Power-Efficient Resource Allocation in a Heterogeneous Network with Cellular and D2D Capabilities

Jordi Pérez-Romero, *Member, IEEE*, Juan Sánchez-González, *Member, IEEE*, Ramón Agustí, *Member, IEEE*, Beatriz Lorenzo, *Member, IEEE*, and Savo Glisic, *Senior Member, IEEE*

Abstract—This paper focuses on a heterogeneous scenario in which cellular and wireless local area technologies coexist and in which mobile devices are enabled with device-to-device communication capabilities. In this context, this paper assumes a network architecture in which a given user equipment (UE) can receive mobile service either by connecting directly to a cellular base station or by connecting through another UE that acts as an access point and relays the traffic from a cellular base station. The paper investigates the optimization of the connectivity of different UEs with the target to minimize the total transmission power. An optimization framework is presented, and a distributed strategy based on Q-learning and *softmax* decision making is proposed as a means to solve the considered problem with reduced complexity. The proposed strategy is evaluated under different conditions, and it is shown that the strategy achieves a performance very close to the optimum. Moreover, significant transmission power reductions of approximately 40% are obtained with respect to the classical approach, in which all UEs are connected to the cellular infrastructure. For multi-cell scenarios, in which the optimum solution cannot be easily known a priori, the proposed approach is compared against a centralized genetic algorithm. The proposed approach achieves similar performance in terms of total transmitted power, while exhibiting much lower computational requirements.

Index Terms—Power efficient communications, D2D, Q-learning, AP selection

I. INTRODUCTION

WITH the proliferation of bandwidth-intensive applications, user data traffic and the corresponding network load are increasing exponentially. As a result, conventional cellular architectures are facing unprecedented challenges to meet user demands, particularly for users located at cell edges or in indoor positions, where a significant portion of the data traffic is being generated. To provide broadband services with satisfactory user experience in these locations, when conventional cellular architectures are used, an increased link budget is required, leading to larger transmit power consumption at both base stations (BSs) and user equipment (UE). As a result, there has been increasing interest in evolving network architectures, functionalities and technologies to better address these challenges.

In particular, the classical cellular network concept is being shifted towards the so-called heterogeneous networks (HetNets) composed of multiple access technologies, such as cellular and wireless local area networks, and multiple cell layers of different sizes [1][2]. The use of device-to-device (D2D) communications, in which UEs are able to directly communicate, is also envisaged as an important component of these future networks because it opens the door for a number of possibilities, such as proximity services and cellular coverage extension by means of relaying other UEs. [3]. Initiatives in this direction are being conducted by the 3rd Generation Partnership Project (3GPP) in Long Term Evolution (LTE) Release 12 [4] and by the Wi-Fi Alliance, which has recently developed Wi-Fi Direct technology [5], which allows a UE to act as an access point (AP) for other UEs. In this way, different UEs communicate between themselves, and one of them can share its cellular connection with others by relaying their traffic to/from a cellular BS.

There are different taxonomies of D2D use cases [3]-[8]. In [7], the D2D use cases are divided into two categories. The first category is simple D2D communication, in which the sender and receiver exchange data with each other, and in the second category, D2D users act as a relay for the other users. In this paper, we focus on the second category, considering a cellular network where the UEs have the capability to act as APs and relay traffic from the cellular infrastructure to other UEs. In [8], a survey on the multiple D2D use cases is presented. The cases are categorized as in-band D2D, in which the D2D link and the cellular link use the same spectrum, and out-of-band D2D, in which the D2D link and the cellular links use different frequency bands or even different technology

(e.g., LTE for the cellular link and IEEE 802.11 for the D2D link). In this paper, we focus on the out-of-band D2D case, which is identified in [8] as an emerging area of research because the majority of current mobile devices are equipped with more than one wireless interface (e.g., Wi-Fi and LTE), which facilitates the implementation and results in advantages compared to in-band D2D, such as the possibility to have simultaneous communication in the cellular and D2D links and the lack of interference between the two.

In [9], different use cases and scenarios of D2D for further research towards Fifth Generation (5G) networks are identified. D2D applications are split into three groups, one of them being network enhancement based services, in which D2D communications are envisaged to improve connectivity, Quality of Service (QoS) and capacity via activation of the appropriate communication modes (i.e., cellular, direct D2D and relay mode). In this paper, we address the last problem, namely, the selection between the cellular and the relay mode to enhance the network performance. Indeed, due to the shorter distances and associated lower propagation losses in the D2D link, it is expected that the higher bit rates associated with mobile broadband services can be more efficiently achieved (e.g., with less power consumption) than when the UEs at the cell edge connect directly to the BS. In this scenario, given the randomness associated with the propagation in mobile environments, the variability in the generation of data traffic and the mobility of UEs and UEs acting as APs, there will be situations in which it may be more efficient for a certain UE to connect to another UE acting as an AP or to connect directly to the cellular BS, leading to a dynamic network architecture in which the UEs can dynamically change the way they connect to the cellular infrastructure. Consequently, it is crucial to have intelligent decision mechanisms to determine the best connection for each UE. Such decisions need to consider aspects such as the propagation conditions of the different links, the load existing in each macrocell and in each AP, the bit rate requirements of each UE and the total power consumption.

In this context, this paper considers the optimization of the UE connectivity with the objective of minimizing the total transmit power, thus targeting an efficient solution from the perspective of energy consumption. The design of strategies that are efficient in providing the desired wireless services with minimum power consumption is relevant not only from an ecological perspective but can also lead to significant economic benefits. As an example, it is stated in [10] that the energy bill for a mobile operator accounted for approximately 18% of the operational expenditures in a mature European market and increased to 32% in other markets, such as India. Following this trend, several initiatives have addressed research towards energy-efficient wireless communications [11]. Transmit power reduction can also be beneficial from the perspective of health because both users and regulators are concerned about the potential undesirable effects of wireless network radiation on the human body. Different national authorities at the worldwide level have conducted intensive studies in this direction, usually recommending the minimization of exposure to citizens as a precautionary measure [12]-[14].

Based on the above, the main contributions of this work are summarized as follows:

1) A new optimization framework is presented to determine the best connectivity option for each UE in a heterogeneous network with out-of-band D2D capabilities used for relaying data. The objective is to minimize the total power consumption in the scenario while satisfying the bit rate requirement of each UE. The main differences from the classical relay selection due to the consideration of D2D are: (i) the UEs acting as APs may have their own data to transmit, (ii) different frequency bands and technologies are used for the cellular link and the D2D link, and (iii) the UEs acting as APs can exhibit mobility. To the authors' best knowledge, no previous work has addressed the optimization of this use case from the perspective of total power consumption.

2) A new distributed strategy based on Q-learning and *softmax* decision making is proposed as a means to implement the presented optimization framework. In this approach, each UE autonomously decides the most appropriate AP or cellular BS to receive the required service based on its previous experience of using the different APs/BSs. The main advantage of this type of distributed approaches is that it allows for a reduction in complexity in comparison to centralized approaches that address the global optimization by jointly considering all APs and UEs. Therefore, the distributed approach can scale better when increasing the network size.

3) The proposed strategy is evaluated under different conditions, revealing that its performance is very close to the optimum and that it can provide significant power consumption reduction with respect to the classical approach, in which the UEs connect directly to the cellular BSs. The proposed approach is also benchmarked against a centralized genetic algorithm, showing similar performance despite the decentralized operation.

The paper is organized as follows. Section II presents a summary of related work, and Section III elaborates the system model and the proposed optimization framework. Section IV presents the proposed Q-learning based solution for AP/BS selection, which is evaluated in Section V. Finally, conclusions are summarized in Section VI.

II. RELATED WORK

Multi-hop cellular networks (MCN) [15], in which the traffic of a UE is relayed to a cellular infrastructure node by means of intermediate relay stations, have received significant interest in recent years as a means to enhance the capacity, data rates and coverage of cellular networks. For example, architectural aspects and routing protocols were studied in [16]-[19], and different relay selection schemes were recently proposed in [20]-[22].

The focus of this work is on the out-of-band relaying D2D use case in which the cellular link and the relay link make use of different technologies. In this respect, in [23], the relay

selection probability is analyzed in the uplink of an LTE-based, multi-hop cellular network with out-of-band relaying. It accounts for the intercell interference in the cellular network, as well as for the fact that both the cellular link and the relay link using IEEE 802.11 can limit the capacity. The method assumes, however, a regular channel allocation of IEEE 802.11 channels to relays in different hexagonal cells that may not be realistic because the deployment of Wi-Fi access points is usually highly irregular. In [24], a network selection scheme is considered in a heterogeneous scenario with LTE and Wi-Fi APs that accounts for the backhaul capacity for each AP. The considered scenario assumes APs deployed at specific positions, in contrast to this work, which assumes that the UEs can act as APs and relay the traffic of other UEs towards the LTE network. Another important difference of this paper with respect to prior studies is that we assume that a UE acting as an AP and relaying data to other UEs may also have its own information to be transmitted, whereas previous works usually assume that a UE can only act as a relay when it does not have its own data to be transmitted. In [25][26], the combination of LTE-A with D2D communications is explored for the provision of multicast services, analyzing the potentialities in terms of energy consumption. Similarly, in [27], the use of Wi-Fi in conjunction with LTE is studied for the provision of in-car communications. A simulation analysis is presented to show that this approach can provide higher bit rates than direct connection to the LTE network. In [28], the so-called user-provided networks are considered, in which mobile hosts with 3G/4G connections are incentivized to forward data for others. Whereas that scenario is similar to the one considered in this paper, the focus of [28] is placed on the incentive mechanisms and not on the optimization of the connectivity options to minimize the transmitted power. Finally, in [29], the opportunistic coverage extension of a cellular network was analyzed to provide service to UEs outside of the direct coverage area of the cellular BS by means of relaying, and a learning-based approach was used to select both the spectrum of the D2D link and the node that provides the coverage extension.

There are also a number of works that have considered different approaches for AP selection in wireless local access networks (WLAN). Apart from the classical approach, in which the AP is selected based on signal strength, different studies have proposed other metrics to achieve a more efficient AP association, such as the packet error rate, the throughput and the bandwidth per user [30]-[32]. Other approaches, such as [33], consider the load balancing problem under max-min fairness considerations, whereas in [34]-[37], game theory concepts are considered for the association of UEs to APs. However, none of the above works assumes the scenario in which the APs can be used to relay traffic from the UEs to the cellular infrastructure, as considered in this work.

III. System Model and Problem Formulation

A. System model

The scenario considered in this work is represented in Fig. 1.

It assumes a cellular network where each UE (e.g., current smartphones) can be turned into an AP and can be used to provide wireless connectivity to other UEs. Let us consider J macrocell BSs denoted as the set $\beta=\{S_1,...,S_J\}$ with cellular technology (e.g., LTE or LTE-A), K UEs acting as APs denoted as the set $\Lambda=\{A_1,....A_K\}$ and N UEs not acting as APs denoted as U=$\{u_1,...,u_N\}$. In the following, the UEs of set Λ will be referred to simply as "APs", whereas those of set U will be referred to as "UEs". The bit rate requirement of UE u_n is R_n. To achieve that bit rate, UE u_n must connect to one BS in set β or one of the APs in set Λ. In this respect, the purpose of this work is to perform an efficient selection of the AP or BS for each UE because this selection will impact the total radio resource consumption. This work assumes communication in the downlink direction, i.e., from the BSs/APs to the UEs, although it could easily be extended to consider the uplink direction.

The BS/AP selection process is executed at a time scale where all short-term effects, such as the frequency-selective fast fading, have been averaged. This time scale prevents the UE from continuously changing the BS or AP due to random channel variations that occur on a very short time scale.

Each AP has the capability to provide wireless Internet access to other UEs. In general, different possibilities exist for this access. This work assumes that a UE acting as an AP is also connected to the cellular infrastructure, so that the AP relays the data traffic of a BS to the UE using a 2-hop approach. Then, in the example of Fig. 1, UE u_1 has different possibilities for obtaining service: direct connection to S_2, 2-hop connection $S_2 \rightarrow A_2 \rightarrow u_1$, 2-hop connection $S_2 \rightarrow A_3 \rightarrow u_1$ and 2-hop connection $S_1 \rightarrow A_1 \rightarrow u_1$. In contrast, UE u_2, out of the coverage area of the macrocells, has only one possibility: $S_2 \rightarrow A_2 \rightarrow u_2$. This work could be easily extended to consider other possibilities for providing access through the APs, for example, in the case that the APs are fixed and have wired connection to the Internet (as in the Dynamic Network Architecture proposed in [38]), in which case no relaying would be needed.

For the APs, it is considered that, in addition to relaying the traffic of other UEs, they may have their own service requirements. Then, the bit rate requirement of AP A_k is $R_{A,k}$.

B. Macrocell link

It is assumed that the J BSs operate with the same LTE carrier composed of M resource blocks (RB) of bandwidth B that can be assigned to the N UEs (if directly connected to one of the BSs) or to one of the K APs. As previously mentioned, BS/AP selection is executed after having averaged the short-term effects (e.g., frequency-selective fast fading), so we are only concerned with the average number of RBs required by each UE or AP to achieve their desired bit rate in a given BS, but not with modelling the scheduling process that will decide which specific RBs are allocated to each UE/AP in the short-term. Then, the average number of RBs required at BS S_j ($j=1,...,J$) to serve its UEs can be expressed as
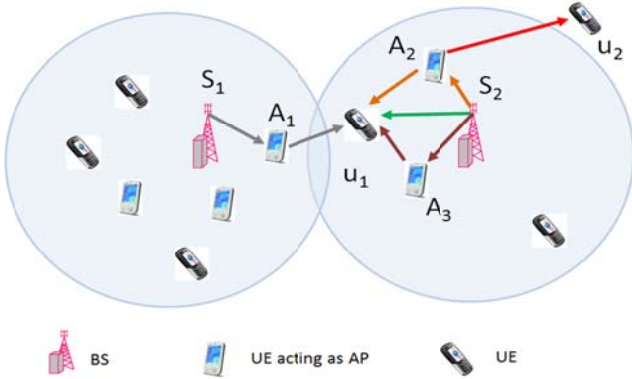
Fig. 1. Considered scenario: Heterogeneous network with D2D capabilities.

$$M_{req,j} = \sum_{n=1}^{N} c_{n,j} \cdot \frac{R_n}{r_{U,n,j}} + \sum_{k=1}^{K} a_{k,j} \cdot \frac{R_{A,k} + \sum_{n=1}^{N} b_{k,n} \cdot R_n}{r_{A,k,j}} \qquad (1)$$

where $c_{n,j}$ is a binary indicator that takes the value 1 if UE u_n is connected to BS S_j and 0 otherwise. Similarly, $a_{k,j}$ takes the value 1 if AP A_k is connected to BS S_j and 0 otherwise, and $b_{k,n}$ takes the value 1 if UE u_n is connected to AP A_k and 0 otherwise. $r_{U,n,j}$ is the capacity that UE u_n can obtain when allocated to one RB of BS S_j, and $r_{A,k,j}$ is the capacity that AP A_k can obtain in one RB when connected to BS S_j. The subscript U in the notation $r_{U,n,j}$ reflects that it is the capacity achieved by a UE when directly connected to a BS, and the subscripts n, j represent the number of the UE and the BS, respectively. Similarly, the subscript A in $r_{A,k,j}$ reflects that it is the capacity achieved by an AP when connected to a BS, and the subscripts k and j represent the number of the AP and the BS, respectively. The first term in (1) corresponds to the average number of RBs required by the UEs that are connected directly to BS S_j to achieve their bit rate requirements R_n, whereas the second term corresponds to the number of RBs required by the APs connected to BS S_j to achieve their own requirements $R_{A,k}$ plus those of the UEs they are serving.

The transmit power per RB at BS S_j is $P_{RB,j}$. Assuming the Shannon bound, the capacity per RB in the link between UE u_n and BS S_j, $r_{U,n,j}$, can be estimated as

$$r_{U,n,j} = B \log_2 \left(1 + \frac{P_{RB,j} / L_{U,n,j}}{P_N + I_{U,n,j}} \right) \qquad (2)$$

where $L_{U,n,j}$ is the propagation loss between UE u_n and BS S_j that includes the distance-dependent losses and the slow fading (shadowing). $P_N = N_o \cdot B$ is the noise power measured over the bandwidth B with N_o as the noise power spectral density. $I_{U,n,j}$ represents the average intercell interference per RB measured at UE u_n if connected at BS S_j, given by

$$I_{U,n,j} = \sum_{\substack{j'=1 \\ j' \neq j}}^{J} \frac{P_{RB,j'}}{L_{U,n,j'}} \cdot \frac{M_{req,j'}}{M} \qquad (3)$$

where $M_{req,j'}/M$ is the fraction of time that an RB will be utilized on average by BS $S_{j'}$.

Similarly, the capacity per RB in the link between AP A_k and BS S_j, $r_{A,k,j}$, is given by

$$r_{A,k,j} = B \log_2 \left(1 + \frac{P_{RB,j} / L_{A,k,j}}{P_N + I_{A,k,j}} \right) \qquad (4)$$

where $L_{A,k,j}$ is the propagation loss between AP A_k and BS S_j, including the distance-dependent losses and shadowing, and $I_{A,k,j}$ is the average intercell interference observed at AP A_k if connected at BS S_j, given by

$$I_{A,k,j} = \sum_{\substack{j'=1 \\ j' \neq j}}^{J} \frac{P_{RB,j'}}{L_{A,k,j'}} \cdot \frac{M_{req,j'}}{M} \qquad (5)$$

C. Device to Device (D2D) link

The communication between an AP of set Λ and a UE of set U makes use of a device-to-device (D2D) technology[1]. The D2D link between the AP and the UE is assumed to have bandwidth B_A that is shared on the time domain between the UEs connected to the AP, e.g., by means of a scheduling algorithm or a medium access control (MAC) protocol. It is also assumed that the APs can simultaneously use the D2D interface and the cellular interface to connect with other UEs and with the infrastructure, respectively, and that both interfaces operate at different frequency bands so that no mutual interference exists. Correspondingly, when the APs relay data from the infrastructure to other UEs, full duplex relay is assumed. It is assumed that a control mechanism exists at the APs so that UE u_n receives at most its bit rate requirement R_n. Then, if the achievable bit rate $r_{D,k,n}$ (where subscript D,k,n denotes the D2D link between AP A_k and UE u_n) is higher than the requirement R_n, the AP will only transmit data for this UE during the fraction of time $\theta_{k,n}$ given by

$$\theta_{k,n} = \min \left(1, \frac{R_n}{r_{D,k,n}} \right). \qquad (6)$$

In this way, the average power transmitted by AP A_k to provide service to UE u_n would be $P_A \cdot \theta_{k,n}$, where P_A is the transmit power of AP A_k, assumed to be constant (and equal for all access points), which is the usual approach in current implementations of Wi-Fi systems that do not apply dynamic transmit power control. In this way, the AP will spend only the minimum power needed to provide the UE with its bit rate requirement. In the case $r_{D,k,n} < R_n$, UE u_n cannot obtain its required bit rate through AP A_k.

Under the above considerations, the total fraction of time

---

[1] We assume the work at this stage to be technology-agnostic. In practice, there could be different possibilities, such as IEEE 802.11, LTE D2D, etc.

that AP A_k is active is denoted as Θ_k and is given by

$$\Theta_k = \sum_{n=1}^{N} b_{k,n} \theta_{k,n}. \tag{7}$$

The criterion of $\Theta_k \leq 1$ should be fulfilled so that all UEs connected to AP A_k are able to reach their bit rate requirements.

The achievable bit rate $r_{D,k,n}$ in the link between AP A_k and UE u_n is given by

$$r_{D,k,n} = B_A \log_2 \left(1 + \frac{P_A / L_{D,k,n}}{P_{N,A} + I_{D,k,n}} \right) \tag{8}$$

where $L_{D,k,n}$ is the average propagation loss between UE u_n and AP A_k, $P_{N,A} = N_o \cdot B_A$ is the noise power at the UE and $I_{D,k,n}$ is the average interference observed at UE u_n coming from the rest of APs $A_{k'}$ that work at the same frequency as AP A_k. It is given by

$$I_{D,k,n} = \sum_{\substack{k'=1 \\ k' \neq k}}^{K} \frac{P_A}{L_{D,k',n}} \cdot \Theta_{k'} \cdot F_{k,k'} \tag{9}$$

where $F_{k,k'}$ is a binary indicator that takes the value 1 if AP A_k operates in the same frequency as $A_{k'}$ and 0 otherwise. The criterion to decide which frequencies are assigned to each AP is out of the scope of this paper, so $F_{k,k'}$ is assumed to be an input.

D. Problem formulation

The possibility of using APs to relay traffic to UEs is intended to achieve a more efficient resource usage and a reduction in the total transmit power in comparison with the case when the UEs are directly connected to the BSs. In particular, UEs with very high bit rate requirements located at the edge of a macrocell require a large amount of RBs and, correspondingly, a large total power if connected directly to the BS. In contrast, if connected through another AP with better propagation conditions to the BS, this may lead to less RBs/power allocated in the BS for the same bit rate requirement at the expense of some additional power transmitted in the link with the AP. Clearly, a trade-off will exist between the usage of resources in the macrocells and the usage of resources in the D2D links, which leads to an optimization problem to identify the best way to associate the UEs with the different BSs/APs, i.e., to find the optimum values of the binary indicators $c_{n,j}$ and $b_{k,n}$ defined in (1).

The focus of this work is on the selection of the BS/AP by the UEs, not on the selection of the BS by the APs. In this respect, the values of $a_{k,j}$, which specify the connections between APs and BSs, are obtained by assuming that each AP is connected to the BS with the lowest propagation losses.

The target for the optimization is to minimize the total average transmitted power. From the perspective of green

communications, total power is considered to be the relevant metric because the power of both the BSs and APs is generated from the electrical grid, so both transmit powers contribute to the CO_2 footprint. The total transmitted power is given by

$$P_{TOT} = \sum_{j=1}^{J} P_{RB,j} \cdot M_{req,j} + \sum_{k=1}^{K} P_A \cdot \Theta_k. \tag{10}$$

The first term is the total power transmitted by the BSs, expressed in terms of the average number of required RBs, $M_{req,j}$, whereas the second term is the total power transmitted by all APs. Therefore, the considered optimization problem can be formulated as follows:

$$\min_{b_{k,n},c_{n,j}} P_{TOT} = \\ = \min_{b_{k,n},c_{n,j}} \left(\sum_{j=1}^{J} P_{RB,j} \cdot M_{req,j} \left(b_{k,n}, c_{n,j} \right) + \sum_{k=1}^{K} P_A \cdot \Theta_k \left(b_{k,n} \right) \right) \tag{11}$$

subject to the following constraints:

$$\sum_{k=1}^{K} b_{k,n} + \sum_{j=1}^{J} c_{n,j} \leq 1 \qquad n=1,\ldots,N \tag{12}$$

$$M_{req,j} \leq M \qquad\qquad j=1,\ldots,J \tag{13}$$

$$r_{D,k,n} \geq b_{k,n} R_n \qquad n=1,\ldots,N, \quad k=1,\ldots,K \tag{14}$$

$$\Theta_k = \sum_{n=1}^{N} b_{k,n} \theta_{k,n} \leq 1 \quad k=1,\ldots,K \tag{15}$$

Constraint (12) reflects the fact that UE u_n can only be connected to one AP/BS, so at most, one of the values of $b_{k,n}$ and $c_{n,j}$ should equal 1. All values of $b_{k,n}$ and $c_{n,j}$ could be 0 if there is no possibility of connection for UE u_n. In turn, the total number of RBs required by a BS, $M_{req,j}$, should be less than the number of available RBs, M, as represented in constraint (13). Constraint (14) reflects that UE u_n can only be connected to an AP A_k in which the available bit rate in the D2D link, $r_{D,k,n}$, is higher than or equal to the required bit rate R_n. Finally, the resource sharing of all UEs connected to AP A_k to allow all of them to receive their required bit rate, as stated in constraint (15).

The fulfilment of constraints (13)-(15) ensures that all the UEs and APs are served with their required bit rates R_n and $R_{A,k}$, respectively. However, depending on the number of available RBs M and the specific propagation conditions, it is possible that no solution exists that fulfils all of the considered conditions. In such a case, either some UEs should not be admitted to the system or their achieved bit rate will be below the minimum requirements.

The problem formulated by (11) and (12)-(15) is a binary nonlinear optimization problem. It falls within the category of integer programming, which is known to be NP-hard [39].

IV. Solution for Distributed AP/BS Selection

The NP-hard problem presented in the previous section can be solved by different methods (e.g., branch and bound and genetic algorithms). However, these solutions would require simultaneously considering all UEs, APs and BSs at a given time and performing the optimization in a centralized way. This method would lead to high complexity as the number of UEs/APs/BSs increases. Moreover, in addition to being nonlinear, the objective function (11) cannot be expressed in a closed form as a function of binary variables $b_{k,n}$ and $c_{n,j}$ because the term $M_{req,j}(b_{k,n},c_{n,j})$ for BS S_j depends on the values of $M_{req,j'}(b_{k,n},c_{n,j'})$ for the BSs other than S_j, which captures the mutual interference existing between BSs, as can be seen in relations (1)-(3). The same occurs for the term $\Theta_k(b_{k,n})$ corresponding to AP A_k. Due to these coupling effects between variables, additional complexity arises when having to compute the total power (11) for a given combination of input variables $b_{k,n}$ and $c_{n,j}$ because it involves iterative numerical analysis.

To overcome the above limitations, in the following, a distributed approach is proposed, in which the different UEs autonomously select the AP or BS that they will be connected to. The main advantage of using distributed approaches is that they allow for a reduction in complexity and signaling overhead because each UE needs to consider only its own selection possibilities. Moreover, to avoid the abovementioned additional complexity needed to explicitly compute the total power according to (11), the considered distributed approaches are based on actual measurements performed by UEs when connected to the different APs/BSs.

The proposed distributed approach is based on Q-learning [40]. Each UE u_n keeps a record of its experience when using each of the APs A_k $k=1,...,K$ stored in a value $Q_{AP,n}(k)$ and each of the BSs S_j $j=1,...,J$ stored in a value $Q_{BS,n}(j)$. Whenever an AP A_k or a BS S_j has been used by UE u_n, the value of $Q_{AP,n}(k)$ or $Q_{BS,n}(j)$, respectively, is updated following a single-state, Q-learning approach with null discount rate given by

$$Q_{AP,n}(k) \leftarrow (1-\alpha)Q_{AP,n}(k) + \alpha \cdot W_{AP,n}(k) \qquad (16)$$

$$Q_{BS,n}(j) \leftarrow (1-\alpha)Q_{BS,n}(j) + \alpha \cdot W_{BS,n}(j) \qquad (17)$$

where $\alpha \in (0,1)$ is the learning rate and $W_{AP,n}(k)$ and $W_{BS,n}(j)$ are the rewards resulting from the use of AP A_k or BS S_j, respectively. The rewards $W_{AP,n}$ and $W_{BS,n}$ reflect the degree of fulfillment of the optimization target as well as the different constraints. In that respect, if we consider that the target to minimize in (11) is the total transmit power by the BSs and the APs to ensure the UE bit rate requirements R_n, the reward will be based on the total power and on the actual achieved bit rate $\hat{r}_n$ since the last AP/BS selection. In this way, those APs/BSs that lead to lower power consumption levels provide larger rewards and correspondingly larger values of $Q_{AP,n}(k)$ or $Q_{BS,n}(j)$. The computation of the reward is detailed in the following sub-sections.

A. Reward computation

1) Reward $W_{AP,n}(k)$ when a UE is connected through an AP

In this case, considering that AP A_k is connected to BS S_j, the total power needed to serve UE u_n results from two contributions:
- $P_A \cdot \theta_{k,n}$ is the total power of AP A_k devoted to serving UE u_n.
- $P_{k,n,j}$ is the power of BS S_j devoted to delivering the traffic of UE u_n through the link between BS S_j and AP A_k. It is given by

$$P_{k,n,j} = P_{RB,j} \cdot M_{k,n,j} \qquad (18)$$

where $M_{k,n,j}$ is the number of RBs in BS S_j required by UE u_n when connected through AP A_k given by

$$M_{k,n,j} = \frac{R_n}{r_{A,k,j}}. \qquad (19)$$

Based on the above discussion, the reward function when UE u_n is connected to AP A_k is defined as

$$W_{AP,n}(k) = \begin{cases} 0 & \text{if } \hat{r}_n < R_n \\ 1 - \dfrac{P_A \cdot \theta_{k,n} + P_{k,n,j}}{P_A + P_{max,j}} & \text{otherwise} \end{cases} \qquad (20)$$

where $P_{max,j}$ is the maximum power of BS S_j given by $P_{max,j}=M \cdot P_{RB,j}$. Note that (20) assigns a value of 0 whenever the achieved bit rate (i.e., measured bit rate) during the connection $\hat{r}_n$ is below the requirement R_n. In contrast, if the service requirement has been successfully fulfilled, the reward is a value between 0 and 1 that decreases when the required power consumption increases. The condition $\hat{r}_n < R_n$ when the UE is not getting its required bit rate can occur for three different reasons: (i) a lack of RBs at BS S_j to provide the service through AP A_k, meaning that constraint (13) is not fulfilled. (ii) The propagation conditions in the D2D link do not allow achieving R_n, meaning that constraint (14) is not fulfilled. (iii) There is an excessive load in AP A_k, meaning that constraint (15) is not fulfilled. Consequently, the formulation of the reward function in (20) takes into account the constraints of the optimization problem.

2) Reward $W_{BS,n}(j)$ when a UE is connected through a BS

In this case, the total power consumption is in the BS. By making similar considerations as before, the transmitted power from BS S_j devoted to UE u_n is given by

$$P_{n,j} = P_{RB,j} \cdot M_{n,j} \qquad (21)$$

where $M_{n,j}$ is the number of RBs that BS S_j would need to serve the requirements of UE u_n and is given by

$$M_{n,j} = \frac{R_n}{r_{U,n,j}}. \qquad (22)$$

Based on this, the reward function when the UE is connected to BS S_j is defined as

$$W_{BS,n}(j) = \begin{cases} 0 & \text{if} \quad \hat{r}_n < R_n \\ 1 - \dfrac{P_{n,j}}{P_{\max,j}} & \text{otherwise} \end{cases}. \quad (23)$$

B. Computation of the $Q_{AP,n}(k)$ and $Q_{BS,n}(j)$ values at initialization

At initialization, i.e., when AP A_k or BS S_j have not been previously used by UE u_n, the values of $Q_{AP,n}(k)$ and $Q_{BS,n}(j)$ can be computed using expressions similar to the reward (20), (23), but replacing the first condition (because there is no measured value of $\hat{r}_n$), as explained in the following.

For the case of an AP, the initial value of $Q_{AP,n}(k)$ is defined as

$$Q_{AP,n,initial}(k) = \begin{cases} 0 & \text{if} \quad (\theta_{k,n} > 1) \ OR \ (M_{k,n,j} > M) \\ 1 - \dfrac{P_A \cdot \theta_{k,n} + P_{k,n,j}}{P_A + P_{\max,j}} & \text{otherwise} \end{cases}. \quad (24)$$

The first condition in (24) reflects the case that AP A_k is not appropriate to serve UE u_n because the propagation conditions in the link between the AP and the UE are not able to provide the service requirements (i.e., $\theta_{k,n} > 1$) or because the link between the AP and the BS would require more RBs to provide the service than are available (i.e., $M_{k,n,j} > M$).

Similarly, for the case of a BS, the initial value of $Q_{BS}(j)$ is given by

$$Q_{BS,n,initial}(j) = \begin{cases} 0 & \text{if} \quad (M_{n,j} > M) \\ 1 - \dfrac{P_{n,j}}{P_{\max,j}} & \text{otherwise} \end{cases}. \quad (25)$$

C. Selection criterion

At the time when UE u_n needs to select an AP/BS for receiving service, it uses the available values of $Q_{AP,n}(k)$ and $Q_{BS,n}(j)$ to apply a *softmax* selection policy [40], in which AP A_k or BS S_j is randomly selected with probabilities $Pr_{AP}(k,n)$ and $Pr_{BS}(j,n)$, respectively, defined as

$$Pr_{AP}(k,n) = \frac{e^{\frac{Q_{AP,n}(k)}{\tau}}}{\sum_{k'=1}^{K} e^{\frac{Q_{AP,n}(k')}{\tau}} + \sum_{j'=1}^{J} e^{\frac{Q_{BS,n}(j')}{\tau}}} \quad (26)$$

$$Pr_{BS}(j,n) = \frac{e^{\frac{Q_{BS,n}(j)}{\tau}}}{\sum_{k'=1}^{K} e^{\frac{Q_{AP,n}(k')}{\tau}} + \sum_{j'=1}^{J} e^{\frac{Q_{BS,n}(j')}{\tau}}} \quad (27)$$

where τ is the so-called temperature parameter. High temperature causes the different options to be nearly equiprobable. In contrast, low temperature leads to a greater difference in selection probability for APs/BSs that differ in their Q value estimates, and the higher the value of Q, the higher the probability of selecting the corresponding AP/BS. *Softmax* decision making is a common means of balancing the exploitation and exploration dilemma in reinforcement learning-based schemes [40]. *Softmax* decision making exploits what the UE already knows to obtain a reward (i.e., selecting APs/BSs that have provided good results in the past), but it also explores ways to take better actions in the future (i.e., the selection must try first a variety of combinations and progressively favor those that appear to be the best ones) [40].

To facilitate the algorithm convergence as the best actions are being identified by the algorithm, a cooling function is also considered in this paper to reduce the value of the temperature τ as time passes. Specifically, the following logarithmic cooling function is considered:

$$\tau = \frac{\tau_0}{\log_2(1+t)} \quad (28)$$

where τ_0 is the initial temperature, and t is the time elapsed since the UE made the first selection.

D. Admission control

Given that the AP/BS selection is performed by the UE, the load in the selected node may already be too high to support the new UE. Consequently, an admission control is used at the selected node to ensure that the number of connected UEs is sufficiently low to ensure that the required bit rates can be provided. This factor is captured in the constraint (13) for the BSs and (15) for the APs. Then, when a UE attempts to connect to BS S_j, if the resulting value of $M_{req,j}$ after including the new UE is higher than M (or in general than a certain threshold), the new UE is not admitted to this BS. When the UE attempts to connect to AP A_k, if the resulting value of Θ_k after including the new UE is higher than 1 (or in general than a certain threshold), the new UE is not admitted to this AP. When this occurs, the reward for the selected AP/BS is set to 0, and another node is selected.

E. Implementation and complexity considerations

Although a detailed analysis of the implementation of the proposed approach for specific technologies is out of the scope of this paper and is left for future work, in this section, we present some high level considerations on how this implementation could be addressed, emphasizing the practical feasibility of the proposed approach.

Each UE needs to store the values $Q_{AP,n}(k)$ and $Q_{BS,n}(j)$ of the candidate BSs and APs and to update them based on the obtained reward each time it is connected to a BS or AP. When the UE has been connected directly to a BS, the UE computes the reward based on (23), which requires that the UE measures the achieved bit rate $\hat{r}_n$ and the power transmitted by the BS, $P_{n,j}$. This power can be calculated by the UE from (21) by measuring the average number of RBs $M_{n,j}$ that the BS has devoted to it. In addition, the power per RB $P_{RB,j}$ and the power available at the BS $P_{\max,j}$ can be sent

by the BS through broadcast channels. When the UE has been connected through an AP, the UE computes the reward from (20). This requires that the UE measures the achieved bit rate $\hat{r}_n$, the power devoted by the D2D link ($P_A \cdot \theta_{k,n}$) and the power devoted by the BS ($P_{k,n,j}$). To obtain $\theta_{k,n}$, the UE can measure the fraction of time that it has received information from the AP. In turn, the power $P_{k,n,j}$ can be determine by the AP using (18) after measuring the number of RBs that the BS has delivered to the UE, $M_{k,n,j}$, and using the powers $P_{RB,j}$ and $P_{max,j}$ from the BS broadcast channels. With this information, the AP can deliver to the UE the values of P_A, $P_{k,n,j}$ and $P_{max,j}$ using a dedicated control signaling message that depends on the technology used for the D2D communication (e.g., Wi-Fi direct). This message will be sent when the reward has to be computed at the end of the data transmission. Alternatively, another implementation option could be that the AP directly computes the reward and sends it to the UE. The initial Q values from (24) and (25) can be obtained following a similar approach. The difference is that the values of $M_{k,n,j}$, $\theta_{k,n}$ and $M_{n,j}$ cannot be directly measured from the actual data transmissions but are estimated from the bit rate requirements R_n and expressions (2)(4)(8) using signal-to-noise-and-interference measurements.

Existing D2D technologies, such as Wi-Fi direct and LTE D2D, already include control messages and procedures that would support the required signaling between UEs and APs. For instance, in the case of Wi-Fi direct [5], there are discovery procedures to facilitate the identification of the UEs acting as APs, and there are probe requests/responses, beacons and association requests/responses through which the required information to compute the reward could be exchanged.

In terms of complexity, the Q-learning approach requires that each UE performs the following operations. First, to update the Q-values for each AP/BS following (16)(17), it requires 2 products and a summation per AP/BS. Second, at the time of selecting the AP/BS, the UE has to compute the probabilities (26)(27) for all BS/APs, which requires $J+K$ exponential functions, $J+K$ summations and $J+K$ divisions. Therefore, the amount of required operations is considered to be very affordable.

V. PERFORMANCE EVALUATION

The performance evaluation of the proposed approach by means of simulations has been carried out in the two baseline scenarios illustrated in Fig. 2. Scenario 1 in Fig. 2(a) assumes a single BS S_1 located in the upper left corner of a square area of 400 m x 400 m. In turn, scenario 2 in Fig. 2(b) is a multi-cell scenario configured initially with $J=3$ BSs deployed in an area of 1000 m x 1000 m. Different positions of the UEs and APs are considered in the simulations, as well as different numbers of UEs and APs. The rationality for the choice of these two scenarios is the following. Scenario 1 is a simple case that allows for the computation of the optimum solution by performing an exhaustive search among all possible combinations of $b_{k,n}$ and $c_{n,j}$. Therefore, scenario 1 allows direct comparison between the proposed algorithm and the
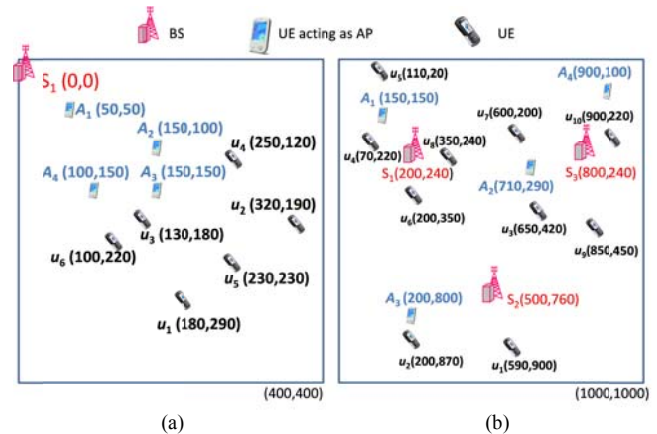

Fig. 2. BS, AP and UE locations in: (a) Scenario 1 (b) Scenario 2. Coordinates are measured in meters.

optimum solution. In turn, scenario 2 is a more realistic multi-cell case, where the choice of $J=3$ BSs has been selected as a reasonable number of BSs that a UE can detect as candidate cells to receive service in a typical macrocell deployment. However, in this scenario, the total number of possible combinations increases dramatically (e.g., up to $3.67 \cdot 10^{70}$ for $J=3$, $K=12$ and $N=60$), so it is not possible to test them all to obtain the optimum solution. For this reason, the proposed algorithm is compared for benchmarking purposes against a genetic algorithm, which is a well-known heuristic search method used to locate near-optimal solutions in complex problems, such as the one considered here. Then, in this case, the genetic algorithm, which operates with full knowledge of all APs/BSs/UEs at each time instant, is taken as a near-optimal performance bound for the proposed decentralized approach, in which each UE makes its own decisions.

The following general propagation model is assumed for computing the propagation losses between the UEs/APs and the BS (i.e., $L_{U,n,j}$, $L_{A,k,j}$) and in the D2D links between the UEs and the APs (i.e., $L_{D,k,n}$):

$$L(dB) = K_p + \beta_p \log f(GHz) + \alpha_p \log d(km) + S \quad (29)$$

Based on [41] and references therein, the considered parameters in (29) are K_p=122.1 dB, β_p=21, and α_p=37.6. Moreover, f=2.6 GHz is used for the propagation loss between the BSs and the UEs or APs and f=2.4 GHz for the propagation loss between the APs and the UEs. S(dB) is the shadowing, which follows a Gaussian distribution with mean 0 and standard deviation σ=6 dB. Spatially correlated shadowing is considered with exponential autocorrelation and decorrelation distance d_{corr}=10 m. The shadowing of the links BS-AP and AP-UE are assumed to be independent.

In scenario 1, it is assumed that the different APs work at different frequencies (i.e., $F_{k,k'}$=0 for all k and k') so that there is no interference in the D2D links. In scenario 2, both the case where all APs work at the same frequency and the case where all APs work at different frequencies are analyzed.

The BSs have M=25 RBs of bandwidth B=180 kHz. The transmit power per RB for all BSs is $P_{RB,j}$=29 dBm. The APs have bandwidth B_A=20 MHz and transmit power P_A=20 dBm.
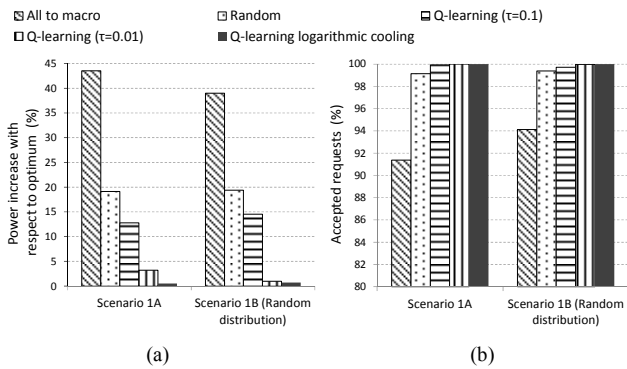
Fig. 3. (a) Increase of the total transmitted power with respect to the optimum solution. (b) Fraction of accepted requests by admission control for the different strategies.

The noise power spectral density is N_o=-164 dBm/Hz, and the required bit rates by the different UEs and APs are R_n=5 Mb/s and $R_{A,k}$=5 Mb/s. The simulation time is measured in generic units denoted as "time steps" that specify when the different events of the simulation occur (e.g., UEs generating and finalizing activity periods). In this way, the results are applicable to different times simply by mapping the time step to a specific time unit.

The UEs and the UEs acting as APs generate activity periods whose duration is geometrically distributed with an average of 30 time steps. At the beginning of each activity period, the UE performs the AP/BS selection process explained in Section IV.C, and at the end of the period, it updates the Q values based on (16)(17). Unless otherwise stated, the time between the end of one activity period and the beginning of the next one is also geometrically distributed with average of 30 time steps.

Each simulation experiment is run for a total of 10000 time steps. The Q-learning algorithm has been configured with learning rate α=0.1, and different values of the temperature parameter τ are analyzed, including the logarithmic cooling given by (28).

A. Performance in terms of transmit power consumption

To gain first insight into the behavior of the proposed strategy, the single cell scenario of Fig. 2(a) is considered, with *K*=4 APs and *N*=6 UEs. In this single cell scenario, with a reduced number of APs and UEs, the proposed algorithm is benchmarked against the optimum solution to the problem obtained by testing all possible combinations. Two variants of the scenario are considered. Scenario 1A considers that the *K*=4 APs and the *N*=6 UEs are static and located at the positions indicated in Fig. 2(a). In turn, to obtain the global performance for different positions of the UEs, scenario 1B considers the case where the positions of the *N*=6 UEs are randomly varied. Specifically, 200 runs of experiments are executed corresponding to different uniform random distributions of the UEs' positions in scenario 1B.

Fig. 3 presents an evaluation of the proposed Q-learning approach for both scenarios 1A and 1B. Two fixed values, τ=0.01 and τ=0.1, are compared against the logarithmic cooling function given by (28) with τ_0=0.1. As a reference for

comparison, Fig. 3 also includes the results with two simpler strategies. The first is the case in which all UEs are connected to the BS with the lowest propagation loss (denoted as "All to Macro"). This situation corresponds to the classical approach, in which no relaying through the APs is used. The second is the random case in which each UE selects randomly, with the same probability, the BS with the lowest path loss or an AP connected to this BS. As a relevant performance metric of the behavior of the proposed algorithm, Fig. 3(a) presents the total transmitted power increase with respect to the optimum solution for the different strategies. The optimum solution has been obtained by performing an exhaustive analysis of all possible combinations at each simulation time step. The optimum solution depends on the UEs that are active at each time step, so it can change during the simulation. The presented results correspond to the average values along the whole simulation time. As another performance metric of interest, Fig. 3(b) plots the rate of accepted requests by the admission control described in Section IV.D.

From Fig. 3(a), it can be observed that the proposed approach with τ=0.01 and with logarithmic cooling achieves performance very close to the optimum (e.g., differences of less than 1% are observed for the logarithmic cooling case). Moreover, significant transmit power reductions are achieved with the proposed strategy in comparison to the classical approach, in which all UEs are connected to the BS with the lowest propagation loss (i.e., "All to macro"). Such approach requires on the order of 40% higher transmit power, which demonstrates the efficiency of the proposed method to reduce the power consumption and thus contribute to the overall energy savings.

The relevance of the temperature parameter τ can also be observed in Fig. 3(a). This parameter controls the trade-off between exploration and exploitation in the learning mechanism. Looking at the *softmax* criterion in section IV.C, low temperature results in a greater difference in selection probability for APs/BSs that differ in their Q value estimates, and the higher the value of Q, the higher the probability of selecting a given AP/BS. As a result, with low values of τ, the system tends to converge quickly towards appropriate solutions with selection probabilities close to 1. In this way, the system can quickly exploit what has been learnt by selecting the BSs/APs that provide the largest reward, at the expense that it will have less exploratory capability to identify other solutions in case the conditions change. In contrast, with large values of temperature, the differences in the selection probabilities for the BSs/APs become smaller, even if their Q values are different. As a result, the UEs require more time to identify the BSs/APs providing the largest reward, so they will have less exploitation capability but higher exploration capability to react to changes. In the results presented in Fig. 3(a), it can be observed that the choice τ=0.01 achieves a much better performance than τ=0.1 because in the latter case, there is excessive exploration, leading to the selection of non-optimal solutions in some cases. In fact, additional results not shown in the figure for the sake of simplicity revealed that

TABLE I
SELECTION PROBABILITIES $Pr_{BS}(j,n)$ AND $Pr_{AP}(k,n)$ FOR THE DIFFERENT
UES IN SCENARIO 1A

	S_1	A_1	A_2	A_3	A_4
u_1	0	0.985	0.005	0.005	0.005
u_2	0	0.990	0.003	0.003	0.004
u_3	0.005	0.980	0.005	0.005	0.005
u_4	0.003	0.988	0.003	0.003	0.003
u_5	0	0.989	0.004	0.004	0.003
u_6	0	0.988	0.004	0.004	0.004

increasing τ to larger values, such as $\tau=1$, results in performance very close to the random case, meaning that the selection probabilities of (26) and (27) are similar for all BSs/APs. In turn, when considering logarithmic cooling, it can be observed in Fig. 3 that the performance improves with respect to the fixed case $\tau=0.01$, mainly because the logarithmic cooling tends to reduce the temperature values as time elapses, so the best solutions are progressively selected with higher probability.

As shown in Fig. 3(b), the proposed approach with $\tau=0.01$ and with logarithmic cooling achieves the best performance, with 100% acceptance. In contrast, for the rest of the strategies, the acceptance ratio degrades significantly, especially for the case in which all UEs are connected to a macrocell.

B. Convergence analysis

This section analyzes the convergence behavior of the proposed approach towards the optimum solution. For this purpose, scenario 1A is considered. Because the optimum solution depends on the number of UEs that are active at each time, the analysis in this section considers that all UEs are continuously generating activity periods with an average duration of 30 time steps without any inactivity period between them. In this way, all UEs are active, and the optimum solution does not change during the simulation. In addition, no traffic is generated by the APs in this study.

Table I presents the AP and BS selection probabilities $Pr_{BS}(j,n)$ and $Pr_{AP}(k,n)$ for the different UEs when the system has converged at the end of a simulation. The Q-learning approach with logarithmic cooling is considered. For all UEs, the probability of connecting to the access point A_1 is greater than 98%. A detailed analysis of the power required with all combinations, not shown here for the sake of brevity, reveals that this is actually the optimum solution in this scenario to minimize the total transmitted power.

Fig. 4 depicts the evolution of the total transmitted power as a function of the total aggregated number of decisions (i.e., AP/BS selections) made by all UEs. The total number of decisions is a representative metric of the convergence because the algorithm progressively learns the optimum solution as new decisions are made by the different UEs. The total power progressively decreases until reaching the optimum value after a total of 17 decisions made by all of the UEs. Then, considering that there are 6 UEs in the scenario, every UE requires, on average, between 2 and 3 decisions to identify the proper AP. Fig. 5 illustrates the evolution of the APs selected by each UE. The figure reflects that after a total of 17 decisions, all UEs are connected to A_1, which
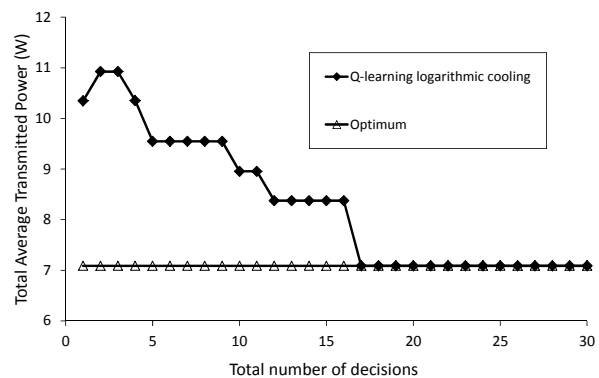

Fig. 4. Evolution of the total transmit power as a function of the total number of decisions made by the UEs in scenario 1A.
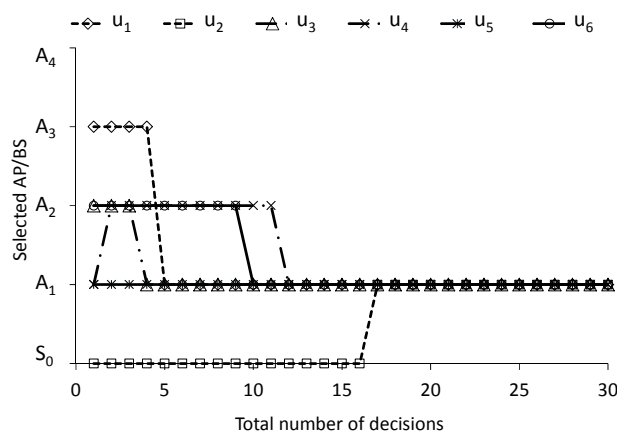

Fig. 5. Evolution of the AP or the BS selected by the different UEs as a function of the total number of decisions in scenario 1A.

corresponds to the optimum solution. Although it is not shown in this paper for the sake of brevity, similar values of the total number of decisions to find the optimum solution are observed in other scenarios where UEs and APs are located at different positions.

C. Comparison with a centralized genetic algorithm for the multi-cell scenario

As previously discussed, due to the dramatic increase in the number of combinations for the multi-cell scenario, it is not feasible to obtain the optimum solution, so the proposed approach is benchmarked against a genetic algorithm that is taken as a near-optimal performance bound. The genetic algorithm jointly considers all UEs, BSs and APs in the optimization process. Therefore, its implementation requires a centralized approach, as opposite to the proposed Q-learning, which is executed in a distributed way at each UE. In this respect, the objective of this study is to benchmark how far from the centralized technique the proposed distributed approach can be and not to discuss the implementation considerations associated with the comparison between the two techniques.

The genetic algorithm used for the benchmark is triggered each time that a UE begins or ends an activity period to consider the possible reconfigurations that may be required as a result of these events. The genetic algorithm operates

iteratively by evaluating in each iteration a population (also known as generation) of N_{pop} individuals or chromosomes, each corresponding to a candidate solution of the optimization problem [42]. The number of genes in each chromosome is N_{ACT}, corresponding to the number of active UEs in the scenario at the time the algorithm is triggered. Then, the g-th gene is associated to UE u_g and takes an integer value depending on the BS or the AP to which the UE is connected. Specifically, the gene takes the value j if the UE is connected to BS S_j and takes value $J+k$ if the UE is connected to AP A_k. The chromosomes considered in each generation correspond to solutions that fulfill the constraints (12)-(15) of the optimization problem (11). Each chromosome is evaluated in terms of a cost or fitness function that captures the total required transmitted power associated with the solution represented by this chromosome. The cost function $C(i)$ corresponding to the i-th chromosome is given by (10).

Based on the above, the operation of the genetic algorithm is as follows:

1) At initialization, a set of N_{pop} chromosomes that fulfill the constraints (12)-(15) are randomly generated.

2) The cost function $C(i)$ is evaluated for each chromosome i.

3) The following operators are applied to the chromosomes to obtain the new set of N_{pop} chromosomes that constitute the next generation:

3.1) Selection: The algorithm selects two chromosomes (parents) to be used to obtain two new chromosomes (children) for the subsequent generation. The parents are selected according to a roulette wheel process, in which chromosomes with lower cost are selected with higher probability. Specifically, the probability of selecting chromosome i is given by

$$P_{Sel}(i) = \frac{\dfrac{1}{C(i)}}{\sum\limits_{h=1}^{N_{pop}} \dfrac{1}{C(h)}}. \tag{30}$$

3.2) Recombination: The two selected chromosomes are recombined following the one-point-crossover methodology (see [42] for details) to obtain a new chromosome.

3.3) Mutation: Consists of changing the value of a gene belonging to the new chromosomes resulting from the recombination step. The probability of mutating one gene is given by $P_{mut}=1/N_{ACT}$. When a gene is mutated, its new value is selected randomly among the values $\{1, ...,J+K\}$, excluding the current value of the gene.

3.4) It is checked whether the resulting chromosome fulfills the constraints (12)-(15). If they are fulfilled, the chromosome is kept. Otherwise, it is discarded. The selection, recombination and mutation steps are repeated until obtaining a total of N_{pop} valid chromosomes for the next iteration/generation.

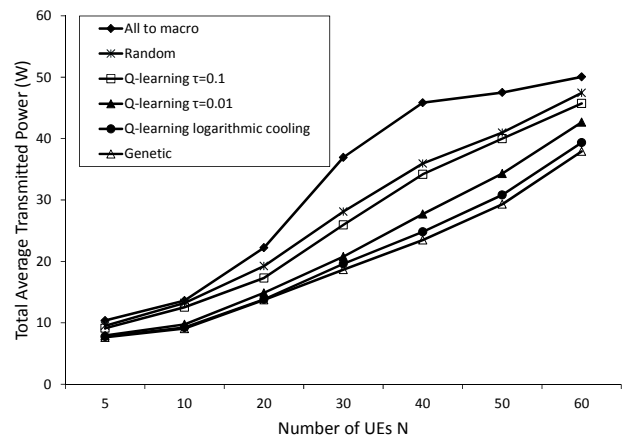4) Steps 2 and 3 are repeated iteratively until reaching a maximum number of iterations/generations. The solution of

Fig. 6. Total average transmitted power aggregated for all BSs and APs when increasing the number of UEs N for K=12 access points.

the algorithm corresponds to the chromosome with the minimum cost that has been found throughout all the generations.

The evaluation is performed in the multi-cell scenario with the J=3 BSs shown in Fig. 2(b). There are K=12 APs and a variable number of UEs N. Twenty runs of experiments with different uniform random distributions of the positions of UEs and APs are executed. UEs are randomly distributed in the whole area, whereas the APs are randomly distributed in square regions of side 100 m centered on each of the BS locations to reflect the fact that the APs that are located far from the BSs are not be useful for relaying the traffic of UEs located closer to a BS. Each execution of the genetic algorithm consists of 100 generations with a population of N_{pop}=30 individuals. The required bit rates of UEs and APs are $R_n=R_{A,k}$=2 Mb/s. The rest of the simulation parameters are the same as in the beginning of section V.

Fig. 6 shows a comparison of the proposed Q-learning approach with logarithmic cooling and τ_0=0.1 with respect to the other considered strategies in terms of the total transmit power as a function of the number of UEs N. In this case, the APs work at different frequencies, so they do not mutually interfere. The results observed with the Q-learning methodology with τ=0.01 and with logarithmic cooling improve the results obtained by the "All to macro" and
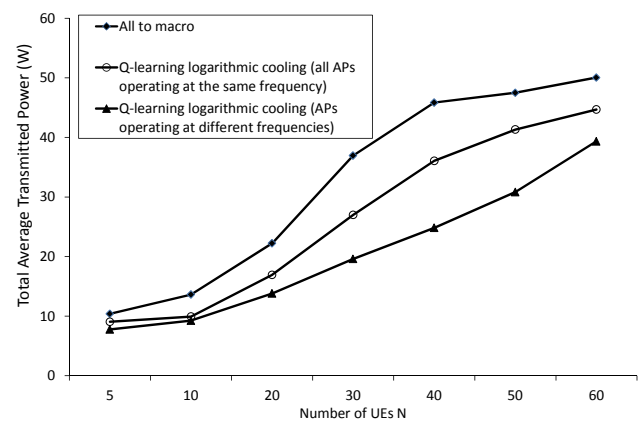

Fig. 7. Total average transmitted power aggregated for all BSs and APs when the APs work at the same and at different frequencies.

"random" strategies by significantly reducing the total transmitted power. Moreover, the proposed Q-learning approach with logarithmic cooling achieved similar performance to the genetic algorithm. Although this does not mathematically prove the guaranteed convergence to the optimum solution, as in Section V.B, because the genetic algorithm could converge to either a global or a local optimum, the results reveal that the proposed distributed approach is able to achieve very close performance to a classical optimization approach, such as the genetic algorithm, in spite of being much less complex. In terms of convergence time, the Q-learning approach achieves convergence to a solution after an average number of 7.3 decisions per UE. In terms of computational complexity, the simulation of 10000 time steps for the case of J=3 BSs, K=4 APs and N=10 UEs lasts approximately 10 s with the Q-learning approach in a state-of-the-art computer. In contrast, the same execution of the simulation with the genetic algorithm lasts approximately 90 minutes. This reflects the dramatic reduction in computational complexity of the proposed distributed approach.

Fig. 7 evaluates the impact of the interference in the D2D links in the case that all APs work at the same frequency, i.e., $F_{k,k'}$=1 in (9). In this case, the interference among the different APs reduces the capacities $r_{D,n,k}$ in the D2D links, which increases the activity Θ_k for the different APs, reducing their availability for relaying traffic and, as a consequence, the UEs tend to connect more frequently to the BSs. As a result, the total transmitted power increases with respect to the case where all APs use different frequencies. However, the power reduction with respect to the reference case, where all UEs connect through the BSs, is still significant.

Finally, to test the behavior when increasing the number of BSs, Fig. 8 considers a multi-cell scenario with J=7 BSs and K=21 APs (using different frequencies) in an area of 1540 m x 1700 m. The figure plots the total average transmitted power by all nodes for the proposed Q-learning scheme and for the case when the UEs are connected only to the macrocells. Similar results as in the previous cases are obtained, revealing that the proposed approach achieves a significant power reduction.

D. Influence of mobility and dynamic changes in the role of APs and UEs

This section presents some illustrative results to provide insight into the capability of the proposed Q-learning methodology to adapt to changes in scenarios where UEs and/or APs move and when the role of the APs and UEs changes dynamically. In the first experiment, we focus on the situation where a moving AP becomes available or unavailable to relay traffic for a particular UE. For that purpose, we consider the multi-cell scenario with the positions of the BSs, APs and UEs shown in Fig. 2(b). At t=2000 time steps, AP A_3 begins to move from its initial position (200,800) following a straight trajectory to the right until reaching position (900,800) at t=9000 time steps. Then, the AP remains at this position until the end of the simulation at t=10000 time
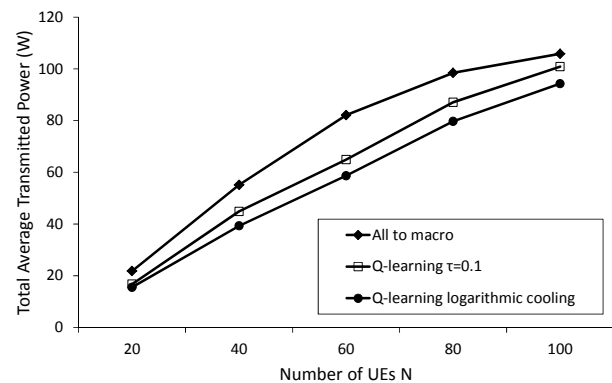

Fig. 8. Total average transmitted power in a scenario with J=7 BSs and K=21 APs.

steps. All UEs are continuously generating activity periods of average duration of 30 time steps without any inactivity period between them. At the beginning of each period, the UEs perform the AP/BS selection. The APs work at different frequencies.

We focus the analysis on the behavior of UE u_1 in Fig. 2(b). Fig. 9(a) shows the evolution of the selection probability $Pr_{BS}(2,1)$ that u_1 connects directly to BS S_2 and the probability $Pr_{AP}(3,1)$ that u_1 selects the 2-hop connection $S_2 \rightarrow A_3 \rightarrow u_1$. The other selection probabilities, $Pr_{BS}(j,1)$ and $Pr_{AP}(k,1)$, are almost zero during the whole simulation and are not represented in Fig. 9a. At the beginning of the simulation, u_1 identifies the direct connection to BS S_2 as the best option, with a selection probability close to 1. This is a reasonable choice because the closest AP, A_3, is located far from this UE. Then, as A_3 moves and approaches the position of u_1, Fig. 9(a) shows that the probability $Pr_{AP}(3,1)$ of selecting this AP begins to increase, and at approximately t=6000 time steps, the connection through AP A_3 is identified as the best option. However, as A_3 moves further to the right and away from u_1, the UE identifies that the direct connection through BS S_2 is again the best option. This occurs at approximately t=7500 time steps, when AP A_3 is located at position (750,800).

The second experiment assesses the capability of the proposed approach to address dynamic changes in the role of the APs and UEs. For that purpose, we consider the positions of the BSs, APs and UEs shown in Fig. 2(b). At t=1000 time steps, AP A_1 decides to switch off its relaying capabilities and becomes a UE. Then, at t=4000 time steps, UE u_4 is configured as an AP, denoted as A_5. These modifications affect the behavior of UE u_5, whose selection probabilities are plotted in Fig. 9(b). At the beginning, the probability $Pr_{AP}(1,5)$ increases to a value close to 1, meaning that u_5 learns to connect through AP A_1. Then, when A_1 becomes a UE at t=1000 time steps, u_5 identifies this situation, and the probability $Pr_{BS}(1,5)$ reaches a high value, indicating that the UE has learnt to use the direct connection to BS S_1, which becomes the best option, as seen in Fig. 2(b). Finally, after t=4000 steps, u_4 becomes configured as AP A_5, and u_5 identifies this new AP as the best connectivity option to receive service, i.e., $Pr_{AP}(5,5)$ reaches a value close to 1. This experiment reveals the robustness of the proposed approach to
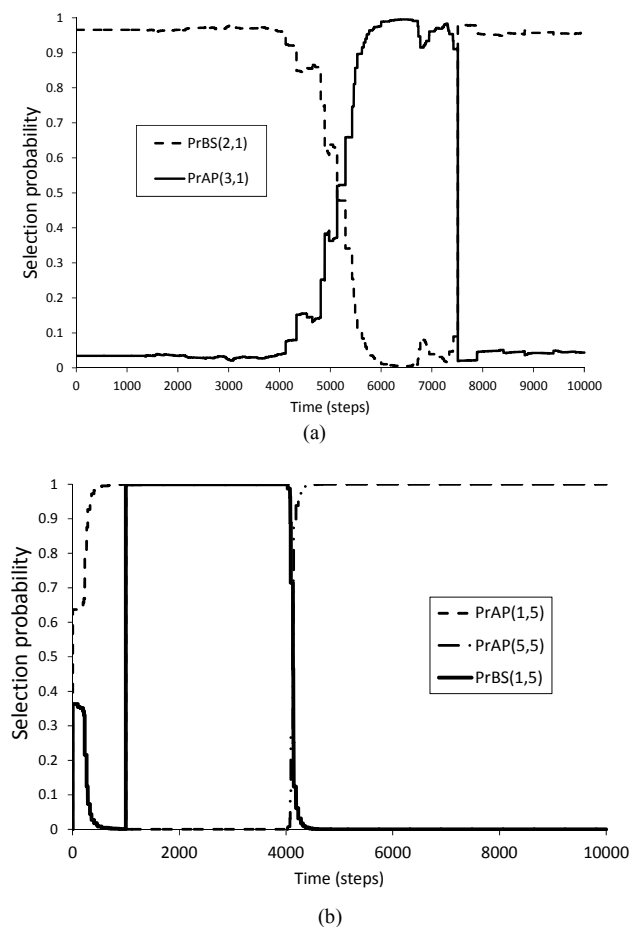
(a)

(b)

Fig. 9. (a) Selection probabilities $Pr_{BS}(2,1)$ and $Pr_{AP}(3,1)$ for u_1 in scenario 2 when there is an AP moving. (b) Selection probabilities $Pr_{BS}(1,5)$, $Pr_{AP}(1,5)$ and $Pr_{AP}(5,5)$ for u_5 when there are dynamic changes in the role of APs and UEs.

adapt to dynamic variations in the operating conditions.

As a third analysis, we consider the scenario of Fig. 2(b) with the APs and BSs located at fixed positions and all UEs moving during the simulation following random trajectories. At each position update, a UE can move forward, move back, turn left or turn right with the same probability. Like in the previous experiment, UEs generate continuous activity periods with average duration of 30 time steps. The mobile speed is such that a UE moves 3 m in each activity period. All APs work at different frequencies. For benchmarking purposes, the proposed Q-learning approach is compared to the centralized genetic algorithm described in Section V.C executed ideally at every time step so that it can be considered as an upper performance bound.

Fig. 10 presents the total transmitted power increase for the different methodologies with respect to the genetic algorithm. Significant power reductions are achieved by the proposed Q-learning approach with respect to the case when all UEs are connected to the macrocell BSs. In turn, the difference between the Q-learning approach with logarithmic cooling and the upper bound given by the centralized genetic algorithm is only 7%, which can be considered satisfactory performance. This result reveals the robustness of the proposed approach to operate under dynamic conditions.
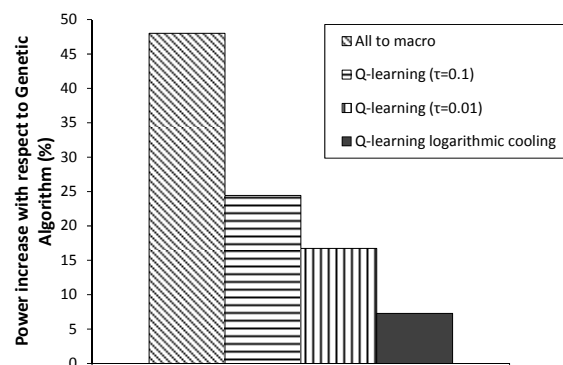

Fig. 10. Increase of the total transmitted power with respect to the centralized genetic algorithm for the different strategies in scenario 2.

VI. CONCLUSIONS AND FUTURE WORK

This paper has investigated the optimization of the connectivity of different UEs in a heterogeneous cellular scenario with D2D capabilities, in which specific UEs can act as access points and provide service to other UEs by relaying their traffic from the base stations, thus leading to a dynamic network architecture in which each UE can change the way it connects to the cellular network. In this respect, a new optimization framework has been presented to determine the most convenient connectivity option for each UE (i.e., one of the BSs or another UE acting as an AP), with the target of minimizing the total transmission power required in the scenario to fulfill the bit rate requirements of the different UEs.

A distributed strategy based on Q-learning and *softmax* decision making has been proposed as a means to implement the considered framework. Due to its distributed nature and to the fact that each UE relies only on its own experience to make decisions, the proposed approach has less complexity than centralized approaches that address the global optimization by jointly considering all APs, BSs and UEs.

The evaluation has demonstrated that the proposed approach can achieve transmitted power reductions of approximately 40% with respect to the classical approach in which the UEs are always connected to the BSs. Moreover, the temperature parameter in the *softmax* decision plays a relevant role for the proposed approach, so a logarithmic cooling technique has been adopted. The obtained performance in terms of transmitted power in a single cell scenario with the proposed approach is very close to the optimum, with differences below 1%. Moreover, a detailed analysis of the convergence properties of the proposed approach has been conducted, showing that the algorithm converges to the optimum solution after an average of 2 or 3 decisions per UE.

For multi-cell scenarios with high numbers of UEs and APs, in which the optimum cannot easily be known a priori, the proposed approach has been benchmarked against a centralized genetic algorithm, demonstrating that the proposed approach achieves similar performance in terms of total transmitted power while exhibiting much lower computational

complexity (e.g., as a reference, the duration of the presented simulations with Q-learning is approximately 10 s, whereas the same simulation with the genetic algorithm lasts approximately 90 minutes). The robustness of the proposed Q-learning methodology to operate in dynamic scenarios, where APs and/or UEs move and where the role of APs and UEs changes dynamically, has also been illustrated.

As future work, the considered framework could be extended to optimally determine which of the UEs are more adequate to act as APs so that the total power is minimized. Similarly, the considered optimization problem could be extended by optimizing the values of the transmitted power of the APs. This would be feasible if the D2D technology allowed some sort of dynamic power control to automatically modify the transmitted power. In addition, the framework could also be extended with consideration of other service requirements, such as the possibility of reducing the bit rate for those UEs that cannot achieve the required bit rate through any AP/BS or by considering other service metrics such as delay. Finally, the detailed implementation of the proposed algorithm for specific technologies is also considered as a future research direction.

REFERENCES

[1] 4G Americas (2012, Oct.). Developing and integrating a high performance HET-NET. [Online]. Available: http://www.4gamericas.org/files/4314/0759/4442/4G_Americas_-Developing_Integrating_High_Performance_HET-NET_October_2012.pdf

[2] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Commun.*, vol.18, no.3, pp.10-21, Jun. 2011.

[3] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communications in 5G cellular networks: challenges, solutions and future directions," *IEEE Commun.Mag.*, vol. 52, no. 5, pp.86-92, May 2014.

[4] "Study on LTE Device to Device Proximity Services; Radio Aspects," 3GPP Tech. Spec. Group Radio Access Network, V12.0.1, Release 12, Tech. Spec. TR36.843, Mar. 2014.

[5] D. Camps-Mur, A. Garcia-Saavedra, and P. Serrano, "Device-to-device communications with WiFi Direct: overview and experimentation," *IEEE Wireless Commun.*, vol. 20, no. 3, pp. 96-104, Jun. 2013.

[6] "Feasibility study for Proximity Services (ProSe)," 3GPP Tech. Spec. Group Services and System Aspects, V12.2.0, Release 12, Tech. Report TR22.803, Jun. 2013.

[7] S. Mumtaz and J. Rodriguez, *Smart Device to Smart Device Communication*, Switzerland: Springer, 2014.

[8] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1801-1819, Fourth Quart., 2014.

[9] S. Hahn *et al.*, Final report on a unified self-management system for heterogeneous radio access networks, Deliverable D6.6 of the SEMAFOUR project, Aug. 2015. [Online]. Available: http://www.fp7-semafour.eu/en/public-deliverables/

[10] D. Lister, "An operator's view on green radio," keynote in *IEEE Int. Workshop on Green Commun.*, Jun. 2009. [Online]. Available: http://www.green-communications.net/icc09/docs/GreenComm-ICC09-Keynote3-Lister.pdf

[11] D. Feng, C. Jiang, G. Lim, L. J. Cimini, G. Feng, and G.Y. Li, "A survey of energy-efficient wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 167-178, First, Quart., 2013.

[12] Health effects of exposure to EMF, Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR), Jan. 2009. [Online]. Available: http://ec.europa.eu/health/ph_risk/committees/04_scenihr/docs/scenihr_o_022.pdf

[13] *IEEE Standard for Safety Levels with Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3 kHz to 300 GHz*, IEEE Std. C95.1-2005, 2005.

[14] Téléphones mobiles: santé et sécurité (in French), Ministère de la Santé, de la Jeunesse et des Sports, Jan. 2008. [Online]. Available: http://www.priartem.fr/IMG/pdf/Ministere_sante.pdf

[15] Y-D. Lin and Y-C. Hsu, "Multihop cellular: a new architecture for wireless communications," in *Proc. IEEE INFOCOM*, Tel Aviv, Israel, Mar. 2000, vol. 3, pp. 1273-1282.

[16] H. Li, M. Lott, M. Weckerle, W. Zirwas, and E. Schulz, "Multihop communications in future mobile radio networks," in *Proc. IEEE PIMRC*, Lisboa, Portugal, Sep. 2002, vol. 1, pp. 54-58.

[17] B. Coll-Perales and J. Gozalvez, "Energy efficient routing protocols for multi-hop cellular networks," in *Proc. IEEE PIMRC*, Tokyo, Japan, Sep. 2009, pp. 1457-1461.

[18] B. Lorenzo and S. Glisic, "Optimal routing and traffic scheduling for multihop cellular networks using genetic algorithm," *IEEE Trans. Mobile Comput.*, vol. 12, no. 11, pp. 2274 - 2288, Nov. 2013.

[19] B. Lorenzo and S. Glisic, "Context aware nano scale modeling of multicast multihop cellular network," *IEEE/ACM Trans. Networking*, vol. 21, no. 2, pp. 359-372, Apr. 2013.

[20] Y. Kim, T. Kim. H. Kim. S. Kim, and Y. Han, "A threshold-based relay switching protocol for enhanced capacity and resource efficiency," *IEEE Commun. Lett.*, vol. 15, no. 10, pp. 1088-1090, Oct. 2011.

[21] H. Zhang, P. Hong, and K. Xue, "Mobile-based relay selection schemes for multi-hop cellular networks," *J. Commun. Netw.*, vol. 15, no. 1, pp. 45-53, Feb. 2013.

[22] L. Le and E. Hossain, "Multihop cellular networks: potential gains, research challenges, and a resource allocation framework," *IEEE Commun. Mag.*, vol. 45, no. 9, pp. 66-73, Sep. 2007.

[23] H. Zhang, P. Hong, and K. Xue, "Uplink performance of LTE-based multi-hop cellular network with out-of-band relaying," in *Proc. IEEE CCNC*, Las Vegas, NV, USA, Jan. 2012, pp. 817-821.

[24] A. Ting, D. Chieng, K. H. Kwong, I. Andonovic, and K.D.Wong, "Dynamic backhaul sensitive network selection scheme in LTE-WiFi wireless HetNet," in *Proc. IEEE PIMRC*, London, UK, Sep. 2013, pp. 3061-3065.

[25] M. Condoluci, L. Militano, G. Araniti, A. Molinaro, and A. Iera, "Multicasting in LTE-A networks enhanced by device-to-device communications," in *Proc. IEEE GLOBECOM Workshops*, Atlanta, GA, USA, Dec. 2013, pp. 567-572.

[26] L. Militano, M. Condoluci, G. Araniti, A. Molinaro, A. Iera, and F. H.P. Fitzek, "Wi-Fi cooperation or D2D-based multicast content distribution in LTE-A: a comparative analysis," in *Proc. IEEE ICC Workshops*, Sydney, NSW, Australia, Jun. 2014, pp. 296-301.

[27] J. Kokkoniemi, Y. Ylitalo, P. Luoto, S. Scott, J. Leinonen, and M. Latva-aho, "Performance evaluation of vehicular LTE mobile relay nodes," in *Proc. IEEE PIMRC*, London, UK, Sep. 2013, pp.1972-1976.

[28] L. Gao, G. Iosifidis, J. Huang, and L. Tassiulas, "Hybrid data pricing for network-assisted user-provided connectivity," in *Proc. IEEE INFOCOM*, Toronto, ON, Canada, Apr.-May 2014, pp. 682-690.

[29] J. Pérez-Romero, O. Sallent, and R. Agustí, "Enhancing cellular coverage through opportunistic networks with learning mechanisms," in *Proc. IEEE GLOBECOM Workshops*, Atlanta, GA, USA, Dec. 2013, pp. 642-648.

[30] Y. Fukuda, T. Abe, and Y. Oie, "Decentralized access point selection architecture for wireless LANs," in *Proc. Wireless Telecommun. Symp.*, Pomona, CA, USA, May 2004, pp. 137-145.

[31] B. Kauffmann, F. Baccelli, A. Chaintreau, V. Mhatre, K. Papagiannaki, and C. Diot, "Measurement-based self organization of interfering 802.11 wireless access networks," in *Proc. IEEE INFOCOM*, Anchorage, AK, USA, May 2007, pp. 1451- 1459.

[32] S. Vasudevan, K. Papagiannaki, C. Diot, J. Kurose, and D. Towsley, "Facilitating access point selection in IEEE 802.11 wireless networks," in *Proc. ACM SIGCOMM IMC*, Oct. 2005, pp. 293-298.

[33] Y. Bejerano, S.-J. Han, and L. Li, "Fairness and load balancing in wireless LANs using association control," *IEEE/ACM Trans. Networking,* vol. 15, no. 3, pp. 560–573, Jun. 2007.

[34] M. Hong, A. Garcia, J. Barrera, and S. G. Wilson, "Joint access point selection and power allocation for uplink wireless networks," *IEEE Trans. Signal Process.*, vol. 61, no.13, pp. 3334-3347, Jul. 2013.

[35] I. Malanchini, M. Cesana, and N. Gatti, "Network selection and resource allocation games for wireless access networks," *IEEE Trans. Mobile Comput.*, vol. 12, no. 12, pp. 2427-2440, Dec. 2013.

[36] K. Mittal, E. M. Belding, and S. Suri, "A game-theoretic analysis of wireless access point selection by mobile users," *J. Comput. Commun.*, vol. 31, no. 10, pp. 2049–2062, Jun. 2008.

[37] L.-H. Yen, J.-J. Li, and C.-M. Lin, "Stability and fairness of AP selection games in IEEE 802.11 access networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 1150–1160, Mar. 2011.

[38] A.S. Shafigh, B. Lorenzo, S. Glisic, J. Pérez-Romero, L.A. DaSilva, A. B. MacKenzie, and J. Röning "A Framework for Dynamic Network Architecture and Topology Optimization", *IEEE/ACM Trans. Networking*, Jan. 2015. [Online]. Available: http://ieeexplore.ieee.org

[39] B. Korte and J. Vygen, *Combinatorial Optimization. Theory and Algorithms*, 2nd ed., Germany: Springer-Verlag, 2002.

[40] R. S. Sutton and A. G. Barto, *Reinforcement Learning: an Introduction*, Cambridge, MA, USA: MIT Press, 1998.

[41] J. Nasreddine, O. Sallent, J. Pérez-Romero, and R. Agustí, "Positioning-based framework for secondary spectrum usage," *Physical Commun.*, vol.1, no.2, pp. 121-133, Jun. 2008.

[42] L. Davis, *Handbook of Genetic Algorithms*, New York, NY, USA: Van Nostrand Reinhold, 1991.

Jordi Pérez-Romero (S'98–M'04) is associate professor at the Dept. of Signal Theory and Communications of the Universitat Politècnica de Catalunya (UPC) in Barcelona, Spain. He received the Telecommunication Engineering degree and the Ph.D. from the same university in 1997 and 2001, respectively. Since then he has been actively working in the field of mobile and wireless communication systems, with particular focus on packet radio techniques, radio resource and QoS management, heterogeneous wireless networks, cognitive radio networks, self-organized networks and network optimization. He has been involved in different European Projects such as EVEREST, E2R, NEWCOM, AROMA, E3, FARAMIR, OneFIT and SESAME with different responsibilities as researcher, workpackage leader and project responsible. He has also participated in different projects for private companies. He has published more than 200 papers in international journals and conferences and has co-authored three books and contributed to 7 book chapters. He holds two international patents and has contributed to 3GPP and ETSI standardization bodies. He is associate editor of two international journals (IEEE Vehicular Technology Magazine and EURASIP Journal on Wireless Communications Networks) and has been TPC in different international conferences. Besides, he participated in the organization of IEEE VTC in Spring 2009 and IEEE PIMRC 2004 international conferences, hold in Barcelona. He has also been TPC Co-chair of EuCNC conference 2014 and of VTC Spring 2015. Moreover, he has organized several workshops, special sessions and special issues in international conferences and journals.

Juan Sánchez-González (M'08) received the Telecommunication Engineering degree in May 2002 and the Doctor Engineer in Telecommunication degree in December 2007 from the Universitat Politècnica de Catalunya (UPC), Spain. In 2002, he joined the Signal Theory and Communications department and he is currently assistant professor in the Escola d'Enginyeria de Telecomunicacions i Aeroespacial de Castelldefels (EETAC). His research interests are in the field of mobile communications systems, especially in heterogeneous wireless networks, self-organizing networks and cognitive networks. He has been involved in different European Projects as well as in projects for private companies.

Ramón Agustí (M'78) received the Engineer of Telecommunications degree from the Universidad Politécnica de Madrid, Spain, in 1973, and the Ph.D. degree from the Universitat Politècnica de Catalunya (UPC), Spain, 1978. He became Full Professor of the Department of Signal Theory and Communications (UPC) in 1987. After graduation he was working in the field of digital communications with particular emphasis on transmission and development aspects in fixed digital radio, both radio relay and mobile communications. For the last fifteen years he has been mainly concerned with aspects related to radio resource management in mobile communications. He has published about two hundred papers in these areas and co-authored three books. He participated in the European program COST 231 and in the COST 259 as Spanish representative delegate. He has also participated in the RACE, ACTS and IST European research programs as well as in many private and public funded projects. He received the Catalonia Engineer of the year prize in 1998 and the Narcis Monturiol Medal issued by the Government of Catalonia in 2002 for his research contributions to the mobile communications field. He is a Member of the Spanish Engineering Academy.

Beatriz Lorenzo (M'08) received the M.Sc degree in Telecommunication Engineering from University of Vigo, Spain, in 2008 and the Ph.D. degree from University of Oulu, Finland, in 2012. From 2008 to 2014 she worked at the Centre for Wireless Communications (CWC) at University of Oulu as researcher and project manager. In June 2014, she joined Atlantic Research Center for Information and Communication Technologies (AtlantTIC) at the University of Vigo where she currently works as senior researcher. Her research interests include multi-hop communications, heterogeneous networks, network economics, opportunistic communications, and complex networks.

Savo Glisic (M'90-SM'94) is a Professor of Telecommunications at University of Oulu, Finland, head of the networking research group, and Director of Globalcomm Institute for Telecommunications. He was Visiting Scientist at Cranfield Institute of Technology, Cranfield, U.K. (1976-1977) and University of California, San Diego (1986-1987). He has been active in the field wireless communications for 30 years and has published a number of papers and books. The latest books "Advanced wireless Networks: 5G/6G Joint design of the technology and business models, 3E, John Wiley and Sons, 2016" cover the enabling technologies for the definition of incoming 5G systems. He is also running an extensive doctoral program in the field of wireless networking (www.telecomlab.oulu.fi/kurssit/networks/).

His research interest is in the area of network optimization theory, network topology control and graph theory, cognitive networks and game theory, radio resource management, QoS and queuing theory, networks information theory, protocol design, advanced routing and network coding, relaying, cellular, WLAN, ad hoc, sensor, active and bio inspired networks with emphasis on genetic algorithms.

Dr. Glisic has served as the Technical Program Chairman of the third IEEE ISSSTA'94, the eighth IEEE PIMRC'97, and IEEE ICC'01. He was Director of IEEE ComSoc MD programs.