

Optimum Multi-Service Access Selection over Heterogeneous Wireless Networks

Johnny Choque, Luis-Francisco Díez, Alberto-Eloy García, Ramón Agüero and Luis Muñoz

Department of Communications Engineering - University of Cantabria, Spain
 Avda Castros s/n, - 39005 Santander, Spain
 {jchoque, ldiez, agarcia, ramon, luis}@tlmat.unican.es

Abstract—This paper proposes a methodology to analyze connectivity over highly heterogeneous wireless networks. We consider a scenario comprising a large number of access elements and end users, who move and initiate different services according to some patterns. The framework that has been implemented takes periodic snapshots, each of them used to pose a different optimization problem. We take into account the intention of end-users to have a connection (those with an active service), disregarding idle users, as well as the outcome of the previous problems (for instance, the base station a particular user was connected to). The feasibility of the methodology proposed in this work is assessed with a scenario over which we study different access selection strategies, including the following criteria: price of resources, service affinity towards particular technologies as well as the willingness to reduce the number of handovers. The results validate the proposed methodology, and highlight the impact that an appropriate design of the access selection strategy may have.

I. INTRODUCTION

It is estimated that mobile traffic will grow to 10 times between 2013 and 2019 [1], and approximately 90% of the world population would be able to use a WCDMA/HSPA connection, while the penetration of LTE is continuously increasing as well; in this sense, the aforementioned report estimates that more than 65% of the world population will be covered by LTE in 2019. On top of this, it is worth highlighting the remarkable increase on the number of advanced devices (smartphones and tablets) with a cellular connection. These usually incorporate other Radio Access Technologies (RAT) and thus the relevance of the so-called multi-RAT networks is likely to increase.

With the above scenario in mind, the relevance of an optimum management of the available resources within the wireless access networks has taken roots again, gathering the interest of the scientific community. Despite the large efforts which were made last decade, just after the appearance of the *Always Best Connected* motto [2], there are still a lot of new challenges and aspects to be looked at.

In this sense, there exists a large number of works which are looking at different techniques, algorithms and protocols to better manage the resources within wireless access networks, by inspecting the new possibilities which are brought about by avant-garde elements which have recently loomed. Some of them base their conclusions on the comparison with different alternatives and approaches, but in some cases it would be also interesting to know the gap with the best possible solution.

This work aims at answering the following question: *What is the best performance that might be expected over a heteroge-*

neous wireless access network? In order to give an reasonable answer we pose a series of optimization problems, in particular belonging to the binary linear programming group. We define a flexible utility function, able to integrate various criteria for both the network operators (for instance, load) or end-users (for example, price), and which provides a certain benefit to each of the access alternatives.

One of the most relevant aspects of this proposal is that, in order to establish the problem to be solved, we take into consideration the willingness of the users to have an active connection, depending on whether the service is active or idle (i.e. not all users want to be *always* connected), as well as how that service was previously handled (for example, the base station the user was connected to).

The paper is structured as follows; first, Section II outlines some of the work sharing the background of this one. Section III discusses the formulation of the problem, paying special attention to service modelling. Section IV briefly describes some of the most relevant aspects of the implementation which has been done. Section V presents the particular access selection strategies which will be challenged in the scope of this work, while Section VI assesses the feasibility of the procedure, by showing a number of results obtained with the developed framework for a particular scenario, over which we study a price-based load balancing scheme. Finally, Section VII concludes the paper, advocating some items that are left for future work and outlining some of the potential possibilities that might appear with the exploitation of the developed framework.

II. RELATED WORK

As said earlier, the *Always Best Connected* motto was originally proposed by Gustafsson and Johnson in 2003 [2]. This research line took roots in the last decade and several proposals were made so as to optimally manage the resources in wireless access networks, providing the best possible *Quality of Service* to the end-users. The reader might refer to [3] and the references therein for a thorough review of some of the approaches which were made at that time.

Afterwards, many new possibilities have been brought about by the novel techniques which have loomed. A clear example of these is the capability to virtualize resources at the access network [4] as a means to offer tailored quality of experience to the end-users. Another technique with a potential impact on the performance of wireless access networks is multi-path (possibility to split a flow between different paths, so as to increase performance or reliability) in multi-homed

enabled devices [5]. In addition we have also seen that new requirements have also gained relevance, such as the need to optimize energy consumption [6]. Considering these aspects and analyzing their influence over the performance and behavior of *multi-RAT* networks is still an open research issue.

All in all, the proposals which were made around 5 years ago, although conceived so as to have a certain degree of flexibility, might not be able to solve all the new challenges and requirements which are continuously appearing. Recently, a framework to promote Open Connectivity Services has been proposed [7] as a new paradigm to manage connectivity in forthcoming communication scenarios.

A common aspect from most of the works which are within the previously described research line, i.e. analyzing algorithms, protocols, etc. to foster a better operation of heterogeneous wireless access networks, is that they compare their performance with that shown by legacy alternatives, but they do not assess how close they are to an optimal behavior [5], [8]. The main goal of this paper is to propose a way to find this optimum solution, considering different parameters of merit over highly heterogeneous networks, in which end-users are able to use different types of services.

This work is an evolution from the framework which we originally presented in [9]. We have added support for an appropriate modeling of services (as opposed to the *always-want-to-be-connected* approach used in [9]), with a clear influence on how the problem needs to be posed, and which led to the integration of a module to keep track of the history of previous connections. This is further discussed in Section III.

As was said earlier, we use Binary Linear Programming to find the optimum solution. To our best knowledge there are not many works that have fostered this approach. For instance, the authors of [10] propose a selection mechanism that uses a utility function to prioritize access alternatives, and they strengthen the application and an energy aware behaviour. They employ an Integer Linear Programming based technique to seek a solution that maximizes the aforementioned utility. On the other hand the whole procedure is only initiated upon a handover event, leaving aside situations in which there might be better access alternatives available, even if a handover was not strictly required.

Other works that use optimization techniques are [11], [12], but they propose utility functions that are quite tight to the scenario they consider, leaving aside relevant parameters, such as the handover. In particular, [11] focuses on load balancing and efficient radio resource management mechanisms, and [12] provides greater relevance to the physical parameters of wireless technologies, and is therefore more limited when it comes to including other figures of merit. All in all, the methodology we present in this paper has a broader view on its scope; in this sense, by using some more abstractions, we gain some additional flexibility in terms of the criteria we can integrate within the access selection strategy. Besides none of the aforementioned works model services as we do herewith.

III. PROBLEM FORMULATION

The access selection problem aims at establishing the optimum association of the current active flows (services) for

all the users amongst the available access alternatives. The scenario comprises N available access networks, which can use various technologies, thus having different characteristics, in terms of coverage and capacity. We also assume that there are U users, who can *simultaneously* start S different flows/services, equipped with a terminal able to establish a connection with any of the involved technologies.

We clarify herewith that in the scope of this work we assume that both services and base stations use a generic and discrete capacity unit (the so called *Traffic Unit*, TU), no matter it refers to time slots (TDMA), codes (CDMA), sub-carriers (OFDMA), etc. Any service would require a number of TUs to be properly handled by the network (if they cannot be assigned, then the service is rejected/dropped) and each of the deployed base stations would have a limit on the number of resources (TU) it can assign to the users¹. Other works, such as [11], [13] also use this abstraction.

We formulate the problem as a binary integer program, in which there are $U \times N \times S$ basic variables (x_{ijk}), which can be defined as follows.

$$x_{ijk} = \begin{cases} 1 & \text{if user } i \text{ uses network } j \text{ for service } k \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

We will use a generic utility function (u_{ijk}) so as to qualify the goodness of a particular connection, according to different criteria. It is important to remark that this utility, although it depends on the particular circumstances and the values the different criteria can take for each of the available access alternatives, can be considered constant for a problem instance, thus ensuring the linearity of the proposed model (i.e. u_{ijk} does not depend on x_{ijk}). In addition, we pose three constraints to the problem: (3a) ensures that the basic variables to be solved are binary; (3b) forces that a single flow can be only connected to one Base Station (BS); while (3c) limits the number of TUs which can be assigned by BS_j (with capacity C_j), according to the capacity required by each of the services (c_k). With all of the above, we can pose the maximization problem as follows.

$$\mathbf{Max.} \quad \sum_{i,j,k} u_{ijk} \cdot x_{ijk} \quad (2)$$

$$\mathbf{s.t.} \quad \forall i, j, k \quad x_{ijk} \in \{0, 1\} \quad (3a)$$

$$\forall i, k \quad \sum_j x_{ijk} \leq 1 \quad (3b)$$

$$\forall j \quad \sum_{i,k} x_{ijk} c_k \leq C_j \quad (3c)$$

Note that not all of the basic variables are part of the problem, since for instance we need to discard those which are not possible due to the lack of physical connectivity between user \mathbf{i} and base station \mathbf{j} , $x_{ijk} = 0 \quad \forall k$. In addition, when a particular service (say \mathbf{k}) at end-user \mathbf{i} is not active, we can

¹In this work we abstract the differences between technologies and we do not consider the impact that the particular physical conditions of the links might have over the capacity provided by each of the resource units assigned by the base stations.

also add some additional constraints to the aforementioned problem: $x_{i;k} = 0 \quad \forall j$.

As mentioned earlier, one of the most relevant contributions of this work, compared to our previous paper [9] is that we do not consider all users for every instance of the optimization problem, but only those who are willing to have a connection. In [9], we looked at the overall connectivity and therefore every user wanted to have a connection, no matter whether he/she had an ongoing service or not. In this sense, the way the problem is formulated herewith is more elaborated. In order to appropriately include service modelling we consider that a particular application flow, once it is rejected or dropped, is not considered in the following optimization problems (until the service is restarted again). For that we need to maintain the history of the previous optimization outcomes, and we define a state machine for all possible *user/service* combinations. As a result, we establish four different circumstances for each of them, as shown below.

- **Idle.** The user *i* is not currently having service *k* active and therefore he/she does not require any connectivity for it.
- **Active.** The user *i* has an ongoing connection for service *k* and this has been accepted by some of the available access alternatives.
- **Rejected.** The user *i* initiated a flow for service *k*, but it was not accepted by the network.
- **Dropped.** The user *i* had an ongoing flow for service *k* which was originally accepted, but it eventually stopped before its correct finalization, due to user mobility or other events.

With the aforementioned states and considering that the optimization, for a single analysis run, consists of the resolution of a series of consecutive *snapshots*, we can establish the evolution of the state for a particular *user-service* pair as shown in Fig. 1. The arrows represent state transitions, which depend on two aspects: the intention of such user for a particular service and the outcome of the current optimization problem. These are the two numbers on top of each of the arrows (state transitions): *intention/optimization_result*. It is straightforward to see that, depending on the previous state and the current intention, there are some cases in which the corresponding variables do not enter the optimization problem, represented in the figure by dashed arrows. No matter the previous state is, when user *i* does not have the intention of having service *k* connected, then the corresponding basic variables do not enter the optimization problem as was discussed earlier ($x_{i;k} = 0 \quad \forall j$) and such pair goes to the *Idle* state. In addition, whenever a service has been rejected or dropped, the corresponding variables are left aside the optimization problem even if the intention is to have a connection. Furthermore, a connection that has been lost cannot be recovered and thus the state is kept, until it goes back to the *Idle* state. This assumption is required so as to make the traces (and therefore the traffic demand) independent from the outcome of the analysis; in this sense we could study various access selection strategies with the same conditions.

In most of the cases discussed above, the *optimization_result*

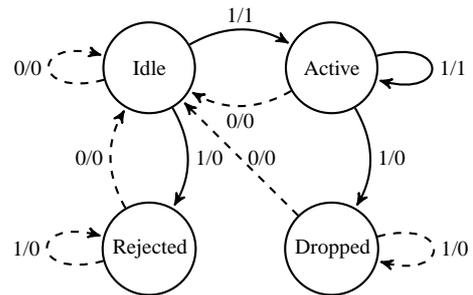


Fig. 1. State machine for a specific user/service pair

equals zero, since it the corresponding basic variable does not belong to the optimization problem. On the other hand, the continuous lines show the transition of the *user/service* pairs towards the final state, depending on the outcome of the current optimization problem.

IV. IMPLEMENTATION

A solution of a single optimization problem refers to the optimum connectivity of a particular situation, which is given by the current position of the end-users, their willingness to have a connection for a particular service, as well as the current status of the network (remaining capacity). Those particular situations can therefore be seen as *snapshots* of a scenario, in which users move and generate flows for their corresponding services within a certain period of time. Those patterns (traffic and movement) are provided by means of traces. In order to solve a particular scenario, a tool has been developed from scratch. It has two well defined modules: the first one solves each of the optimization problems which are posed by the series of snapshots; a process is created per problem and, thus, there is not a ‘real’ connection between the current problem and the previous solutions. As was discussed before, to properly pose the optimization problem, it is important to be able to take into account the history of the scenario (services which were rejected or dropped, load at the base stations, etc.); in order to keep track of this historical evolution and to provide the information to the current snapshot, another module, the monitor, has been added to the tool (this was not needed for the previous stateless approach [9]). It analyzes the outcome of the optimization problem and keeps it so as to be used by the next snapshot; in addition, it is in charge of maintaining the overall statistics of the scenario in order to process them when all the snapshots have been solved.

As said earlier the tool is fed with a set of traces that must reflect: (1) the position of the end-users, according to some predefined mobility pattern, and (2) the intention of all the users to have a connection for each of their services. Figure 2 shows an illustrative example of three users and their services and how these are mapped onto the corresponding snapshots. It is worth mentioning that the time in which a service remains at the idle or active states shall be always higher than the snapshot period, so as to keep track of the new incoming flows, as it is the case of *Service 0* for user *i* on Fig. 2. For a particular user/service (*i/j*) pair, we can thus define the

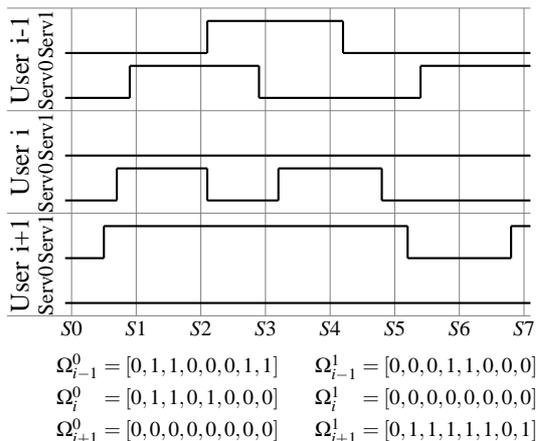


Fig. 2. Illustrative example of service snapshots

vector $\Omega_t^j = [\omega_0, \omega_1, \dots, \omega_N]$, in which ω_k is the status of such service at the k^{th} snapshot. As can be seen, the format which has been selected for the traces processed by the application is rather generic and, therefore, its use can be easily extended (for comparison purposes) to other studies.

The procedure that is being followed, for each of the snapshots, is shown in Procedure 1. We divide it into three different phases: (1) first we establish the scenario, by reading the overall configuration parameters, the position and characteristics of access elements (base station, access point) and the particular position and service status for the end-users; all this information is gathered from external files. (2) Once the scenario is deployed, we establish the physical connectivity between end-users and access elements, and the current state of each user/service pair; we also build the utility function by considering, depending on the particular configuration, the previous access element (if the user/service pair was already connected), the current load of the access elements and the

Procedure 1 Snapshot optimization

Global configuration

- Scenario establishment
- configuration parameters: from-file
- end-users: from-file
- access elements: from-file

Initial processing

- Optimization variables
- Connectivity establishment
- Service intention and previous state
- Utility functions
- Previous access element
- Current loads
- Current price

Optimization

- Problem resolution: GLPK
 - Process result: update state
 - Update statistics
-

price that shall be offered to a particular user/service. (3) Once the utility function is established, the problem is solved by using a solver (in this work, we have used the GLPK library [14]) and the result is processed to update the state of all the user/service pairs and the corresponding statistics.

V. ACCESS SELECTION STRATEGIES

As mentioned earlier, the main idea is to be able to model various criteria with the *utility* function, to assess the goodness of a particular connection. There exist a broad range of different possibilities and, within the scope of this paper, we will focus on two particular aspects: pricing and RAT affinity (which can be described as the preference to connect particular services to specific technologies), as well as their combination. In addition, we will also study their interaction with the cost of changing the base station, i.e. the preference an end-user would have towards maintaining the current access. We will see that different combinations of these criteria may lead to rather distinct performances, so an appropriate selection of the strategy is of utmost relevance.

A. Common criteria

Two criteria are used for the various strategies that will be analysed: the willingness to have a connection as well as the intention to reduce the number of handovers. These are discussed below.

1) *Connectivity*: As mentioned earlier we will give some utility to the connectivity *per se*, prioritizing ongoing services over new calls, since it is sensible to assume that dropping an already established call is worse (from the perspective of the quality of experience) than rejecting a new one. This criterion (α_{ijk}) is defined as follows:

$$\forall j \quad \alpha_{ijk} = \begin{cases} 1 \cdot \frac{c_k}{\max_{v/k} c_k} & \text{if user/service } i/k \\ & \text{was connected} \\ \lambda \cdot \frac{c_k}{\max_{v/k} c_k} & \text{otherwise} \end{cases} \quad (4)$$

where c_k is the capacity required for service k and λ (design parameter) is selected so as to ensure that an ongoing service is always given a higher priority than a new call (for all the considered capacities): $\lambda < \frac{\min_{v/k} c_k}{\max_{v/k} c_k}$.

As can be seen the definition of this criterion considers the capacity required by each of the services; otherwise, those needing less capacity would be given (on an indirect way) a higher priority. In this sense, a connection of a service requiring 2 TUs has the same utility as a connection of a 1 TU service, while it consumes twice the resources; then, the optimization engine will favor two connections of the latter service over one belonging to the former one.

2) *Handover - HO*: We will consider the influence that the cost of change (handover) might have on the assessed performance. In this sense, we model the willingness to keep the current access (so as to include the cost of change). For that, we use the β_{ijk} criterion, defined as follows:

$$\beta_{ijk} = \begin{cases} 1 & \text{if user/service } i/k \text{ was previously} \\ & \text{connected to BS } j \\ 1 - \mu & \text{otherwise} \end{cases} \quad (5)$$

where $\mu < 1$ is a design parameter which would be selected depending on the particular configuration, as will be discussed later.

B. Price criterion

The idea is to favor the monetary preference of the end-users, who would opt for more economical connections. The utility a user would perceive for certain service connections would increase as far as the base stations offer a lower price; we are thus looking for a decreasing function. Since we want to enable relative comparisons (i.e. based on discount percentages) between offered prices, we propose using a logarithm function as follows (see Fig. 3(a)), where p_j is the price offered by base station j given as monetary units per time and capacity unit.

$$\forall k \quad \gamma_{ijk} = \begin{cases} -\log(p_j) & p_j \in [p_{\min}^{\text{BS}}, p_{\max}^{\text{BS}}] \\ 1 & \text{otherwise} \end{cases} \quad (6)$$

where p_{\min}^{BS} and p_{\max}^{BS} correspond to the highest and lowest price offered by BS j to use its resources, respectively. As can be seen, the first one is the fee below which end-users would not perceive any utility gain. We also assume that those base stations that offer a price higher than the maximum an end-user would be willing to pay (user preferences) would be discarded. Relative price units are used, and therefore the maximum price offered by any of the base stations would be 1.0. As will be seen later, we will also assume that $p_{\min}^{\text{BS}} = 0.1$ monetary units per time and capacity unit.

Moreover, we will also consider that base stations are using their offered price so as to encourage or deter users to connect to them; this is reflected in Fig. 3(b), which represents the fee offered by a base station as a function of the currently *relative* available load. As can be seen, when the base station is highly loaded (available capacity lower than $L_{\text{low}}^{\text{th}}$), the offered price is the maximum allowable one (operator policies and rules); on the other hand, if the base station load is low, the offered price decreases to a minimum configured level. For the sake of simplicity, a linear decreasing trend has been used between these two points; in the analysis that will be discussed in Section VI we have used 0.2 and 0.8 for the lower and upper thresholds, respectively.

C. RAT Affinity criterion

This criterion is conceived to favor that particular services are handled by preferred technologies, so as to bring about a better quality of experience. For instance, we could use WiFi accesses for data transfer services, since they benefit from higher bandwidths, but do not require a strict delay, while we would establish a preference of voice services towards cellular base stations, better suited for that. The corresponding component of the utility function (δ_{ijk}) is therefore defined as follows:

$$\forall i \quad \delta_{ijk} = \begin{cases} 1 & \text{if service } k \text{ has affinity towards} \\ & \text{technology of BS } j \\ v & \text{otherwise} \end{cases} \quad (7)$$

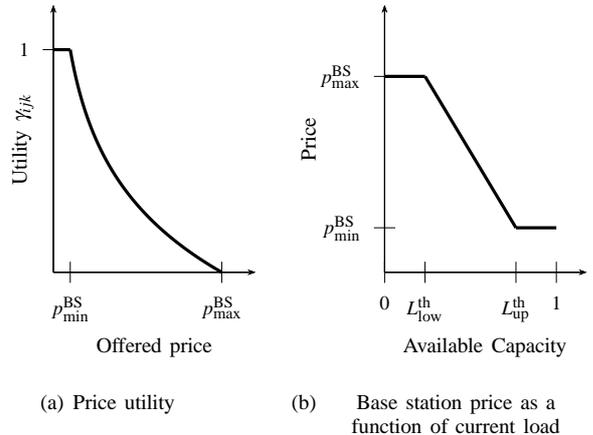


Fig. 3. Price-related utility functions

where $v < 1$ is a design parameter, which modulates the relevance which is given to this particular aspect of the utility function.

D. Utility Function

By linearly combining all the introduced criteria, we can define an overall utility function, which establishes the corresponding access selection strategy.

$$u_{ijk} = A \cdot \alpha_{ijk} + B \cdot \beta_{ijk} + C \cdot \gamma_{ijk} + D \cdot \delta_{ijk} \quad (8)$$

There are two possibilities to favor one of the criteria over the others: (1) changing the corresponding factors (A, B, C, D); or (2) use only a binary version of them (i.e. they are activated or not) and use the corresponding design parameters (λ, μ, v); in the scope of this work we will use the second alternative. In particular, we will fix that $\lambda = 0.4, v = 0.8$, while we discuss the value for μ below.

For that, we do not consider RAT affinity (i.e. $D = 0$). We then compare the utilities for two different access alternatives for an ongoing service (note that the value of α_{ijk} is not relevant, since it will be alike for the two accesses). The end-user would have two choices, the one to which she is currently connected to (1) and another one (2), which is offering a cheaper fee than (1). In particular, we assume that p_2 is $100 \cdot \xi\%$ lower than p_1 , i.e. $p_2 = (1 - \xi)p_1$, with $\xi < 1$. Hence, we can write:

$$\begin{aligned} u^{(1)} = u^{(2)} & \rightarrow 1 - \log(p_1) = (1 - \mu) - \log(p_1(1 - \xi)) \\ & \rightarrow \mu = -\log(p_1(1 - \xi)) + \log(p_1) \rightarrow \\ & \rightarrow \mu = -\log(1 - \xi) \end{aligned} \quad (9)$$

Therefore, if we consider that a discount of 20% should be enough for an end-user to change her current access, the value of μ to be used is ≈ 0.1 .

With all the above in mind, we can define the six different strategies that will be used during the analysis presented in the next Section, as defined in Table I. Each of them corresponds to a different instance of the utility function that was introduced earlier.

TABLE I
ACCESS SELECTION STRATEGIES

	Price		RAT affinity		Combined	
	HO	No-HO	HO	No-HO	HO	No-HO
Connectivity (<i>A</i>)	1	1	1	1	1	1
Handover (<i>B</i>)	1	0	1	0	1	0
Price (<i>C</i>)	1	1	0	0	1	1
RAT Affinity (<i>D</i>)	0	0	1	1	1	1

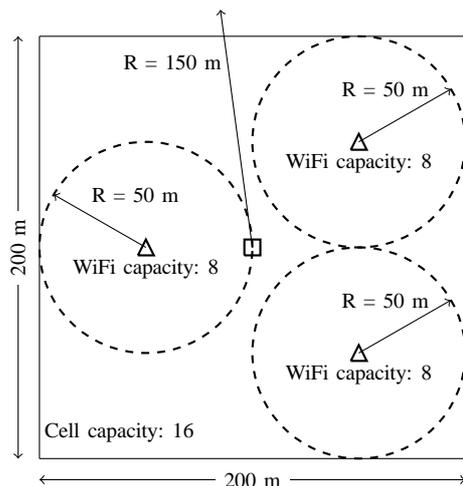


Fig. 4. Network deployment

VI. RESULTS

This section presents some results obtained by applying the method that has been previously presented over a particular network deployment. The goal is twofold: first, to assess its feasibility and, furthermore, to discuss the impact of an appropriate selection of the utility function.

In particular, we will consider a $200 \times 200 \text{ m}^2$ scenario, in which we deploy two different types of base stations (see Fig. 4). The first one corresponds to a traditional cellular technology, with a coverage of 150 m , effectively covering the whole scenario, and a capacity of 16 TUs, while the second one mimics a WiFi access router, with a range of 50 m and a capacity of 8 TUs. Over such scenario we deploy a number of users which we increase from 20 to 200; they move according to a *Random Waypoint* model, whose parameters are given in Table II. Each user generates flows belonging to two different service types, following *ON-OFF* models, having some preference towards a particular technology. All the base stations use the pricing policy which was depicted before (Fig. 3(b)), and thus they offer a higher fee when the current carried load is higher. We run 10 independent simulations per scenario (each of them comprises 360 different optimization problems, corresponding to a simulation time of one hour, and taking snapshots every 10 seconds) and we represent the corresponding average results. We consider the six strategies which were introduced in Table I, and we analyze the figures of merit which are enumerated below, for the two types of service which are being used (0, 1).

- **Success Rate (SR).** Probability that a service is success-

TABLE II
ANALYSIS PARAMETERS

Base stations	
<i>Cellular</i>	
Coverage (m)	150
Capacity (TU)	16
Pricing policy	$P \in [0.1, 1.0]$, Thresholds = $\{0.2, 0.8\}$
<i>WiFi</i>	
Coverage (m)	50
Capacity (TU)	8
Pricing policy	$P \in [0.1, 1.0]$, Thresholds = $\{0.2, 0.8\}$
Mobility model	
<i>Random Waypoint</i>	
Speed (m/s)	$U(1, 3)$
Movement time (s)	$U(800, 1000)$
Pause time (s)	$U(80, 100)$
Service model	
<i>Service 0: Voice</i>	
Model	On-Off
Idle (<i>off</i>) time (s)	420
Service (<i>on</i>) time (s)	180
Capacity (TU)	1
Affinity	Cellular
<i>Service 1: Data transfer</i>	
Model	On-Off
Idle (<i>off</i>) time (s)	480
Service (<i>on</i>) time (s)	120
Capacity (TU)	2
Affinity	WiFi
General parameters	
Simulation time (s)	3600
# of snapshots	360
# of runs	10

fully finished (i.e. it is neither rejected nor dropped).

- **Handovers (HO).** Average number of handovers which are carried out during a service lifetime.
- **Price per service (PS).** Average service price paid per time and traffic unit.
- **RAT affinity (RA).** We study the percentage of the time that the service was using the technology towards which it has some preference.

Before discussing the performance metrics that were obtained, Figure 5 shows the traffic demand that was generated by all the users, relative to the overall network capacity. Note that the traffic demand does not exactly correspond to the offered traffic (that would show a linear increase with a slope of 0.7 - traffic offered per user according to the parameters shown in Table II), but to the one that is used to pose the corresponding optimization problem. As was mentioned earlier, whenever a service is rejected or dropped, it is not longer considered in the future problems until it is restarted. This would effectively reduce the traffic demand, explaining the saturation behavior that is observed in Figure 5(a). We also represent (Figure 5(b)) the relative carried load per RAT type; we can see that the cellular base station gets almost fully loaded (for the three strategies²), since it covers the whole scenario. The load of the WiFi access points increases as long

²In this case we have used the strategies that considered the handover criterion, since the results obtained for the others are alike.

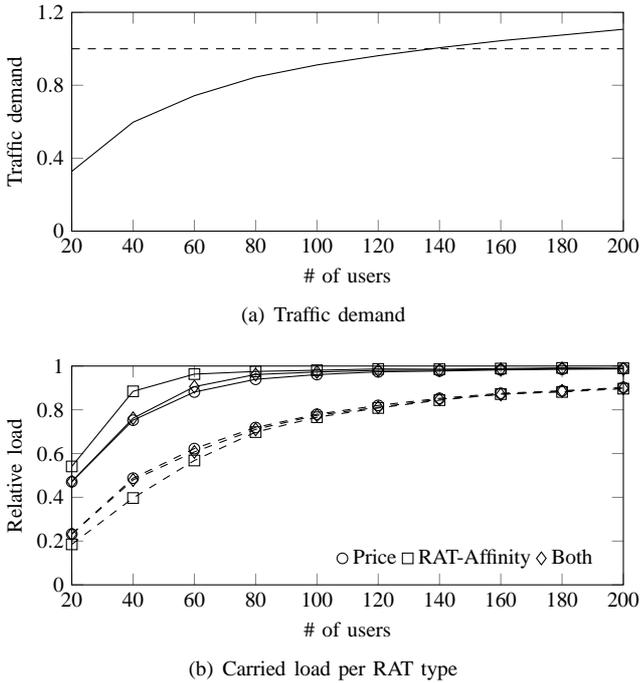


Fig. 5. Traffic demand and carried load per RAT type. Continuous lines are for Cellular and dashed lines for WiFi

as the number of users gets higher. Despite the traffic demand surpasses the network capacity, there are still some available resources in these access points, since they do not fully cover the area under analysis.

First, Figure. 6 shows the probability for a service to be successful. As can be seen when there is enough available capacity (i.e. # of users 20 or 40), the utility function definition (in particular the α parameter) leads to having similar results for the two types of services (if we had not considered the capacity in the corresponding definition, the success rate for *Service 1* would have been lower). Afterwards, the probability for a service to be successful is higher for *Service 0*, since it requires fewer resources. There is an exception, for the *Price w/o HO* strategy, in which *Service 1* has a better performance than *Service 0*. In this case, cheaper connections will be favored (without looking at the possible handover side-effects) and therefore *Service 0* calls would try to use WiFi accesses as soon as they become available³, opposed to the other strategies, where the RAT affinity criterion would foster that *Service 0* calls are handled by the cellular base station. Afterwards, due to user mobility, that connection might not be longer available, and the call would probably need to be dropped, since the cellular base station might be fully loaded (see Figure 5). In addition, there is as well an increase on the number of rejected services, since it is rather unlikely that a new call from *Service 0* would cause a call from *Service 1* to be dropped (the additional utility of the price criterion is low, since the load/offered price is high); on the other hand, in those strategies strengthening the RAT-affinity criterion, a

³Those accesses are likely to be cheaper since the corresponding access points are less loaded, as can be seen on Figure 5.

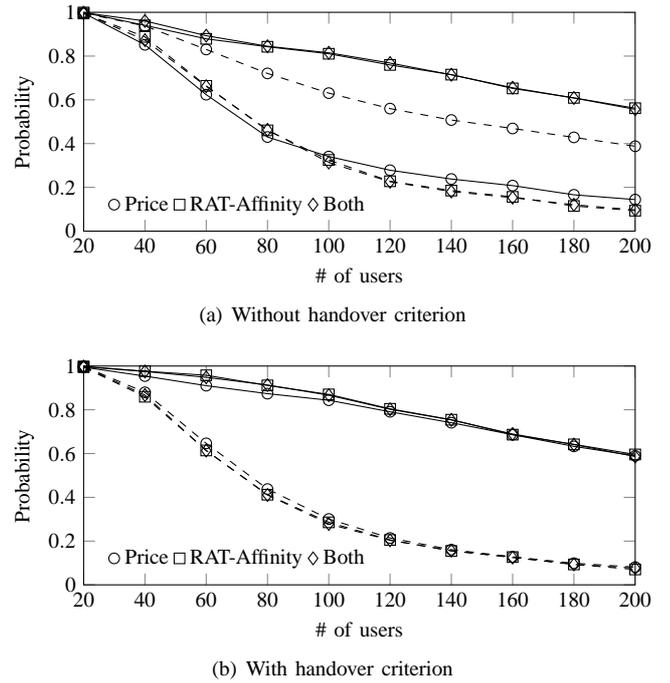


Fig. 6. Average success rate per service Vs. the number of users. Continuous lines are for *Service 0* and dashed lines for *Service 1*

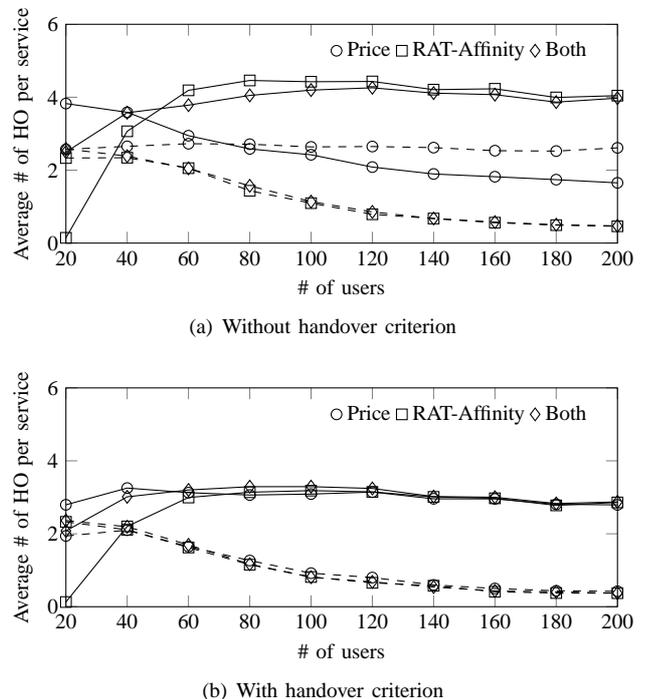


Fig. 7. Average number of handovers per service Vs. the number of users. Continuous lines are for *Service 0* and dashed lines for *Service 1*

new call from *Service 0* would quite likely cause an ongoing *Service 1* call to be dropped (if it is using the cellular base station), due to the stronger weight given to this particular criterion. This, together with the fact that the *Service 0* calls are longer than those corresponding to *Service 1*, explain the aforementioned exception.

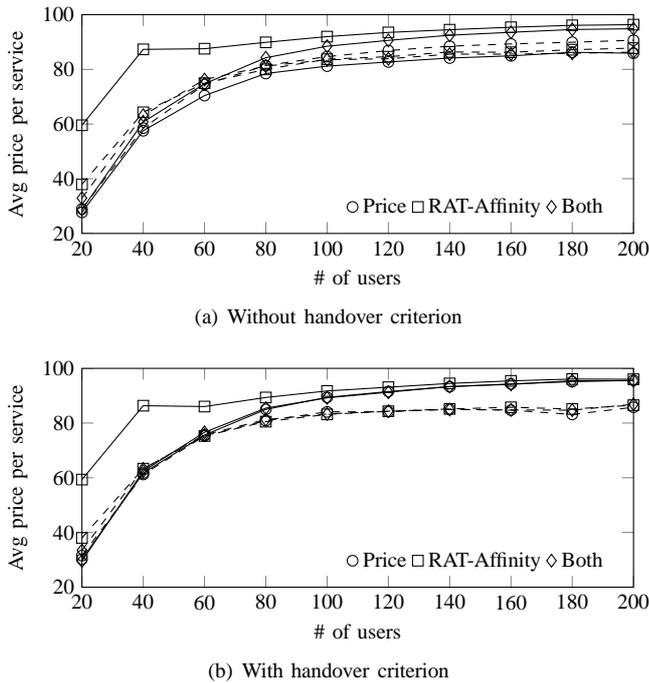


Fig. 8. Average price per service and TU Vs. the number of users. Continuous lines are for *Service 0* and dashed lines for *Service 1*

On the other hand, Figure 7 yields a clear impact of considering the handover criterion. We can see a decrease of the number of handovers (in particular for *Service 0*) when the handover criterion is considered in the utility function ($B = 1$). The impact is less relevant for *Service 1*, especially when the number of users is large, since the networks get highly loaded (which is more relevant for the cellular base station, see Figure 5) and there might be not many connection alternatives. On the other hand, when $B = 0$ (i.e. the cost of change is not considered) we can also see (again for *Service 1*) that the RAT Affinity strategy leads to a lower number of handovers, while the behaviour for *Service 0* is different. This is in part due to the *ping-pong* effect which might loom in this particular strategy (a double change of access will not have any impact over the overall utility function), which is more relevant for *Service 0* flows, since they are longer.

Regarding the price that a user needs to pay (per traffic and time unit) per service, Figure 8 yields an interesting result. For *Service 0* we get the expected behaviour, and the Price strategy leads to lower prices (as compared to the RAT affinity one). However, for *Service 1* we can see that the RAT affinity strategy leads to prices which are slightly lower than those obtained for the price-based one. The reason is that the WiFi access points (due to their coverage) are less loaded than the cellular base station and thus the price they offer (following the previously presented pricing policy) is lower. Since services belonging to type 0 would stick to the cellular base station (no matter the price is), according to the RAT affinity criterion, the prices offered by WiFi access points are lower and therefore the price paid per service (type 1) is consequently lower. On the other hand, this result also reflects the fact that the Price-based strategy just seeks a global price reduction, without

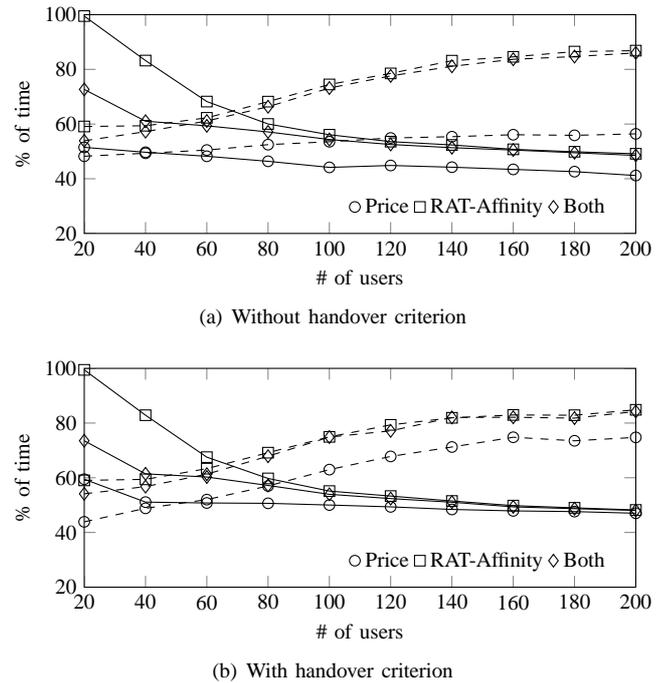


Fig. 9. Average RAT affinity per service Vs. the number of users. Continuous lines are for *Service 0* and dashed lines for *Service 1*

distinguishing between service types; the prices for both of them tend to be alike. In this case it is also worth highlighting that there is not a clear dependency on the handover criterion, since the results do not change between the utility functions that did not take it into account and those which did so. Besides, the figure also yields that the higher the carried load, the higher the price per service. It is also interesting to note that, for *Service 0*, the RAT Affinity strategy always leads to a higher price, since it aims at handling all the corresponding calls by the cellular base station, which gets fully loaded, even when the number of users is still small (see Figure 5(b)).

Figure 9 shows the probability for a service to be connected to the technology it has certain affinity to. We shall see that the use of the appropriate strategy does really have a relevant effect regarding this figure of merit, since the results obtained for the RAT Affinity strategy are much higher than the ones which were observed for the Price one. We can also observe the little influence of the handover criterion. Last, but not least, the obtained results also yield that *Service 1* gets better results than *Service 0* regarding the RAT affinity criterion for all the considered strategies. This is due to the fact that *Service 1* prefers using WiFi access points and as we have already discussed, these access alternatives are less loaded than the cellular base station; in addition, it is also worth bearing in mind that those services are shorter and it is therefore less likely that they would need to connect to the cellular base station due to user mobility after being connected to a WiFi access router.

In order to get a global view on the overall performance of the various strategies, Figure 10 uses a spider graph in which we represent the four figures of merit (for the two types of service). The edges of the different axes represent

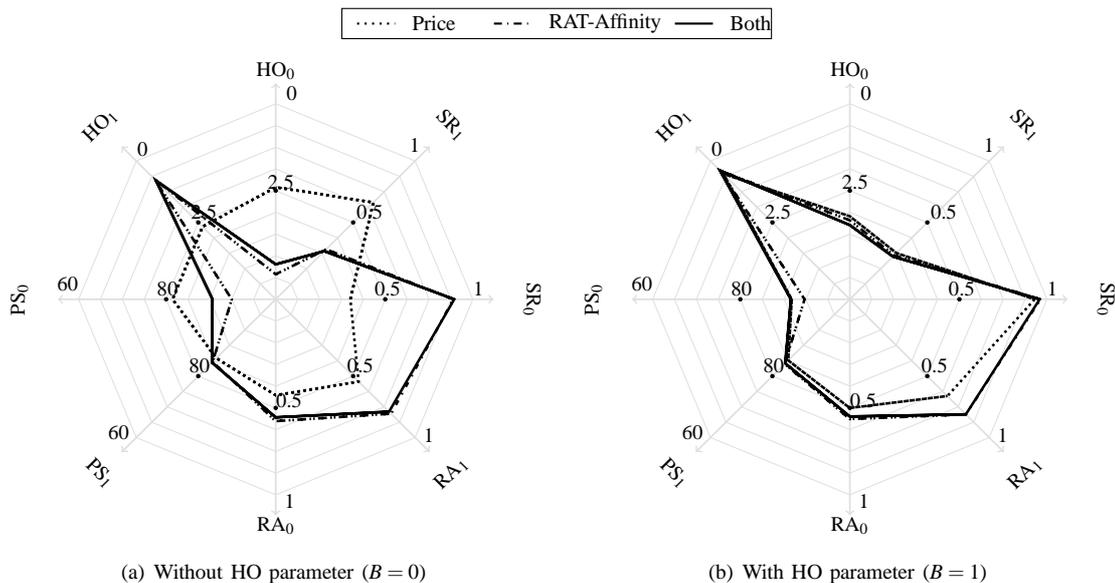


Fig. 10. Performance for the different utility functions. # of users was fixed to 100.

the best potential performance for the four of them, while the center of the spider can be considered as the worst result (for the handover parameter we used the empirically observed values to establish the performance bounds). The results were obtained for the scenario with 100 users. As was already discussed, we can see the improvement brought about by the integration of the handover criterion on the utility function, since the rest of criteria do not get clearly affected, while there is an enhancement in the number of handovers that are required per service. In the case of the Price strategy, we can also see that the use of the handover criterion brings about an improvement on the success rate for *Service 0*, although it is compensated by the decrease on the corresponding value for *Service 1*.

In general, all figures yield that the strategies considering the two goals (i.e. price and RAT affinity) lead to performances that are rather similar to those obtained for the one strengthening *just* RAT affinity (without jeopardizing the price per service too much), while for the Price-based strategy the performance in terms of RAT affinity are remarkably worse.

VII. CONCLUSIONS

In this paper we have proposed the use of binary linear programming techniques to assess the best performance which might be expected on highly heterogeneous access environments. The developed framework, which considers the time-evolution of services, is generic enough so as to extend its usage to different use cases. We have assessed its feasibility by means of a particular scenario, over which we studied strategies considering price, RAT affinity and cost of change (handover). The results show that an appropriate selection of the utility function criteria leads to rather different performance results, so this is an aspect that needs to be thoroughly considered when establishing the *goodness* of the assessed performance. In particular the performance figures that were previously reported yield that the integration of the *cost of*

change (handover) within the utility function brings about relevant benefits, since it leads to a significant improvement in terms of the average number of handovers, while it does not jeopardize the rest of figures of merit (in fact there were also some slight enhancements for certain cases).

Regarding future work, the main objective is to exploit the potential of the developed framework so as to assess the goodness of different distributed algorithms and procedures for access selection in highly heterogeneous networks, including the use of virtualized resources and flow management techniques. The idea would be to analyze how far their behaviour is from the optimum performance. Besides, in the analysis that has been discussed in Section VI, we have seen that there is an impact when considering different combinations of the utility function criteria; we have provided a discussion about the cost of change (handover) and its relationship with the price, but there are still other interrelations that require further analysis; this will be also considered in our future research. On the other hand, the design of the tool is rather generic and it is not rigid torn to the solver; it can be therefore used so as to mimic any potential scenario, as soon as the input traces follow the appropriate format. In this sense, we will also study the impact of considering other types of traffic, having more complicated models (for instance, elastic traffic), which might probably lead to non-linear optimization problems. Furthermore we will also enhance the abstractions that have been performed regarding the capacity; in this sense, we will modulate the capacity that a user gets from a resource based on the particular conditions of the link with the corresponding base station.

ACKNOWLEDGEMENTS

We would like to express our gratitude to the Spanish government for the funding in the following project: “Connectivity as a Service: Access for the Internet of the Future”, COSAIF (TEC2012-38574-C02-01). This work is an extension of the short paper that was presented in the *Wireless Days 2013*

conference. The authors are deeply grateful to the anonymous referees for the time they put on revising this work; their detailed review has certainly helped us to improve the overall quality of the paper.

REFERENCES

- [1] Ericsson, “Mobility report - on the pulse of the networked society,” June 2014.
- [2] E. Gustafsson and A. Jonsson, “Always best connected,” *Wireless Communications, IEEE*, vol. 10, no. 1, pp. 49 – 55, feb. 2003.
- [3] K. Pentikousis, R. Agüero, J. Gebert, J. A. Galache, O. Blume, and P. Pääkkönen, “The Ambient Networks heterogeneous access selection architecture,” in *Proceedings of the 1st Ambient Networks Workshop on Mobility, M2NM*, Octubre 2007.
- [4] Y. Zaki, L. Zhao, C. Goerg, and A. Timm-Giel, “LTE mobile network virtualization,” *Mob. Netw. Appl.*, vol. 16, no. 4, pp. 424–432, Aug. 2011.
- [5] U. Toseef, Y. Zaki, A. Timm-Giel, and C. Görg, “Uplink QoS Aware Multi-homing in Integrated 3GPP and non-3GPP Future Networks ,” in *Proceedings of the 4th ICST International Conference on Mobile Networks & Management, MONAMI’12*, September 2012.
- [6] L. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Goanddor, G. Auer, and L. Van Der Perre, “Challenges and enabling technologies for energy aware mobile radio networks,” *Communications Magazine, IEEE*, vol. 48, no. 11, pp. 66 –72, november 2010.
- [7] R. Agüero, L. Caerio, L. Correia, L. Ferreira, M. García-Arranz, L. Suciu, and A. Timm-Giel, “OConS: towards open connectivity services in the Future Internet,” in *Proc. of the 3rd ICST International Conference on Mobile Networks & Management, MONAMI’11*, Sept. 2011.
- [8] O. E. Falowo, S. Zeadally, and H. A. Chan, “Dynamic pricing for load-balancing in user-centric joint call admission control of next-generation wireless networks,” *International Journal of Communication Systems*, vol. 23, no. 3, pp. 335–368, 2010.
- [9] J. Choque, R. Agüero, and L. Munoz, “Optimum selection of access networks within heterogeneous wireless environments based on linear programming techniques,” *ACM Springer Mobile Networks and Applications*, vol. 16, no. 4, pp. 412 – 423, Aug 2011.
- [10] L. Pirmez, J. C. Carvalho Jr, F. C. Delicato, F. Protti, L. F. Carmo, P. F. Pires, and M. Pirmez, “Sutil–network selection based on utility function and integer linear programming,” *Computer Networks*, vol. 54, no. 13, pp. 2117–2136, 2010.
- [11] M. C. Lucas-Estañ, J. Gozávez, and J. Sanchez-Soriano, “Integer linear programming optimization of joint rrm policies for heterogeneous wireless systems,” *Computer Networks*, vol. 56, no. 1, pp. 112–126, 2012.
- [12] P. Xue, P. Gong, J. H. Park, D. Park, and D. K. Kim, “Radio resource management with proportional rate constraint in the heterogeneous networks,” *Wireless Communications, IEEE Transactions on*, vol. 11, no. 3, pp. 1066–1075, March 2012.
- [13] N. Nasser and H. Hassanein, “Dynamic threshold-based call admission framework for prioritized multimedia traffic in wireless cellular networks,” in *Global Telecommunications Conference, 2004. GLOBECOM ’04. IEEE*, vol. 2, Nov 2004, pp. 644–649 Vol.2.
- [14] “The GNU linear programming kit (GLPK),” www.gnu.org/software/glpk, 2012.