Computer Communications 97 (2017) 96–110

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Mid-Term frequency domain scheduler for resource allocation in wireless mobile communications systems

Álvaro Pachón[a,*], Ubaldo M. García-Palomares[b,1]

[a] ICT Department, Universidad Icesi, Calle 18 # 122-135, Cali-Colombia
[b] Telematics Engineering Department, Vigo University, Lagoas-Marcosende 36310, Vigo-Spain

ARTICLE INFO

Article history:
Received 12 January 2016
Revised 6 August 2016
Accepted 18 August 2016
Available online 23 August 2016

Keywords:
Optimization
Resource allocation
LTE
OFDMA

ABSTRACT

This article approaches the dynamic resource allocation problem for the downlink of a wireless mobile communication system (WMCS). The article defines the architecture and functions of the global resource scheduler, as well as the quality index for scheduling, the signal to interference-plus-noise ratio (SINR). The proposed approach divides the scheduling task into two components: a distributed one, with local, short-term scope; and a centralized one, with global, medium-term scope. The optimization model considers a set of slack variables for guaranteeing feasibility. This allows the service provider to fully satisfy the users' service demands. The model type is mixed, non-linear, which demands large computational power for an exact solution. So an approximate strategy is used, in order to decouple the search space. The time limit imposed to reach a solution forces to define a reduced neighborhood structure. Thus, the obtained results are the best solution obtained in the allotted time interval, evaluating a suitable set of neighbors, and using an objective, effective criterion for searching. The solution offers high levels of full service satisfaction (greater than 97%), low levels of service denial (less than 2%), and efficient power usage (30% in average).

1. Introduction

A WMCS consists of a set of cells covering a geographical area, with defined topographical features that determine the signal propagation environment. Each cell has a base station (BS), in charge of managing resources (power and frequency) allowing offering services to users, that interact with the system using terminal equipment (TE). Communication channels are characterized by the signal to interference-plus-noise ratio (SINR), a quality index allowing to establish service satisfaction levels (relevant to users) and resource utilization levels (relevant to service providers). A WMCS incorporates several technologies, LTE (Long Term Evolution) among them. LTE is a standard proposed by 3GPP [1] to develop and consolidate 3 G/4 G networks. The most salient feature of LTE, with respect to this article, is OFDMA (Orthogonal Frequency Division Multiple Access), the multiple access technique for the radio interface (physical level), which has become a *de facto* standard for the current generation of WMCS.

The article structure is as follows: First section introduces the problem and its context; second section formulates the problem; third section describes the resource scheduling task, its background, and proposed solution strategies. Fourth section formulates the optimization model; fifth section presents the model's solution; sixth section presents the resource scheduler design in its structural, operative, spatial and temporal dimensions. Seventh section presents the proposed scenarios and experiments to evaluate the goodness of the model; finally, eighth section presents conclusions and future work.

2. Problem formulation

The problem solved by this article is the dynamic resource allocation to a user population demanding a set of services, on a WMCS using OFDMA at its physical level. The context for solution is the downlink of a multi-user, multi-cell, multi-service WMCS, with either a homogeneous or heterogeneous architecture. This task is executed by the base station resource scheduler, which accounts for channel variations in the time and frequency domains, and for changes in user number, position and service requests. In order to optimally allocate resources, the scheduler faces a key constraint: a limited time interval for task execution, as a

* Corresponding author: +Fax: +57 25552345.
E-mail addresses: alvaro@icesi.edu.co (Á. Pachón), ubaldo@det.uvigo.es (U.M. García-Palomares).
1 fax: +34 986812166

http://dx.doi.org/10.1016/j.comcom.2016.08.007

consequence of the system's dynamics. The scheduler must identify all required resources for offering the different services and harmonize their coexistence in such time interval.

3. Resource scheduling task

Resource scheduling is a complex task, facing multiple hurdles: efficient resource usage, time limit (see previous section), and satisfying users with different service requirements [2]. The BS scheduler must consider the services, characterized by the quality of service class identifier (QCI); and the channel condition, characterized by the channel quality indicator (CQI). These indicators must be reported by every TE to the BS. This process has several challenges: overhead (both in processing and communications), accuracy and timeliness of reports, and the length of the time interval for guaranteeing an accurate report of the system status [3,4]. There is another difficulty, arising from frequency reuse: inter-cell interference, the most important of OFDMA's shortcomings [5]. Power allocation to a certain frequency block in one cell becomes interference to the neighboring cells. In order to minimize its impact, inter-cell interference coordination (ICIC) strategies [6], including constraints for resource usage, are employed.

3.1. Problem formulations

The treatment of this problem has evolved over time. After reviewing related work from over the last fifteen years, authors classified their evolution into five phases: Orientation to performance, quality of service, interference management, heterogeneous networks (HetNets) and energy efficiency (EE) management (details of some of these phases can be found in [7]).

Oriented-performance proposals are centered in the WMCS and the satisfaction of some system performance indicator (maximize data rate or minimize power use), ignoring users and fair resource distribution. Later, researchers recognize the importance of giving service to all users, without regard to their condition. In order to achieve this, the models incorporate a set of constraints guaranteeing proportional fairness at the service levels, hence recognizing the existing compromise between fairness and performance. Then, WMCS faced the challenge of providing and satisfying several services to their users. In order to do so, researchers resort to Economical Sciences theory, and formulate the problem using the utility concept [8]. The optimization problem aims to maximize profit for best-effort traffic users, subject to full satisfaction of guaranteed user traffic requirements. Several utility functions have been proposed [9,10]. Popularization of offered services caused an exponential increase in the number of users. Thus, the operator was forced to intensively reuse available spectrum, taking interference to critical levels. Interference management became necessary and SINR was proposed as a quality metric for such purpose. Cicalo et al. [11] suggest that adequate interference management is a key factor in the Long-Term Evolution (LTE) context. Kim et al. [12] suggest using an intelligent strategy for resource allocation, which reduces interference in order to improve SINR and increase the data rate.

The need of covering large geographical areas prompted deployment of multiple cells, to increase the system coverage area. Distributed solutions for resource allocation, featuring cooperation and coordination mechanisms, are proposed for this scenario. Fehske et al [13] consider resource allocation under a distributed, cooperative approach, by using the historical rate in place of the system's global vision; Bolla et al [14] simultaneously allocate modulation type, coding rate and resources (power and frequency) by using a coordination-based self-organized approach that aims to keep a stable frequency reuse pattern; Mokari et al. [15] formulate a proposal for dynamic spectrum sharing by using cognitive radio. In this proposal, the secondary infers the primary's behavior and nature of its environment. This knowledge keeps the secondary from generating excessive interference on the primary's transmissions. Bai et al. [16] proposal allocates resources in an environment with different quality of service demands (streams with different sizes and latencies); Keerthana & Vinoth [17] propose to allocate resources in two phases: in the first one, they use the geographical location of users to build a graph to mitigate interference, and in the second one, they allocate frequency resources by using said graph. With respect to formulating the optimization model: Feshke et al. [13], propose optimizing spectral efficiency by using a profit function; Keertana & Vinoth [17], Rajamannar & Vijaya [18], and Karthik & Kumaran [19], formulate a model to maximize throughput for users at the edge of the cell, subject to throughput fulfillment of users close to the base station. This formulation punishes users at the edge of the cell, because they are not offered full service guarantees, thus its fairness is questionable. Swapna et al. [20] formulate a proposal for resource allocation if OFDMA using cooperation and aiming for energy efficiency; Abdelhadi & Clancy [21] define a context, time and location-aware architecture for resource allocation in next-generation WMCS.

Increase in the number of users and demand for rich multimedia services impact the system architecture, motivating the inclusion of small cells (cells with less capacity and coverage area). WMCS are evolving towards a heterogeneous architecture, which features cells with different coverage and capacity, yielding a multi-level hierarchical structure (HetNet). Inclusion of small cells allows to: Improve the spectral efficiency (SE) by taking advantage of the spatial diversity; increase the offered data rates and reduce the amount of radiated power [22]. However, this approach also poses challenges, i.e. interference increase [23] and the backhaul's capacity for allowing exchange of coordination information for resource allocation [24]. López-Pérez et al [25] formulate a proposal for resource allocation, based upon minimizing power in OFDMA networks, by arbitrarily deploying femtocells in homes or businesses. Femtocells constitute a distributed system, and make decisions in an independent, self-organized fashion; Fan et al. [23], analyze the strategies for resource allocation in HetNets, and propose an algorithm for resource allocation in a cluster featuring a macrocell (in the center) and a set of small cells (randomly distributed in the macrocell's coverage area). The proposed model maximizes the transmission rate, subject to power constraints, and setting priorities for accessing and using frequency resources, favoring the macrocell.

Increase of computing power requirements in next-generation WMCS also increases the energy consumption of its components. CO_2 emissions lead to consider energy efficiency as an important design parameter for WMCS [26]. Initially, EE considered only the power used for transmission by the base station; however Li et al. [27] and Zappone et al. [28], among others, extend the concept, and present a more general expression including power consumed by the components' circuits. With respect to proposals for better EE in the WMCS, Miao et al. [29] maximize it through adaptive transmission, considering the channel state, and aiming to balance power used for transmission with power used by the components' circuits; Abdulkafi et al. [24] analyze energy-aware proposals in the HetNet context, establishing compromise relationships between architecture, base station design, and quantity/location of deployed sites in the macrocell-microcell scenario, which is the object of study of the present article; Devarajan et al. [30] and Gurupandi & Vadivel [31], formulate a proposal for guaranteeing EE in the cluster formed by the base station and the set of associated relays in a HetNet; Yu et al. [26] use a different strategy, by minimizing the energy consumption of the user terminal, subject to meeting the per-user rate requirements and the availability of power; Ren et al. [32] include the equity criterion and consider EE in their