong candidate to be included in the 5G ecosystem. Consequently, the challenges associated with its efficient integration in the RAN architecture have drawn a lot of attention. Some of them, will also be useful for the technical solution proposed in this chapter.

NOMA is also considered a very powerful tool for massive MIMO technologies. What is more, recently there has been also proposed in one of the most promising communication paradigms: the visible light communications (VLC) 0. Finally, regarding the LDM implementation, the error cancelation impact on the overlaid layers and the low complexity implementations should also be studied.

In the case of the subgrouping techniques, the main challenge is to adapt the current metrics to the requirements expected on the 5G ecosystem, for supporting not only bandwidth hungry services but also machine/IoT group oriented applications. That is to say, new architectures, protocols and metrics will be needed in order to guarantee high quality services to the different groups of receivers.

Therefore, the integration of NOMA with subgrouping techniques looks a very promising solution in order to boost datarate when groups of users require broadband video applications.

Conclusions

In this research, we explore the fifth generation (5G) mobile networks. Particular attention have been addressed on the Mobile Video, an application that will reach the 80% of the total amount of mobile traffic within 2019. 5G system aims to manage the ever-growing traffic demand, the huge number of connected devices e to improve the performance of current cellular networks both in terms of received data rates and latency. Group-oriented communications, well known as Multicasting, are an important solution that could meet such issues. They exploit point-to-multipoint communications, in order to efficiently manage the limited available radio resources. It means that a video content could be delivered at the same time to many users exploiting the same radio transmission. This results in higher bandwidth efficiency and increasing capacity. Nevertheless, Radio Resource Management (RRM) is a challenging issue that will be exacerbated by the increasing number of connected devices. Indeed, the presence of many users in different channel conditions affects the multicast transmission.

Thus, this work focuses on RRM in multicast transmissions for supporting emerging 5G systems. Forthcoming 5G systems will exploit different enabling technologies to achieve the required target of high capability and increased data rates. The current LTE-A cellular system as well as Dense Heterogeneous Networks, Single Frequency Networks, are some of the enabling technologies, which will support the 5G deployment.

During this research the RRM in multicast transmissions have been analysed from different perspectives according to such technologies. More in detail the contribution of the thesis can be summarized as follow.

The first contribution is provided in the Chapter 1. Here, a survey of 5G challenges in the view of an effective management of multicast applications is presented. The surveyed literature provides interesting results, though these are mostly aimed at improving the data rates experienced by multicast users and the spectrum utilization. Nevertheless, we then present the issues still needed to be addressed, in order to meet the multicast 5G requirements. Moreover, the current Multimedia Broadcast/Multicast Service (MBMS) have been considered, in order to show its limitations for the multicast implementations over the 5G systems. Indeed, by accounting for the presence of both human and machine-related traffic, the needed enhancement to the mobile network architecture have been proposed. Then, future research directions have been identified, especially in the support of machine-type communications, as the control messaging, control overhead. Open issues related to human-oriented applications, like video applications, are finally summarized posing a starting point for the next of the research work.

In the second chapter, a network selection approach that exploits the benefit of co-existing unicast and multicast transmissions during video deliveries is presented. The proposed algorithm, named Hybrid Unicast-Multicast utility-based Network Selection (HUMANS), considers bandwidth utilization and the trade-off between quality and energy consumption when delivering video in DenseNet scenarios. The considered scenario is a Dense Heterogeneous environment, where users try to access to video content. In such a scenario, each user has to select the most appropriate network to which connect, in order to balance a trade-off between some network parameters. The network selection is carried out by computing, for each neighbour network a utility-score that considers the estimated energy consumption of the mobile device running a real-time video application, the estimated achievable data-rate, the utilized resources and the expected user satisfaction level. A major contribution of HUMANS is the consideration of joining a multicast group as a possible option in the network selection process, thus allowing for smart bandwidth management. HUMANS serves users with good channel conditions via unicast transmissions and the remaining users via multicast. Performance evaluation carried out in low- and high-density scenarios, demonstrate how the proposed hybrid unicast-multicast approach provides a significant improvement in terms of capacity and radio resource utilization in comparison with other unicast-only solutions. The performance gain is much higher when user density increases within a system, thus providing an interesting solution for the future dense networks. Furthermore, it has been shown cases unicast transmission is activated to users with high CQI levels (i.e., in good channel condition) only. This is because users in good channel condition require a few RBs, whereas users experiencing bad channel conditions need more resources to obtain the required datarate. In HUMANS, the users with lower CQI levels are served via multicast transmissions and this has a double advantage: (i) they do not waste additional resources and (ii) make use of multicast flows (i.e. they receive all the RBs dedicated to the multicast group). Therefore, thanks to this approach, users requiring many resources that cannot be served if an only-unicast oriented algorithm is implemented, can always receive the video service, especially when the system is in high load conditions.

Future extensions of this work will account for the variation of both the background traffic in the network and the weights assigned to each utility. Finally, the proposed algorithm will be deployed in future dense wireless networks scenarios by also exploiting other solutions, such as Device-to-Device (D2D) communications for network traffic overloading and innovative management of the least channel gain users.

Following the above consideration on RRM in DenseNet, in chapter 3 the attention moved towards Single-Frequency Networks (SFN). MBMS for SFN (MBSFN) has not been deeply investigated in literature. The area formation is one the main issues that need to be addressed for better exploiting the full potential of MBSFN. Therefore, here we proposed a Dynamic MBSFN Area Formation algorithm for multicast service delivery in 5G Wireless Networks. The proposed DMAF algorithm dynamically creates MBSFN Areas by exploiting the multicast subgrouping approach. Each area is formed according to four-step algorithm. First, the minimum MCS is increasing, in order to enhance the aggregate data rate of the system. Then, users with are not able to decode such MCS, are grouped in a different MBSFN area. Since each cell could be part of more overlapping MBSFN Areas, each of them delivering a video with a different quality level, the SVC technique have been also exploited. Therefore, the Base Layer is delivered to all users, whereas users with better channel conditions can receive also Enhancement Layers. It means that the Base

Layer is delivered within areas with users in bad channel conditions, where the Enhancement Layer is delivered to users in better channel conditions. The results confirm that DMAF algorithm enhances the overall performance of all users, thanks to the multi-rate approach of subgrouping; improves the perceived video quality for users with higher CQIs, thanks to SVC technique; increases the ADR by choosing the best MBSFN Area configuration; reduces resource waste, thanks to the dynamic radio resource allocation and also guarantees total coverage by serving 100% of users.

The last research contribution is presented in the Chapter 4. Here, the RRM has been faced from a different point of view. The LDM technique, a Non-Orthogonal Multiplexing Access (NOMA), which is typical for broadband system as DVB-H, has been exploited jointly with the multicast subgrouping approach in order to enhance users' data rates. The LDM exploit the hierarchical spectrum re-use and is able to deliver two different video flows in the same time using the same radio frequencies. Such an approach has been integrated with the subgrouping multicast algorithm. Therefore, the two multicast subgroups, of users with god channel conditions and bad channel condition, respectively, corresponds to the two LDM layers. Preliminary analysis demonstrated that this approach achieve better performance with respect to standard orthogonal access technique (i.e., OFDMA, used in LTE-A cellular system). The gain is much higher as greater is the asymmetry between the two groups. On the other hand, this technique requires more complexity from the hardware point of view, which is still unfeasible for mobile devices. This is a for sure one of future research starting point of this joint approach LDM-multicast, together with a further study regarding both the number of possible layer that could be deployed and the transmission power level of each layer.

References

[1] D. Lecompte, F. Gabin, «Evolved Multimedia Broadcast/Multicast Service (eMBMS) in LTE-Advanced: Overview and Rel-11 Enhancements,» *IEEE Communications Magazine,* vol. 50, n. 11, pp. 68-74, Nov 2012.

[2] 3GPP, «TS 23.246, Multimedia Broadcast/Multicast Service (MBMS); Architecture and functional descrition, Rel. 13,» Mar. 2015.

[3] F. Boccardi, R.W. Heath, A. Lozano, T.L. Marzetta, P. Popovski, «Five disruptive technology directions for 5G,» *IEEE Communications Magazine,* vol. 52, n. 2, pp. 74-82, Feb. 2014.

[4] B. Zhou, H. Hu, S.-Q. Huang, H.-H. Chen, «Intracluster Device-to-Device Relay Algorithm With Optimal Resource Utilization,» *IEEE Transactions on Vehicular Technology,* vol. 62, n. 5, pp. 2315-2326, Jun. 2013.

[5] L. Song, D. Niyato, Z. Han, E. Hossain, «Wireless Device-to-Device Comunications and Networks,» Cambridge University Press, 2015.

[6] W. Lee, H. Park, H.-B. Kong, J. S. Kwak, I. Lee, «A New Beamforming Design for Multicast Systems,» *IEEE Transactions on Vehicular Technology,* vol. 62, n. 8, pp. 4093-4097, Oct. 2013.

[7] O. Mehanna, N. Sidiropoulos, G. Giannakis, «Joint Multicast Beamforming and Antenna Selection,» *IEEE Transactions on Signal Processing,* vol. 61, n. 13, pp. 2660-2674, May 2013.

[8] A. Tassi, I. Chatzigeorgiou, D. Vukobratovic, «Resource-Allocation Frameworks for Network-Coded Layered Multimedia Multicast Services,» *IEEE Journal on Selected Areas in Communications,* vol. 33, n. 2, pp. 141-155, Feb, 2015.

[9] T. Zhou, B. Xu, T. Xu, H. Hu, L. Xiong, «User-specific link adaptation scheme for device-to-device network coding

multicast,» *IET Communications,* vol. 9, n. 3, pp. 367-374, Feb. 2015.

[10] White Paper on, «5G: Challenges, Research Priorities, and Recommendations,» [Online]. Available: http://www.networld2020.eu/sria-and-whitepapers/.

[11] G. Araniti, M. Condoluci, P. Scopelliti, A. Molinaro and A. Iera, «Multicasting over Emerging 5G Networks: Challenges and Perspectives,",» *IEEE Networks,* vol. 31, n. 2, pp. 80-89, March/April 2017.

[12] J. Liu, N. Kato, J. Ma, N. Kadowaki, «Device-to-Device Communication in LTE-Advanced Networks: A Survey,» *IEEE Communications Surveys & Tutorials,* vol. 17, n. 4, pp. 1923-1940, Fourth Quarter 2015.

[13] H. Nishiyama, M. Ito, N. Kato, «Relay-by-smartphone: realizing multihop device-to-device communications,» *IEEE Communications Magazine,* vol. 52, n. 4, pp. 56-65, Apr. 2014.

[14] L. Militano, M. Condoluci, G. Araniti, A. Molinaro, A. Iera, «When D2D communication improves group oriented services in beyond 4G networks,» *Wireless Networks,* pp. 1-15, Nov. 2014.

[15] Y. Zhou, H. Liu, Z. Pan, L. Tian, J. Shi, G. Yang, «Two-Stage Cooperative Multicast Transmission with Optimized Power Consumption and Guaranteed Coverage,» *IEEE Journal on Selected Areas in Communications,* vol. 32, n. 2, pp. 274-284, Feb. 2014.

[16] L. Militano, M. Condoluci, G. Araniti, A. Molinaro, A. Iera, 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, n. 2, pp. 263-278, Jun. 2015.

[17] J.-B. Seo, T. Kwon, V. Leung, «Social Groupcasting Algorithm for Wireless Cellular Multicast Services,» *IEEE Communications Letters,* vol. 17, n. 1, pp. 47-50, Jan. 2013.

[18] X. Zhang, R. Yu, Y. Zhang, Y. Gao, M. Im, L. Cuthbert, W. Wang, «Energy-efficient multimedia transmissions through base station cooperation over heterogeneous cellular networks exploiting user behavior,» *IEEE Wireless Communications,* vol. 21, n. 4, pp. 54-61, Aug. 2014.

[19] Y. Wu, B. Rong, K. Salehian, G. Gagnon, «Cloud Transmission: A New Spectrum-Reuse Friendly Digital Terrestrial Broadcasting Transmission System,» *IEEE Transactions on Broadcasting,* vol. 58, n. 3, pp. 329-337, Sep. 2012.

[20] H. Park, K. Singh, K. Piamrat, J.-M. Bonnin, C. Viho, «Adaptive beam scheduling for scalable video multicast in wireless networks,» *Electronics Letters,* vol. 50, n. 16, pp. 1167-1169, Jul. 2014.

[21] «The 1000x Mobile Data Challenge: More Spectrum, More Small Cells,,» Qualcomm, San Diego, CA, USA, 2012. [Online]. Available: http://www.qualcomm.com/.

[22] D. López-Pérez, M. Ding, H. Claussen and A. H. Jafari, «Towards 1 Gbps/UE in Cellular Systems: Understanding Ultra-Dense Small Cell Deployments,» *IEEE Communications Surveys & Tutorials,* vol. 17, n. 4, pp. 2078-2101, Fourth Quarter 2015.

[23] X. Chu, D. López-Pérez, Y. Yang, and F. Gunnarsson,, «Heterogeneous Cellular Networks: Theory, Simulation and Deployment,» Cambridge University Press, 2013.

[24] J. G. Andrews, H. Claussen, M. Dohler, S. Ragan, and M. C. Reed,, «Femtocells: Past, present, and future,» *IEEE Journal on Selected Areas in Communications - Special Issue: Femtocell Netw.,* vol. 30, n. 3, p. 497–508, Apr. 2012.

[25] 3GPP, «Evolved Universal Terrestrial Radio Access (E-UTRA); Carrier aggregation; Base Station (BS) radio transmission and reception,» Sophia Antipolis, France, 2013.

[26] 3GPP, «Feasibility study on licensed-assisted access to unlicensed spectrum,» Sophia Antipolis, France, 2014.

[27] Y. A. L. W. a. M. J. L. L. Hanzo, «MIMO-OFDM for LTE,WiFi and WiMAX: Coherent Versus Non-Coherent and Cooperative Turbo Transceivers,» Hoboken, NJ, USA, Wiley, 2011.

[28] M. Ding and L. Hanwen, «Multi-point Cooperative Communication Systems: Theory and Applications,» Berlin, Germany, Springer-Verlag, 2013.

[29] Z. Shen, A. Khoryaev, E. Eriksson, and X. Pan, «Dynamic uplink-downlink configuration and interference management in

TD-LTE,",» *IEEE Communications Magazine,* vol. 50, n. 11, p. 51–59, Nov. 2012.

[30] M. Ding, D. López-Pérez, R. Xue, W. Chen, and A. Vasilakos,, «Small cell dynamic TDD transmissions in heterogeneous networks,» in *2014 IEEE International Conference on Communications (ICC)*, Sidney, NSW, Australia.

[31] «Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update,» White paper, Cisco, San Jose, CA, USA,, 2015.

[32] A. Damnjanovic et al., «A survey on 3GPP heterogeneous networks,» *EEE Wireless Communications,* vol. 18, n. 3, pp. 10-21, Jun. 2011.

[33] D. Lopez-Perez, I. Guvenc, G. de la Roche, M. Kountouris, T. Q. S. Quek and J. Zhang, «Enhanced intercell interference coordination challenges in heterogeneous networks,» *IEEE Wireless Communications,* vol. 18, n. 3, pp. 22-30, Jun. 2011.

[34] H. Claussen, L. T. W. Ho, and L. G. Samuel, «An overview of the femtocell concept,» *Bell Labs Tech. Journal,* vol. 13, n. 1, p. 221–245, May 2008.

[35] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, «Femtocell networks: A survey,» *IEEE Communications Magazine,* vol. 46, n. 9, p. 59–67, Sep. 2008.

[36] D. López-Pérez, A. Valcarce, G. de la Roche, and J. Zhang, «OFDMA femtocells: A roadmap on interference avoidance,» *IEEE Communications Magazine,* vol. 47, n. 9, p. 41–48, Sep. 2009.

[37] D. López-Pérez, X. Chu, and I. Guvenc, «On the Expanded Region of Picocells in Heterogeneous Networks,» *IEEE Journal of Selected Topics in Signal Processing,* vol. 6, n. 3, pp. 281-294, Jun. 2012.

[38] C. E. Shannon, «A mathematical theory of communication,» *The Bell System Technical Journal,* vol. 27, n. 4, p. The Bell System Technical Journal, Oct. 1948.

[39] E. MCC, «Draft report of 3GPP TSG RAN WG1 #75,» Sophia Antipolis, France, 2013.

[40] 3GPP, «TS 36.300 V10.7.0, E-UTRA and E-UTRAN Overall Description,» 2012.

[41] S. Sesia, I. Toufik, M. Baker, «LTE - The UMTS Long Term Evolution: From Theory to Practice,» John Wiley & Sons, 2009.

[42] J. Moon, D. Cho, «Efficient handoff algorithm for inbound mobility in hierarchical macro/femto cell networks,» *IEEE Communications Letters,* vol. 13, n. 10, pp. 755-757, Oct. 2009.

[43] D. Xenakis, N. Passas, «An energy-centric handover decision algorithm for the integrated LTE macrocell–femtocell network,» *Computer Communications,* vol. 35, n. 14, pp. 1684-1694, 2012.

[44] J. M. Moon and D. H. Cho, «Novel Handoff Decision Algorithm in Hierarchical Macro/Femto-Cell Networks,» in *2010 IEEE Wireless Communication and Networking Conference*, Sidney, NSW, Australia, 2010.

[45] D. Lopez-Perez, A. Valcarce, A. Ladanyi, G. de la Roche, J. Zhang, «Intracell Handover for Interference and Handover Mitigation in OFDMA Two-Tier Macrocell-Femtocell Networks,» *EURASIP Journal on Wireless Communications and Networking,* vol. 10, n. 1, Jan. 2010.

[46] Z. Becvar and P. Mach, «Adaptive Hysteresis Margin for Handover in Femtocell Networks,» in *2010 6th International Conference on Wireless and Mobile Communications*, Valencia, 2010.

[47] H. Zhang, W. Ma, W. Li, W. Zheng, X. Wen and C. Jiang, «Signalling Cost Evaluation of Handover Management Schemes in LTE-Advanced Femtocell,» in *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, Yokohama, 2011.

[48] H. Zhang, X. Wen, B. Wang, W. Zheng and Y. Sun, «A Novel Handover Mechanism Between Femtocell and Macrocell for LTE Based Networks,» in *2010 Second International Conference on Communication Software and Networks*, Singapore, 2010.

[49] A. Ulvan, R. Bestak, M. Ulvan, «Handover Scenario and Procedure in LTE-based Femtocell Networks,» in *UBICOMM 2010 : The Fourth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, 2010.

[50] S. Wu, S. Lo,, «Handover Scheme in LTE-based Networks with Hybrid Access Mode Femtocells,» *Journal of Convergence Information Technology,* vol. 6, n. 7, pp. 68-78, 2011.

[51] D. Xenakis, N. Passas, L. Merakos and C. Verikoukis, «Mobility Management for Femtocells in LTE-Advanced: Key Aspects and Survey of Handover Decision Algorithms,» *IEEE Communications Surveys & Tutorials,* vol. 16, n. 1, pp. 64-91, First Quarter 2014.

[52] D. Xenakis, N. Passas and C. Verikoukis, «A novel handover decision policy for reducing power transmissions in the two-tier LTE network,» in *2012 IEEE International Conference on Communications (ICC)*, Ottawa, ON, 2012.

[53] G. Yang, X. Wang and X. Chen, «Handover control for LTE femtocell networks,» in *2011 International Conference on Electronics, Communications and Control (ICECC)*, Zhejiang, 2011.

[54] K. S. B. Reguiga, F. Mhiri, R. Bouallegue, «Handoff Management in Green FEMTOCELL Network,» *International Journal of Computer Applications,* vol. 27, n. 4, Aug. 2011.

[55] D. Lee, G. Gil, D. Kim, «A Cost-Based Adaptive Handover Hysteresis Scheme to Minimize the Handover Failure Rate in 3GPP LTE System,» *EURASIP Journal on Wireless Communications and Networking,* Dec. 2010.

[56] P. Xu, X. Fang, J. Yang and Y. Cui, «P. Xu, X. Fang, J. Yang and Y. Cui,» in *2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)*, Chengdu, 2010.

[57] B. Jeong, S. Shin, I. Jang, N. W. Sung, H. Yoon, «A Smart Handover Decision Algorithm Using Location Prediction for Hierarchical Macro/Femto-Cell Networks,» in *2011 IEEE Vehicular Technology Conference (VTC Fall)*, San Francisco, CA, 2011.

[58] R. Trestian, O. Ormond, G.-M. Muntean, «Energy–Quality–Cost Tradeoff in a Multimedia-Based Heterogeneous Wireless Network Environment,» *IEEE Transaction on Broadcasting,* vol. 59, n. 2, pp. 340-357, Jun. 2013.

[59] T. Bi, Z. Yuan, R. Trestian and G. M. Muntean, «URAN: Utility-based reputation-oriented access network selection strategy for

HetNets,» in *2015 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, Ghent, 2015.

[60] G. Araniti, J. Cosmas, A. Iera, A. Molinaro, A. Orsino and P. Scopelliti, «Energy efficient handover algorithm for green radio networks,» in *2014 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, Beijing, 2014.

[61] K. Mahmud, M. Inoue, H. Murakami, M. Hasegawa and H. Morikawa, «Measurement and usage of power consumption parameters of wireless interfaces in energy-aware multi-service mobile terminals,» in *2004 IEEE 15th International Symposium on Personal, Indoor and Mobile Radio Communications*, 2004.

[62] R. Trestian, O. Ormond, and G.-M. Muntean, «Performance evaluation of MADM-based methods for network selection in a multimedia wireless environment,» *Wireless Networks,* vol. 21, n. 5, pp. 1745-1763, Jul. 2015.

[63] P. Scopelliti, G. Araniti, G. M. Muntean and A. Iera, «Mobility-aware energy-quality trade-off for video delivery in dense heterogeneous networks,» in *2016 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Nara, 2016.

[64] R. Trestian, O. Ormond and G. M. Muntean, «Enhanced Power-Friendly Access Network Selection Strategy for Multimedia Delivery Over Heterogeneous Wireless Networks,» *IEEE Transactions on Broadcasting,* vol. 60, n. 1, pp. 85-101, Mar. 2014.

[65] A. Asadi, Q. Wang and V. Mancuso, «A Survey on Device-to-Device Communication in Cellular Networks,» *IEEE Communications Surveys & Tutorials,* vol. 16, n. 4, pp. 801-1819, Fourth Quarter 2014.

[66] L. Wang, H. Tang, H. Wu and G. L. Stber, «Resource Allocation for D2D Communications Underlay in Rayleigh Fading Channels,» *IEEE Transactions on Vehicular Technology,* vol. 66, n. 2, pp. 1159-1170, Feb. 2017.

[67] 3GPP, «TS 36.440, "General aspects and principles for interfaces supporting Multimedia Broadcast Multicast Service (MBMS) within EUTRAN,» 2012.

[68] J. B. Rao and A. O. Fapojuwo, «A Survey of Energy Efficient Resource Management Techniques for Multicell Cellular Networks,» *IEEE Communications Surveys & Tutorials,* vol. 16, n. 1, pp. 154-180, First Quarter 2014.

[69] O. Ormond, G.-M. Muntean, J. Murphy, «Economic Model for Cost Effective Network Selection Strategy in Service Oriented Heterogeneous Wireless Network Environment,» in *IEEE/IFIP Network Operations and Management Symposium*, Vancouver, Canada, April 2006.

[70] O. Ormond, G.-M. Muntean, J. Murphy, «Network Selection Strategy in Heterogeneous Wireless Networks,» in *IT&T Conference*, Cork, Ireland, 2005.

[71] E. B. Rodrigues and F. Casadevall, «Control of the trade-off between resource efficiency and user fairness in wireless networks using utility-based adaptive resource allocation,» *IEEE Communications Magazine,* vol. 49, n. 9, pp. 90-98,, 2011.

[72] M. Anedda, G.-M. Muntean and M. Murroni, «Adaptive real-time multi-user access network selection algorithm for load-balancing over heterogeneous wireless networks,» in *2016 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Nara, 2016.

[73] C. Desogus, M. Anedda and M. Murroni, «Real-time load optimization for multimedia delivery content over heterogeneuos wireless network using a MEW approach,» in *2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Cagliari, 2017.

[74] A. Raschellà, F. Bouhafs, M. Seyedebrahimi, M. Mackay and Q. Shi, «A centralized framework for smart access point selection based on the Fittingness Factor,» in *2016 23rd International Conference on Telecommunications (ICT)*, Thessaloniki, 2016.

[75] A. Orsino, G. Araniti, A. Molinaro and A. Iera, «Effective RAT Selection Approach for 5G Dense Wireless Networks,» in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, Glasgow, 2015.

[76] 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, n. 1, pp. 240-254, First Quarter 2013.

[77] T. P. Low, M. O. Pun, Y. W. P. Hong and C. C. J. Kuo, «Optimized opportunistic multicast scheduling (OMS) over wireless cellular networks,» *IEEE Transactions on Wireless Communications,* vol. 9, n. 2, pp. 791-801, Feb. 2010.

[78] G. Araniti, M. Condoluci, L. Militano and A. Iera, «Adaptive Resource Allocation to Multicast Services in LTE Systems,» *IEEE Transactions on Broadcasting,* vol. 59, n. 4, pp. 658-664, Dec. 2013.

[79] C. Borgiattino, C. Casetti, C. F. Chiasserini and F. Malandrino, «Efficient area formation for LTE broadcasting,» in *2015 12th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, Seattle, WA, 2015.

[80] S. Pizzi, M. Condoluci, G. Araniti, A. Molinaro, A. Iera and G. M. Muntean, «A Unified Approach for Efficient Delivery of Unicast and Multicast Wireless Video Services,» *IEEE Transactions on Wireless Communications,* vol. 15, n. 12, pp. 8063-8076, Dec. 2016.

[81] R. Trestian, A. N. Moldovan, C. H. Muntean, O. Ormond and G. M. Muntean, «Quality Utility modelling for multimedia applications for Android Mobile devices,» in *IEEE international Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Seoul, 2012.

[82] L. Zou, A. Javed and G. M. Muntean, «Smart mobile device power consumption measurement for video streaming in wireless environments: WiFi vs. LTE,» in *2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Cagliari, 2017.

[83] 3GPP, «TR 36.210, LTE physical layer; General description, Rel. 11,» 2012.

[84] 3GPP, «TS 36.331, LTE; Evolved Universal Terrestrial Radio Access (EUTRA); Radio Resource Control (RRC); Protocol specification,» 2009.

[85] 3GPP, «TR 25.905, Technical Specification Group Radio Access Network; Improvement of the Multimedia Broadcast Multicast Service (MBMS) in UTRAN,» 2007.

[86] Qualcomm, «Lte broadcast - evolving and going beyond mobile,» [Online]. Available: https://www.slideshare.net/qualcommwirelessevolution.

[87] M. Moiz Anis, X. Lagrange, R. Pyndiah,, «Overview of evolved Multimedia Broadcast Multicast Services (eMBMS),» Institut Mines Telecom - Telecom Bretagne, 2016.

[88] T. Lohmar, M. Slssingar, S. Puustinen, V. Kenehan,, «Delivering content with LTE broadcast,» Ericsson Review, Feb. 2013.

[89] J. F. Monserrat, J. Calabuig, A. Fernandez-Aguilella and D. Gomez-Barquero, «Joint Delivery of Unicast and E-MBMS Services in LTE Networks,» *IEEE Transactions on Broadcasting,* vol. 58, n. 2, pp. 157-167, Jun. 2012.

[90] A. De La Fuente Iglesias, R. P. Leal, and A. G. Armada,, «Performance Analysis of eMBMS in LTE: Dynamic MBSFN Areas,» in *OPNETWORK*, 2013.

[91] F. Malandrino, C. Casetti, C. F. Chiasserini and S. Zhou, «Real-Time Scheduling for Content Broadcasting in LTE,» in *2014 IEEE 22nd International Symposium on Modelling, Analysis & Simulation of Computer and Telecommunication Systems*, Paris, 2014.

[92] S. Talarico and M.C. Valenti, «An accurate and efficient analysis of a MBSFN Network,» in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, 2014.

[93] 3GPP, «Radio Access Networks - LTE progress report,» 3GPP Seminar in Moscow, 2010.

[94] 3GPP, «SC-PTM, an agile broadcast/multicast mechanism,» 2016.

[95] 3GPP, «TS 38.913, "5G; Study on Scenarios and Requirements for NExt Generation Access Technologies", Rel. 14.,» 2017.

[96] S.Y. Baek, Y.J. Hong and D.K. Sung, «Adaptive transmission scheme for mixed multicast and unicast traffic in cellular systems,» *IEEE Transaction on Vehicular Technology,* vol. 58, n. 6, pp. 2899-2908, Jul. 2009.

[97] M. Konrad, W. Koch and J. Huschke, «Throughput based switching threshold for multicast and unicast in cellular OFDM

networks,» in *2007 IEEE 66th Vehicular Technology Conference*, Baltimore, MD, 2007.

[98] C. Singhal, C. F. Chiasserini and C. CasettI, «Efficient Multimedia Broadcast for heterogeneous users in cellular networkS,» in *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)*, Paphos, 2016.

[99] H. Schwarz, D. Marpe and T. Wiegand, «Overview of the Scalable Video Coding Extension of the H.264/AVC Standard,» *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 17, n. 9, pp. 1103-1120, Sep. 2007.

[100] 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, n. 2, pp. 149-152, Apr. 2015.

[101] M. Condoluci, G. Araniti, A. Molinaro and A. Iera, «Multicast Resource Allocation Enhanced by Channel State Feedbacks for Multiple Scalable Video Coding Streams in LTE Networks,» *IEEE Transactions on Vehicular Technology,* vol. 65, n. 5, pp. 2907-2921, May 2016.

[102] 3GPP, «TS 36.213, Evolved Universal Terrestrial Radio Access (EUTRA) and Evolved Universal Terrestrial Radio Access Network (EUTRAN),» 2012.

[103] 3GPP, «TS 36.942, LTE; Evolved Universal Terrestrial Radio Access (E-UTRA);Radio Frequency (RF) system scenarios,» 2010.

[104] J. Calabuig, J. Monserrat, D. Gómez-Barquero, «5th generation of mobile networks: a new opportunity for the convergence of mobile broadband and broadcast services,» *IEEE Communications Magazine,* vol. 53, n. 2, pp. 198-205, 2015.

[105] M. Fadda, M. Murroni, and V. Popescu, «Interference Issues for VANET Communications in the TVWS in Urban Environments,» *IEEE Transactions on Vehicular Technology,* vol. 65, n. 7, pp. 4952-4958, Jul. 2016.

[106] J. Montalban, L. Zhang, U. Gil, Y. Wu, I. Angulo, K. Salehian, S.I. Park, B. Rong, W. Li, H.M. Kim, P. Angueira, and M. Velez, «Cloud Transmission: System Performance and Application

Scenarios,» *IEEE Transactions on Broadcasting,* vol. 60, n. 2, pp. 170-184, Jun. 2014.

[107] L. Zhang et al., «Layered-Division-Multiplexing: Theory and Practice,» *IEEE Transactions on Broadcasting,* vol. 62, n. 1, pp. 216-232, Mar. 2016.

[108] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li and K. Higuchi, «Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access,» in *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, Dresden, 2013.

[109] S. I. Park et al., «Low Complexity Layered Division Multiplexing for ATSC 3.0,» *IEEE Transactions on Broadcasting,,* vol. 62, n. 1, pp. 233-243, Mar. 2016.

[110] Y. Wu, E. Pliszka, B. Caron, P. Bouchard, and G. Chouinard, «Comparison of terrestrial DTV transmission systems: TheATSC 8-VSB, the DVB-T COFDM and the ISDB-T BST-OFDM,» *IEEE Transaction on Broadcasting,* vol. 46, n. 2, p. 101–113, Jun. 2000.

[111] B. Rong, S. I. Park, Y. Wu, H. M. Kim, J. Kim, G. Gagnon, and X. Wang,, «Signal cancellation techniques for RF watermark detection in ATSC mobile DTV system,» *IEEE Transaction on Vehicular Technology,* vol. 60, n. 8, p. 4070–4076, Oct. 2011.

[112] S. I. Park, H. Lim, H. M. Kim, Y. Wu, and W. Oh,, «Augmented data transmission for the ATSC terrestrial DTV system,» *IEEE Transaction on Broadcasting,* vol. 58, n. 2, p. 261–268, Jun. 2012.

[113] B. Liu, Y. Li, B. Rong; L. Gui, and Y. Wu,, «LDPC-RS product codes for digital terrestrial broadcasting transmission system,» *IEEE Transcation on Broadcasting,* vol. 69, n. 1, Mar. 2014.

[114] S. I. Park, H. M. Kim, Y. Wu, and J. Kim, «A newly designed quarterrate QC-LDPC code for the cloud transmission system,» *IEEE Transaction on Broadcasting,* vol. 59, n. 1, p. 155–159, Mar. 2013.

[115] J. Montalban, B. Rong, Y. Wu, L. Zhang, P. Angueira and M. Velez, «Cloud Transmission frequency domain cancellation,» in *2013 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, London, 2013.

[116] I. Fernandez et al., «Field trials based validation of the suitable configuration parameters for mobile urban reception, using the

new generation broadcasting systems,» in *2014 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Beijing, 2014.

[117] M. R. Islam, N. Avazov, O. A. Dobre and K. s. Kwak, «Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges,» *IEEE Communications Surveys & Tutorials,* vol. 19, n. 2, pp. 721-742, Second Quarter 2017.

[118] 3GPP, «Technical Specification Group Radio Access Network; Study on New Radio (NR) Access Technology; Physical Layer Aspects v(Release 14).».

List of Publications

[1] A. Orsino, **P. Scopelliti** and M. Condoluci, “A Multi-Criteria Approach for Multicast Resource Allocation over LTE and Beyond Cellular Systems*,” European Wireless 2016*, Oulu, Finland, Jun. 2016.

[2] **P. Scopelliti**, G. Araniti, G.-M. Muntean and A. Iera, “Mobility-aware Energy-Quality Trade-off for Video Delivery in Dense Heterogeneous Networks,” *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB) 2016*, Nara, Japan, Jun. 2016.

[3] G. Araniti, M. Condoluci, **P. Scopelliti**, A. Molinaro and A. Iera, "Multicasting over Emerging 5G Networks: Challenges and Perspectives," in *IEEE Network*, vol. 31, no. 2, pp. 80-89, March/April 2017.

[4] **P. Scopelliti**, G. Araniti, G. M. Muntean, A. Molinaro and A. Iera, "A hybrid unicast-multicast utility-based network selection algorithm," *2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Cagliari, 2017, pp. 1-6.

[5] A. Orsino, G. Araniti, **P. Scopelliti**, I. A. Gudkova, K. E. Samouylov and A. Iera, "Optimal subgroup configuration for multicast services over 5G-satellite systems," *2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Cagliari, 2017, pp. 1-6.

[6] V. Beschastnyi, V. Naumov, **P. Scopelliti**, I. Gudkova, C. Campolo, G. Araniti, I. Dzantiev, and K. Samouylov, “Discrete Time Bulk Service Queue for Analyzing LTE Packet Scheduling for V2X Communications”, *Internet of Things, Smart Spaces, and Next Generation Networks and Systems: 17th International Conference, NEW2AN 2017, 10th Conference, ruSMART 2017, Third Workshop NsCC 2017*, St. Petersburg, Russia, August 28--30, 2017, Proceedings, Springer International Publishing, 2017, 395-407.

[7] J. Montalban, **P. Scopelliti**, M. Fadda, E. Iradier, C. Desogus, P. Angueira, M. Murroni and G. Araniti, "Multimedia Multicast Services in 5G Networks: Subgrouping and Non-Orthogonal Multiple Access Techniques", *IEEE Communication Magazine,* vol. 56, no. 3, pp. 91-95, Mar 2018.

[8] G. Araniti, **P. Scopelliti**, G.-M. Muntean and A. Iera, "A Joint Unicast-Multicast Approach for Supporting Network Selection in Dense Heterogeneous Networks", *IEEE Transaction on Broadcasting* (in Press).