Optimization Models for the Joint Power-Delay Minimization Problem in Green Wireless Access Networks

Farah Moety, Samer Lahoud, Bernard Cousin, Kinda Khawam

Optimization Models for the Joint Power-Delay Minimization Problem in Green Wireless Access Networks

Farah Moety[a], Samer Lahoud[a], Bernard Cousin[a], Kinda Khawam[b]

[a] *IRISA labs, University of Rennes I, Rennes, France*
[b] *PRISM Labs, University of Versailles Saint-Quentin-en-Yvelines, France*

Abstract

In wireless access networks, one of the most recent challenges is reducing the power consumption of the network, while preserving the quality of service perceived by users. Hence, mobile operators are rethinking their network design by considering two objectives, namely, saving power and guaranteeing a satisfactory quality of service. Since these objectives are conflicting, a tradeoff becomes inevitable. We formulate a multi-objective optimization with aims of minimizing the network power consumption and transmission delay. Power saving is achieved by adjusting the operation mode of the network base stations from high transmit power levels to low transmit levels or even sleep mode. Minimizing the transmission delay is achieved by selecting the best user association with the network base stations. In this article, we cover two different technologies: IEEE 802.11 and LTE. Our formulation captures the specificity of each technology in terms of the power model and radio resource allocation. After exploring typical multi-objective approaches, we resort to a weighted sum mixed integer linear program. This enables us to efficiently tune the impact of the power and delay objectives.

We provide extensive simulations for various preference settings that enable to assess the tradeoff between power and delay in IEEE 802.11 WLANs and LTE networks. We show that for a power minimization setting, a WLAN consumes up to 16% less power than legacy solutions. A thorough analysis of the optimization results reveals the impact of the network topology, particularly the inter-cell distance, on both objectives. For an LTE network, we assess the impact of urban, rural and realistic deployments on the achievable tradeoffs. The power savings mainly depend on user distribution and the power consumption of the sleep mode. Compared with legacy solutions, we obtained power savings of up to 22.3% in a realistic LTE networks. When adequately tuned, our optimization approach reduces the transmission delay by up to 6% in a WLAN and 8% in an LTE network.

Keywords: Wireless access networks, Optimization, Power Consumption, Transmission Delay, User Association.

1. Introduction

In recent years, green radio has been increasingly emphasized for not only ecological concerns but also for significant economic incentives. Information and Communication Technology (ICT) accounts for around 3% of the world's annual electrical energy consumption and 2% of total carbon emissions. Moreover, it is estimated that ICT energy consumption is rising at 15 to 20 %, doubling every five years [1]. In 2008 this corresponded to about 60 billion kWh of electrical energy consumption and about 40 million metric tons of CO_2 [2]. As a branch of the ICT sector, mobile networks are responsible for 0.2% of these emissions [3]. In addition to the environmental impacts, mobile operators are interested in reducing energy consumption for economic reasons, especially with increasing energy costs becoming a significant portion of mobile operator expenditure. Moreover, the recent explosive growth of the number of mobile devices and the consequent mobile internet traffic all produce continually high energy consumption. This calls for green solutions to address the challenges in energy-efficient communications. Operators have focused on

Email addresses: farah.moety@irisa.fr (Farah Moety), samer.lahoud@irisa.fr (Samer Lahoud), bernard.cousin@irisa.fr (Bernard Cousin), kinda.khawam@prism.uvsq.fr (Kinda Khawam)

technological developments in the past years to meet capacity and Quality of Service (QoS) demands for User Equipment (UE). Pushed by the needs to reduce energy, mobile operators have recently been rethinking network design for optimizing energy efficiency and satisfying user QoS requirements.

Currently, over 80% of the power in mobile telecommunications is consumed by the radio access network, more specifically at the base station (BS) level [4]. Hence, many research activities focus on improving the energy efficiency of wireless access networks. In the following, we give an overview of these activities and classify them according to different approaches that run at different timescales.

Planning and deployment: The planning of energy-efficient wireless networks and the deployment of energy-aware BSs deal with the problem of determining the positioning of BSs, the type (*e.g.*, macro, micro, pico or femto) and the number of BSs needed to be deployed. In this context, we find that heterogeneous networks have gained great attention in current research. Precisely, deploying small and low-power BSs co-localized with macro cells is believed to decrease power consumption compared to high-power macro BSs. Moreover, it extends the coverage area of the macro BS where signals fail to reach distant UEs. Furthermore, small cells increase the network capacity in areas with very dense data usage. Planning and deployment tasks are performed at very coarse temporal levels, ranging from a few months to possibly years.

Cell sizing: Also known as cell breathing, cell sizing is a well-known concept that enables balancing traffic load in cellular telephony [5, 6]. When the cell becomes heavily loaded, the cell zooms in to reduce its coverage area, and the lightly loaded neighboring cells zoom out to accommodate the extra traffic. Many state-of-the-art techniques are used to implement cell sizing, such as adjusting the transmit power of a BS, cooperating between multiple BSs, and using relay stations and switching BSs for sleep/off mode. Cell sizing is performed at medium temporal levels, ranging from hours to days.

User Association: User association is the functionality devoted to deciding the BS (macro, micro, pico or femto) with which a given user will be associated in a heterogeneous network [7, 8]. The challenge is to optimize for example the delay, throughput, or network cost. User association is impacted by the cell sizing tasks: when an active BS is switched off or changes its transmit power level in a homogeneous network, users may need to change their associations. Many metrics are considered for selecting the serving BS, such as the received signal quality (Signal-to-Noise Ratio (SNR) or the corresponding achievable rate), the traffic load, or the distance between BS and UE, etc. User association is performed at small temporal levels, ranging from seconds to minutes.

Scheduling: Scheduling algorithms allocate the radio resources in wireless access networks, where the objectives consist of improving the network throughput, satisfying the delay constraints of real-time traffic, or achieving fair resource distribution among users. Scheduling is performed at very short temporal levels of an order of few milliseconds. In Figure 1, we illustrate the different green approaches studied in the state-of-the art as well as their operating timescales.

Reducing power consumption in wireless networks is coupled with satisfying the QoS requirements (delay, blocking probability, etc.). As these objectives are conflicting, a tradeoff becomes ineluctable. Chen *et al.* [9] identified four key tradeoffs of energy efficiency with network performance: *i)* deployment-energy efficiency to balance the deployment cost, throughput, and energy consumption in the network as a whole; *ii)* spectrum-energy efficiency to balance the achievable rate and energy consumption of the network; *iii)* bandwidth-power to balance the utilized bandwidth and the power needed for transmission; and *iv)* delay-power to balance the average end-to-end service delay and average power consumed in transmission.

The delay-power tradeoff has not been studied deeply in the literature except for in a few recent cases [10]. In this article, we address the multi-objective optimization problem of power saving and transmission delay minimization in wireless access networks. Specifically, power saving is achieved by adjusting the operation mode of the network BSs from high transmit power levels to low transmit power levels, or even sleep mode. In this context, changing the operation mode of the BSs is coupled with optimized user association. Such coupling makes solving the problem more challenging. Minimizing the transmission delay is achieved by selecting the best user association with the network BSs.

State-of-the-art power saving approaches studied can also be performed in Wireless Local Area Networks (WLANs). Although power consumption of a cellular network BS is much higher than that of a WLAN Access Point (AP), the large number of APs deployed in classrooms, offices, airports, hotels and malls, contributes to a rapid increase in the power consumption in wireless access networks. Hence, efficient management of the power consumed by a WLAN is an interesting challenge. In the present article, we cover two different wireless networks: WLANs with IEEE 802.11g

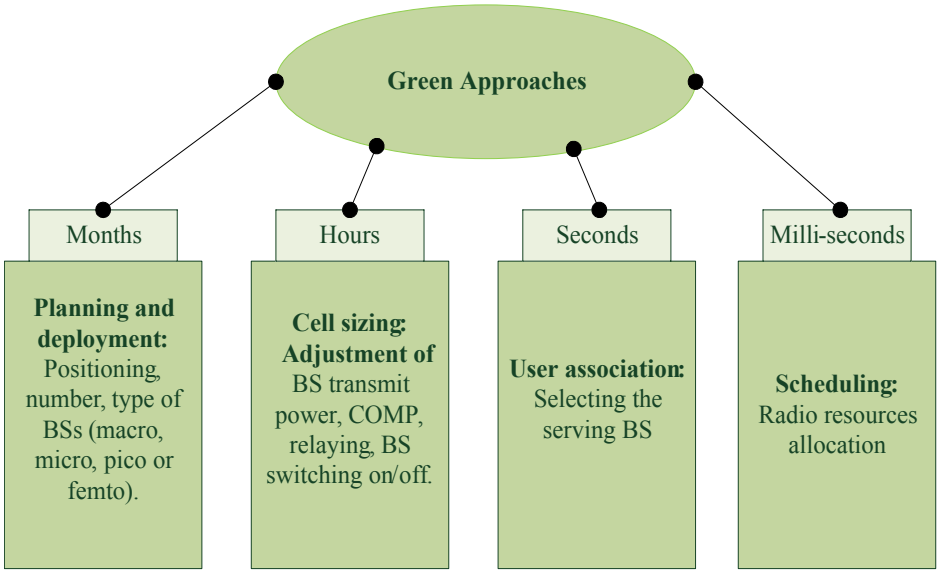


Figure 1: Green approaches at different timescales.

technology and cellular networks with Long Term Evolution (LTE) technology.

Our approach presents multiple novelties compared to the state-of the art, and our main contributions are summarized as follows:

- We formulate the multi-objective optimization problem of power saving and delay minimization in wireless access networks, going beyond prior work which has focused on either minimizing energy without considering the delay minimization [4, 11, 12, 13, 14], or on delay analysis without accounting for energy minimization [7, 8]. Hence, the novelty in our approach is that it does not only strive to save energy by reducing the network power consumption, but also considers the minimization of the transmission delay.

- Unlike most of the literature studies, we combine different green approaches (BS on/sleep mode, adjustment of BS transmit power, user association) retaining advantages of each approach to provide power savings.

- We cover IEEE 802.11 and LTE technologies. To the best of our knowledge, our formulation is the first one that captures the specificity of each technology in terms of the power model and radio resource allocation (fair-rate sharing and fair-time sharing). In the WLAN scenario, considering the fair-rate scheduler, the delay model provided is a unique feature of our work, and it is a realistic model used in IEEE 802.11 WLAN [7, 8, 15] compared to the pessimistic bound of the queuing delay model used in the literature. In the LTE scenario, we consider a flat channel model with the fair-time scheduler.

- To solve the multi-objective optimization problem, we resort to the weighted sum method by combining the multiple objectives into a single objective scalar function. This method allows us to investigate the power-delay tradeoffs by tuning the weights associated with each objective.

- Starting from a binary non-linear formulation of the problem, we provide a MILP formulation of our problem that makes it computationally tractable. We obtain optimal results for both WLAN and LTE scenarios.

The aim of the present paper is to study the different optimization models for the joint Power-Delay minimization problem in green wireless access networks. Such study brings a value to help in designing distributed solutions for the joint problem of power saving and delay minimization. Precisely, the optimal solution of the problem allows for an assessment and evaluation of any distributed solution. Moreover, the proposed formulation allows investigating the power-delay tradeoffs by tuning the weights associated with the power and delay objectives. This is an important feature of our model allowing it to reflect various preferences (i.e., saving power, minimizing delay or balance the tradeoff between minimizing power and delay). Thus, we can integrate the objective function of the joint Power-Delay

minimization problem in any practical algorithm (based on known heuristic algorithms such as simulated annealing, etc.) [16, 17].

The rest of the article is organized as follows. The performance metrics in green wireless networks are provided in Section 2. The different studied approaches to improve the energy efficiency of wireless access networks are presented in Section 3. The definitions and notations, used throughout the article, are introduced in Section 4. The Power-Delay multi-objective problem is formulated in Section 5. The network model is described in Section 6 with the models adopted for traffic, perceived rates, delay and power in IEEE 802.11 WLANs and LTE networks. The multi-objective optimization approach for WLANs and LTE cases are presented in Section 7. The methods for solving the multi-objective optimization problem are provided in Section 8. The non-linear Power-Delay minimization problem is formulated as a Mixed Integer Linear Programming (MILP) problem in Section 9. Reference models for power and user association are introduced in Section 10.1. Extensive simulation results for both IEEE 802.11 WLANs and LTE scenarios are provided in Section 10. Finally, we conclude in Section 11.

2. Performance Metrics in Green Wireless Networks

The relevant state-of-the-art performance metrics are the *energy consumption metrics*, the *QoS performance metrics* and the *tradeoff metrics*. The *energy consumption metrics* are Power consumption (P [W]), Energy consumption (E [J]), and Area Power Consumption (APC [W/m^2]). APC [4] is defined as the average power consumed in a cell divided by the corresponding cell area. Reducing energy consumption in wireless networks is coupled with satisfying required *QoS performance metrics*. The different *QoS performance metrics* are Throughput (Th [bit/s]), Delay or Transmission Delay (D [s]), Blocking Probability (BP [%]), Area Throughput (ATh [bit/s/m^2]) [11], and Area Spectral Efficiency (ASE [bit/s/Hz/m^2]). ASE [18] is defined as the summation of the spectral efficiency over a reference area. Coverage (Cov) and capacity (Cap) are also used in the literature. For instance, when planning an energy-efficient wireless network, the coverage constraint can be expressed in terms of the minimum achievable bit-rate at the cell edge, and the capacity constraint can be expressed in terms of the maximum load at a BS. Finally, the *tradeoff metrics* are used to evaluate the tradeoff between energy consumption and QoS performance and include Energy Efficiency (EE [bit/Joule] or [bit/s/Watt]) and Area Energy Efficiency (AEE [bit/Joule/m^2]). EE is defined as the average data rate provided by the network over the power consumption of the BSs. AEE [12] is defined as the EE over the area covered by the network BSs.

3. Research Approaches in Green Wireless Networks

In this section, we provide different studies on cellular networks and WLANs according to the classification presented in Section 1. We start with the planning and deployment approaches. Then, we present studies on the cell sizing approach coupled with the user association approach. Finally, we put forward the scheduling approach.

3.1. Planning and Deployment

In the planning and deployment approach, topology-specific design perspectives and improved planning methodologies were developed to improve power efficiency. Different network deployment strategies have been investigated [4, 12, 11]. The idea of deploying small, low-power micro BSs alongside with macro sites was exploited to reduce the energy consumption of cellular radio networks[4, 12]. Simulation results [4] showed that the power savings obtained from such deployments depend strongly on the offset of site power (both macro and micro). In fact, this offset accounts for the power consumed in BSs independently of the average transmit power. Traffic is assumed to be uniformly distributed [4, 12]. Tombaz *et al.* [11] introduced WLAN APs at the cell border and investigated the improvements in energy efficiency improvements through different heterogeneous networks for both uniform and non-uniform traffic distribution scenarios. Simulations showed that the heterogeneous network composed of macro BSs and WLAN APs gives the best energy efficiency results due to the low power consumption of APs. Moreover, an energy-efficient deployment strategy highly depends on the area throughput demand. For instance, for a high area throughput target, heterogeneous deployments are more energy efficient than a network composed of only macro BSs.

However, in these deployment strategies, the network configuration is fixed, even if the network may be composed of various types of BS. Precisely, at the planning/deployment stage of the network, cell size and capacity are usually

fixed based on the estimation of peak traffic load. However, traffic load in wireless networks can have significant spatial and temporal fluctuations due to user mobility and the bursty nature of mobile applications [19]. Therefore, many studies have investigated the effects of switching off BSs in consideration of traffic fluctuations in wireless networks. Eunsung *et al.* [13] proposed a basic distributed algorithm for dynamically switching off BSs to reduce network energy consumption, considering the variation of traffic characteristics over time. The results showed that the energy saving not only depend on the traffic fluctuations, but also on the BS density. Traffic is considered to be homogeneous among all BSs. In another approach [14], the coverage planning in cellular networks was studied while taking into account a sleep mode for saving energy. The results showed that for a careful design of the inter-cell distance, the network energy consumption can be reduced. The authors only took into account the coverage constraints without capacity constraints. A joint design and management optimization approach of cellular networks allowed for the adjustment of tradeoffs between the installation cost, operational cost, and connection quality cost by tuning weighting factors for each cost [20]. Moreover, BSs are switched on and off to dynamically adapt the network capacity to the traffic load without violating coverage constraints. The results showed that including energy cost in the operational cost and considering energy management strategies at the design stages produce more energy-efficient topologies than when only installation cost is considered.

3.2. Cell Sizing and User Association

The concept of cell sizing (zooming or breathing) by integrating BS switching was introduced in [21]. The authors proposed centralized and decentralized cell zooming algorithms based on the transmission rate requirements of the UEs and the capacity of the BSs. The results showed that the algorithms save a large amount of energy when traffic load is light, and they can leverage the tradeoff between energy consumption and outage probability. Bahl *et al.* [22] proposed cell breathing algorithms for WLANs that attempt to maximize the overall system throughput where results showed that this throughput is improved for both uniform and non uniform distribution.

Lorincz *et al.* [23] derived ILP models to minimize the network power consumption in WLANs while ensuring coverage of the active UEs and sufficient capacity for guaranteeing QoS. The optimization consists of switching on/off APs and adjusting their transmit power according to the traffic pattern during the day. Moreover, UEs are associated with BSs according to bandwidth requirements. The results showed significant savings in the monthly network energy consumption when optimized network management based on UE activity is implemented. By assuming that the inter-cell interference is static, Son *et al.* [10] formulated a minimization problem that allows for a flexible tradeoff between flow-level performance and energy consumption. UEs are associated with BSs so as to minimize the average flow delay, and greedy algorithms were proposed for switching the network BSs on and off. The results showed that the user association and greedy algorithms can reduce the total energy consumption, depending on the arrival rate of the traffic and its spatial distribution and the density of BS deployment. The case where BSs switch between on and off modes without adjusting their transmit power was investigated.

A distributed pricing-based algorithm was proposed that assigns UEs to BSs and adjusts the transmission power in a way that minimizes the network energy expenditure [24]. The main idea of the algorithm is to decrease the power price until all of the UEs are associated with the network BSs. The algorithm provides significant energy savings compared with a nearest-BS algorithm. For the LTE-Advanced standard, a greedy heuristic algorithm was proposed to switch off a BS according to the average distance of its UEs, thus neglecting the actual traffic load [25]. The algorithm minimizes the energy consumption of the network without compromising the QoS referred by the outage probability of the UEs.

An energy-efficient algorithm was proposed for cellular networks based on the principle of cooperation between BSs [26]. In this algorithm, the BSs dynamically switch between active/sleep modes depending on the traffic situation. Another study [27] used deterministic patterns for switching BSs through mutual cooperation among BSs. QoS is guaranteed by focusing on the worst-case transmission/reception location of the UE situated in the switched-off cell. Given the amount of time required to switch on/off a BS, focus was directed toward the design of base-station sleep and wake-up transitions, which led to a progressive BS switch off and on, respectively [28]. The results under realistic test scenarios showed that these transitions are promptly operated, allowing BSs to be switched on and off within a short time.

3.3. Scheduling

In the scheduling approach, energy-efficient schedulers were developed to reduce the network energy consumption while maintaining a satisfactory service for the end UEs. Videv *et al.* [29] developed a scheduler with aims of solving the problem of energy-efficient resource allocation in Orthogonal Frequency Division Multiple Access (OFDMA) cellular systems. The results showed that energy savings are achieved with no detriment to UE satisfaction in terms of achieved data rate. Chen *et al.* [30] proposed an energy-efficient coordinated scheduling mechanism to reduce the energy consumption in cellular networks. This is done by dynamically switching off the component carrier feature specified in LTE-A systems and BSs according to load variations, while the maintaining service continuity of UEs.

Tables 1 and 2 show a survey of recent papers that studied greening approaches and algorithms in wireless networks, with focus on the metrics used for energy consumption and QoS performance.

Table 1: Classification of approaches in green wireless networks (a)

		Richter et al. [4]	Wang et al. [12]	Tombaz et al. [11]	Eun-sung et al. [13]	Wu et al. [14]	Boiardi et al. [20]	Niu et al. [21]	Lorincz et al. [23]
Green Approaches	Planning	-	-	-	-	✓	✓	-	-
	Deployment (Macro and micro, pico, or AP)	micro	pico	micro, pico, AP	-	-	micro, pico	-	-
	BS On/Sleep/Off	-	-	-	✓	✓	✓	✓	✓
	Adjustment of BS transmit power	-	-	-	-	✓	✓	✓	✓
	CoMP	-	-	-	-	-	-	-	-
	User association	-	✓	✓	-	-	✓	✓	✓
	Scheduling	-	-	-	-	-	-	-	-
Metrics	Energy Consumption *E, P or APC*	*APC*	*APC*	*APC*	*E*	*APC*	*E*	*E*	*E,P*
	QoS Performance *Th, D, BP, ATh, ASE*	*ASE*	-	*ATh*	*BP*	-	-	*BP*	-
	Cov or	✓	✓	✓	-	✓	✓	✓	✓
	Cap (cf. Section 2)	✓	✓	✓	✓	-	✓	✓	✓
	Tradeoffs *EE* or *AEE*	-	*EE*, *AEE*	-	-	-	-	-	-
Solution Approach	Optimization	-	-	-	✓	✓	✓	-	✓
	Heuristic	-	-	-	✓	✓	-	✓	-
	Analytic	-	-	-	-	✓	-	-	-
Network and Technology Application	Cellular	✓	✓	✓	✓	✓	✓	✓	-
	LTE or LTE-A	✓	✓	-	-	-	-	-	-
	WiMAX	-	-	-	-	-	-	-	-
	2G/3G	-	-	3G	-	-	GSM	-	-
	WLAN	-	-	-	-	-	-	-	✓

4. Definitions and Notations

Let us introduce some definitions to formally characterize some concepts used in this article.
The *transmission delay* of a given UE is defined as the inverse of the throughput perceived by this UE.
The *peak rate* of a given UE is defined as the throughput experienced by the UE when served alone in the cell. The peak rate of each UE depends on its received SNR from the serving BS. The latter depend on many factors such as the transmit power of the serving BS, the pathloss model, the bandwidth, etc.

Table 2: Classification of approaches in green wireless networks (b)

		Son et al. [10]	Yildiz et al. [24]	Bousia et al. [25]	Hossain et al. [26]	Han et al. [27]	Videv et al. [29]	Chen et al. [30]
Green Approaches	Planning	-	-	-	-	-	-	-
	Deployment (Macro and micro, pico, or AP)	micro	-	-	-	-	-	-
	BS On/Sleep/Off	✓	✓	✓	✓	✓	-	✓
	Adjustment of BS transmit power	-	✓	✓	✓	✓	✓	✓
	CoMP	-	-	-	✓	✓	-	-
	User association	✓	✓	✓	-	✓	✓	✓
	Scheduling	-	-	-	-	-	✓	✓
Metrics	Energy Consumption *E, P or APC*	*P*	*P*	*P*	-	*P*	*P*	*P*
	QoS Performance *Th, D, BP, ATh, ASE*	*D*	-	*BP*	-	*BP*	*Th*	*Th, ASE*
	Cov or	✓	✓	✓	-	✓	✓	✓
	Cap (cf. Section 2)	✓	✓	-	✓	-	✓	✓
	Tradeoffs *EE or AEE*	-	-	*EE*	-	-		*EE*
Solution Approach	Optimization	✓	-	-	-	-	-	-
	Heuristic	✓	✓	✓	✓	✓	✓	✓
	Analytic	✓	-	-	-	-	-	-
Network and Technology Application	Cellular	✓	✓	✓	✓	✓	✓	✓
	LTE or LTE-A	-	-	-	-	-	✓	✓
	WiMAX	✓	-	-	-	-	-	-
	2G/3G	-	-	-	-	-	-	-
	WLAN	-	-	-	-	-	-	-

The *coverage area* of a given BS is defined as the geographical area where the received SNR of each UE is above a given minimum threshold. As the peak rate perceived by a given UE depends on its SNR, we thus consider that a UE is covered if it perceives a peak rate, from at least one BS, higher than a given peak rate threshold ($\chi_{threshold}$).

We consider a wireless access network with N_{bs} BSs. we assume that each BS operates in two modes: active mode and sleep mode. N_l denotes the number of transmit power levels of a BS. Transmitting at different power levels leads to different coverage area sizes. The indexes $i \in I = \{1, \ldots, N_{bs}\}$, and $j \in J = \{1, \ldots, N_l\}$, are used throughout the paper to designate a given BS and its transmit power level, respectively. Note that for $j = 1$, we consider that the BS transmits at the highest power level, and for $j = N_l$, the BS is in sleep mode. We term by $k \in K = \{1, \ldots, N_u\}$, the index of a given UE where N_u is the total number of UEs in the network. Let $T_{i,j,k}$ denote the transmission delay of UE k associated with BS i transmitting at level j. Let $\chi_{i,j,k}$ denote the peak rate perceived by UE k from BS i transmitting at level j.

5. Multi-objective Optimization Formulation

Our approach is formulated as an optimization problem that consists of minimizing the power consumption of the network and the sum of the transmission delays of all UEs. A key tradeoff in our problem is between these two objectives. On the one hand, reducing the transmit power level of the BSs or switching them to sleep mode to save energy, may result in increasing the transmission delay. Precisely, if there are no coverage constraints, then all BSs could be in sleep mode, no UE is served, and the transmission delay becomes infinite. On the other hand, to minimize

the transmission delay, each BS should transmit at the highest power level possible. We thus formulate the joint Power-Delay minimization problem that enables tuning the predominance of each objective. The design variables in our Power-Delay minimization problem are as follows:

- The operation mode of the network BSs (on/sleep) and the corresponding transmit power level for active BSs.
- The users association with the network BSs.

Let Λ be a matrix, with elements $\lambda_{i,j}$, defining the operation mode of the network BSs; and $\lambda_{i,j}$ be a binary variable that indicates whether or not BS i transmits at level j.

$$\lambda_{i,j} = \begin{cases} 1 & \text{if BS } i \text{ transmits at power level } j, \\ 0 & \text{otherwise.} \end{cases}$$

Let Θ be a matrix, with elements $\theta_{i,k}$, defining the users association with the network BSs; and $\theta_{i,k}$ be a binary variable that indicates whether or not a UE k is associated with BS i.

$$\theta_{i,k} = \begin{cases} 1 & \text{if UE } k \text{ is associated with BS } i, \\ 0 & \text{otherwise.} \end{cases}$$

The constraints on the decision variables are:

$$\sum_{j \in J} \lambda_{i,j} = 1, \quad \forall i \in I, \tag{1}$$

$$\sum_{i \in I} \theta_{i,k} = 1, \quad \forall k \in K, \tag{2}$$

$$\lambda_{i,N_l}\, \theta_{i,k} = 0, \quad \forall i \in I, \forall k \in K, \tag{3}$$

$$\lambda_{i,j} \in \{0,1\}, \quad \forall i \in I, \forall j \in J, \tag{4}$$

$$\theta_{i,k} \in \{0,1\}, \quad \forall i \in I, \forall k \in K. \tag{5}$$

Constraints (1) state that each BS transmits at only one power level. Constraints (2) ensure that a given UE is associated with only one BS. In practice, when BSs are in sleep mode, some UEs will be out of coverage. Thus, to prevent UEs from being associated with a BS in sleep mode, we add constraints (3). These equations ensure that λ_{i,N_l} and $\theta_{i,k}$ are not both equal to one. Indeed, when BS i is in sleep mode, λ_{i,N_l} is equal to one, so $\theta_{i,k}$ of all UEs cannot be equal to one. Constraints (4) and (5) are the integrality constraints for the decision variables $\lambda_{i,j}$ and $\theta_{i,k}$.
To eliminate some trivial cases that must not be included in the solution, we add the following constraints:

- If UE k is not covered by BS i transmitting at the first (highest) power level, then

$$\theta_{i,k} = 0. \tag{6}$$

  The equation (6) prevents a given UE from being associated with a BS if that UE is not in the BS's first power level coverage area.

- If UE k is not covered by BS i transmitting at power level j, $j \in \{2, .., N_l - 1\}$, then

$$\lambda_{i,j}\, \theta_{i,k} = 0, \quad \forall j \in \{2, .., N_l - 1\}. \tag{7}$$

  Equations (7) ensure that $\lambda_{i,j}$ and $\theta_{i,k}$ are not both equal to one, which prevents a given UE from being associated with a BS if the former is not in the BS's j^{th} power level coverage area.

The goal of our approach is to jointly minimize the total network power and the total network delay. Thus, the two objectives are:

1. The *total network power* is defined as the total power consumption of active BSs in the network. $p_{i,j}$ denotes the average power consumed per BS i transmitting at power level j. Thus, the total network power, denoted by C_p, is given by:
$$C_p(\Lambda) = \sum_{i \in I,\ j \in J} p_{i,j}\ \lambda_{i,j}. \tag{8}$$
The minimization of the total network power aims at reducing power consumption of the network.
2. The *total network delay* is defined as the sum of transmission delays (cf. Section 4) of all UEs in the network. $T_{i,j,k}$ being the transmission delay of UE k associated with BS i transmitting at level j, the total network delay, denoted by C_d, is thus given by:
$$C_d(\Lambda, \Theta) = \sum_{i \in I,\ j \in J,\ k \in K} T_{i,j,k}\ \lambda_{i,j}\ \theta_{i,k}. \tag{9}$$
The minimization of the total network delay aims at selecting the best user association that incurs the lowest sum of data unit transmission delays.

The proposed multi-objective optimization problem denoted by Multiobj-Power-Delay-Min aims at computing the transmit power level of the BSs deployed in the network as well as associating UEs with these BSs in a way that jointly minimizes the total network power and the total network delay. Therefore, Multiobj-Power-Delay-Min is given by:

$$\begin{aligned} \underset{\Lambda,\Theta}{\text{minimize}} \quad & C_p(\Lambda), \\ & C_d(\Lambda, \Theta), \\ \text{subject to} \quad & (1) \text{ to } (7). \end{aligned}$$

6. Network Model

In this section, we investigate the problem of joint Power-Delay minimization for two types of networks. Firstly, we study the case of an IEEE 802.11g WLAN, where we consider a fair-rate sharing scheme because it is the resource sharing model that stems from the Carrier Sense Multiple Access (CSMA) protocol adopted in WLANs. Secondly, we study the case of an LTE network, where we consider the fair-time sharing scheme as it corresponds to the widely used OFDMA in LTE with a round robin scheduler.

6.1. Traffic and Delay Model

As the current downlink traffic on mobile networks is still several orders higher than the uplink traffic, we only consider the downlink traffic [31]. Moreover, elastic traffic currently constitutes the majority of Internet traffic [32, 33]. We thus consider an elastic traffic model. Furthermore, we assume that:

- The network is in a static state where users are stationary.
- The network is in a saturation state. A saturation state is a worst-case scenario where every BS has persistent traffic toward UEs.
- For the WLAN case, the inter-cell interference is mitigated by assigning adjacent WLAN BSs[1] to the different IEEE 802.11 channels [34]. Particularly, in IEEE 802.11, the 2.4 GHz band consists of 14 overlapping channels, each occupying a bandwidth of 22 MHz. The three non-overlapping channels (channels 1, 6, and 11) are commonly used when designing a WLAN. Thus, one can assign one of these three frequencies to each network BS in a way that minimizes co-channel overlap. Assignment of frequencies is essentially a map coloring problem with three colors [35].

[1]For the case of WLANs, we use the term BS in this paper to designate an access point.

- In LTE networks, OFDMA is adopted as the downlink access method, which allows multiple UEs to transmit simultaneously on different subcarriers. As subcarriers are orthogonal, intra-cell interference is highly reduced. However, inter-cell interference is a key issue in OFDMA networks that greatly limits the network performance, especially for users located at the cell edge. One of the fundamental techniques to deal with the inter-cell interference problem is to control the use of frequencies over the various channels in the network [36]. There are three major frequency reuse patterns for mitigating inter-cell interference: hard frequency reuse (such as frequency reuse 1 and 3) , fractional frequency reuse and soft frequency reuse fractional [37]. Hard frequency reuse splits the system bandwidth into a number of distinct sub-bands according to a chosen reuse factor and neighboring cells transmit on different sub-bands. For instance, Frequency Reuse 3 scheme consists of dividing the frequency band into three sub-bands and allocating only one sub-band to a given cell, in such a way that the adjacent cells use different frequency bands. Compared with frequency reuse 1, this scheme leads to low interference with at the cost of a capacity loss because only one third of the resources are used in each cell [38].

6.1.1. Data Rate Model in IEEE 802.11 WLANs

With IEEE 802.11, neglecting the uplink traffic leads to a fair access scheme on the downlink channel. Accordingly, when a low-rate UE captures the channel, this UE will penalize the high-rate UEs. This also reduces the fair access strategy to a case of fair rate sharing of the radio channel among UEs [15] with the assumption of neglecting the 802.11 waiting times (*i.e.*, DIFS[2], SIFS[3]). Thus, all UEs will have the same mean throughput. When UE k is associated with BS i transmitting at level j, its mean throughput $R^W_{i,j,k}$ depends on its peak rate $\chi_{i,j,k}$ and the peak rates of other UEs associated with this same BS ($\chi_{i,j,k'}, k' \neq k$). $R^W_{i,j,k}$ is given by [7, 8]:

$$R^W_{i,j,k} = \frac{1}{\frac{1}{\chi_{i,j,k}} + \sum_{k'=1,k'\neq k}^{N_u} \frac{\theta_{i,k'}}{\chi_{i,j,k'}}}, \tag{10}$$

where $\theta_{i,k'}$ is the binary variable indicating whether or not UE k' is associated with BS i.

6.1.2. Data Rate Model in LTE

In OFDMA, the system spectrum is divided into a number of consecutive orthogonal OFDM subcarriers. The Resource Block (RB) is the smallest resource unit that can be scheduled. The RB consists of 12 consecutive subcarriers for one slot (0.5 msec) in duration. In this paper, we consider a flat channel model where each UE has similar radio conditions on all the RBs. Moreover, we consider a fair-time sharing model where RBs are assigned with equal time to UEs within a given cell. These UEs are given the same chance to access the RBs. Based on these considerations and on UEs being stationary, the scheduler is equivalent to one that allocates periodically all RBs to each UE at each scheduling epoch. Hence, when UE k is associated with BS i transmitting at level j, its mean throughput $R^L_{i,j,k}$ depends on its peak rate $\chi_{i,j,k}$ and on the number of UEs associated with the same BS. $R^L_{i,j,k}$ is given by [8]:

$$R^L_{i,j,k} = \frac{\chi_{i,j,k}}{1 + \sum_{k'=1,k'\neq k}^{N_u} \theta_{i,k'}}, \tag{11}$$

where $\theta_{i,k'}$ is the binary variable indicating whether or not UE k' is associated with BS i.

6.1.3. Delay Model in IEEE 802.11 WLANs and in LTE

$T^W_{i,j,k}$ and $T^L_{i,j,k}$ denote the transmission delay of UE k from BS i transmitting at level j in the case of a WLAN and an LTE network, respectively. As mentioned in Section 4, the transmission delay for a given UE is the inverse of the throughput perceived by this UE. Thus, for the WLAN, $T^W_{i,j,k}$ is given by:

$$T^W_{i,j,k} = \frac{1}{\chi_{i,j,k}} + \sum_{k'=1,k'\neq k}^{N_u} \frac{\theta_{i,k'}}{\chi_{i,j,k'}}, \tag{12}$$

[2]DIFS: Distributed Coordination Function Interframe Space

[3]SIFS: Short Interframe Space

and for the LTE network, $T^L_{i,j,k}$ is given by:

$$T^L_{i,j,k} = \frac{1 + \sum_{k'=1,k'\neq k}^{N_u} \theta_{i,k'}}{\chi_{i,j,k}}. \tag{13}$$

In fact, in our model, the transmission delay of a given user relies on the peak rate perceived by this user. The peak rate of each UE depends on its received SNR from the serving BS as mentioned in Section 4.

6.2. Power Consumption Model

6.2.1. Power Consumption Model in IEEE 802.11 WLANs

We consider the power consumption of an IEEE 802.11g WLAN, with BSs working in infrastructure mode. In practice, the transmission power of a WLAN BS is discrete and the maximum number of transmit power levels is equal to 5 or 6 depending on the BS manufacturer. Following the model proposed in [11], the power consumption of a WLAN BS is modeled as a linear function of the average transmit power:

$$p^W_{i,j} = L\ (a\pi^W_j + b), \tag{14}$$

where $p^W_{i,j}$ and π^W_j denote the average consumed power per WLAN BS i and the transmit power at level j respectively. The coefficient a accounts for the power consumption that scales with the transmit power due to radio frequency amplifier and feeder losses. The coefficient b models the power consumed independently of the transmit power due to signal processing, power supply consumption and cooling. Recall that for $j = 1$, we consider that the BS transmits at the highest power level, and for $j = N_l$, the BS is in sleep mode. L reflects the activity level of the WLAN BSs. Since we assume that the network is in a saturation state, L is equal to one; *e.g.*, each active WLAN BS has at least one UE requesting data and to which all resources are being allocated [4, 11].

6.2.2. Power Consumption Model in LTE

Following the model proposed in the Energy Aware Radio and neTwork tecHnologies (EARTH) project [39], the power consumption of an LTE BS is also modeled as a linear function of the average transmit power:

$$\forall i \in I, p^L_{i,j} = \begin{cases} N_{TRX}\ (v\pi^L_j + w_j), \ \ 0 < \pi^L_j \leq P^L_1, \ \ j = 1, \ldots, (N_l - 1); \\ N_{TRX}\ w_{N_l}, \ \ \pi^{N_l}_j = 0. \end{cases} \tag{15}$$

where $p^L_{i,j}$ and π^L_j denote the average consumed power per LTE BS i and the transmit power at level j respectively. For $j = 1$, we consider that the BS transmits at the highest power level, and for $j = N_l$, the BS is in sleep mode. The coefficient v is the slope of the load-dependent power consumption and it accounts for the power consumption that scales with the transmit power due to radio frequency amplifier and feeder losses. The coefficients w_j ($j = 1, .., (N_l - 1)$) represent the power consumption at zero output power (it is actually estimated using the power consumption calculated at a reasonably low output power, assumed to be 1% of p^L_1). These coefficients model the power consumed independently of the transmit power due to signal processing, power supply consumption and cooling. w_{N_l} is a coefficient that represents the sleep mode power consumption. N_{TRX} is the number of BS transceivers.

7. Optimization Approach in WLANs and LTE Scenarios

In this section we present the multi-objective optimization approach for WLANs and LTE cases. For WLANs, the proposed multi-objective optimization problem denoted by Multiobj-Power-Delay-Min-WLAN is thus obtained from the problem Multiobj-Power-Delay-Min by replacing $p_{i,j}$ and $T_{i,j,k}$ by the expressions of $p^W_{i,j}$ and $T^W_{i,j,k}$. Let C^W_p and C^W_d denote the total network power and the total network delay for this case, respectively. Therefore, the Multiobj-Power-Delay-Min-WLAN problem is given by:

$$\underset{\Lambda,\Theta}{\text{minimize}} \quad C^W_p(\Lambda) = \sum_{i\in I, j\in J} (a\pi^W_j + b)\ \lambda_{i,j}, \tag{16}$$

$$C^W_d(\Lambda, \Theta) = \sum_{i\in I, j\in J, k\in K} \left(\frac{\lambda_{i,j}\ \theta_{i,k}}{\chi_{i,j,k}} + \sum_{k'=1,k'\neq k}^{N_u} \frac{\lambda_{i,j}\ \theta_{i,k}\ \theta_{i,k'}}{\chi_{i,j,k'}}\right), \tag{17}$$

subject to: (1) to (7).

For LTE, the proposed multi-objective optimization problem denoted by Multiobj-Power-Delay-Min-LTE is thus obtained from the problem Multiobj-Power-Delay-Min by replacing $p_{i,j}$ and $T_{i,j,k}$ by the expressions of $p^L_{i,j}$ and $T^L_{i,j,k}$. Let C^L_p and C^L_d denote the total network power and the total network delay for this case, respectively. Therefore, the Multiobj-Power-Delay-Min-LTE problem is given by:

$$\underset{\Lambda,\Theta}{\text{minimize}} \quad C^L_p(\Lambda) = \sum_{i\in I, j\in J} N_{TRX}\,(v\pi^L_j + w_j)\,\lambda_{i,j}, \tag{18}$$

$$C^L_d(\Lambda,\Theta) = \sum_{i\in I, j\in J, k\in K} \frac{\lambda_{i,j}\,\theta_{i,k} + \sum_{k'=1,k'\neq k}^{N_u} \lambda_{i,j}\,\theta_{i,k}\,\theta_{i,k'}}{\chi_{i,j,k}}, \tag{19}$$

$$\text{subject to:} \quad (1)\text{ to }(7).$$

8. From Multi-objective Optimization to Single-objective Optimization

Solving a multi-objective optimization problem is a very challenging task. In this section, we provide two solution methods to multi-objective optimization: the ϵ-constraints method and the weighted sum method. Using these techniques, we obtain new optimization problems with a single objective function, which are easier to solve than the original problems.

8.1. ϵ-Constraint Method

The ϵ-constraint method is based on minimizing one objective function and considering the other objectives as constraints bound by some allowable level ϵ_n. Hence, a single objective minimization is carried out subject to additional constraints on the other objective functions. In our multi-objective optimization, since we have two objective functions, this method may be formulated in two variants presented in the following.

Power Minimization subject to Delay Constraints problem. The power minimization subject to delay constraints problem, denoted by Power-Min-Delay-Const, is given by:

$$\underset{\Lambda}{\text{minimize}} \quad C_p(\Lambda), \tag{20}$$

$$C_d(\Lambda,\Theta) \leq \epsilon_1, \tag{21}$$

$$\text{subject to:} \quad (1)\text{ to }(7).$$

ϵ_1 is a value of the total network delay which we do not wish to exceed. The Power-Min-Delay-Const problem can be literally expressed as: given some delay bound (constraint (21)), is there a BS operation mode and a user association satisfying constraints (1) to (7) such that the total network power C_p is minimized?

Delay Minimization subject to Power Constraints problem. The delay minimization subject to power constraints problem, denoted by Delay-Min-Power-Const is given by:

$$\underset{\Lambda,\Theta}{\text{minimize}} \quad C_d(\Lambda,\Theta), \tag{22}$$

$$C_p(\Lambda) \leq \epsilon_2, \tag{23}$$

$$\text{subject to:} \quad (1)\text{ to }(7).$$

ϵ_2 is a value of the total network power which we do not wish to exceed. In other words, the aforementioned problem can be expressed as: given some power bound (constraint (23)), is there a BS operation mode and a user association satisfying constraints (1) to (7) such that the total network delay is C_d is minimized?
The major drawback of such problems is that the decision maker (*i.e*, the network operator) cannot estimate the total network delay or the total network power. Thus, it is hard to choose the adequate bounds on the delay or the power.

8.2. Weighted Sum Method

The weighted sum method consists of summing the objective functions combined with different weighting coefficients. The multi-objective optimization problem is transformed into a scalar optimization problem, denoted by Weighted-Sum-Power-Delay-Min:

$$\begin{aligned} \underset{\Lambda,\Theta}{\text{minimize}} \quad & C_t(\Lambda,\Theta) = \alpha C_p(\Lambda) + \beta\beta' C_d(\Lambda,\Theta), \\ \text{subject to:} \quad & (1) \text{ to } (7). \end{aligned} \tag{24}$$

where C_t denotes the *total network cost* defined as the weighted sum of the total network power and the total network delay. β' is a normalization factor that will scale the two objectives properly. α and β are the weighting coefficients representing the relative importance of the two objectives. It is usually assumed that $\alpha + \beta = 1$ and that α and $\beta \in [0,1]$. In particular, when α equals 1 and β equals 0, we only focus on the power saving. As α decreases and β increases more importance is given to minimizing the delay. By tuning the weighting coefficients, we obtain different points located on the Pareto frontier presenting all the compromises between the two objectives. The weighting coefficients are also called tuning factors, as decision makers use them to fine-tune the model to reflect their decision preferences.

In this article, we choose the weighted sum method in order to study the tradeoffs between minimizing the power consumption of the network and minimizing the sum of UE transmission delays in the network for the WLAN and LTE cases. Let Weighted-Sum-Power-Delay-Min-WLAN and Weighted-Sum-Power-Delay-Min-LTE denote the scalar optimization problems for the considered cases, respectively. Consequently, the objective functions of these problems are obtained by replacing C_p and C_d by C_p^W and C_d^W for the WLAN case and by C_p^L and C_d^L for the LTE case, respectively. Let C_t^W and C_t^L denote the total network cost for WLAN and LTE cases, respectively.
Therefore, the objective function of the Weighted-Sum-Power-Delay-Min-WLAN problem is given by:

$$\underset{\Lambda,\Theta}{\text{minimize}}\ C_t^W(\Lambda,\Theta) = \alpha \sum_{i\in I, j\in J} (a\pi_j^W + b)\ \lambda_{i,j} + \beta\beta' \sum_{i\in I, j\in J, k\in K} \left(\frac{\lambda_{i,j}\ \theta_{i,k}}{\chi_{i,j,k}} + \sum_{k'=1,k'\neq k}^{N_u} \frac{\lambda_{i,j}\ \theta_{i,k}\ \theta_{i,k'}}{\chi_{i,j,k'}}\right); \tag{25}$$

and the objective function of the Weighted-Sum-Power-Delay-Min-LTE problem is given by:

$$\underset{\Lambda,\Theta}{\text{minimize}}\ C_t^L(\Lambda,\Theta) = \alpha \sum_{i\in I, j\in J} N_{TRX}\ (v\pi_j^L + w_j)\ \lambda_{i,j} + \beta\beta' \sum_{i\in I, j\in J, k\in K} \frac{\lambda_{i,j}\ \theta_{i,k} + \sum_{k'=1,k'\neq k}^{N_u} \lambda_{i,j}\ \theta_{i,k}\ \theta_{i,k'}}{\chi_{i,j,k}}. \tag{26}$$

The scalar optimization problems are binary non-linear. Such problems can be optimally solved using an exhaustive search algorithm [40]. However, the complexity of searching only for the operation mode of the BS is in $O(N_l^{N_{bs}})$. This makes the exhaustive search very computational intensive, and rapidly becomes intractable for modest sized networks. Thus, in the next section we convert the optimization problems into MILP problems to make them computationally tractable.

9. Mixed Integer Linear Programming Formulation

In this section, we explain how to convert our non-linear optimization problems Weighted-Sum-Power-Delay-Min-WLAN and Weighted-Sum-Power-Delay-Min-LTE into MILP problems. A MILP problem consists of a linear objective function, a set of linear equality and inequality constraints and a set of variables with integer restrictions. The number of constraints and variables are important factors when estimating if a problem is tractable. Generally, MILP problems are solved using a linear-programming based branch-and-bound approach. The idea of this approach is to solve Linear Program (LP) relaxations of the MILP and to look for an integer solution by branching and bounding on the decision variables provided by the LP relaxations. Thus, in a branch-and-bound approach the number of integer variables determines the size of the search tree and influences the running time of the algorithm.

Based on our work [41], to linearize the non-linear optimization problems we replace the non-linear terms by new variables and additional inequality constraints which ensure that the new variables behave according to the non-linear

terms they are replacing. Particularly, in the objective functions ((25) and (26)), we replace each quadratic term $\lambda_{i,j}\,\theta_{i,k}$ by a new linear variable $y_{i,j,k}$ and add the following three inequalities to the set of constraints:

$$y_{i,j,k} - \lambda_{i,j} \leq 0, \quad \forall i \in I, \forall j \in J, \forall k \in K, \tag{27}$$

$$y_{i,j,k} - \theta_{i,k} \leq 0, \quad \forall i \in I, \forall j \in J, \forall k \in K, \tag{28}$$

$$\lambda_{i,j} + \theta_{i,k} - y_{i,j,k} \leq 1, \quad \forall i \in I, \forall j \in J, \forall k \in K. \tag{29}$$

The inequalities (27) and (28) ensure that $y_{i,j,k}$ equals zero when either $\lambda_{i,j}$ or $\theta_{i,k}$ equals zero, while the inequalities (29) force $y_{i,j,k}$ to be equal to one if both $\lambda_{i,j}$ and $\theta_{i,k}$ equal one. Moreover, constraints (3) and (7) will be replaced respectively by (30) and (31):

$$y_{i,N_l,k} = 0, \quad \forall i \in I, \forall k \in K, \tag{30}$$

$$y_{i,j,k} = 0, \quad \forall i \in I, \forall j \in \{2, .., N_l - 1\}, \forall k \in K / \chi_{i,j,k} \leq \chi_{threshold}. \tag{31}$$

Similarly, we replace each term $\lambda_{i,j}\,\theta_{i,k}\,\theta_{i,k'}$ in the objective functions ((25) and (26)) by a new variable $z_{i,j,k,k'}$ and add the following inequalities to the set of constraints:

$$z_{i,j,k,k'} - \lambda_{i,j} \leq 0, \quad \forall i \in I, \forall j \in J, \forall k < k' \in K, \tag{32}$$

$$z_{i,j,k,k'} - \theta_{i,k} \leq 0, \quad \forall i \in I, \forall j \in J, \forall k < k' \in K, \tag{33}$$

$$z_{i,j,k,k'} - \theta_{i,k'} \leq 0, \quad \forall i \in I, \forall j \in J, \forall k < k' \in K, \tag{34}$$

$$\lambda_{i,j} + \theta_{i,k} + \theta_{i,k'} - z_{i,j,k,k'} \leq 2, \quad \forall i \in I, \forall j \in J, \forall k < k' \in K, \tag{35}$$

$$z_{i,j,k,k'} - z_{i,j,k',k} = 0, \quad \forall i \in I, \forall j \in J, \forall k < k' \in K. \tag{36}$$

The inequalities (32), (33) and (34) ensure that $z_{i,j,k,k'}$ is equal to zero when either $\lambda_{i,j}$ or $\theta_{i,k}$ or $\theta_{i,k'}$ equals zero, while the inequalities (35) force $y_{i,j,k}$ to be equal to one if $\lambda_{i,j}$, $\theta_{i,k}$ and $\theta_{i,k'}$ are equal to one. Furthermore, as $\lambda_{i,j}\,\theta_{i,k}\,\theta_{i,k'} = \lambda_{i,j}\,\theta_{i,k'}\,\theta_{i,k}$, constraints (36) force $z_{i,j,k,k'}$ to be equal to $z_{i,j,k',k}$. In addition, we give the bound constraints for the variables $y_{i,j,k}$ and $z_{i,j,k,k'}$ as follows:

$$0 \leq y_{i,j,k} \leq 1, \quad \forall i \in I, \forall j \in J, \forall k \in K, \tag{37}$$

$$0 \leq z_{i,j,k,k'} \leq 1, \quad \forall i \in I, \forall j \in J, \forall k < k' \in K. \tag{38}$$

The MILP Weighted-Sum-Power-Delay-Min-WLAN problem is given by:

$$\begin{aligned} \underset{\Lambda,Y,Z}{\text{minimize}} \quad & C_t^W(\Lambda, Y, Z) = \alpha \sum_{i \in I, j \in J} (a\pi_j^W + b)\,\lambda_{i,j} + \beta\beta' \sum_{i \in I, j \in J, k \in K} \Big(\frac{y_{i,j,k}}{\chi_{i,j,k}} + \sum_{k' \in K, k' \neq k} \frac{z_{i,j,k,k'}}{\chi_{i,j,k'}}\Big), \\ \text{subject to:} \quad & (1) \text{ to } (7) \text{ and } (27) \text{ to } (38); \end{aligned} \tag{39}$$

where $y_{i,j,k}$ and $z_{i,j,k,k'}$ are respectively the elements of the matrices Y and Z. Similarly, the MILP Weighted-Sum-Power-Delay-Min-LTE problem is given by:

$$\begin{aligned} \underset{\Lambda,Y,Z}{\text{minimize}} \quad & C_t^L(\Lambda, Y, Z) = \alpha \sum_{i \in I, j \in J} (N_{TRX}\,(v\pi_j^L + w_j))\,\lambda_{i,j} + \beta\beta' \sum_{i \in I, j \in J, k \in K} \Big(\frac{y_{i,j,k} + \sum_{k' \in K, k' \neq k} z_{i,j,k,k'}}{\chi_{i,j,k}}\Big), \\ \text{subject to:} \quad & (1) \text{ and } (2), (4) \text{ to } (6), \text{ and } (27) \text{ to } (38); \end{aligned} \tag{40}$$

The main notations used in our paper are reported in Table 3.

Table 3: Notation Summary

Notation	Definition
I	Set of network BSs
J	Set of transmit power levels of a given BS
K	Set of UEs in the network
N_{bs}	The total number of BSs
N_l	The total number of transmit power levels
N_u	The total number of UEs
$p_{i,j}$	The average consumed power per BS i transmitting at power level j
π_j^L	The transmit power at level j for LTE BSs
π_j^W	The transmit power at level j for WLAN BSs
$\chi_{i,j,k}$	The peak rate perceived by UE k from BS i transmitting at level j
$T_{i,j,k}$	The transmission delay of UE k associated with BS i transmitting at level j
$\theta_{i,k}$	A binary variable that indicates if UE k is associated with BS i
$\lambda_{i,j}$	A binary variable that indicates if BS i transmits at power level j
$y_{i,j,k}$	A real variable that indicates if UE k is associated with BS i transmitting at power level j
$z_{i,j,k,k'}$	A real variable that indicates if UE k and UE k' are associated with BS i transmitting at power level j

Table 4: Five studied settings for WLAN and LTE scenarios

Settings	Weighting coefficients value	Description
S1	α = 0.99, β=0.01	Preference is given to saving power
S2	α = 0.75, β=0.25	
S3	α = 0.5, β=0.5	Balance the tradeoff between minimizing power and delay
S4	α = 0.25, β=0.75	Preference is given to minimizing delay
S5	α = 0.01, β=0.99	

10. Performance Evaluation

To study the tradeoff between minimizing the power consumption of the network and minimizing the sum of UE transmission delays in the network, we tune the values of the weights α and β (in (39) and (40)) associated with the total network power and total network delay respectively, and investigate the obtained solutions for WLANs and LTE networks. We consider five settings illustrated in Table 4. Settings S1 and S2 match the case where preference is given to power saving. Setting S3 matches the case where the power and delay are equally important. Settings S4 and S5 match the case where preference is given to minimizing delay.

Moreover, the normalization factor β' is calculated in each simulation so as to scale the total network power and the total network delay [42]. Furthermore, we adopt the Monte Carlo method by generating 50 snapshots with different random uniform UE distribution. After doing the calculations for all the snapshots, we provide the 95% Confidence Interval (CI) for each simulation result. In the simulation results, the optimal solutions are provided for the different settings (S1, S2, S3, S4, S5).

10.1. Power and User Association Reference Models

In legacy WLANs or cellular networks, BSs transmit at a fixed power level, and UEs are associated with the BS delivering the highest SNR [12, 43]. Based on these legacy networks, we devise a reference model composed of *i)* the Highest Power Level (HPL) as the reference power model, which assumes that all BSs transmit at the highest power level, and *ii)* the Power-based User Association (Po-UA) as the reference user association model, which associates a UE with the BS where it obtains the highest SNR. In what follows, we denote the reference model by Po-UA/HPL. The total network power and the total network delay of this model will serve as reference values for comparison of

Table 5: Covering BSs per UE vs. inter-cell distance in WLAN scenarios.

Inter-cell distance [m]	120.8	134.2	147.6	161.1	174.5	187.9	201.3	214.8
Number of covering BSs per UE	2.02	1.76	1.53	1.38	1.25	1.15	1.05	1.00

the results. In the following, we present the evaluation method and the simulation results for the WLAN and LTE scenarios.

10.2. IEEE 802.11g WLAN Scenarios

10.2.1. Evaluation Method

To evaluate the tradeoff between power and delay in WLANs, we compute the optimal solution of our MILP Weighted-Sum-Power-Delay-Min-WLAN problem with the GLPK (GNU Linear Programming Kit) solver [44] over a network topology composed of nine cells ($N_{bs} = 9$) using the IEEE 802.11g technology and six UEs in each cell ($N_u = 9 * 6 = 54$). The positioning of the WLAN BSs in the network is performed following a grid structure and the positioning of UEs is generated randomly following a uniform distribution.

In the BS power model, for simplicity, we set the number of transmit power levels to three (N_l=3). Precisely, an active BS is able to transmit at two different power levels, and when the power level equals $N_l = 3$, the BS is in sleep mode. For the WLAN, we consider that when the BS is in sleep mode, it consumes only power due to signal processing neglecting the cooling. It is estimated that the power consumption of signal processing circuits accounts for only 10% of the total consumed power [45]. Therefore, we assume that in sleep mode, the WLAN BS power consumption is negligible, and it is considered to be switched-off. We aim to compute the optimal solution of the MILP problem. Thus, if we increase N_l, the granularity will be finer but the problem becomes intractable. The input parameters of the power consumption model in (14) are as follows: *i)* as proposed in [11], the values of a and b used in this paper are $a = 3.2$ and $b = 10.2$; *ii)* as proposed in [46], the transmit power at levels one and two is $\pi_1^W = 0.03$ W and $\pi_2^W = 0.015$ W, respectively. Hence, the average power consumed per BS $i \in I$ at the first, second and third power levels is given by $p_{i,1} = 10.296$ W, $p_{i,2} = 10.248$ W, and $p_{i,3}$=0, respectively.

Peak rate and coverage area computation. In NS2 [47], we implement a benchmark scenario that enables the computation of the peak rate perceived by the UE from the BS and the coverage area of the WLAN BS. Particularly, the benchmark scenario consists of a free propagation model to characterize the WLAN radio environment, an IEEE 8021.11g BS working at 2.4 GHz, and a single UE at different positions. This UE receives Constant Bit Rate (CBR) traffic from the BS with a packet size of 1000 bytes and an inter-arrival time of 0.4 ms corresponding to a rate of 20 Mbit/s. This leads to a saturation state of the network according to the assumption presented in section 6.1. In these conditions, the throughput experienced by the single UE is the maximum achievable throughput (peak rate) for the current SNR. We run this scenario for each BS transmit power level ($\pi_1 = 0.03$ W and $\pi_2 = 0.015$ W) to obtain $\chi_{i,1,k}$ and $\chi_{i,2,k}$, respectively, for the corresponding UE. Figure 2 shows the peak rate perceived by the UE from the BS, transmitting at the first and the second power level, as a function of the distance. In addition, the coverage radius for the first and second power levels are $R_1 = 107{,}4$ m and $R_2 = 75{,}8$ m, respectively. These radii correspond to an SNR threshold that equals -0.5 dB at the cell boundary. This SNR is the minimum value to be maintained in order to consider that a given UE is covered by a BS. It corresponds to a cell edge peak rate that equals 1 Mbit/s ($\chi_{threshold}$=1 Mbit/s) on the downlink. We note that, considering a more realistic propagation model will only affect the values of the user peak rate. The considered peak rate will be lower than that considering a Free propagation model.

In the following, the simulation results are plotted as a function of the inter-cell distance D. Particularly, this parameter has a large impact on not only the computational complexity of the algorithm but also on the quality of the solution. For small inter-cell distances, we obtain a dense coverage area, while large inter-cell distances produce a sparse coverage area. Table 5 shows the average number of covering BSs per UE as a function of the inter-cell distance. For $D = 120.8$ m, we obtain a dense coverage area where the average number of covering BSs per UE is 2.02. As D increases, the average number of covering BSs per UE decreases to be equal to one when there is no overlap between cells ($D = 2R_1$). Figure 3 shows an example of the network topology with an inter-cell distance equals 120.8 m.

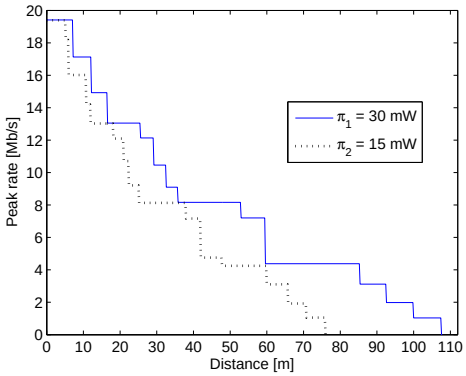


Figure 2: Peak rates in IEEE 802.11g for different transmit power levels.

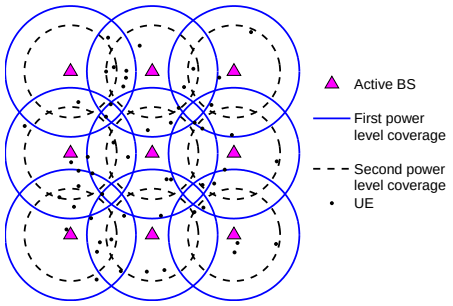


Figure 3: An example of WLAN topology with inter-cell distance D = 120.8 m.

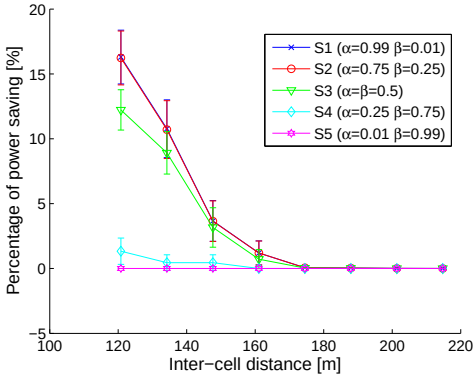


Figure 4: Power saving for the considered settings in WLAN scenarios.

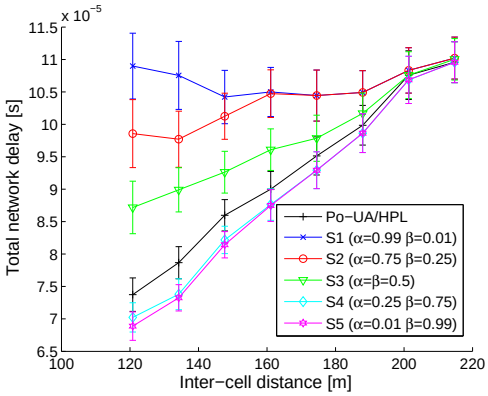


Figure 5: Total network delay for the considered settings and for Po-UA/HPL in WLAN scenarios.

10.2.2. Simulation Results

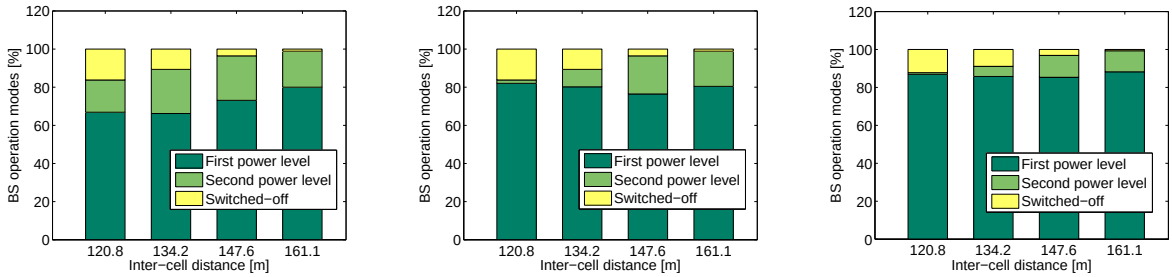
Let us start by examining the power saving achieved for the five considered settings, which is computed as follows:

$$100 \times (1 - \frac{\text{total network power for the considered setting}}{\text{total network power for Po-UA/HPL model}}), \tag{41}$$

where the total network power is computed according to the value of C_p^W in (16). Recall that in the Po-UA/HPL reference model, all BSs transmit at the highest power level.

Figure 4 plots the percentage of power saving for the five considered settings as a function of the inter-cell distance D. S1 and S2 have the highest percentage of power saving, followed successively by S3 then S4 for D ranging from 120.8 m to 161.1 m, while no power saving is obtained for $D \geq 161.1$ m. For instance, when D = 120.8 m, we obtain power saving of up to 16% in S1 and S2, followed by S3 at 12.22% and by S4 at 1.33%. S5 has no power saving for any distance. In other words, in S5, we obtain a BS operation mode where all BSs transmit at the highest power level (similar to the Po-UA/HPL model). Precisely, in this setting preference is given to minimizing the sum of UE delays, so when all BSs transmit at the highest level, UEs experience lower delays in comparison with the case where some of the BSs transmit at the second power level or are switched off.

In order to examine the cause of significant power savings in settings S1, S2 and S3, we plot Fig. 6, which illustrates the percentage of the BS operation modes for D ranging from 120.8 m to 161.1 m. We see that S1 has the highest percentages of BSs transmitting at the second power level and switched-off, followed by S2 then by S3 for the different values of D. Moreover, for these three settings, we note that when D increases, the percentage of



(a) S1 (α=0.99 β=0.01) (b) S2 (α=0.75 β=0.25) (c) S3 (α=0.5 β=0.5)

Figure 6: Percentage of BS operation modes in settings S1, S2 and S3 in WLAN scenarios.

switched-off BSs decreases, and the percentage of BSs transmitting at the second power level increases. On the one hand, this explains the decreasing curves for the corresponding inter-cell distances in Fig. 4. On the other hand, for low values of D (*i.e.*, 120.8 m), this behavior is due to the relatively high number of covering BSs per UE (*i.e.*, 2.02 as shown in Tab. 5). Thus, the possibility of switching off the BS or transmitting at low power level is high. However, for large D values, the number of covering BSs per UE decreases, and thus the possibility to switch off a BS or to transmit at low power level decreases due to the coverage constraints.

We now investigate the total network delay for the considered settings compared to the Po-UA/HPL model while varying the inter-cell distance (D). The total network delay is computed according to the value of C_d^W in (17). For the comparison of the five settings, Fig. 5 shows that S5 has the lowest total network delay, followed successively by S4, S3, S2, and finally S1. Particularly, in S5, more weight is given to minimizing the delay (β=0.99), thus we obtain a network operation mode where all BSs transmit at the highest power level (as shown in Fig. 4). The problem becomes a user association problem that aims to minimize the sum of network UE transmission delays. With the decrease of β, more BSs are switched off or transmit at the second power level (as shown in Fig. 6), and thereby UEs will experience higher delay. Compared to the Po-UA/HPL model, we obtain a delay reduction for all the inter-cell distances in S4 and S5. For instance, the delay reduction is 4.5% and 6.64% in S4 and S5, respectively, for D=120.8 m. However, we obtain a higher total network delay compared to Po-UA/HPL for all D in S1, S2 and S3. Further, we see that in S4 and S5, the total network delay has an increasing curve. Precisely, for a given UE distribution, when D increases, the SNR of the UE will decrease, causing the delay to increase. Similarly, we also see that in S2 and S3, the total network delay has an increasing curve but with lower slope at the first inter-cell distances. In S1, the total network delay has a decreasing curve for D between 120.8 m and 161.1 m and then it increases for $D \geq 161.1$ m. In particular, for D between 120.8 m and 161.1 m, more BSs transmit at either the highest power level or the second power level (Fig. 6(a)). UEs will thereby experience a lower delay. Note that all the curves converge to the same point. Indeed, when D increases, the cell overlap decreases, and thus, the optimal solution for the five settings turns on the BSs to achieve a point where all the BSs transmit at the highest power level. Therefore, the problem boils down to a user association problem that minimizes the sum of UE delays.

In Fig. 7, we plot the power-delay tradeoff curves for different inter-cell distances D ranging from 120.8 m to 161.1 m. The five points of the illustrated curves are obtained by plotting the values of the 95% CI of the total network power as a function of the 95% CI of the total network delay for the five considered settings. For all the inter-cell distances, we obtain a reduction in the network power consumption at the cost of delay increase. In particular, for D=120.8 m, in S5 (α = 0.01 β=0.99), we obtain the solution with the lowest total network delay (6.88×10^{-5} s) and the highest total network power (92.664 Watt); while in S1 (α = 0.99 β=0.01), we obtain the solution with the lowest total network power (77.62 W) and the highest total network delay (10.90×10^{-5} s). Moreover, we note that when D increases, the tradeoff curves become flat. For instance, for D=161.1 m, we obtain similar total network power in the five settings. Indeed, for sparse coverage area, the problem in the five settings becomes similar to a user association problem where there is no longer an interesting power-delay tradeoff. In fact, these curves represent the Pareto frontier at different inter-cell distances. Hence, a network operator has the option to choose the operation point of the network. For instance, the operator can choose the optimal inter-cell distance of his network. Moreover, the operator has the choice to privilege power saving, minimize delay, or balance the tradeoff between the two objectives.

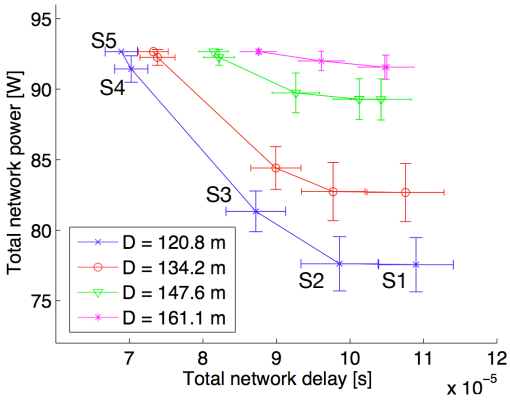
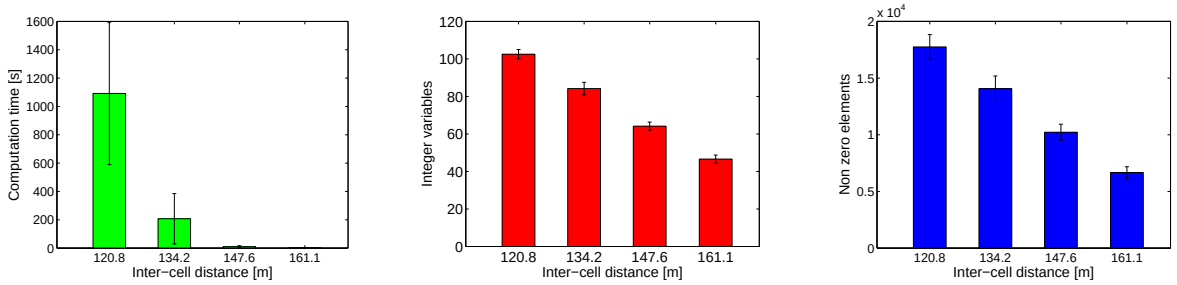


Figure 7: Pareto frontier at different inter-cell distances in WLAN scenarios.

10.3. Computational complexity

In order to assess the computational complexity of the optimal solution, we calculate in the following its computation time and the number of binary integer variables. Also, we compute the number of non-zero elements of the matrix defining the constraints of the minimization problem. Figure 8 shows the 95% CI of the computational com-



(a) Computation time of the optimal solution (b) Number of binary integer variables (c) Number of non-zero elements

Figure 8: Optimal solution computational complexity measurements.

plexity measurements as a function of the inter-cell distance (D). We note that the computation time of the optimal solution decreases when D increases as shown in Fig. 8(a). Precisely, when D increases, the number of binary integer variables and non-zero elements decreases as shown in Fig. 8(b) and Fig. 8(c). In fact, when the inter-cell distance increases, the number of UEs covered by each BS decreases. This causes the related solution space (for selecting the BS transmit power level, and the user association) to be small. Thereby, this decreases the number of binary integer variables and non-zero elements. Moreover, we note that we cannot obtain solutions for very dense networks (*e.g.*, $D \leq 120.8\ m$), as the problem becomes intractable. Therefore, in order to overcome such issue, we introduce in [16] a heuristic algorithm that computes satisfactory solutions for the problem while keeping low computation complexity.

10.4. LTE Scenario

10.4.1. Evaluation Method

To evaluate the tradeoff between power and delay in LTE, we compute the optimal solution of our MILP Weighted-Sum-Power-Delay-Min-LTE problem using the CPLEX solver [48]. The input data for the CPLEX solver are generated using MATLAB [49]. Thus, in MATLAB, we implement an LTE network topology where the LTE BSs are transmitting using omni-directional antennas in three deployment cases: urban, rural and realistic LTE deployment. For both cases (urban and rural deployment), the network topology is composed of nine cells ($N_{bs} = 9$) and the positioning of the LTE BSs in the network is performed following a grid structure. For the realistic LTE deployment, we

Table 6: Simulation parameters for LTE

Parameter	**Value**
Input parameters of power consumption model	N_{TRX}=1, v = 4.7, $w_1 = w_2$ = 130 W w_3 = 75 W
Transmit power	π_1=10 W, π_2=5 W π_3=0
Average power consumed per BS i	$p_{i,1}$=177 W, $p_{i,2}$=153.5 W, $p_{i,3}$=75 W ($i \in I$)
Transmit antenna gain	15 dBi
Receiver antenna gain	0
Coverage radius for the first and the second power levels	R_1 = 500 m R_2 = 250 m
Inter-cell distance	Urban: 700 m Rural: 900 m
Pathloss model	Cost 231 extended Hata model
Shadowing standard deviation	10 dB
Carrier frequency	2000 MHz
Bandwidth	5 MHz
Frequency Reuse scheme	3
Number of RB per cell	8
Bandwidth per RB	180 KHz
Traffic model	Full buffer
Noise figure	9 dB
Thermal noise density	-174 dBm/Hz
Thermal noise power	-103.4 dBm

consider the positioning of the LTE network BS for the 14^{th} district of Paris-France, where the network topology is composed of 18 cells (N_{bs} = 18).
The simulated LTE system bandwidth is 5 MHz. Therefore we have 25 RBs available in each cell. We assume a frequency reuse 3 scheme in the network to mitigate the inter-cell interference. Thus, the system bandwidth is divided into three equal sub-bands, each of these sub-bands is allocated to cells in a manner that no other surrounding cell is using the same sub-band. Consequently, we have eight RBs available in each cell. The fair-time sharing model is used, and the scheduler allocates periodically all RBs to each UE at each scheduling epoch as explained in Section 6.1.2. Moreover, we assume a full buffer traffic model. The simulation parameters and the pathloss model follow those in [39, 50, 51], which are summarized in Tab. 6.

Propagation model. The Cost 231 extended Hata model is used for modeling the radio channel of the LTE BS in urban [51] and rural [50] environments, with a carrier frequency f of 2000 MHz. The shadowing is represented by a random variable following a normal distribution with a mean of 0 dB and a standard deviation of 10 dB. For urban deployment cases, the antenna height h_a equals 30 m, the UE height h_u equals 1.5 m, the area type correction factor C_m equals 3 dBm, and the UE-BS separation is denoted by d [Km]. Therefore, the urban path loss L_1 is given by:

$$L_1 = 46.3 + 33.9 \times \log_{10}(f) - 13.82 \times \log_{10}(h_a) - a + (44.9 - 6.55 \times log10(h_a)) \times \log_{10}(d) + C_m + shadowing \ \ [dB]; \quad (42)$$

where $a = (1.1 \times \log_{10}(f) - 0.7) \times h_u - (1.56 \times \log_{10}(f) - 0.8)$.
For rural deployment cases, the antenna height equals 45 m. Therefore, the rural path loss L_2 is given by:

$$\begin{aligned} L_2 = 69.55 + 26.16 \times \log_{10}(f) - 13.82 \times \log_{10}(h_a) + (44.9 - 6.55 \times \log_{10}(h_a)) \times \log_{10}(d) \\ - 4.78 \times (\log_{10}(f))^2 + 18.33 \times \log_{10}(f) - 40.94 + shadowing \ \ [dB]; \end{aligned} \quad (43)$$

Peak rate computation. Knowing the path loss, the signal strength $S_{i,j,k}$ received by UE k from BS i transmitting at power level j is calculated according to:

$$S_{i,j,k} = 10 \times \log_{10}(\pi_j^L \times 1000) - (PathLoss - TransmitAntennaGain - ReceiverAntennaGain) \ \ [dBm]. \quad (44)$$

The SNR detected by UE k from BS i transmitting at power level j is thus given by:

$$SNR_{i,j,k} = S_{i,j,k} - ThermalNoisePower \ \ [dB], \tag{45}$$

where the thermal noise power is given by:

$$\begin{aligned} ThermalNoisePower = 10 \times \log_{10}(1000 \times ThermalNoiseDensity \times BandwidthperRB \\ \times Numberof RBperCell) + NoiseFigure \ \ [dBm]. \end{aligned} \tag{46}$$

Knowing the SNR, the spectral efficiency (in bit/s/Hz) is computed according to Fig. 9 in the 3GPP TR 36.942 [51]. As mentioned earlier, the scheduler allocates all RBs to one UE at each scheduling epoch. Therefore, to compute the

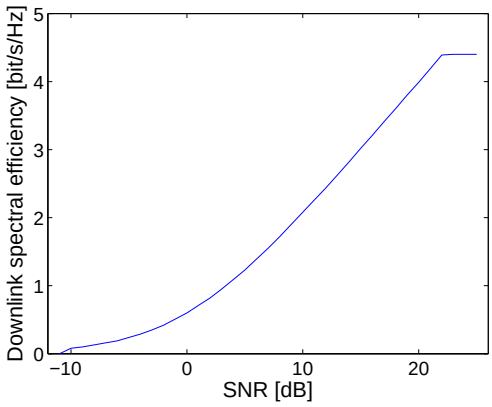


Figure 9: Spectral efficiency in LTE as a function of SNR [51].

peak rate $\chi_{i,j,k}$ perceived by UE k from BS i transmitting at power level j in bit/s, we multiply the value obtained from Fig. 9 by the bandwidth per RB and by the number of RBs per cell. Next, we consider urban and rural deployments, and assess the impact of the end UEs position in the cell on the achievable tradeoffs.

10.4.2. Simulation Results for a Urban Deployment

For the urban deployment, we consider an inter-cell distance of 700 m, and we provide simulation results for both uniform and non-uniform UE distributions.

Uniform Distribution of UEs. We consider six UEs in each cell, and a total of 54 UEs ($N_u = 9{\times}6 = 54$) in the network. The positioning of UEs is generated randomly following a uniform distribution. In the present urban deployment, the number of covering BSs per UE equals 1.5. As illustrated in Tab. 4, we recall that settings S1 and S2 match the case where preference is given to power saving. Setting S3 matches the case where the power and delay are equally important. Settings S4 and S5 match the case where preference is given to minimizing delay.

We illustrate in Fig. 10 the percentage of power saving for the considered settings, which is computed as follows:

$$100 \times (1 - \frac{\text{total network power for the considered setting}}{\text{total network power for Po-UA/HPL model}}), \tag{47}$$

where the total network power is replaced by the expression of C_p^L in (18). Figure 10 shows that settings S1 and S2 exhibit the highest percentage of power saving at 3.5%, followed by S3 at 2% and by S4 and S5 at 0.4%. In order to examine the cause of power savings in settings S1, S2 and S3, we provide the percentage of the operation modes of the BSs in Tab. 7. S1 has the highest percentages of BSs transmitting at the second power level and in sleep mode, followed by S2 and S3.

The percentage of power saving is relatively low. This is because the power consumption of sleep mode (75 W) constitutes 42% of the power consumption of the highest power level (177 W).

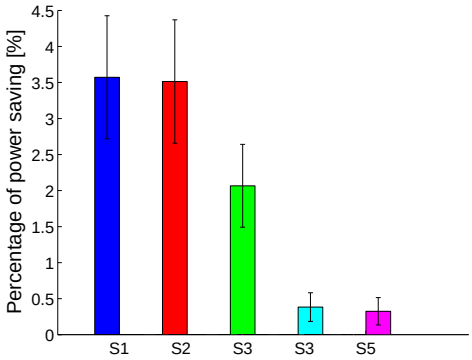


Figure 10: Power saving for the considered settings in LTE scenario for a urban deployment.

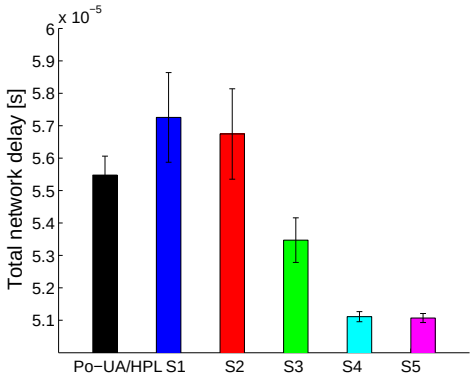


Figure 11: Total network delay for the considered settings and for Po-UA/HPL in LTE scenario for a urban deployment.

Table 7: Percentage of the BS operation modes [%]

	First power level	Second power level	Sleep
S1	82	15.33	2.67
S2	82.44	14.89	2.67
S3	88.89	9.78	1.33

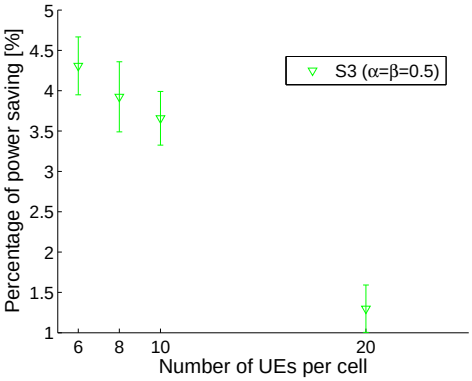


Figure 12: Power saving for setting S3 for non-uniform UEs distribution in LTE scenario for a urban deployment.

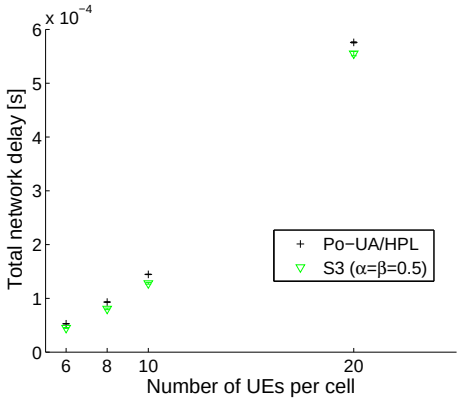


Figure 13: Total network delay for setting S3 and for Po-UA/HPL for non-uniform UEs distribution in LTE scenario for a urban deployment.

Let us investigate the total network delay for the considered settings compared to the Po-UA/HPL model. The total network delay is computed according to the expression of C_d^L given in (19). Figure 11 shows that S5 has the lowest total network delay, followed successively by S4, S3, S2, and S1. Particularly, in S5 preference is given to minimizing the delay (β=0.99). As for scenarios with decreasing values of β, more BSs transmit at the second power level or are in sleep mode (as shown in Tab. 7). Thus, UEs experience higher delays. Compared with the Po-UA/HPL model, we obtain a reduction in the total network delay that equals 3.6% in S3, while in S4 and S5, the delay reduction equals 8.1%.

Non-Uniform Distribution of UEs. In this case, the positioning of the UEs is generated in the cell following a Gaussian distribution centered at the BS positioning with a mean 0 m and a standard deviation of 200 m. The simulated results are plotted as a function of the number of UEs per cell (*e.g*, we consider the cases of 6, 8, 10 and 20 UEs per cell), and we only study the performance of setting S3. In Fig. 12, the percentage of power saving decreases as the

number of UEs per cell increases. Precisely, with the increase of the number of UEs per cell, the BS cannot operate at low power level or sleep mode due to the coverage constraints. Moreover, considering the case of six UEs per cell, the percentage of power saving for non-uniform UE distribution (4.3%) is higher than that of the uniform distribution for the same setting S3 (2%, as shown in Fig. 10). This is because, in the former case UEs are located near the BS, the BS can lower its transmit power level.
Fig. 13 shows that the total network delay increases as the number of UEs per cell increases. Precisely, the transmission delay of a UE associated with a given BS is proportional to the number of UEs associated with the same BS (as given in (13)). Moreover, the obtained total network delay is lower than in the Po-UA/HPL case.
The percentage of power saving for both uniform and non-uniform distributions is relatively low. In fact, the power saving depends on the power consumption of sleep mode and on the power consumption of the second transmit power level. Particularly, in the case studied, the former (75 W) represents 42% of the power consumption at the highest power level (177 W), and the latter (153.5 W) represents 86.7% of the power consumption at the highest level.

10.5. Simulation Results for a Rural Deployment

For the rural deployment, we consider an inter-cell distance of 900 m. Moreover, we consider six UEs in each cell, and a total of 54 UEs in the network. The positioning of UEs is generated randomly following a uniform distribution. In this rural deployment, the number of covering BSs per UE equals 1.1. The simulation results show that we obtain no power saving for any studied setting. Moreover, we obtain the same average total network delay (5.1439×10^{-5}) for the considered settings, with a delay reduction of 0.01% compared to Po-UA/HPL. Therefore, in rural environments, where UEs are usually covered by one BS, there is no substantial gain compared with a reference model, and thus power saving and delay reduction are superfluous. Table 8 shows the total network delay for the considered settings and for Po-UA/HPL in a rural environment.

Table 8: Total network delay [s] for the considered settings and for Po-UA/HPL in a rural environment

Po-UA/HPL	Mean	5.2027×10^{-5}
	95% CI	$[5.186\times10^{-5}, 5.218\times10^{-5}]$
Settings S1 to S5	Mean	5.1439×10^{-5}
	95% CI	$[5.1357\times10^{-5}, 5.1522\times10^{-5}]$

10.6. Simulation Results for a Realistic LTE Deployment

We consider the realistic positioning of the LTE network BS for the 14^{th} district of Paris-France [52]. The network topology is composed of 18 cells (N_{bs}=18) and the positioning of UEs follows a random uniform distribution, as shown in Fig. 14. We only study the performance of setting S1 where preference is given to saving power. The

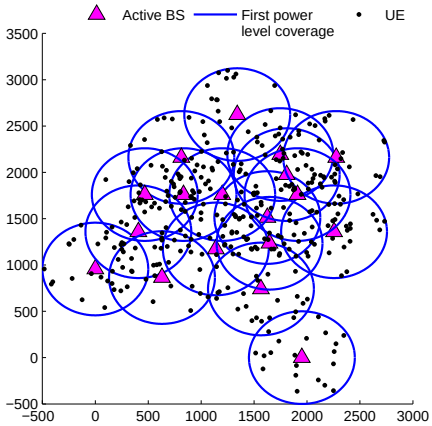


Figure 14: LTE network topology of the 14^{th} district of Paris.

simulated results are plotted as a function of the number of UEs per cell. Due to the large scale test scenario, the memory space limitation and the high computational complexity of the considered problem, we only consider the

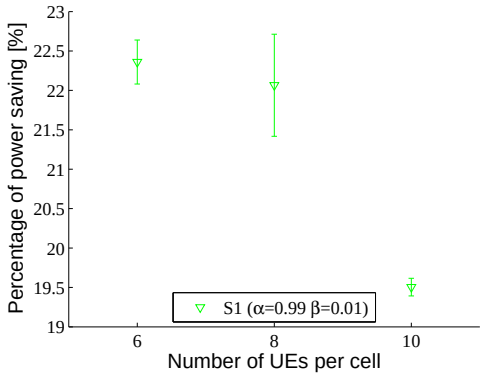


Figure 15: Power saving for setting S1 for uniform UEs distribution in LTE scenario for the 14^{th} district of Paris.

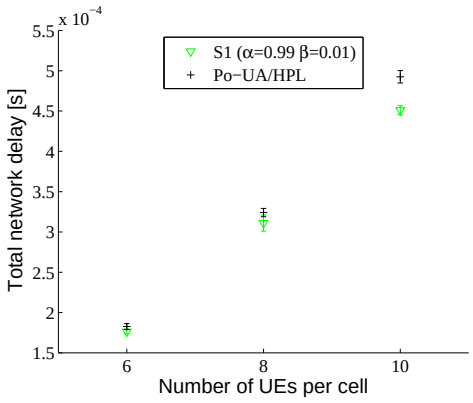


Figure 16: Total network delay for setting S1 and for Po-UA/HPL for uniform UEs distribution in LTE scenario for the 14^{th} district of Paris.

cases of 6, 8, 10 and 20 UEs per cell. Fig. 15 shows that we obtain power saving up to 22.3% compared with legacy solution. Moreover, the percentage of power saving decreases as the number of UEs per cell increases. Precisely, with the increase of the number of UEs per cell, the BS cannot operate at low power level or sleep mode due to the coverage constraints. Fig. 16 shows that the obtained total network delay is close to the Po-UA/HPL case. Moreover, the total network delay increases as the number of UEs per cell increases. Precisely, the transmission delay of a UE associated with a given BS is proportional to the number of UEs associated with the same BS (as given in (13)). We note that the obtained results are similar to the one obtained in the case of urban deployment (as shown in Fig. 12), but with a higher percentage of power saving. In fact, in the present case, we consider setting S1 where preference is given to power saving.

11. Conclusion

In this paper, we considered the joint optimization problem of saving power and minimizing transmission delay in wireless access networks. We formulated a multi-objective optimization problem that consists of finding a trade-off between reducing the network power consumption and selecting the best user association that incurs the lowest transmission delay. Our formulation was studied for IEEE 802.11g and LTE technologies. We presented different approaches to multi-objective optimization and used the weighted sum method to solve the problem. We provided a MILP formulation of the problem to make it computationally tractable. Moreover, different settings reflecting various preferences were carried out by tuning the weights of the power and delay objectives. Simulation results showed that for a power reduction setting, we obtained significant power savings of up to 16% in WLANs compared with legacy solutions. Moreover, our optimization results revealed the impact of the network topology, particularly the inter-cell distance, on both objectives. In an LTE network, we studied the impact of urban, rural and realistic deployments on the achievable tradeoffs. The power savings mainly depend on user distribution and on the power consumption of the sleep mode. Compared with legacy solutions, we obtained power savings of up to 22.3% in a realistic LTE network in Paris-France. Finally, our optimization approach reduced the transmission delay by 6% in a WLAN and 8% in an LTE network when adequately tuned.

The MILP formulation cannot deliver solutions in a reasonable amount of time due to computational complexity issues. In our work [16, 17], we propose heuristic algorithms for the joint Power-Delay minimization problem that overcome such issues. The heuristic algorithms compute satisfactory solutions for the problem while keeping the computation complexity suitably low for practical implementations. Therefore, in the aforementioned papers, a detailed comparison between the analytical models and the heuristics is provided. However, the aim of the present paper is to study only the different optimization models for the joint Power-Delay minimization problem.

For future work, we plan to extend our model to take into account the inter-cell interference as it will make transmitting at the highest power level less attractive. Besides, the expected percentage of power saving and delay

reduction will be more interesting. Another interesting perspective is to study the dynamics of the network and its impact on power saving and delay minimization. In particular, we need to take into consideration the mobility of UEs, the arrival and departure of the UEs in the network.

References

[1] G. Fettweis, E. Zimmermann, ICT energy consumption-trends and challenges, in: Proc. International Symposium on Wireless Personal Multimedia Communications (WPMC), 2008.

[2] E. Oh, B. Krishnamachari, X. Liu, Z. Niu, Toward dynamic energy-efficient operation of cellular network infrastructure, IEEE Communications Magazine 49 (6) (2011) 56–61.

[3] H. Toure, ICTs and Climate Change-the ITU perspective, Climate Action, Sustainable Development International and United Nations Environment Programme, 2008, pp. 116–117, www.climateactionprogramme.org.

[4] F. Richter, A. J. Fehske, G. P. Fettweis, Energy efficiency aspects of base station deployment strategies for cellular networks, in: Proc. IEEE Vehicular Technology Conference Fall (VTC-Fall), 2009.

[5] T. Togo, I. Yoshii, R. Kohno, Dynamic cell-size control according to geographical mobile distribution in a DS/CDMA cellular system, in: IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 1998.

[6] A. Jalali, On cell breathing in CDMA networks, in: Proc. IEEE International Conference on Communications (ICC), 1998.

[7] F. Moety, M. Ibrahim, S. Lahoud, K. Khawam, Distributed heuristic algorithms for rat selection in wireless heterogeneous networks, in: Proc. IEEE Wireless Communications and Networking Conference (WCNC), 2012.

[8] K. Khawam, M. Ibrahim, J. Cohen, S. Lahoud, S. Tohme, Individual vs. global radio resource management in a hybrid broadband network, in: Proc. IEEE International Conference on Communications (ICC), 2011.

[9] Y. Chen, S. Zhang, S. Xu, G. Li, Fundamental trade-offs on green wireless networks, IEEE Communications Magazine 49 (6) (2011) 30–37.

[10] K. Son, H. Kim, Y. Yi, B. Krishnamachari, Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks, IEEE Journal on Selected Areas in Communications 29 (8) (2011) 1525–1536.

[11] S. Tombaz, M. Usman, J. Zander, Energy efficiency improvements through heterogeneous networks in diverse traffic distribution scenarios, in: Proc. IEEE International ICST Conference on Communications and Networking in China (CHINACOM), 2011.

[12] W. Wang, G. Shen, Energy efficiency of heterogeneous cellular network, in: Proc. IEEE Vehicular Technology Conference Fall (VTC-Fall), 2010.

[13] O. Eunsung, B. Krishnamachari, Energy savings through dynamic base station switching in cellular wireless access networks, in: Proc. IEEE Global Telecommunications Conference (GLOBECOM), 2010.

[14] Y. Wu, G. He, S. Zhang, Y. Chen, S. Xu, Energy efficient coverage planning in cellular networks with sleep mode, in: Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), 2013.

[15] M. Heusse, F. Rousseau, G. Berger-Sabbatel, A. Duda, Performance anomaly of 802.11b, in: Proc. IEEE INFOCOM, 2003.

[16] F. Moety, S. Lahoud, B. Cousin, K. Khawam, A heuristic algorithm for joint power-delay minimization in green wireless access networks, in: International Conference on Computing, Networking and Communications (ICNC), Workshop on Computing, Networking and Communications (CNC), 2015.

[17] F. Moety, S. Lahoud, B. Cousin, K. Khawam, Joint power-delay minimization in 4g wireless networks, in: Wireless Days (WD), 2014.

[18] M.-S. Alouini, A. Goldsmith, Area spectral efficiency of cellular mobile radio systems, IEEE Transactions on Vehicular Technology 48 (4) (1999) 1047–1066.

[19] D. Willkomm, S. Machiraju, J. Bolot, A. Wolisz, Primary user behavior in cellular networks and implications for dynamic spectrum access, IEEE Communications Magazine 47 (3) (2009) 88–95.

[20] S. Boiardi, A. Capone, B. Sanso, Radio planning of energy-aware cellular networks, Computer Networks 57 (13) (2013) 2564–2577.

[21] Z. Niu, Y. Wu, J. Gong, Z. Yang, Cell zooming for cost-efficient green cellular networks, IEEE Communications Magazine 48 (11) (2010) 74–79.

[22] P. Bahl, M. T. Hajiaghayi, K. Jain, S. V. Mirrokni, L. Qiu, A. Saberi, Cell breathing in wireless LANs: Algorithms and evaluation, IEEE Transactions on Mobile Computing (TMC) 6 (2) (2007) 164–178.

[23] J. Lorincz, A. Capone, D. Begusic, Optimized network management for energy savings of wireless access networks, Computer Networks 55 (3) (2011) 514 – 540.

[24] A. Yildiz, T. Girici, H. Yanikomeroglu, A pricing based algorithm for cell switching off in green cellular networks, in: Proc. IEEE Vehicular Technology Conference Fall (VTC-Spring), 2013.

[25] A. Bousia, A. Antonopoulos, L. Alonso, C. Verikoukis, Distance-aware base station sleeping algorithm in LTE-Advanced, in: Proc. IEEE International Conference on Communications (ICC), 2012.

[26] M. Hossain, K. Munasinghe, A. Jamalipour, A protocooperation-based sleep-wake architecture for next generation green cellular access networks, in: Proc. International Conference on Signal Processing and Communication Systems (ICSPCS), 2010.

[27] F. Han, Z. Safar, W. Lin, Y. Chen, K. Liu, Energy-efficient cellular network operation via base station cooperation, in: Proc. IEEE International Conference on Communications (ICC), 2012.

[28] A. Conte, A. Feki, L. Chiaraviglio, D. Ciullo, M. Meo, M. Marsan, Cell wilting and blossoming for energy efficiency, IEEE Wireless Communications 18 (5) (2011) 50–57.

[29] S. Videv, J. S. Thompson, H. Haas, P. M. Grant, Resource allocation for energy efficient cellular systems, EURASIP Journal on Wireless Communications and Networking 2012 (2012) 181.

[30] H. Chen, Y. Jiang, J. Xu, H. Hu, Energy-efficient coordinated scheduling mechanism for cellular communication systems with multiple component carriers, IEEE Journal on Selected Areas in Communications 31 (5) (2013) 959–968.

[31] Ericsson mobility report on the pulse of the networked society, in: ERICSSON, Report, 2012.

[32] N. Benameur, S. Ben Fredj, F. Delcoigne, S. Oueslati-Boulahia, J. Roberts, Integrated Admission Control for Streaming and Elastic Traffic, Vol. 2156, Springer Berlin Heidelberg, 2001, pp. 69–81.
[33] S. Methley, Fundamentals of Mesh Technology, Vol. 2156, Cambridge University Press, 2009.
[34] S. Chieochan, E. Hossain, J. Diamond, Channel assignment schemes for infrastructure-based 802.11 WLANs: A survey, IEEE Communications Surveys Tutorials 12 (1) (2010) 124–136.
[35] A. Hills, C. Mellon, Large-scale wireless lan design, IEEE Communications Magazine 39 (11) (2001) 98–107.
[36] A. S. Hamza, S. S. Khalifa, H. S. Hamza, K. Elsayed, A survey on inter-cell interference coordination techniques in ofdma-based cellular networks, IEEE Communications Surveys & Tutorials 15 (4) (2013) 1642–1670.
[37] R. Ghaffar, R. Knopp, Fractional frequency reuse and interference suppression for ofdma networks, in: Proc. IEEE International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2010, pp. 273–277.
[38] S. E. Elayoubi, O. Ben Haddada, B. Fourestie, Performance evaluation of frequency planning schemes in OFDMA-based networks, IEEE Transactions on Wireless Communications 7 (5).
[39] G. Auer, O. Blume, V. Giannini, I. Godor, M. Imran, Y. Jading, E. Katranaras, M. Olsson, D. Sabella, P. Skillermark, W. Wajda, Energy efficiency analysis of the reference systems, areas of improvements and target breakdown, Tech. Rep. D2.3, ALUD, DOCOMO, EAB, ETH, IMEC, TI, UNIS, available online (2010).
[40] F. Moety, S. Lahoud, B. Cousin, K. Khawam, Power-delay tradeoffs in green wireless access networks, in: Proc. IEEE Vehicular Technology Conference Fall (VTC-Fall), 2013.
[41] F. Moety, S. Lahoud, K. Khawam, B. Cousin, Joint power-delay minimization in green wireless access networks, in: Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), 2013.
[42] O. J. Grodzevich, O. Romanko, Normalization and other topics in multi-objective optimization, in: Workshop of the Fields MITACS Industrial Problems, 2006.
[43] The benefits of centralization in WLANs via the cisco unified wireless network, in: Cisco Systems Inc., White paper, 2006.
[44] GLPK Solver, `http://www.gnu.org/software/glpk/`.
[45] D. Ferling, T. Bohn, D. Zeller, P. Frenger, I. Gdor, Y. Jading, W. Tomaselli, Energy efficiency approaches for radio nodes, in: Future Network and Mobile Summit, 2010, pp. 1–9.
[46] Channels and maximum power settings for cisco aironet lightweight access points, in: Cisco Systems Inc., White paper, 2008.
[47] The Network Simulator -NS-2, `http://www.isi.edu/nsnam/ns/`.
[48] CPLEX Optimizer, `http://www.ilog.com/products/cplex/`.
[49] Matlab, `http://http://www.mathworks.fr/`.
[50] 3GPP, Spacial channel model for Multiple Input Multiple Output (MIMO) simulations v6.1.0 (release 6), TR 15.996, 3rd Generation Partnership Project (3GPP) (2003-09).
[51] 3GPP, Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Frequency (RF) system scenarios; v8.0.0 (Release 8), TR 36.942, 3rd Generation Partnership Project (3GPP) (2008-09).
[52] Antennes mobiles, `http://www.antennesmobiles.fr`.