Optimization of Access Points Selection and Resource Allocation in Heterogeneous Wireless Network

Iaad BEN DHIA*, Mustapha BOUHTOU*, Taoufik EN-NAJJARY*,
Samer LAHOUD†, Xavier LAGRANGE‡
*Orange Labs Network, France, Chatillon
† ESIB, Saint-Joseph University of Beirut
‡ IMT Atlantique - IRISA D2, France, Rennes
Email: {iaad.bendhia, mustapha.bouhtou, taoufik.ennajjary}@orange.com
samer.lahoud@usj.edu.lb , xavier.lagrange@imt-atlantique.fr

***Abstract*—In this paper, we tackle the problem of access point selection and resource allocation in heterogeneous networks including 3G/4G and Wi-Fi. We study the user satisfaction optimization in terms of throughput and blocking rate. More precisely, the aim is to associate users with the optimal Radio Access Technology (RAT) and to allocate them the optimal number of Resource Units (RUs) based on their requested services and contracts. Starting with a non-linear problem, we re-formulate it into a Mixed Integer Linear Programming (MILP) and assess it on a realistic network configuration. We also compare it with a legacy strategy of selection and allocation. We show that our approach improves the user satisfaction while maintaining a reasonable blocking rate.**

I. Introduction

Heterogeneous networks are increasingly becoming an important feature of current wireless networks. With 5G, traffic forecasts predict an exponential growth of the throughput demand and all of new smart-phones are multi-technologies, which make the question of the best way of resource management for future heterogeneous networks one of the important issues for years. Indeed, in this context, for a giver User Equipment (UE) it is important to determine to which Base Station (BS) over which Radio access Technology (RAT) a UE will be associated, and how many Resource Units (RUs) are allocated. Moreover, harnessing additional spectrum in un-licensed bands by integrating Wi-Fi in the network could lead to a significant capacity gain.

In the state-of-art, network selection has received a lot of attention and several approaches are presented. In [1] authors present a survey of heterogeneous networks selection issues. Mainly, there are two classical approaches: the network-centric and the user centric. The network-centric approach takes decisions in a way to optimize overall network performance. For instance, in [2] authors modelled the user association issue as a linear optimization problem using the network-centric approach. In the user-centric approach, users select the appropriate access point without requiring any signalling overhead or coordination among the different access networks. For example, authors in [3] model the RAT selection problem in wireless heterogeneous networks as a non-cooperative game. Users try to maximize their own throughputs without regard for other clients. Another way to solve this classic problem is by using a hybrid approach. This takes into account both the user needs and network performance. In [4] authors use a Semi-Markov Decision Process to choose the best policy. Our work is based on [5], in this previous work the problem of selection is formulate as a maximization of the user satisfaction and solved by linear programming techniques. Several works addressed the problem of collaboration between the cellular network and the Wi-Fi. In [6] authors establish an empirical study of the throughput with a collaboration LTE/Wi-Fi. Using the theory of optimal control, authors in [7] propose a model for interface selection in heterogeneous network (LTE/Wi-Fi/Femto) and measured the benefit on the network congestion.

In the present work, we address the joint problem of user association and resource allocation in wireless heterogeneous networks. We investigate benefits of the optimization by a maximization of the overall user satisfaction and the number of users connected. User satisfaction is modelled by different utility functions that vary according to different service types and contracts. The optimization determines to which Base Station / Wi-Fi hotspot (HS) a given UE is connected and how many resources will be allocated to this UE. We formulate this problem as a Mixed Integer Linear Programming (MILP).

978-1-5386-3531-5/17/$31.00

The key contributions of our work are:

- Formulate a non-linear optimization problem to solve jointly the access point selection and resources allocation problem. We maximize both the overall user satisfaction and the number of connected users. This optimization considers the network performance and users preferences.
- Propose a formulation allowing collaboration between cellular networks and Wi-Fi and quantify this collaboration in terms of user satisfaction, throughput perceived and load balancing.
- Re-formulate the non-linear problem as a MILP. We solve it and give experimental results on a realistic network. Performances are compared with legacy approach.

This paper is organized as follows. Firstly, we describe the network model in Section II. Secondly, we present our user satisfaction model in Section III. Then, the mathematical formulation of the problem and results are presented in sections IV and V. Finally, conclusions and perspectives are given in Section VI.

II. Wireless Network Model: Resources and Throughput

A. Network Architecture and Resources

We consider a heterogeneous wireless network architecture that consists of N_{bs} base stations with N_T co-localized RATs, N_{bh} Wi-Fi hotspots and N_u users' equipment. The indexes $i \in I_{bs} = [1, \ldots, N_{bs}]$, $j \in J = [1, \ldots, N_T]$ and $k \in K = [1, \ldots, N_u]$ are used throughout the paper to designate a given BS, a given RAT and a given UE, respectively. The index $i \in I_{hs} = [1, \ldots, N_{bh}]$ is also used throughout the paper to designate a given hotspot (HS).
Cellular network resources are divided into Resource Units (RUs). Each RAT_j has a fixed number of RU denoted by R_j . For 3G networks, we assume that all codes have the same power and only codes are thus treated as RUs. In 4G, a resource block is the smallest RU that can be scheduled. This allows us to work with a linear formulation of the throughput perceived by the UE_k noted γ_k

B. Throughput Model

1) Cellular Throughput: Let γ_k^{cel} be the perceived throughput by UE_k using one or several cellular technologies. Let $\varphi_{i,j,k}$ be the perceived throughput of UE_k from BS_i over RAT_j per RU and $\lambda_{i,j,k}$ be the number of RUs assigned to UE_k associated with BS_i over RAT_j. The expression of γ_k^{cel} is given by:

$$\gamma_k^{cel} = \sum_{i \in I_{bs}\ j \in J} \lambda_{i,j,k} \varphi_{i,j,k}. \tag{1}$$

The theoretical value of unit throughput $\varphi_{i,j,k}$ is based on the Shannon formula. Let $\nu_{i,j,k}$ be the Signal-to-Interference-plus-Noise Ratio (SINR) of UE k from BS i over RAT j, and w_j the bandwidth per RU. The throughput that can be delivered to UE k from BS_i over RAT_j per RU is given by:

$$\varphi_{i,j,k} = w_j \log_2(1 + \nu_{i,j,k}). \tag{2}$$

As established in [5] the SINR is given by

$$\nu_{i,j,k} = \frac{G_j}{G_j(a + ISR_{i,j,k}) + L_{i,k}\frac{P_j}{P}} ,$$

where G_j is the transmit antenna gain for technology j and a the orthogonality factor ($a = 0$ in 4G), $L_{i,k}$ is the path loss between UE_k and BS_i, P_j is the noise power for a given technology and P is the power per RU. $ISR_{i,j,k}$ is the Interference to Signal Ratio of UE k from BS_i over RAT_j:

$$ISR_{i,j,k} = \sum_{i' \in I_{bs}\ ,\ i' \neq i} \pi_{i',j} \frac{L_{i,k}}{L_{i',k}} , \tag{3}$$

where $\pi_{i',j}$ is the percentage of resource used by the interfering $BS_{i'}$ over RAT_j.

2) Wi-Fi Throughput: As described in [8], with 802.11 protocol, the mechanism to acces the medium causes a decrease of the throughput according to the number of terminals. Let's consider a given HS_i. All users associate to to HS_i have the same throughput denoted by γ_k^{wifi} which is given by:

$$\gamma_k^{wifi} = \frac{1}{\sum_k \frac{1}{\chi_{i,k}}} , \tag{4}$$

with $\chi_{i,k}$ the peak rate of user k, namely the user's rate if he were the only one connected to the HS_i. With hypotheses of saturation, same back-off and same length of packet for all users, the equal sharing of the Wi-Fi 802.11 throughput is established in [9]. Mathematically, $\forall\ k \in K$ connected to HS_i we have $\chi_{i,k} = \chi_i$. Equation (4) becomes:

$$\gamma_k^{wifi} = \frac{\chi_i}{N_i} , \tag{5}$$

where N_i is the number of UE connected to HS_i. Finally, the perceived throughput γ_k is the sum of γ_k^{cel} and γ_k^{wifi}

III. The User Utility

As described in [10], several criteria or parameters (throughput, delay ...) could be chosen as indicator of satisfaction. In this paper, we assume that satisfaction is only a function of the throughput. We also differentiate users in two ways: their class of data traffic (elastic or non-elastic) and their contract (low cost or premium). As in [5] we use sigmoid and concave utility functions which map the UE perceived throughput with the level of user satisfaction.

A. Class of data traffic

We consider two types of traffic classes: non real-time (for elastic data use) and real-time (for non-elastic data use). Index s is used throughout the paper to designate a given class of a service. Let U_k^s be the user's k utility function with class s service.

Non real-time services ($s = \text{NRT}$) are generated by traditional data applications such as mail download, web surfing, etc. Thus, the elasticity of these services can be modelled by concave utility functions:

$$U_k^{\text{NRT}}(\gamma_k) = 1 - e^{-\frac{\gamma_k}{\gamma^c}} \tag{6}$$

where γ^c is the comfort throughput demand of the user (i.e., user satisfaction exceeds 63 % of maximum satisfaction). We note that the satisfaction increases slowly when the throughput exceeds the comfort throughput demand.

Real-time services ($s = \text{RT}$) that are generated by voice applications or video streaming. These services are non-elastic; it can be modelled by a sigmoid function:

$$U_k^{\text{RT}}(\gamma_k) = \frac{1 + e^{b\gamma^a}}{e^{b\gamma^a}}\left(\frac{1}{1 + e^{b(\gamma^a - \gamma_k)}} - \frac{1}{1 + e^{b\gamma^a}}\right) \tag{7}$$

where γ^a represents the average throughput demand of class B service, b is a positive constant that determines the shape of the sigmoid.

B. Type of Contract

We consider that the network operator provides two differentiated types of contracts, which are represented by index t. The Regular contract ($t = R$) and Premium contract ($t = P$) differ in the user's satisfaction. As we can see on figure 1, for a given class of service, to achieve the same level of satisfaction, a premium user ($t = P$) will require a higher throughput than a regular user ($t = R$). Thereafter, our model has to take into account these differences.

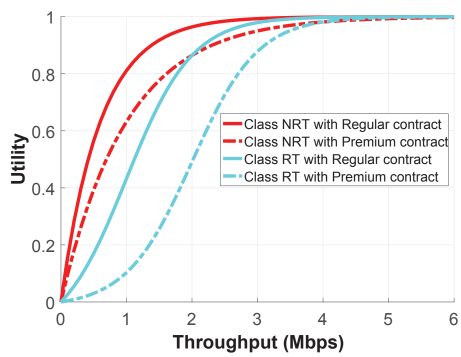

Figure 1. Utility function with different service and class (with parameters as shown in Table I)

IV. Optimization Problem

A. General Formulation

1) Decision Variables: We formulate a maximization problem of the overall satisfaction that is the sum of each user satisfaction. We optimize the access point selection (with binary variable of association $\theta_{i,j,k}$ and $\bar{\theta}_{i,k}$) and allocate the optimum number of resources to each users (with the decision variable $\lambda_{i,j,k}$). More precisely we have the decisions variables:

$$\theta_{i,j,k} = \begin{cases} 1 & \text{if } UE_k \text{ is associated with} \\ & BS_i \text{ over } RAT_j \\ 0 & \text{otherwise,} \end{cases} \tag{8}$$

$$\bar{\theta}_{i,k} = \begin{cases} 1 & \text{if } UE_k \text{ is associated with} \\ & \text{the hotspot } HS_i \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

2) Objective Function: To maximize the global user utility we need to take into account the level of UEs rejected. Therefore, we choose to penalise the objective function by a constraint of blockage. This penalized objective function is composed by the utility part and the blocking rate part:

$$\sum_{s,t,k} \alpha^{s,t} U_k^{s,t}(\gamma_k) - \beta(N_u - (\sum_{i,j,k} \theta_{i,j,k} + \sum_{i,k} \bar{\theta}_{i,k})). \tag{10}$$

The first term is the weighted utility sum which represents the global users' satisfaction. Weights $\alpha^{s,t}$ are fixed according to the operator strategy, that consists to choose which group of users are prioritised for the optimization.
The second term allows having a solution with a low level of blockage. Parameter $\beta \geq 0$ is a penalty coefficient. Theoretically when $\beta \to \infty$ the blocking rate decreases to the minimal feasible level. In section V, we do some sensitivity analysis on β in order to find an acceptable tradeoff. For more detail about penalty theory, one can refer to [11].

3) Constraints: This maximization is under constraints on decision variables and utility functions. Let $u_{min}^{s,t}$ and $u_{max}^{s,t}$ be the minimal and maximal required utility, respectively, and R_{max} the maximal number of user connected to a given HS. Let $\rho_{i,j,k}$ and $\bar{\rho}_{i,k}$ be the coverage parameters. These are equal to 1 if user k is covered by BS_i over RAT_j or by HS_i, respectively and 0 otherwise. Constraints are:

$$\sum_{i \in I_{bs}, j \in J} \theta_{i,j,k} + \sum_{i \in I_{hs}} \bar{\theta}_{i,k} \leq 1 \ , k \in K \tag{11}$$

$$\sum_{k \in K} \lambda_{i,j,k} \leq R_j \ , \ \forall i \in I_{bs} \ j \in J \tag{12}$$

$$1 \leq \sum_{k} \bar{\theta}_{i,k} \leq R_{max} \ , \ i \in I_{HS} \tag{13}$$

$$\theta_{i,j,k} \leq \lambda_{i,j,k} \leq R_j \theta_{i,j,k} \ , i \in I_{bs}, j \in J, k \in K \quad (14)$$

$$u^{s,t}_{min}(\theta_{i,j,k} + \bar{\theta}_{i,k}) \leq U^{s,t}_k \leq u^{s,t,}_{max}(\theta_{i,j,k} + \bar{\theta}_{i,k}) \quad (15)$$

$$i \in I_{bs} \cup I_{hs}, j \in J, k \in K, s \in \{\mathrm{RT}, \mathrm{NRT}\}, t \in \{\mathrm{R}, \mathrm{P}\}$$

$$\gamma_k = \sum_{i,j} \lambda_{i,j,k}\varphi_{i,j,k} + \sum_i \chi_i \frac{\bar{\theta}_{i,k}}{\sum_{k'} \bar{\theta}_{i,k'}} \ , \ k \in K \quad (16)$$

$$\theta_{i,j,k} \leq \rho_{i,j,k} \ , \ i \in I_{bs} \ , j \in J \ , k \in K \quad (17)$$

$$\bar{\theta}_{i,k} \leq \bar{\rho}_{i,k} \ , \ i \in I_{HS} \ , k \in K \quad (18)$$

$$\theta_{i,j,k} \in \{0,1\} \ , \ \bar{\theta}_{i,k} \in \{0,1\} \ , \ \lambda_{i,j,k} \in \mathbb{N}. \quad (19)$$

Constraints (11) to (14) are about user association and resources allocation on the cellular network. Constraints (11) state that a given user k can be connected to at most one BS over one RAT or to one HS. These constraints introduce the coupling between the cellular and the Wi-Fi in the selection. Constraints (12) are capacity constraints: they ensure that a UE cannot be assigned to more resource than available. Constraints (13) allow setting a minimal and a maximal number of users on each HS. Constraints (14) guarantee that at least, one RU is given when a terminal is connected to a RAT and at last R_j RUs. Constraints (15) guarantee a level of satisfaction for users connected on the network. Then constraints (16) give the throughput expression with the cellular part and the Wi-Fi part. Constraints (17) and (18) allow connecting only UE if they are covered by a BS/RAT or a HS. In conclusion, we have to solve a non-linear combinatorial optimization problem (P) that is given by:

$$(P) : \mathrm{Max} \sum_{s,t,k} \alpha^{s,t} U^{s,t}_k - \beta(N_u - (\sum_{i,j,k} \theta_{i,j,k} + \sum_{i,k} \bar{\theta}_{i,k}))$$

subject to: (11) to (19)

This is a combinatorial and non-linear problem. One of the aims of this work is to study the problem of selection and allocation on a large and realistic network with a lot of users. That causes a heavily increase of combinatorial of the problem. Furthermore, using sigmoid and concave function for objective function and our Wi-Fi throughput modelling (constraint (16)), are at the origin of the non-linearity of (P). In the next section we re-formulate the problem into a linear equivalent formulation. Having a linear formulation is an important task to solve effectively the problem (P) with classic solvers. In general, MILP are solved using a linear-programming based on branch-and-bound approach [12].

B. Linear Equivalent Formulation

Transforming problem (P) into a MILP problem (P_1) consists in developing an equivalent expression of (P) where objective function and all constraints are linear. In order to obtain a linear objective function we use the discrete throughput formulation. Equation (1) gives a discrete formulation of the cellular throughput. Let $r \in R^j = \{1, \ldots, R_j\}$ be the number of RUs allocated to a UE over a RAT_j. All utilities' possibilities are computed in parameter $u^{s,t}_{i,j,k,r}$ which is the utility value if UE_k is associated with BS_i over RAT_j with n RU. We define a new binary variable $\theta_{i,j,k,r}$ that equals to one if UE_k is connected to BS_i over RAT_j with r RUs and 0 otherwise. Thereby, $\lambda_{i,j,k}$ is given by:

$$\lambda_{i,j,k} = \sum_{n \in R^j} n \ \theta_{i,j,k,r} \ , i \in I_{bs}, j \in J, k \in K \ . \quad (20)$$

This formulation gives a linear expression of the users' satisfaction if they are connected to the cellular network. Equation (5) gives a discrete formulation of the Wi-Fi throughput. Indeed, with hypothesis of equal sharing we just need to have the number of UEs connected to a given HS_i to have the throughput perceived by all UEs connected. Let be $n \in R^{HS} = \{1, \ldots, R_{max}\}$ the number of UEs which can be connected to a HS. All utilities' possibilities are computed in parameter $\bar{u}^{s,t}_{k,n}$, which is the utility value if UE_k is connected to an HS with $n-1$ other UEs and perceived the throughput $\frac{\chi_i}{n}$. Let also introduce a new binary variable $\bar{\theta}_{i,k,n}$ equalling to 1 if UE_k is connected to HS_i with $n-1$ other UEs and 0 otherwise. Thereby, the number of users connected to HS_i is given by:

$$\sum_{k'} \bar{\theta}_{i,k'} = \sum_{n \in R^{HS}} n \ \bar{\theta}_{i,k,n} \ , i \in I_{hs}, k \in K \ . \quad (21)$$

This formulation gives a linear expression of the users' satisfaction if they are connected to the Wi-Fi network. Finally, we obtain the formulation of the MILP problem (P_1) given by:

$$(P_1) : \mathrm{Max} \sum_{s,t,i,j,k,n} \alpha^{s,t} \theta_{i,j,k,n} u^{s,t}_{i,j,k,n}$$

$$+ \sum_{s,t,i,k,n} \alpha^{s,t} \bar{\theta}_{i,k,n} \bar{u}^{s,t}_{k,n} - \beta \ (N_u - \sum_{i,j,k} \theta_{i,j,k} - \sum_{i,k} \bar{\theta}_{i,k})$$

subject to:

$$\sum_i \bar{\theta}_{i,k} + \sum_{i,j,r} \theta_{i,j,k,r} \leq 1 \quad , \ k \in K \quad (22)$$

$$\lambda_{i,j,k} = \sum_{r \in R^j} r \ \theta_{i,j,k,r} \ , i \in I_{bs}, j \in J, k \in K \quad (23)$$

$$\sum_k \bar{\theta}_{i,k} = \sum_n n \ \bar{\theta}_{i,k,n} \qquad , \ i \in I_{hs} \quad (24)$$

$$u^{s,t}_{min}\theta_{i,j,k} \leq \sum_r \theta_{i,j,k,r} u^{s,t}_{i,j,k,r} \leq u^{s,t}_{max}\theta_{i,j,k} \quad (25)$$

$$u^{s,t}_{min}\bar{\theta}_{i,k} \leq \sum_n \bar{\theta}_{i,k,n}\bar{u}^{s,t}_{k,n} \leq u^{s,t}_{max}\bar{\theta}_{i,k} \quad (26)$$

$$\theta_{i,j,k,r} \leq \theta_{i,j,k} \ , i \in I_{bs}, j \in J, k \in K, r \in R^j \quad (27)$$

$$\bar{\theta}_{i,k,n} \leq \bar{\theta}_{i,k} \ , \ i \in I_{hs} \ k \in K, \ n \in R^{HS} \quad (28)$$

$$\theta_{i,j,k} \leq \rho_{i,j,k} \quad , i \in I_{bs} \ , j \in J \ , k \in K \quad (29)$$

$$\bar{\theta}_{i,k} \leq \bar{\rho}_{i,k} \quad , i \in I_{HS} \quad , k \in K \quad (30)$$

$$\theta_{i,j,k} \in \{0,1\} \quad , \bar{\theta}_{i,k} \in \{0,1\} \quad , \lambda_{i,j,k} \in \mathbb{N} \quad (31)$$

$$\theta_{i,j,k,r} \in \{0,1\} \quad , \bar{\theta}_{i,k,n} \in \{0,1\} \quad (32)$$

Constraints (22) state that a given UE_k can be connected to at most one BS over one RAT with n RUs or to one HS. Constraints (23) ensure coupling of the number of resources allocated to UE_k associated to BS_i over RAT_j which is $\lambda_{i,j,k}$ and the index r. Constraints (24) ensure coupling of the number of UEs associate to a given HS_i and the index n. Constraints (25) and (26) guarantee a level of satisfaction for users connected on the network. Then constraints (27) and (28) ensure coupling between $\theta_{i,j,k,r}$, $\bar{\theta}_{i,k,n}$ and $\theta_{i,j,k}$, $\bar{\theta}_{i,k}$. Constraints (29) and (30) allow to connected only UE if they are covered by a given BS_i/RAT_j or a given HS_i.
We have to solve (P_1), the linear formulation of the problem using operational research algorithm as branch-and-bound algorithm. We solve the issue of non-linearity of (P) but that increases the size of the problem thus increases the combinatorial too.

V. Performance Evaluation

A. Network Topology and Setting

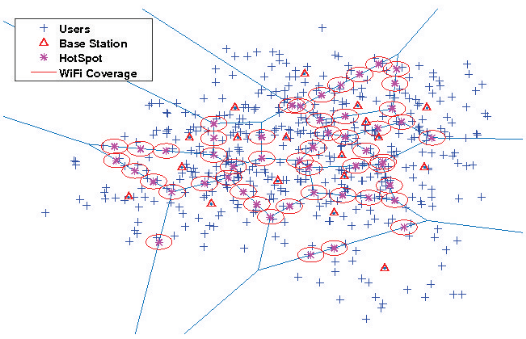

Figure 2. Network Topology

As in [13], we use the positioning of Orange's operator base stations of the 14^{th} district of Paris. It is composed of 18 base stations of two co-localized technologies namely, LTE and HSDPA. In each cell, UEs are placed using a random uniform distribution. As shown in figure 2, we consider a network topology as Voronoi cells; Wi-Fi hotspots are placed on the edge of the cells in order to cover users with low SINR. For our simulation, we use the same setting as in [5]. Table I, Table III and Table III summarize the simulation parameters.

	Class NRT	Class RT
Weighting $\alpha^{s,t}$		
Regular	0.15	0.30
Premium	0.20	0.35
User Repartition		
Regular	40%	10%
Premium	20%	30%
Bounds of utility		
$u^{s,t}_{min}$	0.0%	10 %
$u^{s,t}_{max}$	77.70%	95.00 %
Comfort/Average Rate		
Regular	$\gamma^a = 0.6$	$\gamma^c = 1.0$
Premium	$\gamma^a = 1.0$	$\gamma^c = 2.0$

Table I - Simulation parameters

	4G	3G
Number of RUs	48	14
Carrier frequency (in MHz)	2000	1900
Bandwidth (in MHz)	10	5
Orthogonality factor (a)	0	0.5
Occupied resources (π_j)	80 %	90%
Transmit power (in Watt)	10	10
Antenna gain G_t (in dBi)	15	15
Noise Figure (in dB)	9	9
Shadowing Deviation (in dB)	10	10
Cells Radius (in meter)	700	500

Table II - Simulation parameters for LTE, HSDPA

Parameter	Cells Radius	R_{max}	χ_i
Value	100 m	10	7.7 Mb/s

Table III - Simulation parameters for Wi-Fi

B. Legacy Approach

We use the Highest Received Power (HRP) strategy as reference to assess our model of optimization. As described in [14] with HRP, UEs are connected to BS on RAT that deliver the highest power and starting the selecting with the last network generation. Furthermore, with HRP, RUs are distributed equitably between all UE connected to a BS/RAT. The HRP model works as follows:

- UE put in a descendent order queue each RAT_j in term of the received power for the covering BSs/RATs
- UE is associated with the first BS/RAT_1 that accept its request.
- If no BSs in the RAT_1 queue accept to connect the UE, the same procedure is executed on RAT_2 queue.

C. Experimentation and Results

Different instance of the MILP problem (P_1) are solved using CPLEX V12.6.0.0 solver. An instance is a random distribution of the users' position. All of the presented solutions are average solutions of the different instance. Through simulations we present a comparison of our approach with the legacy approach (HRP) and discuss the impact of Wi-Fi in term of user satisfaction, percentage of served UEs, and resource management. We discuss the load balancing between the different technologies and the computation time. Throughout this section we use *moderately loaded network* to design network with less of 420 UEs and *very loaded network* for network with more than 420 UEs.
The choice of the penalisation parameter β can have a large impact on the global satisfaction. We determine and choose the range of its acceptable values, the one that at the same time minimizes the blockage and maximizes the global satisfaction.

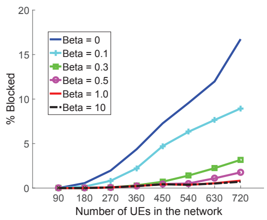

Figure 3. Blockage for different values of β

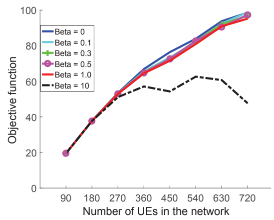

Figure 4. Objective function for different values of β

We perform a sensitivity analysis to measure the impact of β on the blockage and the objective function. As we shown in figure 3, for small values of β, there are a high percentage of blockages. When β value is more than 1, we reduce the blockage to a value very close to 0 %. In figure 4 we show that for all small values of β (less than 1), there is no important degradation of the objective function. For the remaining analysis we use $\beta = 1$, because it is a good trade-off between the average satisfaction and the average blockage.

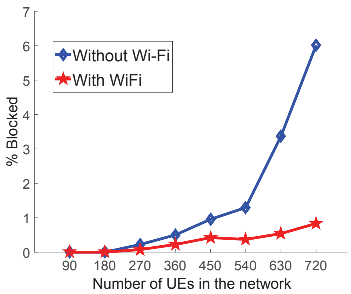

Figure 5. Blocking percentage with $\beta = 1$

1) Percentage of served UEs: Figure 5 shows that the optimization problem without Wi-Fi provides satisfactory results in term of blockage for moderately loaded network but for very loaded network the blockage have a high increase. For all configurations of load of network, having Wi-Fi allow to reduce blockage very close to 0 %.

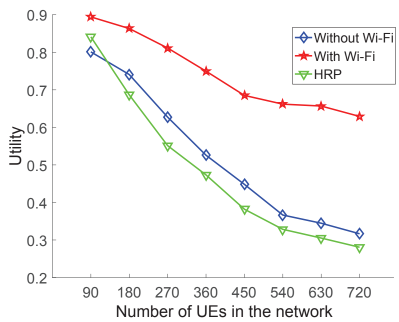

Figure 6. Average utility with $\beta = 1$

2) User satisfaction: Figure 6 shows that our model allows us to achieve a higher level of satisfaction compare to HRP. Without Wi-Fi, this improvement is approximately of 10 % compare to HRP results. However, we observe that the average utility decreases sharply for very loaded network. Introduction of Wi-Fi allows a high increase of the average satisfaction. Indeed, with Wi-Fi, there is an average utility above 65% whereas without Wi-Fi, the utility can go down to 30%.
To refine our analysis, we plot the average throughput and utility by group (class and contract). As results, we can see on figure 7 that we have an equitable repartition of the throughput. As we see on figure 8, optimization gives precisely the need of throughput for a given group. Indeed, average utility by group is very close between the premium streaming group and the regular streaming group. For these simulations we choose a strategy (with $\alpha^{s,t}$) which privileges premium contract in term of throughput to guarantees an equitable level of satisfaction between all users connected.

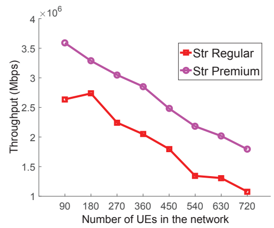

Figure 7. Average Throughput for Elastic Service

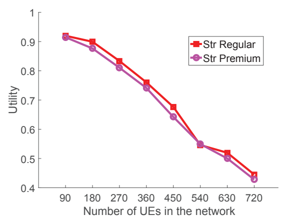

Figure 8. Average Utility for Elastic Service

3) Load balancing: Table IV shows the distribution of users over different technologies. As expected, HRP method associate more part of user to 4G (90%), a small part to 3G (10%) and no UE to Wi-Fi. That creates a congestion of the 4G. With the optimization we have a better load balancing between technologies.

	HRP	Without Wi-Fi	With Wi-Fi
4G	89%	77%	49%
3G	10%	22%	12%
Wi-Fi	0%	0%	38%

Table IV - User repartition with 540 users

4) Computation time: We compute the optimal solution of the MILP problem (P_1) using the CPLEX solver running on a computer equipped with an Intel(R)Xenon(R)CPU L5630, 4 cores. This tool uses the branch and cut approach [15].

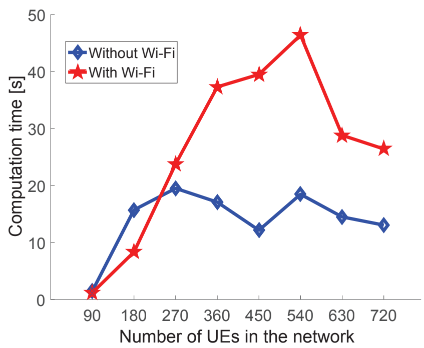

Figure 9. Computation time, with and without Wi-Fi

As we can see in figure 9, the average computation time varies between 10 to 20 seconds without Wi-Fi and between 20 to 40 seconds with it. This is an encouraging result. Indeed, CPU time is relatively stable with respect of number of UE. In spite of the important combinatorics due to the introduction of Wi-Fi and the linear reformulation of the problem, computations don't have an exponential behaviour.

VI. Conclusion

In this work, we presented an optimization approach for RAT selection (3G,4G,Wi-Fi) while taking into account user needs and network resources. We formulated the access point selection and resource allocation as a MILP problem. This MILP is a maximization problem considering a double objective: maximize the average utility and the number of users connected. Through simulations using a realistic network, our approach demonstrated a better performance in term of users' satisfaction, blockage and load balancing between 3G/4G and Wi-Fi. For future work, we will study the problem of selection in a dynamic network taking into account user's arrival, departure and mobility. The question of the handover will be a very big task for the rest of this work. We also plan to reduce the computation and consider the energy efficiency of the network.

References

[1] Lusheng Wang and Geng-Sheng GS Kuo. Mathematical modeling for network selection in heterogeneous wireless networks—a tutorial. *IEEE Communications Surveys & Tutorials*, 15(1):271–292, 2013.

[2] Johnny Choque, Ramón Agüero, and Luis Muñoz. Optimum selection of access networks within heterogeneous wireless environments based on linear programming techniques. *Mobile Networks and Applications*, 16(4):412–423, 2011.

[3] Parnia Naghavi, Seyed Hamed Rastegar, Vahid Shah-Mansouri, and Hamed Kebriaei. Learning rat selection game in 5g heterogeneous networks. *IEEE Wireless Communications Letters*, 5(1):52–55, 2016.

[4] Melhem El Helou, Marc Ibrahim, Samer Lahoud, Kinda Khawam, Dany Mezher, and Bernard Cousin. A network-assisted approach for rat selection in heterogeneous cellular networks. *IEEE Journal on Selected Areas in Communications*, 33(6):1055–1067, 2015.

[5] Farah Moety, Mustapha Bouhtou, Taoufik En-Najjary, and Riadh Nasri. Joint optimization of user association and user satisfaction in heterogeneous cellular networks. ITC, 2016.

[6] Y. Khadraoui, X. Lagrange, and A. Gravey. Very Tight Coupling between LTE and WiFi: a Practical Analysis. In *CoRes 2016 : Rencontres Francophones sur la Conception de Protocoles, l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication*, pages 1–4, 2016.

[7] Yue Li, Yassine Hadjadj-Aoul, Philippe Bertin, and Gerardo Rubino. A control theoretic strategy for intelligent interface selection in heterogeneous network environments. In *2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, pages 1006–1011. IEEE, 2016.

[8] Martin Heusse, Franck Rousseau, Gilles Berger-Sabbatel, and Andrzej Duda. Performance anomaly of 802.11 b. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, volume 2, pages 836–843. IEEE, 2003.

[9] Anurag Kumar, Eitan Altman, Daniele Miorandi, and Munish Goyal. New insights from a fixed point analysis of single cell ieee 802.11 wlans. In *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, volume 3, pages 1550–1561. IEEE, 2005.

[10] Andrew Campbell, Geoff Coulson, and David Hutchison. A quality of service architecture. *ACM SIGCOMM Computer Communication Review*, 24(2):6–27, 1994.

[11] Joseph-Frédéric Bonnans, Jean Charles Gilbert, Claude Lemaréchal, and Claudia A Sagastizábal. *Numerical optimization: theoretical and practical aspects*. Springer Science & Business Media, 2006.

[12] Jens Clausen. Branch and bound algorithms-principles and examples. *Department of Computer Science, University of Copenhagen*, pages 1–30, 1999.

[13] Farah Moety, Samer Lahoud, Bernard Cousin, and Kinda Khawam. Joint power-delay minimization in 4g wireless networks. In *Wireless Days (WD), 2014 IFIP*, pages 1–8. IEEE, 2014.

[14] Evolved Universal Terrestrial Radio Access. and evolved universal terrestrial radio access network (e-utran). *Overall description*, 126, 2008.

[15] John E Mitchell. Branch-and-cut algorithms for combinatorial optimization problems. *Handbook of applied optimization*, pages 65–77, 2002.