A Load-Based and Fair Radio Access Network Selection Strategy with Traffic Offloading in Heterogeneous Networks

Juan Montoya
Department of Informatic and Systems
Universidad EAFIT
Email: jcmontoy@eafit.edu.co

Adarshpal Sethi
Department of Computer and Information Sciences
University of Delaware
Email: sethi@udel.edu

Natalia Gaviria Gómez
Department of Electronics Engineering
Universidad de Antioquia
Email: natalia.gaviria@udea.edu.co

***Abstract*—Next Generation Mobile Networks are envisioned to integrate and coordinate Heterogeneous Networks with the aim to cope with the new mobile traffic demands by taking advantage of the features of each wireless network. In this context, it is accepted that a centralized management resource framework is required by Mobile Network Operators in order to efficiently manage the scarce and limited radio resources from each wireless network. One of the main challenges for radio resource management architecture is the network selection function. We investigate the problem of Radio Access Network selection for a Heterogeneous Networks scenario with the objective of distributing the traffic among several Radio Access Networks in a fair way. The problem is theoretically modeled as a sequential decision-making problem using Semi Markovian Decision Process, where the optimization problem seeks to maximize the long-term discounted reward of the Heterogeneous Network system. In addition to this, taking advantage of the departure of sessions, we have considered a load distribution process that allows offloading of traffic to alleviate the load of the macro cell. In order to solve the model and obtain an optimal policy, we have used Value Iteration algorithm. From the resulting policy, the blocking probability for each possible event in the system is calculated. Several simulations were carried out, and the obtained results indicate that the proposed network selection strategy exhibits good performance for distributing the traffic load in a fair way among several wireless networks.**

Keywords*—*Radio Access Network Selection, HetNet, Optimal Policy, Traffic Offloading

I. Introduction

The accelerated growth in mobile communication services that we have experienced over the past decade has led to the evolution and creation of new mobile devices as well as new applications and services. These have been accompanied by the evolution of new and multiple wireless network technologies, which must be carefully designed and deployed by Mobile Networks Operators (MNOs) in order to satisfy new user requirements and to cope with increasing traffic demands.

Particularly, with regards to the evolution of wireless networks, the concept of New Generation Mobile Networks (NGMN) implies the coexistence, integration and cooperation of multiple and overlapping Heterogeneous Networks (HetNet). In this context, the base stations of the system can differ in terms of the technology and the size [1], i.e., combination of a macro cell (MC) (using a technology like UMTS, HSPA, LTE) with small cells (SCs), which can operate in licensed or unlicensed bands.

In the aforementioned scenario, one of the big challenges inherent to the features of each Radio Access Network (RAN), is the implementation of centralized and optimal resource management strategies with the aim of guaranteeing aa efficient use the resources. For this reason, academic and industry communities have been making big efforts in research to design a full set of functionalities to achieve a real and true integration of the multiple RANs in a successful HetNet environment. One of the major concerns in this architecture is the design of the RAN selection function [2], which is in turn a difficult problem to solve due to the number of variables and networks involved.

In general, RAN, network or cell selection can be understood as the ability to choose the best RAN to allocate an incoming session taking into consideration one or multiple attributes [3]. In this sense, the design of optimal network selection policies is a major issue for MNOs. The main reason behind this is that the right choice with regards to which RAN should be used for hosting an incoming session, has an important effect over the performance of the system in terms of the quality of service, as well as in the efficiency of radio resource utilization. In the same way, in the context of HetNets, pushing traffic from the MC to the SCs has become an attractive strategy to avoid an overloaded MC and underutilized SCs [4]. This strategy has been shown to enhance the overall performance of the system, and therefore it must be carefully designed in order to achieve this goal and thus prevent a misuse of the resources (for a survey see [5]).

Taking these issues into consideration, it is critical for MNOs to implement a RAN selection and offloading strategy, which leads to the enhancement of the global performance of the system in terms of the overall utilization of radio resources, while guaranteeing a level of quality of service. Even though the RAN selection problem has been widely studied, and traffic offloading strategies have gained a lot of attention, there are still open issues arising from the technological challenges involved in the implementation of a fully converged and centralized framework in HetNets to prevent overload scenarios. It is hence necessary and justified to continue carrying out research works that allow us to formally study and analyze the problem.

A. Purpose

Our purpose in this article is to study and analyze the RAN selection problem in a HetNet scenario. Since the network selection service is considered a key functionality in this context, our work focuses on this process from an MNO's perspective. In this way, a network-centric strategy is proposed to evaluate and decide which RAN of a HetNet system will be used to host an arriving session. Specifically, we are interested in modeling the RAN selection problem as a sequential decision-making process, where the goal of the optimization problem is to maximize the long-term discounted reward by dynamically allocating the incoming sessions to any of the available RANs. To achieve this, the reward perceived by the system is calculated by considering the type of event taking place and its priority, as well as the level of fairness of the overall usage of resources. The goal of the strategy is to guarantee an efficient use of each of the RANs in the system.

In our proposal, this goal is met by taking advantage of the session departures, such that we perform a network-based load distribution procedure (i.e., traffic offloading), resulting in a fair allocation of the load among the RANs, as well as an efficient usage of the radio resources over the long-term. Unlike conventional strategies, the process allows offloading traffic from the MC to the set of SCs if and only if the load-based fairness is increased, by using a network-based load distribution process.

Semi Markovian Decision Process (SMDP) has been chosen as the tool for modeling the problem. In this manner, the network selection entity is modeled as a decision-maker, who is in charge of selecting the RAN to allocate an incoming session. With the aim to solve the SMDP model, and in order to obtain the optimal policy for the RAN selection problem, a dynamic programming algorithm, particularly Value Iteration (VI), is used. Finally, an analysis is made for the model proposed and a performance comparison is carried out with random and greedy strategies in terms of the system performance for the problem stated.

B. Contributions

The main contributions of this work are summarized as follows:

- The RAN selection problem for HetNet is studied and formulated as a sequential decision-making process. In our case, SMDP has been chosen as the decision-theoretic framework, where the general goal is to maximize the long term discounted-reward criterion. Thus, for every incoming session, the fairness of the system in terms of the normalized load is evaluated in order to manage the radio resources in an efficient way.
- We extend the SMDP model for the RAN selection process by taking advantage of the session departures and evaluating whether a load distribution process can be carried out when these events occur. The goal is to improve the efficiency of the system in terms of the fair usage of the radio resources as well as the quality of service, which is measured in terms of the blocking probability. This process is implemented in order to avoid an overloaded MC situation by alleviating its load, and thus achieving a similar or lower level of utilization as compared to the SCs.
- One of the main interests with this research is to offer an analysis as well as a performance evaluation of the optimal policy for the network selection problem. We use the optimal policy obtained in a simulated environment in order to evaluate the performance of this policy.

Numerical and simulation results indicate that the obtained optimal network selection policy exhibits a lower blocking probability and better radio resource utilization in comparison to greedy and random strategies.

C. Organization of this paper

The rest of the paper is organized as follows. In Section II, we offer a summarized review of prior works in the field of RAN selection problem. The modelling of the system is presented in Section III. Section IV presents the SMDP formulation for the RAN selection problem in HetNet. The solution for the SMDP model is presented in Section V. In Section VI, the performance metrics to be used in the evaluation process are defined. Numerical and simulation results are presented in Section VII. Finally, Section VIII concludes the paper.

II. Related Work and Motivations

The RAN selection problem in a HetNet scenario has been studied in the literature using different approaches, not only from the modeling, but also from the architectural point of view. In this sense, the decision-making architecture can be either network-centric or user-centric. Since we are proposing a network-centric strategy, we present a summary of the works related to this architecture.

Some previous works have reported the use of MDP as the decision-theoretic framework to model network-centric approaches for the network selection problem. In [6], the authors consider the optimal radio selection of a heterogeneous wireless environment composed by two RANs and transmitting voice and data traffic. The optimization functions are expressed in terms of the blocking probability and throughput of each service. Based on a numerical analysis, the authors derive heuristic policies, which are compared to the optimal one.

Considering real-time and non-real-time traffic in a HetNet, the authors in [7] formulate an optimization problem for network selection based on an SMDP framework. The goals in this work are to maximize the system capacity and to distribute the traffic in an efficient way. To achieve this, the optimization function is defined in terms of the blocking cost function as well as an alternative acceptance cost function with the aim to perform the optimal decision. Similarly, in [8], the Joint Call Admission Control problem is formulated as an SMDP problem and an interior point method is used to solve it. For this network-centric approach, where two different RANs are considered, the maximization average network revenue is defined as the main goal taking into account several constraints, such as handoff call-blocking probability, the fairness among heterogeneous terminals, as well as the fairness and priority of different traffic classes. Likewise, in [9], MDP is also used to model the network selection problem for a HetNet scenario with two RANs (i.e., LTE and WiFi). The main goal of the problem is to maximize the revenue of the overall system.

Unlike [6], we use SMDP as the decision-theoretic framework. In this sense, our work is similar to [7] [8] [9], because we use the same mathematical tool to analyze the system in continuous-time. In contrast to these works, however, we pursue a different objective and perform a load distribution or traffic offloading process.

In the field of HetNet, there is a set of proposals that involve the network selection problem with load balancing process or traffic offloading strategies, which have been reported in the literature. In [10], the authors study the problem of Joint Radio Resource Management in HetNet with the aim of obtaining an energy-efficient control policy. They propose two schemes that leat to optimal and near-optimal policies using SMDP and Continuous Time Markov Chains (CTMC), respectively. The strategy is based on defining a threshold related to the load level of the MC in order to turn the SCs on/off and then perform a load balancing process between the available RANs for saving energy. Our work differs from that one in the goal and the optimization function. We are only interested in offloading ongoing sessions from the MC to one of the SCs for achieving a better and fairer resource utilization of the set of RANs, whereas they seek to reduce the power level consumption.

In [11], a centralized cell association scheme with a load balancing scheme is proposed in a HetNet. The main objective is to avoid a misuse of the radio resources through cell association, taking into account the load of each RAN. A load balancing process is thus defined, and only performed when the system is unbalanced, which occurs when the value of the load fairness index is lower than a pre-established threshold. The effectiveness of this proposal is verified through simulations.

In [12], the authors formulate the network selection problem jointly considering user and base station information. The work combines information such as Channel Quality Indicator, traffic load and available bandwidth to perform the decision for the network selection problem. Simulation results show that their proposal outperforms, in terms of service blocking ratio and average throughput, when compared to schemes that only take into account either user or base station information.

Although previous research works in the field of network selection strategies have made significant contributions, our proposal introduces a new network-centric scheme for the problem, which allows the distribution of the incoming sessions in a fair way among several RANs, whereas the long-term reward is maximized. In addition to this, at departure times, we propose to perform a load distribution process that allows traffic offloading from the MC to one of the SCs. To the best of our knowledge, previous works do not consider jointly fair distribution of incoming sessions among the RANs together with a load distribution procedure at departure times for alleviating the load level of the MC. Particularly, in this regard, our proposal differs from the previous ones because both the admission and offloading decision criteria are conditioned for increasing the fairness of the system with the aim of improving the overall efficiency of radio resource utilization and QoS offered to the users.

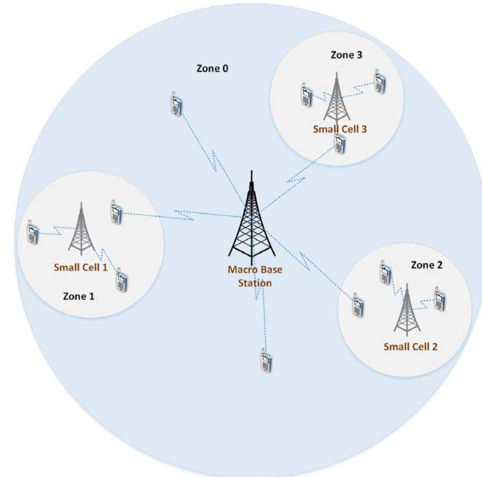

Fig. 1. Illustration of HetNet system.

III. System Model and Description

A. Network model

In general, a HetNet is a system composed of multiple and different wireless networks, which are deployed in a specific area. In this context, we are focusing on modeling the downlink of a HetNet scenario, where a single macro cell (MC) is overlaid by multiple small cells (i.e., micro, metro, pico, femto cells) located in the range of the MC. Let $B = M \cup K$ denote the set of all RANs in the system where $M = \{0\}$ is a singleton that represents the only MC in the system and $K = \{1, 2, ..., k\}$ denotes the set of SCs that belong to the HetNet, $|K| = k$ and $|B| = k + 1$. In this work, $i \in \{1, 2, ..., k\}$ is used to denote the i^{th} SC and the index $i = 0$ will be used to denote the MC. It is important to note that the set of base stations are located in fixed locations and each one has a high-speed backhaul link to the core network. Likewise, the MC is operating in lower cellular bands and, in order to avoid the cross-tier interference, the set of SCs are operating either in unlicensed bands or separate bands.

In this scenario, illustrated in Figure 1, multiple zones can be identified as follows: i) a zone Z_0 in which users can only access resources from the MC, and ii) k zones Z_i (for $i \neq 0$) in which users can access both the MC and the i^{th} SC. Users are assumed to be uniformly distributed in the space and they are totally free to be connected to the MC or SCs as long as they are in an area with coverage from both base stations. Hence, they can get resources from either the MC or SC_i. It is also assumed that two or more SCs are deployed in such a way that they do not intersect each other.

RAN_i has a finite and limited capacity MS_i which is the number of sessions that can be supported by the RAN at any time. Thus, for all possible incoming sessions that arrive to the system, a radio or channel will be allocated to host the session in the system either in the MC or i^{th} SC.

We are also assuming that all networks in the system have a circular coverage area with radius r_i, which is calculated as

$A_{B_i} = \pi r_{B_i}^2, \forall i \in B$.

B. Traffic model

It is assumed that the HetNet system supports two types of traffic: new sessions that arrive to the system to be allocated over one specific RAN, and handoff traffic, which consists of petitions generated when an ongoing session is considered to be moved from one RAN to another one (i.e. from MC to one of the SCs).

In order to model the RAN selection as an SMDP, it is assumed that incoming requests follow a Poisson process with parameter λ_e, where e denotes the specific type of event. As mentioned before, the incoming requests can be either a new call or a handoff, and the corresponding mean rates are defined as follows. $\lambda^n_{B_0 z_0}$ denotes the arrival rate of new requests generated by users located in Z_0, and $\lambda^n_{B_0 z_i}, \forall i > 0$ the arrival rate by users located in Z_i. All these petitions arrive to the system via the MC (e.g., B_0). In a similar way, $\lambda^n_{B_i z_i}$ denotes the arrival rate for new traffic petitions in Z_i but coming from the i^{th} SC (e.g., B_i). On the other hand, $\lambda^h_{B_i B_0}$ indicates the arrival of handoff petitions from ongoing sessions allocated in B_i to B_0, and $\lambda^h_{B_0 B_i}$ represents the arrival of handoff requests from ongoing sessions in B_0 to B_i. In addition, due to the additive property of Poisson process, it is feasible to express the total arrival rate $\Lambda = \sum_{e \in E} \lambda_e$, where E denotes the set of possible events.

The holding time for each session is assumed to be exponentially distributed with mean of $^1/_\mu$ for any session in the system.

C. Load metric

With regards to the load of each RAN, and given that each one has a specific capacity, the load of the MC cannot be directly compared to the load of the SCs. In this work, we consider the normalized load, which is defined by taking into account the number of users connected to a RAN at a specific time. Thus, the load θ of a RAN $i \in B$ is given by:

$$\theta_i(t) = \frac{s_i(t)}{\text{MS}_i} \tag{1}$$

where s denotes the number of ongoing sessions at a specific time (t) for each RAN$_i$, and MS$_i$ indicates the maximum capacity of RAN$_i$ in terms of the number of channels.

D. Load Distribution and Fairness Index

In order to evaluate the fairness level of the load distribution among different RANs, Jain's Fairness Index (JFI) has been employed. This index is considered a quantitative measure of fairness, which was introduced in [13], and has been widely used and studied in the field of wireless networks [14]. JFI is defined as follows:

$$\gamma(t) = \frac{\left[\sum_{i=0}^{k} \theta_i(t)\right]^2}{(k+1)\sum_{i=0}^{k} \theta_i^2(t)}, \forall \theta_i > 0 \tag{2}$$

TABLE I. SYMBOLS AND NOTATIONS

Symbol	Description
B	Set of RANs
MS$_i$	Maximum system capacity (channels) of RAN$_i$
$\lambda^n_{B_i Z_i}$	Request's arrival rate of new events
$\lambda^h_{B_i B_i}$	Request's arrival rate of handoff events
μ	Departure rate for a session
θ_i	Traffic load for RAN$_i$
γ	Jain's Fairness Index for state x
s_i	Number of current sessions in RAN$_i$
e	New event in the HetNet system
η	Proportion of users which can perform a handoff
α	Continuous-time discount factor

where $\gamma \in [\frac{1}{k+1}, 1]$ and $k+1$ denotes the number of RANs.

Table I summarizes the notations used in this work.

IV. Problem Formulation of Load-Based Access Network Selection Strategy

In this section, we analyze the RAN selection problem using the SMDP framework. The model and the notation are based on the ideas presented in [15].

A. Discounted-Reward SMDP Model

An SMDP model is defined by considering the following elements: i) decision-epochs, ii) states, iii) actions, iv) state dynamics, and v) reward.

1) Decision epoch: In this approach, a decision epoch is defined as the time when an event occurs. In this way, all posible events are determined in terms of session arrival (i.e, new or handoff request) and session departure from the HetNet system. It means that the system will be analyzed when either an arrival or a departure takes place.

2) Space state: A discrete and finite state space has been considered in this work. Thus, the space state is based on the radio resource status and is expressed as $X = \{1, 2, ...x\}$, where $x(t) \in X$ represents a state of the HetNet system at decision epoch t. In this scenario, the space state is defined in the following way:

$$X = \{x | x = (s_0, s_1, ..., s_i, ..., s_k, e); s_i \geq 0; e \in E\}; \tag{3}$$

Specifically, the state $x \in X$ is expressed in terms of the number of ongoing sessions at time t where $s_i \geq 0$ denotes the number of active sessions in B_i. In addition to this, the space state includes the variable e. This variable denotes an event which can occur in a B_i at different instants of time and represents the type of event: new session, handoff, or session termination.

Thus, let E denote the set of events in the HetNet system, which is defined as: $E = \{I, L, S, T, V, D\}$. Subset $I = \{1\}$ represents a new arriving petition in Z_0 (i.e., it means that the request can only be assigned to the MC) and L denotes the subset of events that represent new session arrivals to the system via the MC in zones Z_i (i.e., potentially can be allocated to the MC or to the i^{th} SC). On the other hand,

S denotes the subset of arrivals via the SCs in zones Z_i. T indicates the subset of events representing a handoff operation from B_0 to B_i and V represents the subset of events of handoff requests from B_i to B_0. Finally, D denotes the subset of events that indicate a departure from a B_i. Thus, E can be expressed by:

$$\begin{aligned} E = \{&\{1\}, \{2, ..., k+1\}, \\ &\{k+2, ..., 2k+1\}, \{2k+2, ..., 3k+1\}, \\ &\{3k+2, ..., 4k+1\}, \{4k+2, ..., 5k+2\}\}. \end{aligned} \tag{4}$$

According to this notation, each event $e \in E$ is assigned an integer in the range $(1, 5k+2)$ that denotes what type of event it is.

3) Action: Whenever one of the defined events occurs, the system has to take an action a. In this sense, if a session arrives when the HetNet system is in state $x \in X$ at time t, the decision-maker (DM) has the responsibility to choose the most suitable RAN to host the session, taking into account the general goal for this process. In our case, the driver is maximizing the reward while using the radio resources in the fairest possible way. It is important to highlight that there is a possibility that the DM can not choose a RAN to serve the incoming request, because there might not be enough resources to support the service, in which case the request must be rejected.

In general, for each posible state $x = (\hat{s}, e)$ where $(\hat{s} = s_0, s_i, ..., s_k)$, $\forall e \in E$ and $i \in [1, k]$, the following possible actions have been identified for the RAN selection problem: continue, reject, and accept and allocate the event in the MC or i^{th} SC. In the following, we define the set of actions. Let $A = \{-1, 0, 1, ...k+1\}$ be the set of all possible actions where $a \in A$ denotes an action. $a = -1$ means that the system will continue without performing any procedure, $a = 0$ means that the request is rejected, and $a \in [1, k+1]$ indicates that the request is allocated to RAN_{a-1}.

Since the set of actions depend on the state, we define $A_x \subset A$ as the subset of possible actions for state $x \in X$ at each epoch decision time.

$$A_x = \begin{cases} \{0, 1\}, & \text{if } x = (\hat{s}, e), e = I \\ \{0, 1, i+1, ..., k+1\}, & \text{if } x = (\hat{s}, e), e \in L \\ \{0, 1, i+1, ..., k+1\}, & \text{if } x = (\hat{s}, e), e \in S \\ \{0, i+1, ..., k+1\}, & \text{if } x = (\hat{s}, e), e \in T \\ \{0, 1\}, & \text{if } x = (\hat{s}, e), e \in V \\ \{-1\}, & \text{if } x = (\hat{s}, e), e \in D \end{cases} \tag{5}$$

When action $a = 0$ the decision maker will reject the petition for all possible events, except for the departures (i.e., $e \in D$). For departures, the only possible action is $a = -1$, which means to continue. Action $a = 1$ will be available for events related with the MC (i.e., $e \in I, L, S, V$), which means that there is a possibility for the decision entity to allocate the incoming session in the MC. On the other hand, when the decision-entity receives handoff petitions from the MC to the i^{th} SC (i.e., $e \in T$), the request can be either rejected (with $a = 0$) or accepted (with $a = i+1$).

4) State dynamics: The state dynamics of the system are determined by its transition probability matrix, which in turn depends on the action $a \in A_x$ chosen by the DM in state $x \in X$.

Let $\beta(x, a)$ be the mean rate to leave state x and $\tau(x, a)$ the sojourn time, defined as $\tau(x, a) = \beta(x, a)^{-1}$. $\beta(x, a)$ is expressed in terms of the summation of the arrival and departure rates of all possible events in the state, and is given by (8). The term η_i is used to calculate the proportion of ongoing sessions that could perform a handoff petition from B_0 to B_i at a specific instant of time. In addition, the operator $\mathbb{1}$ represents the indicator function that equals one if the condition $\{\eta_i s_0 \geq 1\}$ is satisfied, and zero otherwise. Thus, in state $x \in X$, the handoff rate of sessions from B_0 to B_i is taken into account, as long as the possibility of performing a handoff exists. This parameter depends on the the geometric area of each RAN, and hence η_i is determined as follows:

$$\eta_i = \frac{A_{B_i}}{A_{B_0}} = \frac{\pi r_{B_i}^2}{\pi r_{B_0}^2} = \frac{r_{B_i}^2}{r_{B_0}^2} \tag{6}$$

Now, let $q(x'|x, a)$ denote the probability that in the next decision epoch, the system will be in state x' considering that the current state is x and action a is taken. In the following, we describe how the transition probabilities are computed for the states $x \in X$ in which the event is the arrival of a new session in the zone where there is only coverage of MC (i.e., event $e \in I$ that occurs in Z_0). In addition, if the event in state $x \in X$ is either a handoff or a departure, the probabilities are computed in a similar manner.

Assuming that the upcoming event in state $x' \in X$ is $e \in I$, the system will change with probability $\lambda^n_{B_0 Z_0}\tau(x, a)$. In case of $e \in \{L, S\}$ in state $x' \in X$, the transition probabilities are given by $\lambda^n_{B_0 Z_i}\tau(x, a)$ and $\lambda^n_{B_i Z_i}\tau(x, a)$, respectively. In these cases, if action $a = 1$ is taken, $s_0^{'} = s_0 + 1$. However, if $a = 0$, the variable s_0 maintains the same value in state $x' \in X$.

Now, when the next event is $e \in T$, the transition occurs with probability $\mathbb{1}_{\{\eta_i(s_0+1)\geq 1\}}\eta_i(s_0+1)\lambda^h_{B_0 B_i}\tau(x, a)$ if the session is accepted with $a = 1$. If the upcoming event in state x' is $e \in V$, the transition occurs with probability $s_i\lambda^h_{B_i B_0}$ when the action chosen is $a = 1$. In both cases, if action $a = 0$ is selected, state variables s_i, $\forall i \geq 0$ maintain the same value. On the other hand, if the next event $e \in D_0$ occurs in state x', the transition happens with probability $(s_0+1)\mu\tau(x, a)$. If $e \in D_i$, $i \geq 1$, the probability is $s_i\mu\tau(x, a)$. Here, when action $a = 1$ is taken, s_0 in x' is increased by one unit. If action $a = 0$, takes place, the variable s_i remains unchanged.

5) Reward: In this formulation, $r(s, a)$ denotes the total reward received when the system is in state $x \in X$ and action $a \in A_x$ is selected. The reward function is determined in the following way:

$$r(x, a) = f(x, a) - c(x, a) \tag{7}$$

where $f(x, a)$ is the lump sum income of the HetNet system received by the decision entity when action a has been chosen in state $x \in X$. On the other hand, $c(x, a)$ denotes the

$$\beta(x,a) = \sum_{i=0}^{k} \lambda^{n}_{B_0 Z_i} + \sum_{i=1}^{k} \lambda^{n}_{B_i Z_i} + \sum_{i=1}^{k} \mathbb{1}_{\eta_i s_0 \geq 1} \eta_i s_0 \lambda^{h}_{B_0 B_i} + \sum_{i=1}^{k} s_i \lambda^{h}_{B_i B_0} + \mu \sum_{i=0}^{k} s_i \tag{8}$$

system cost function for allocating a session over any RAN that belongs to the HetNet system.

The income $f(x,a)$, in turn, is defined in the following way:

$$f(x,a) = \begin{cases} Y_e - C_{e,B_i}\ , & \text{if } \gamma_x \leq \gamma_{x'} \\ Y_e - C_{e,B_i} + Y_f, & \text{if } \gamma_x > \gamma_{x'} \\ 0, & \text{Otherwise} \end{cases} \tag{8}$$

where Y_e and Y_f denote figurative incentives related to the type of the event and the increase of the fairness for the resource utilization respectively, and γ_x represents the value of fairness in state $x \in X$. These values are established by the MNO according to its objective and preferences, and they are given by:

$$Y_e = \begin{cases} Y^{h}_{B_0 B_i}, & a(x) = i+1, ..., k+1 \\ Y^{h}_{B_i B_0}, & a(x) = 1 \\ Y^{n}_{B_i}, & a(x) = i+1, ..., k+1 \\ Y^{n}_{B_0}, & a(x) = 1 \\ 0, & \text{Otherwise} \end{cases} \tag{9}$$

In this work, the figurative incentives for the decision entity are expressed in terms of the type of event and its priority for the network operator. In our model, we consider that handover sessions have a higher priority than new incoming requests and hence this incentive is larger. In this sense, $Y^{h}_{B_0 B_i} > Y^{h}_{B_i B_0} > Y^{n}_{B_i} > Y^{n}_{B_0}$.

On the other hand, since the interest is also to improve the fairness of the allocation by the MNO, it is necessary to carry out an evaluation process to establish whether action $a_x \in A_x$ will result in an increase of the fairness. In this case, the income associated with the fairness evaluation is given by:

$$Y_f = \begin{cases} Y_{IF} * \gamma_{x'} & \text{if } \gamma_x < \gamma_{x'} \\ 0 & Otherwise \end{cases} \tag{10}$$

Here γ_x and $\gamma_{x'}$ are the values of the fairness index given by equation (2) evaluated for the states x and x' respectively. A higher value of the fairness index in next state x', indicates a fairer load distribution among the RANs. For this reason, the value Y_{IF} is assigned in the process of getting the reward. This value is multiplied by the fairness value associated to state x' because of the following reason: if two or more possible actions in state x increase the fairness, a higher fairness-incentive will be earned by the action that allows to get the largest fairness index. Lastly, it is important to highlight, that no income is assigned when a session abandons the system or when the fairness remains equal or decreases.

The value of C represents the cost associated with the use of the specific RAN to allocate the session. In our case, $C_{e,B_0} > C_{e,B_i}$. This value is also set up according to MNO's preferences. The cost function of the system is given by:

$$c(x,a) = o(x,a) * \tau(x,a) \tag{11}$$

where $o(x,a)$ is the cost rate of the HetNet system defined in terms of the number of channels occupied at a specific time:

$$o(x,a) = \sum_{i=0}^{k} s_i \tag{12}$$

As we can observe, the cost is multiplied by the time between decision epochs $\tau(x,a)$. Thus, the total discounted-reward for the SMDP HetNet model is given by [15]:

$$\begin{aligned} r(x,a) &= f(x,a) - c(x,a) E^{a}_{x} \left\{ \int_{0}^{\tau} \mathrm{e}^{-\alpha t} \mathrm{d}t \right\} \\ &= f(x,a) - c(x,a) E^{a}_{x} \left\{ \frac{[1 - \mathrm{e}^{-\alpha \tau}]}{\alpha} \right\} \\ &= f(x,a) - \frac{c(x,a)}{[\alpha + \beta(x,a)]}, \end{aligned} \tag{13}$$

where $\beta(x,a)$ represents the rate and α denotes the discount factor.

B. Load Distribution at Departure Time

To illustrate this, it is necessary to consider the remaining load in each RAN in the next state x' after the departure event has occurred, in order to decide whether to perform a load balance process. If we assume that the system is in state $x = (s_0, s_1, ..., s_i, ..., s_k, e), \forall e \in D_i,\ i = 0$, a normal departure operation will indicate that the next state will be $x' = (s_0 - 1, s_1, ..., s_i, ..., s_k, e)$ or $x' = (s_0, s_1, ..., s_{i-1}, ..., s_k, e)$ when $i \neq 0$.

Next, it is necessary to add the new action $a_x = -2$ for events $e \in D$ which represents a normal departure with the offloading of a session from the MC to one of the SCs. For events action, the transition probabilities as well as the reward are calculated in the same manner as explained in section IV-A.

1) Offloading Procedure: With the aim to make an efficient use of the radio resources, we use the strategy described in this subsection to evaluate whether a session should be moved or not among RANs (i.e., offloading session from MC to one SC). Let x denote the current state, x' the state after the departure without performing the offloading procedure, and x'' the final state after the departure and the load distribution process is invoked. It should be noted that the traffic offloading is a network-based initiated process, which takes place at departure times, specifically, after a session has abandoned the system. In addition, the offloading is supported by moving a session from the MC to one of the SCs, with the goal to increase the fairness of the system from state x' to state x''. Lastly, since the SCs do not overlap each other, sessions can only be moved from MC to i^{th} SC.

Assuming that the system is in state $x = (s_0, s_1, ..., s_i, ..., s_k, e), e \in D$, if action $a = -2$ is taken, after the departure of a session ($s_0 - 1$ or $s_i - 1$), the offloading procedure is invoked and carried out as described in algorithm 1.

Algorithm 1 Offloading decision strategy at departures time.

Input: $x' \in X$
Output: Return if offloading should be performed or not.

```
γ_{x'} ← CalculateFairness(x')
ii ← 1
for i = 1 to k do
    Simulate offloading to estimate x'' from x'
    A_{γ_{x''}}[ii] ← CalculateFairness(x'')
    ii ← ii + 1
end for
if (max(A_{γ_{x''}}) > γ_{x'}) then
    Choose offloading that increase fairness in state x''.
    TOFF ← true
else
    TOFF ← false
end if
return TOFF
```

To sum up, the traffic offloading procedure will be only carried out if and only if there is a possibility to increase the fairness by moving session among RANs taking into account the load of each one after the departure event.

C. Optimization Function

In this approach, the optimization problem seeks to obtain a stationary deterministic policy that maximizes the discounted-reward criterion ψ under the policy ρ. Thus, the well-known expected total discounted-reward criterion is defined by [15]:

$$\psi_\alpha^\rho(x) = E_x^\rho\left[\sum_{t=1}^{\infty} \alpha^{t-1} r(x_t, a_t) | x_0 = x\right], \quad (14)$$

In this sense, for our problem, the Bellman equation is expressed in the following way:

$$\psi_\alpha^\rho(x) = \max_{a \in A_x}\left[r(x,a) + \alpha \sum_{x' \in X} q(x'|x,a) v(x')\right], \quad (15)$$

where α represents the discount factor. Now, a dynamic programming algorithm is required to solve Bellman equation (15) such that we can find which action should be taken by the decision-making entity for every state $x \in X$.

V. Solution of the SMDP Problem and Optimal Stationary Policy

The SMDP model needs to be solved in order to obtain a stationary deterministic policy. In the scenario of an SMDP, however, the VI algorithm [16] can not be directly applied, and a uniformization process has to be carried out in the following way:

a) Uniformization: This technique allows the conversion of the SMDP to an equivalent Discrete Time Markovian Decision Process (DTMDP) model with a constant transition rate. To achieve this, it is necessary to choose a number $c < \infty$ which satisfies the following condition [15]:

$$[1 - q(x'|x)]\beta(x,a) \le c < \infty; \forall x \in X, a \in A_x. \quad (16)$$

where the c parameter is defined as $max(\beta(x,a))$. Now, in the scope of our proposal, the uniformization process is defined in the following way according to [15]:

$$\begin{gathered} \tilde{X} = X, \tilde{A}_x = A_x, \forall x \in \tilde{X} \\ \tilde{r}(x,a) = r(x,a)\frac{\alpha + \beta(x,a)}{\alpha + c}, a \in \tilde{A}_x \text{ and } x \in \tilde{X} \\ \tilde{\alpha} = \frac{c}{c + \alpha} \\ \tilde{q}(x'|x,a) = \begin{cases} 1 - \frac{[1 - q(x|x,a)\beta(x,a)]}{c}, x = x' \\ \frac{q(x'|x,a)\beta(x,a)}{c}, x \neq x' \end{cases} \end{gathered} \quad (17)$$

where $\tilde{X}$, $\tilde{A}_x$, $\tilde{\alpha}$, $\tilde{q}(x'|x,a)$ $\tilde{r}(x,a)$ denote the equivalent set of states, actions, discount factor, discounted-reward, transition probabilities in the DTMDP model, respectively.

b) Value Iteration (VI) algorithm: Once the uniformization is completed, VI can be used to find the optimal policy for the SMDP model. The following pseudoalgorithm represents the steps followed by the VI [16]:

Step 0: (Initialization). Set $\tilde{\psi}(x) = 0, \forall x \in X$. Specify $\epsilon > 0$. Let $n := 1$

Step 1: (Value iteration step). For each state $x \in X$, compute:

$$\tilde{\psi}_{n,\phi}^*(x) = \max_{a \in \tilde{A}_x}\left[\tilde{r}(x,a) + \tilde{\alpha} \sum_{x' \in \tilde{X}} \tilde{q}(x'|x,a)\tilde{\psi}_{n-1}(x')\right], \quad (18)$$

Let $\hat{\rho}(n)$ be a stationary policy whose action $a = \hat{\rho}_x(t)$ maximizes the right side of (18).

Step 2: (Stop condition) If

$$||\tilde{\psi}_n(x') - \tilde{\psi}_{n-1}(x)|| < \epsilon(1 - \tilde{\alpha})/2\tilde{\alpha} \quad (19)$$

VI will stop with policy $\hat{\rho}(n)$. Otherwise, go to Step 3.

Step 3: (Iteration) $n := n + 1$ and go to Step 1.

VI. Performance Metrics

In this section, the performance metrics for the RAN selection scheme based on SMDP approach are defined. Blocking probability per event, average blocking probability of the system (ABP), utilization and average fairness index are defined in order to evaluate the performance of the proposed model. At this point, it should be noted that the stationary policy $\hat{\rho}$ obtained in Section V induces an embedded Markov Chain with transition probability matrix $q(x'|x, \hat{\rho}(x))$. Once the markov chain has been obtained, we proceed to build the infinitesimal generator Q in the following way [15]:

$$Q(x'|x) = \begin{cases} -[1 - q(x|x)]\beta(x); & x = x' \\ q(x'|x)\beta(x); & x \neq x' \end{cases} \tag{20}$$

Now, we can compute the steady-state probability vector π by solving the linear equation system $\pi Q = 0$ under the normalization condition $\pi 1 = 1$. In this work, we have used Gauss Seidel as the iterative numerical method to solve the aforementioned linear equation system [17].

A. Blocking and Average Blocking Probability

Taking into account that the steady-state probability for each state π_x has been calculated, blocking probabilities for each possible event in the HetNet system can be easily computed. Let P_{b_e} denote the blocking probability for event e, where $e \in E - \{D\}$

$$P_{b_e} = \sum_{x \in X, \hat{\rho}(x)=0} \pi_x; \quad x = (\hat{s}, e), \tag{21}$$

where $\hat{s} = (s_0, s_1, ..., s_{i+1}, ., s_k)$ and $e \in I, L, S, T, V$.

Once the blocking probability for each event has been defined, it is possible to formulate the systems's average blocking probability (ABP) as:

$$ABP = \frac{\sum_{e \in E} P_{b_e} \lambda_e}{\Lambda} \tag{22}$$

B. Utilization

In order to analyze the long-term utilization of the resources, it is important to take into consideration the utilization of each RAN. This metric gives us a way of visualizing the behavior of the system in terms of how the radio resources are being utilized over time. In this paper, we are interested in assessing whether the optimal policy obtained results in a fair allocation of the incoming sessions among the different RANs. Thus, the utilization of each RAN can be easily determined as follows:

$$U_{B_i} = \frac{\sum_{x \in X} s_{i,x} \pi_x}{MS_i}; \quad x = (\hat{s}, e) \tag{23}$$

where $\hat{s} = (s_0, s_1, ..., s_{i+1}, ., s_k)$ and $e \in E$.

TABLE II. Summary of Parameters

Parameter	Value	Parameter	Value
α	0.1	μ	0.25
Radius B_0	500m	Simulation Time	5 x 10^5
Radius B_1	160m	$Y^n_{B_0}$	15
Radius B_2	160m	$Y^n_{B_i}$	20
$Y^h_{B_0 B_i}$	30	$Y^h_{B_i B_0}$	25
Y_F	25	$C_{B_0} : C_{B_i}$	6:2

C. Average Fairness Index

Based on the load level of the RANs, it is possible to estimate the fairness of the system γ_x for each possible state x. It is important, however, to find the average fairness of the HetNet, taking into consideration the steady state probability, as follows:

$$\mathbb{E}(\gamma) = \sum_{x \in X} \gamma_x \pi_x; \quad x = (\hat{s}, e) \tag{24}$$

where $\hat{s} = (s_0, s_1, ..., s_{i+1}, ., s_k)$ and $e \in E$.

VII. Numerical and Simulation Results

This section presents the results for the performance evaluation of the optimal RAN selection policy obtained previously. A discrete-event simulation has been carried out in order to compare the theoretical results obtained in the scope of this work.

A. Experimental Setup

We consider a simple but representative scenario where an MNO deploys a HetNet with three RANs. Thus, a single MC (B_0) is defined as the access network with a broad coverage with capacity MS = 36. Two small cells SC1 (B_1) and SC2 (B_2) each one with capacity MS = 12 are defined as the access networks with less radio coverage that overlap the MC. Thus, the state of the system is stated in the following way: $x = (s_0, s_1, s_2, e)$.

On the other hand, three different zones are differentiated in this scenario. Z_0 is defined as the area where users can only be assigned to B_0. Z_1 and Z_2 are the areas where the users have the possibility to get resources from either B_0 or B_1/B_2. Users in Z_1 and Z_2 are able to send traffic over both networks but not in a simultaneous way at a specific instant of time. Lastly, taking into account the capacity of each RAN, an Average Blocking Probability (ABP) of around 2% or less is expected for the highest value of the Offered Traffic Load (OTL). This OTL is expressed in erlang units and ranges from 0 to 49.64.

The list of parameters used in the numerical analysis and in the simulation are summarized in Table II.

B. Perfomance Evaluation

In this subsection, we evaluate the efficiency of the proposed RAN selection policy, which will be compared to other policies obtained with random and greedy approaches. In this respect, it is important to note that while several previous works have proposed RAN selection strategies in the HetNet

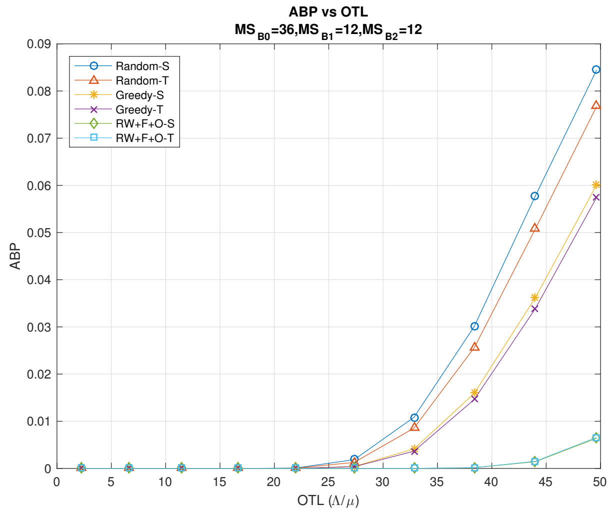

Fig. 2. Comparison of ABP vs OTL for Random, Greedy and RW+F+O policies. Theoretical results (T) and simulated results (S).

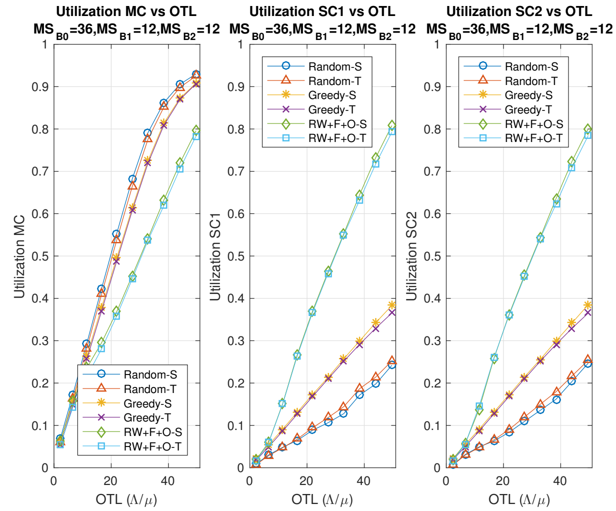

Fig. 3. Comparison of RAN Utilization vs OTL for Random, Greedy and RW+F+O policies. Theoretical results (T) and simulated results (S).

context, it is difficult to make a true comparison between these strategies and our proposal because of the differences in terms of modeling assumptions and the set of parameters used in each work (even for those that use SMDP). For this reason, we provide a comparison of our proposal with random and greedy approaches. In this comparison, we have labeled our policy as RW+F+O, since it includes reward, fairness and offloading. The random strategy consists of choosing an available action in a random way. It is important to note in this strategy that the decision entity only chooses $a_x = 0$ (i.e. reject) if and only if there are no available resources to allocate the incoming session. On the other hand, in the greedy approach, the decision-maker chooses the access network that offers the highest immediate reward i.e., $\text{argmax}\, r(x,a)$ when the system is in state $x \in X$. In the following, we offer an analysis in terms of the system's ABP, the utilization of each RAN, the steady-state probability vs the Jain's Fairness Index and the average Jain's Fairness Index.

Figure 2 shows the system ABP for different values of Offered Traffic Load (OTL) and the three different strategies. We can observe that the RW+F+O policy exhibits better performance levels in comparison to the other ones, i.e. lower blocking probability for the same OTL. Clearly, it is evidenced that the proposed scheme outperforms the others and guarantees an ABP lower than 2%, achieving the quality of service requirements established for the traffic conditions and network capacity. We can also observe that the results obtained from the simulations resemble the theoretical ones, validating the model presented in Section IV.

In Figure 3, the variation of the utilization of each RAN in the system is shown as the traffic load increases. The utilization of each RAN obtained by our solution is much better in comparison to the other schemes considered. Note that the utilization of the MC, SC1 and SC2 are very similar for the SMDP-based optimal policy RW+F+O, i.e., around 78% for the highest OTL. Meanwhile, for the random and greedy strategies, the utilization of each RAN are very different from each other, i.e. the MC utilization ($92\% - 90\%$) is higher than the utilization of SC1 or SC2 (25%-36%) for the highest OTL.

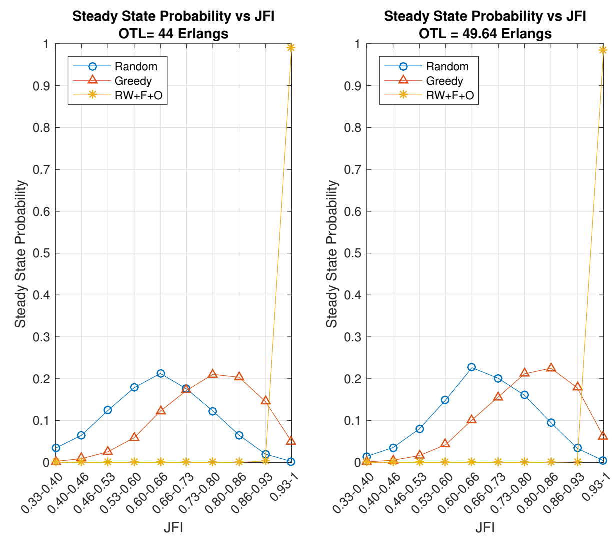

Fig. 4. Comparison of cumulative steady-state probabilities vs Jain's Fairness Index for different strategies such as Random, Greedy and RF+F+O.

We can also observe that the RW+F+O policy exhibits better performance as the load increases, since the utilization of the MC tends to be equal or even lower than the utilization of the SCs.

With the aim of evaluating the fairness of the resource usage in the long-term, Figure 4 shows the cumulative steady-state probability versus the Jain Fairness Index (JFI). Thus, we have organized the states in ten groups considering the value of the fairness of each state and, for each group, the cumulative steady-state probabilities have been determined. The length of the fairness index interval has been set up to 0.06 in the interval $[\frac{1}{k+1}, 1]$. This plot can be explained as follows: in the case of the RW+F+O policy, the cumulative steady-state probability of states with higher JFI (i.e., $(0.933 - 1]$), is larger than the cumulative steady-state probabilities of states with lower JFI. This implies that, in the long-term, the resources are being used in the fairer possible way.

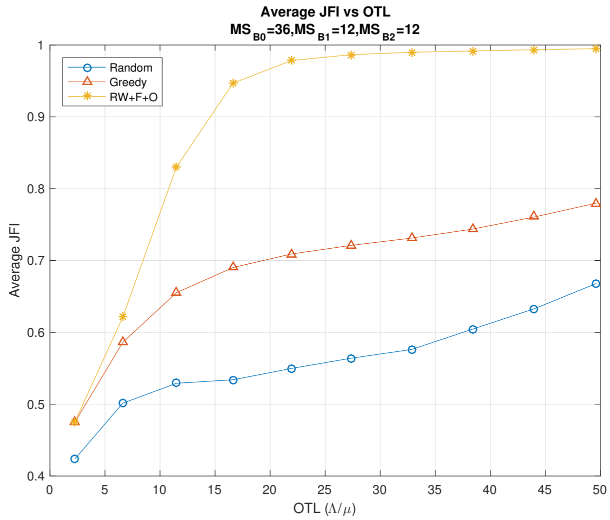

Fig. 5. Comparison of Average Jain's Fairness Index for different strategies such as Random, Greedy and RW+F+O.

In Figure 5, the fairness of the analyzed schemes are compared, taking into consideration the Average JFI (A-JFI) for different values of OTL. As the load increases, we can observe that the RW+F+O policy leads to A-JFI value closer to one (i.e., for OTL = 49.64 we have A-JFI = 0.99). In this sense, we evidence that our scheme outperforms the other two with respect to the fair usage of the radio resources. This behavior is due to the fact that our scheme introduces the fair allocation for incoming events and the traffic offloading at departures.

In summary, our proposed optimal network selection policy exhibits better performance that the other ones in terms of the ABP. Likewise, it is observed that we obtain a better resource usage of radio resources in the long-term, where it is clearly possible to alleviate the level of load for the MC.

VIII. Conclusions

In this paper, we have investigated the RAN selection problem in a HetNet system where the main goal is to distribute the traffic among the RANs in a fair way. Since network selection is considered a sequential decision-making problem, SMDP has been chosen as the decision-theoretic framework to model it and Value Iteration algorithm was used to obtain the optimal policy.

By using the induced Markov Chain that was built employing the stationary optimal policy, a mathematical analysis was carried out in order to estimate several performance metrics such as: the blocking probability for each possible event, the average blocking probability in the system, the average number of users served and the utilization of each RAN as well as the average Jain's Fairnes index.

The results indicate that the policy obtained allows us to distribute the traffic in a fair way among the RANs. Likewise, the level of quality of service in terms of ABP is also guaranteed.

Acknowledgment

The authors would like thank Universidad EAFIT for the grant to support this research.

References

[1] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Communications*, vol. 21, no. 3, pp. 118–127, June 2014.

[2] S. Navaratnarajah, A. Saeed, M. Dianati, and M. A. Imran, "Energy efficiency in heterogeneous wireless access networks," *IEEE Wireless Communications*, vol. 20, no. 5, pp. 37–43, October 2013.

[3] R. Trestian, O. Ormond, and G. M. Muntean, "Game theory-based network selection: Solutions and challenges," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 1212–1231, 2012.

[4] S. Singh, H. S. Dhillon, and J. G. Andrews, "Offloading in heterogeneous networks: Modeling, analysis, and design insights," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 2484–2497, May 2013.

[5] F. Rebecchi, M. D. de Amorim, V. Conan, A. Passarella, R. Bruno, and M. Conti, "Data offloading techniques in cellular networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 580–603, 2015.

[6] D. Pacheco-Paramo, V. Pla, V. Casares-Giner, and J. Martinez-Bauset, "Optimal radio access technology selection on heterogeneous networks," *Physical Communication*, vol. 5, no. 3, pp. 253 – 271, 2012.

[7] G. H. Carvalho, I. Woungang, A. Anpalagan, R. W. Coutinho, and J. C. Costa, "A semi-Markov decision process-based joint call admission control for inter-RAT cell re-selection in next generation wireless networks," *Computer Networks*, vol. 57, no. 17, pp. 3545 – 3562, 2013.

[8] G. Liu, M. Sheng, X. Wang, Y. Li, and Y. Li, "Fairness-based joint call admission control for heterogeneous wireless networks: An SMDP approach," *Science China Information Sciences*, vol. 57, no. 8, pp. 1–12, 2014.

[9] E. Khloussy, X. Gelabert, and Y. Jiang, "Investigation on MDP-based radio access technology selection in heterogeneous wireless networks," *Computer Networks*, vol. 91, pp. 57 – 67, 2015.

[10] G. H. S. Carvalho, I. Woungang, A. Anpalagan, and E. Hossain, "QoS-aware energy-efficient joint radio resource nanagement in multi-RAT heterogeneous networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6343–6365, Aug 2016.

[11] H. Du, Y. Zhou, L. Tian, X. Wang, Z. Pan, J. Shi, and Y. Yuan, "A load fairness aware cell association for centralized heterogeneous networks," in *2015 IEEE International Conference on Communications (ICC)*, June 2015, pp. 2178–2183.

[12] Y.-D. Lin, C.-Y. Ku, Y.-C. Lai, and Y.-H. Liang, "Wi-Fi offloading between LTE and WLAN with combined UE and BS information," *Wireless Networks*, pp. 1–10, 2016.

[13] R. K. Jain, D.-M. W. Chiu, and W. R. Hawe, "A Quantitative measure of fairness and discrimination for resource allocation in shared computer systems," Technical Report DEC-TR-301, Digital Equipment Corporation, Tech. Rep., 1984.

[14] H. SHI, R. V. Prasad, E. Onur, and I. G. M. M. Niemegeers, "Fairness in wireless networks: Issues, measures and challenges," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 5–24, 2014.

[15] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, 1st ed. New York, NY, USA: John Wiley and Sons, 1994.

[16] H. C. Tijms, *A First Course in Stochastic Models*. John Wiley and Sons, 2003.

[17] G. Bolch, S. Greiner, H. de Meer, and K. S. Trivedi, *Queueing networks and Markov Chains - modeling and performance evaluation with computer science applications*, 2nd ed. John Wiley and Sons, 2006.