

## Research Article

# Revenue-Maximizing Radio Access Technology Selection with Net Neutrality Compliance in Heterogeneous Wireless Networks

**Elissar Khloussy**  and **Yuming Jiang**

*Department of Information Security and Communication Technology, Norwegian University of Science and Technology, Trondheim, Norway*

Correspondence should be addressed to Elissar Khloussy; [khloussy@item.ntnu.no](mailto:khloussy@item.ntnu.no)

Received 4 October 2017; Revised 21 December 2017; Accepted 8 January 2018; Published 31 January 2018

Academic Editor: Vicente Casares-Giner

Copyright © 2018 Elissar Khloussy and Yuming Jiang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The net neutrality principle states that users should have equal access to all Internet content and that Internet Service Providers (ISPs) should not practice differentiated treatment on any of the Internet traffic. While net neutrality aims to restrain any kind of discrimination, it also grants exemption to a certain category of traffic known as specialized services (SS), by allowing the ISP to dedicate part of the resources for the latter. In this work, we consider a heterogeneous LTE/WiFi wireless network and we investigate revenue-maximizing Radio Access Technology (RAT) selection strategies that are net neutrality-compliant, with exemption granted to SS traffic. Our objective is to find out how the bandwidth reservation for SS traffic would be made in a way that allows maximizing the revenue while being in compliance with net neutrality and how the choice of the ratio of reserved bandwidth would affect the revenue. The results show that reserving bandwidth for SS traffic in one RAT (LTE) can achieve higher revenue. On the other hand, when the capacity is reserved across both LTE and WiFi, higher social benefit in terms of number of admitted users can be realized, as well as lower blocking probability for the Internet access traffic.

## 1. Introduction

Heterogeneous Wireless Networks (HWNs), where two or more Radio Access Technologies (RATs) coexist in the same geographical area, offer several opportunities to the Internet Service Provider (ISP) such as multiple connectivity options and a low-cost coverage expansion [1]. The ISPs, facing the fast increase in traffic demands, have interest in making the best utilization of all available resources in the HWN in order to increase the capacity of the network and meet, as much as possible, their customers' expectations and demands.

Managing resources in HWNs involves setting up policies that regulate how the arriving traffic is distributed and served among the available RATs. A well-known key mechanism for resource management in HWNs scenarios is RAT selection. It consists of taking a decision, at each arrival of a new call request, on whether to accept this call or not and the RAT to which it can be admitted.

The decision taken by the RAT selection policy is based on the objective set by the ISP such as the maximization of

the generated revenue. However, it is important to set rules that regulate how the traffic is served in order to avoid that the ISP exercises any kind of traffic discrimination. Hence, the principle of net neutrality has gained lots of attention recently.

The main idea behind net neutrality is that ISPs should treat all Internet traffic equally regardless of the content, application, device, sender, or receiver. While net neutrality principle states that all Internet traffic has to receive equal treatment, an exemption is granted to some non-Internet access services that require high transmission quality, known as specialized services (SS) [2]. Some examples of SS include VoLTE, linear broadcasting IPTV, and real-time health services [3]. In order to secure higher Quality of Service (QoS) to SS, the ISP is allowed to dedicate a certain amount of bandwidth to those services, without causing a degradation of the QoS experienced by the Internet access services (IAS) traffic.

However, in order to follow the net neutrality regulations, the ISPs might lose some of the generated revenue. Having revenue maximization as our main focus, we address the

following problem: *how the bandwidth reservation for SS traffic would be made in a way that allows maximizing the revenue while being in compliance with net neutrality and how the choice of the ratio of reserved bandwidth would affect the revenue?*

In the present work, we derive RAT selection policies that allow maximizing the revenue while being net neutrality-compliant at the same time. We consider an integrated Long-Term Evolution (LTE)/Wireless Fidelity (WiFi) heterogeneous network. The optimal RAT selection policies are derived with the help of Markov Decision Process (MDP). Two types of traffic are considered, namely, SS and IAS traffic. To reserve bandwidth for SS traffic, two cases are proposed: bandwidth reserved in LTE only and bandwidth reserved in the whole HWN. Our aim is to figure out which way of bandwidth reservation is better to adopt and to investigate how the revenue would be affected by the choice of the ratio of reserved bandwidth to SS traffic.

The main contributions of this paper can be summarized as follows:

- (1) Investigation of the MDP-based approach for RAT selection, with focus on revenue maximization as objective
- (2) Integration of net neutrality in the RAT selection policy, with two variants of bandwidth reservation for SS traffic
- (3) The impact of the ratio of reserved bandwidth for SS traffic being studied with both variants
- (4) The coverage probability of WLAN being analytically modeled with the help of Poisson Point Process (PPP)
- (5) The spatial distributions of the cellular base stations (BS), WiFi access points (AP), and the users being also captured with PPP.

The remaining of the paper consists of the following parts: Section 2 presents the motivation and related work in the literature. Section 3 describes the system model. In Section 4, the components of the MDP problem are presented. Section 5 presents and analyzes the obtained results. Finally, we conclude this study in Section 6.

## 2. Motivation and Related Work

Net neutrality has been heavily discussed in the past decade as a potential way to prevent the ISPs from exercising any type of discrimination on the Internet traffic. Content providers, in general, support net neutrality especially in monopolistic regimes where an ISP might have pricing power over the Internet access market. The ISPs, on the other hand, argue that service differentiation is crucial for QoS enhancement [4].

The use of capacity increase as an alternative to deal with QoS concerns resulting from applying net neutrality has been addressed in the literature. While this seems plausible, early study, for example, [5], already showed that the relationship between the net neutrality regulation and investment incentives is subtle and that it is difficult to draw general

unambiguous conclusions regarding this issue. In addition, recent study, for example, [6], further shows that when strict net neutrality is applied, the ISPs will no longer have incentive to invest in expending their infrastructure and enhance the QoS.

Wu in [7] expanded the net neutrality debate by suggesting that policymakers ought to consider how to apply net neutrality regulations to wireless networks. This was opposed by a number of economists (e.g., [8]) who argued that, unlike the wired market, the competition will remain high in the wireless one.

Martínez et al. in [9] provided an initial analysis of the impact of net neutrality on quality of experience-based differentiation in mobile networks. In [10], the authors shed light on the content provider discrimination and discussed the impact of some of the disruptive network applications on net neutrality. In some other work, for example, [4, 11], alternative regulations to net neutrality have been proposed. Authors in [6] studied the paid prioritization where the content providers decide to pay for this priority in monopolistic access market. They showed that, with ISP's optimal pricing, the service differentiation becomes efficient and the social welfare among the different content providers is close to its maximum. Altman et al. in [12] presented a bargaining framework to decide how much the ISP should charge the content provider.

In a previous work [13], we investigated MDP as a tool for modeling revenue-maximizing RAT selection policies. However, net neutrality was not taken into account in the derived model. The way the traffic was handled provided privileges to the high-priority traffic in getting admission to LTE which offers better QoS guarantees, at the expense of blocking or handing over part of the low-priority traffic from one RAT to another. This allows achieving higher revenue but violates the net neutrality regulations.

Net neutrality and its integration in RAT selection policies have been addressed in [14], where the performance of various RAT selection strategies that are net neutrality-compliant was compared. The objective was to give insight into the effect of applying net neutrality regulations on the revenue and the QoS.

In the present work, we model two variants of RAT selection policy that differ by the way the bandwidth reservation for SS traffic is exercised. Both are net neutrality-compliant and aim to maximize the generated revenue. By comparing their performance, we try to find the most appropriate way of bandwidth reservation for SS traffic and investigate the impact of the ratio of this reserved bandwidth on the generated revenue.

## 3. System Model

*3.1. Network Architecture.* We consider the case of an LTE/WiFi overlay network [15]. The traffic arrivals to the different base stations are independently distributed. Hence, without loss of generality, we can shift our focus to a single cell  $C_{\text{targ}}$  that corresponds to the coverage area of one cellular BS. LTE has global coverage, overlaying the WLAN; that is, within the coverage of the considered BS, there exists one or more WiFi AP(s).

Two types of traffic are served, namely, SS and IAS traffic, where IAS is charged a price  $P_i$ , while SS is charged a price  $P_h > P_i$ . Naturally, this pricing differentiation affects the traffic distribution among SS and IAS. In this paper, we adopt that and the ratio of traffic that is being sent as SS traffic out of the total traffic can be computed with the help of the following demand function that was proposed in [16] and has been adopted in the literature, for example, [17]:

$$D[P_h] = e^{-(P_h/P_i-1)^2}, \quad (1)$$

which implies that, out of the total traffic, the ratio of IAS traffic is  $(1 - e^{-(P_h/P_i-1)^2})$ .

**3.2. Spatial Distribution.** Because of the overlay nature of our studied HWN scenario, a connection request might occur either in an area that is covered by the cellular RAT only or in a dual coverage area. In the latter case, an arriving session request can be admitted to LTE or to WiFi depending on the decision provided by the RAT selection policy. Here arises the need of getting knowledge regarding the spatial distribution of the BSs and the APs. The considered network architecture can be seen as a 2-tier heterogeneous network, where tier-1 is LTE and tier-2 is WiFi. A spatial point process, such as PPP, provides a concise and tractable model for HWNs, by offering a statistical modeling for the spatial distribution of the BSs and APs. In fact, PPP model has been used extensively for modeling unplanned networks [18] which is typically the case of WLAN APs' deployment. In our considered scenario, the different aspects of the PPP model can be described as follows:

- (i) The positions of BSs/APs belonging to tier- $k$  are modeled according to a homogeneous PPP  $\phi^{(k)}$  with intensity  $\lambda^{(k)}$ , where  $\lambda^{(k)}$  is defined as the number of BSs/APs per area unit, and  $k \in \{1, 2\}$  with  $k = 1$  refers to LTE and  $k = 2$  refers to WiFi.
- (ii) Users are also scattered in the plane according to a homogeneous PPP  $\phi^{(u)}$  with intensity  $\lambda^{(u)}$  users per area unit, independently of  $\phi^{(k)}$ .

Through PPP modeling, different metrics can be captured. In the following, we derive the probability for a user to be under tier- $k$ 's coverage and the traffic arrival rates.

**3.2.1. Coverage Probability.** The cellular system has global coverage; that is, all users in the considered HWN fall under the coverage of the cellular RAT. Hence, the coverage probability of LTE is  $P_{c,1} = 1$ .

As for the coverage probability of WiFi, it can be derived with the help of PPP as follows. First, we assume that each AP covers a circular area of known radius  $R$ ; that is, the transmission of each AP can be received clearly by users residing at a distance not exceeding  $R$ . Second, the interference from neighboring APs is considered negligible. Hence, a typical user is said to be under the coverage of WLAN if the distance  $r$  separating this user from the nearest AP is less than  $R$ . Therefore, the probability that a user is under WLAN coverage is equivalent to the cumulative distribution function

of  $r$ , namely,  $\mathbb{P}[r < R]$ . Without loss of generality, we consider that the typical user is located at the origin of the plane under consideration [18]. Then, knowing that the null probability of a 2D Poisson process in an area  $Z$  is  $\exp(-\lambda Z)$  [19], we can derive the coverage probability of WiFi  $P_{c,2}$  as follows:

$$\mathbb{P}[r > R] = \mathbb{P}[\phi^{(2)} \cap b(0, R) = 0] = e^{-\pi\lambda^{(2)}R^2}, \quad (2)$$

where  $b(0, R)$  is the Euclidean ball of radius  $R$  centered at origin. Hence, the coverage probability of tier-2 is given by

$$P_{c,2} = \mathbb{P}[r < R] = 1 - \mathbb{P}[r > R] = 1 - e^{-\pi\lambda^{(2)}R^2}. \quad (3)$$

**3.2.2. Traffic Arrivals and Holding Times.** With the assumption that each user of class  $i$ ,  $i \in \{1, 2\}$  ( $i = 1$  represents SS and  $i = 2$  represents IAS) generates traffic following a Poisson distribution with average  $\sigma_i$  calls/second, the traffic arrival rates  $\lambda_1$  and  $\lambda_2$  of classes 1 and 2, respectively, can be easily derived as follows:

$$\begin{aligned} \lambda_1 &= \sigma_1 \cdot D[P_h] \cdot \lambda^{(u)} \cdot |C_{\text{targ}}| \text{ arrivals/second} \\ \lambda_2 &= \sigma_2 \cdot (1 - D[P_h]) \cdot \lambda^{(u)} \cdot |C_{\text{targ}}| \text{ arrivals/second,} \end{aligned} \quad (4)$$

where  $|C_{\text{targ}}|$  is the area covered by the targeted cell  $C_{\text{targ}}$  (in area unit) and  $D[P_h]$  is found from (1).

Note that, in (4),  $\lambda^{(u)} \cdot |C_{\text{targ}}|$  appears in both  $\lambda_1$  and  $\lambda_2$ , together with  $\sigma_1$  and  $\sigma_2$ , respectively. To simplify the representation, in later analysis and results, we will simply use  $\lambda_1 = \sigma_1 D[P_h]$  and  $\lambda_2 = \sigma_2 (1 - D[P_h])$  with  $\sigma_1$  and  $\sigma_2$  normalized against  $\lambda^{(u)} \cdot |C_{\text{targ}}|$ .

As for the call holding time for class  $i$ , the traffic of each class is assumed to be inelastic; that is, the average duration of the service is independent of the allocated number of channels and following exponential distribution with mean  $1/\mu_i$ ,  $i \in \{1, 2\}$ .

## 4. Markov Decision Process Formulation

An MDP model is provided to derive the optimal RAT selection policy which maximizes our objective function. This model can be uniquely identified by five components: the state space, decision epochs, action space, state dynamics, and the reward function. We define each of these components in the following subsections.

**4.1. State Space.** The state space represents the number of ongoing sessions in the HWN, that is, the number of SS sessions being served in LTE and WiFi and, similarly, number of IAS sessions being served in both LTE and WiFi. For ease of representation in MDP, we model the problem with one particular AP in WLAN that we call the targeted AP. Hence, a 4D-MDP serves to build our model. On the other hand, we assume a fixed total capacity for both RATs, each being partitioned into a fixed number of basic bandwidth units (bbu) as in, for example, [15, 20, 21]. This implies that a limited number of sessions can be served simultaneously by each RAT. The total capacities of LTE and WiFi can be defined as integers that we denote by  $C_1$  and  $C_2$ , respectively. Any

newly arriving session that cannot be granted its required amount of bbu is blocked. Thus, by restricting the number of ongoing connections in the system, the delivered QoS to the different connections can be maintained at an acceptable level. To simplify the notation, we refer to each type of served traffic as class  $i$ , with  $i = 1$  denoting SS traffic and  $i = 2$  denoting IAS traffic.

We define the following row vectors:

(i) State vector of LTE is

$$\mathbf{s}_1 = [n_{1,1}, n_{1,2}] \in \mathbb{Z}_+^2. \quad (5)$$

(ii) State vector of WiFi is

$$\mathbf{s}_2 = [n_{2,1}, n_{2,2}] \in \mathbb{Z}_+^2. \quad (6)$$

(iii) State vector of the system is

$$\mathbf{s} = [s_1, s_2] = [n_{1,1}, n_{1,2}, n_{2,1}, n_{2,2}], \quad (7)$$

where  $n_{j,i}$  denotes the number of sessions of class  $i$  in RAT  $j$ ;  $\mathbb{Z}_+$  represents the set of nonnegative integer numbers.

The state space  $S$  of the system, which is the set of all feasible states, differs according to the RAT selection policy. The following two cases are to be distinguished: reserved bandwidth for SS traffic in LTE only and reserved bandwidth for SS in both LTE and WiFi.

*State Space: Reserved Bandwidth in LTE Only.* In this case, the state space can be defined as follows:

$$S = \left\{ \mathbf{s} = [s_1, s_2] = [n_{1,1}, n_{1,2}, n_{2,1}, n_{2,2}] \in \mathbb{Z}_+^4, n_{1,1} + n_{1,2} \leq C_1, n_{2,1} + n_{2,2} \leq C_2, n_{1,2} \leq (C_1 - C_{1,\text{res}}) \right\}, \quad (8)$$

where  $C_{1,\text{res}}$  represents the number of reserved bbu in LTE for the usage of SS traffic.

*State Space: Reserved Bandwidth in LTE and WiFi.* In this case, the reserved bandwidth for SS traffic is spread across the available RATs, namely, LTE and WiFi. The state space in this case becomes

$$S = \left\{ \mathbf{s} = [s_1, s_2] = [n_{1,1}, n_{1,2}, n_{2,1}, n_{2,2}] \in \mathbb{Z}_+^4, n_{1,1} + n_{1,2} \leq C_1, n_{2,1} + n_{2,2} \leq C_2, n_{1,2} + n_{2,2} \leq (C_1 + C_2 - C_{\text{res}}) \right\}, \quad (9)$$

where  $C_{\text{res}}$  denotes the number of bbu reserved for SS traffic in both LTE and WiFi.

*4.2. Decision Epochs and Actions.* At each arrival of a class  $i$  session request,  $i \in \{1, 2\}$ , the RAT selection policy makes a decision concerning the admission of this new session. A decision epoch occurs at each new session request. It is defined as the time following immediately an arrival event. As for the events of call completion, they do not require any decision to be taken by the system.

The action taken following each decision epoch can be defined as a vector  $\mathbf{a} = [a_1, a_2]$  where  $a_i$  denotes the action resulting from the arrival of a class  $i$  session. A decision can be either to admit the arriving session to LTE or to admit it to WiFi or block it.  $a_i$  can be defined as follows:

$$a_i = \begin{cases} -1, & \text{if the session is admitted to LTE,} \\ 1, & \text{if the session is admitted to WiFi,} \\ 0, & \text{if the session is blocked.} \end{cases} \quad (10)$$

The action space of the MDP is defined as the set of vectors  $\mathbf{a}$  as follows:

$$A = \{ \mathbf{a} = [a_1, a_2], a_1 \in \{-1, 0, 1\}, a_2 \in \{-1, 0, 1\} \}. \quad (11)$$

However, for a given state  $\mathbf{s} \in S$ , the decision should always lead to a state  $\mathbf{s}'$  that is also in  $S$ . Moreover, when the system is in state  $(0, 0, 0, 0)$ , the action  $(0, 0)$  should be avoided in order for the system to keep evolving. Hence, for a given state  $\mathbf{s} = [n_{1,1}, n_{1,2}, n_{2,1}, n_{2,2}] \in S$ , the state action space  $A_s \subset A$  is given by the following:

$$A_s = \{ \mathbf{a} \in A : a_i \neq -1 \text{ if } [s_1 + e_i^u, s_2] \notin S, a_i \neq 1 \text{ if } [s_1, s_2 + e_i^u] \notin S, a_i = 0 \text{ if } [s_1 + e_i^u, s_2] \notin S, [s_1, s_2 + e_i^u] \notin S, \mathbf{a} \neq (0, 0) \text{ if } s = (0, 0, 0, 0) \}, \quad (12)$$

where  $e_i^u \in \{0, 1\}^2$  is a vector of zeros except for the  $i$ th element which is equal to 1.

*4.3. State Dynamics.* The state dynamics of the MDP are defined by two parameters, namely, the expected sojourn time and the transition probabilities.

*4.3.1. Expected Sojourn Time.* The sojourn time  $\tau(s, a)$  is defined as the expected time for the system to stay in state  $\mathbf{s} \in S$  given that action  $\mathbf{a} \in A_s$  is chosen, until a new state is entered. The sojourn time is used to compute the transition probabilities for a continuous-time MDP, and its value can be expressed as follows [22, 23]:

$$\tau(s, a) = [\lambda_1 |a_1| + (n_{1,1} + n_{2,1})\mu_1 + \lambda_2 |a_2| + (n_{1,2} + n_{2,2})\mu_2]^{-1}, \quad (13)$$

where  $\lambda_i$  is the arrival rate for class  $i$  traffic,  $i \in \{1, 2\}$  is found from (4), and  $1/\mu_i$  is the mean value of the call holding time of class  $i$ .

*4.3.2. Transition Probabilities.* Let  $P_{\mathbf{s}\mathbf{s}'}(a)$  denote the transition probability from state  $\mathbf{s} = [s_1, s_2] \in S$  to state  $\mathbf{s}' \in S$ ,



TABLE I: System parameters.

Parameter	Symbol	Value
Capacity of LTE	$C_1$	30 bbu
Capacity of WiFi	$C_2$	5 bbu
Price charged for SS traffic	$P_h$	1.6 monetary unit (MU)
Price charged for IAS traffic	$P_l$	1.0 MU
Reserved bandwidth for SS traffic in LTE	$C_{1,res}$	5 bbu
Reserved bandwidth for SS traffic in the HWN	$C_{res}$	5 bbu
Total traffic intensity	$\rho$	2 to 12 Erlang (E)
Ratio of SS traffic	$D[P_h]$	0.69
Average Session holding time, SS traffic	$1/\mu_1$	200 s
Average Session holding time, IAS traffic	$1/\mu_2$	150 s

$\mathbf{s} \neq \mathbf{s}'$ , provided that action  $\mathbf{a} \in A_s$  is chosen. The state transition probabilities can thus be written as follows:

$$P_{\mathbf{s}\mathbf{s}'}(\mathbf{a}) = \begin{cases} \lambda_1 \delta(-a_1) \tau(\mathbf{s}, \mathbf{a}), & \text{if } \mathbf{s}' = [\mathbf{s}_1 + e_1^u, \mathbf{s}_2], \\ \lambda_2 \delta(-a_2) \tau(\mathbf{s}, \mathbf{a}), & \text{if } \mathbf{s}' = [\mathbf{s}_1 + e_2^u, \mathbf{s}_2], \\ \lambda_1 P_{c,2}^* \delta(a_1) \tau(\mathbf{s}, \mathbf{a}), & \text{if } \mathbf{s}' = [\mathbf{s}_1, \mathbf{s}_2 + e_1^u], \\ \lambda_2 P_{c,2}^* \delta(a_2) \tau(\mathbf{s}, \mathbf{a}), & \text{if } \mathbf{s}' = [\mathbf{s}_1, \mathbf{s}_2 + e_2^u], \\ \mu_1 n_{1,1} \tau(\mathbf{s}, \mathbf{a}), & \text{if } \mathbf{s}' = [\mathbf{s}_1 - e_1^u, \mathbf{s}_2], \\ \mu_2 n_{1,2} \tau(\mathbf{s}, \mathbf{a}), & \text{if } \mathbf{s}' = [\mathbf{s}_1 - e_2^u, \mathbf{s}_2], \\ \mu_1 n_{2,1} \tau(\mathbf{s}, \mathbf{a}), & \text{if } \mathbf{s}' = [\mathbf{s}_1, \mathbf{s}_2 - e_1^u], \\ \mu_2 n_{2,2} \tau(\mathbf{s}, \mathbf{a}), & \text{if } \mathbf{s}' = [\mathbf{s}_1, \mathbf{s}_2 - e_2^u], \\ 0, & \text{otherwise.} \end{cases} \quad (14)$$

where  $P_{c,2}^*$  is the coverage probability of the targeted AP:  $P_{c,2}^* = P_{c,2}/|C_{\text{targ}}|\lambda^{(2)}$  and  $\delta(x)$  is a function defined as follows:

$$\delta(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ 1, & \text{if } x > 0. \end{cases} \quad (15)$$

**4.4. Policy and Reward Function.** For each state  $\mathbf{s} = [s_1, s_2] \in S$ , an action  $\mathbf{a} \in A_s$  is chosen according to a policy  $\pi_s \in \Pi$ , where  $\Pi$  is a set of admissible policies defined as follows:

$$\Pi = \{\pi : S \rightarrow A \mid \pi_s \in A_s, \forall \mathbf{s} \in S\}. \quad (16)$$

The reward function for choosing action  $\mathbf{a} \in A_s$ , when the system is in state  $\mathbf{s} \in S$ , can be defined as follows:

$$r(\mathbf{s}, \mathbf{a}) = w_{1,1} \cdot \delta(-a_1) + w_{1,2} \cdot \delta(-a_2) + w_{2,1} \cdot \delta(a_1) + w_{2,2} \cdot \delta(a_2), \quad (17)$$

where  $w_{j,i} \in \mathbb{R}_+$  is the weight associated with the admission of a class  $i$  session into RAT  $j$ ,  $\mathbb{R}_+$  being the set of nonnegative real numbers. Since our objective is to maximize the revenue, the weights  $w_{j,i}$  in the reward function are assigned the value  $P_h$  for  $i = 1$  and  $P_l$  for  $i = 2$ .

By solving the MDP, an optimal policy  $\pi^*$  that maximizes the reward function can be found. The RAT selection module will then, based on the optimal policy provided by the MDP, decide on the admission or rejection for each arriving session. A summary of the notations used in the paper is presented in Notations.

## 5. Numerical Results

In this section we present and analyze the results obtained from the implementation of the two variants of the net neutral revenue-maximizing RAT selection policy, namely, bandwidth reservation for SS in LTE only and bandwidth reservation for SS in the HWN as a whole. In addition, the results obtained from a *non-net neutral* revenue-maximizing RAT selection policy (introduced in [13, 14]) are presented as reference. The non-net neutral policy prioritizes the admission of SS services traffic and allows the handover of IAS traffic between LTE and WiFi when there is need to free resources for SS traffic.

To solve the MDP problem and find the optimal policy, we used the relative value iteration algorithm, defined in the MDP toolbox (developed by [24]). If not otherwise stated, the values of the system parameters used in our analysis are as shown in Table 1.

**5.1. Revenue.** The term revenue, as used in this paper, designates the charges paid by the customers in exchange for the services they are receiving. In our case, it is the amount paid by them for transmitting their traffic over the ISP's network in monetary unit (MU).

The revenue achieved when applying the two variants of the net neutral RAT selection policy along with that achieved by the non-net neutral one is depicted in Figure 1, for the values of average session holding times ( $1/\mu_1$  and  $1/\mu_2$ ) stated in Table 1, and by varying the average call arrival rates  $\sigma_1$  and  $\sigma_2$  (i.e. varying  $\lambda_1$  and  $\lambda_2$  (c.f. (4))). The  $x$ -axis represents the total traffic intensity  $\rho = \lambda_1/\mu_1 + \lambda_2/\mu_2$ . As expected, the non-net neutral policy is the one that achieved the highest revenue. In addition, the results show that both net neutral variants have comparable performance in terms of revenue with an advantage to reserving capacity in LTE. Because LTE has global coverage and better QoS guarantees than WiFi,

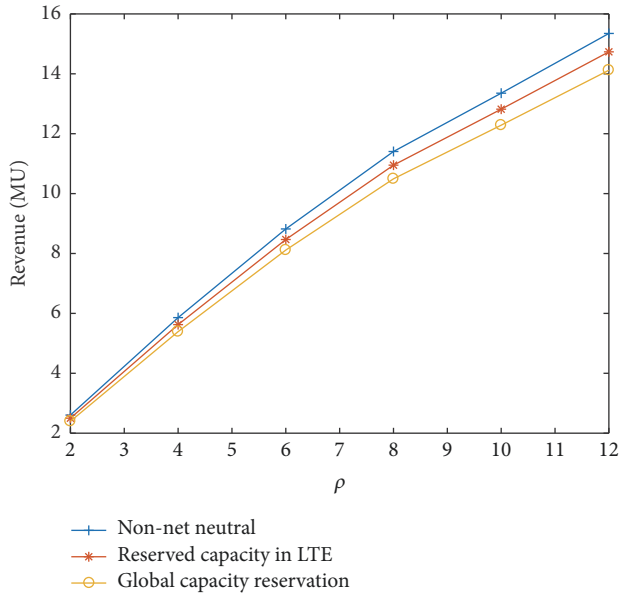


FIGURE 1: Revenue achieved in monetary unit [MU], fixed session durations.

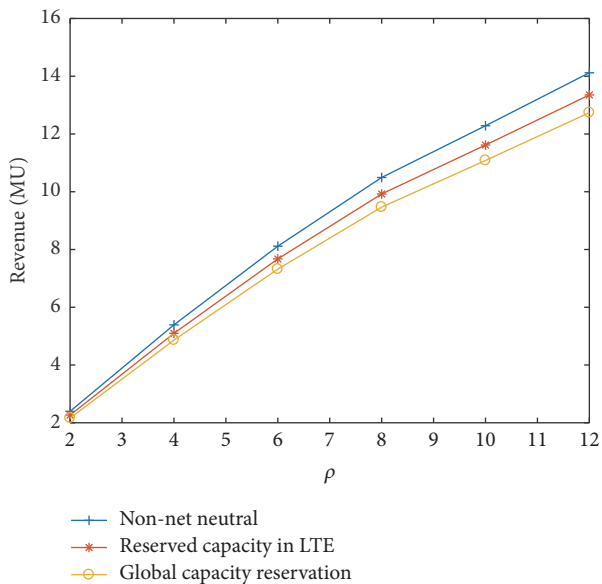


FIGURE 2: Revenue achieved in monetary unit [MU], fixed arrival rates.

reserving capacity in LTE to SS traffic allows larger number of SS sessions to be admitted to the system, resulting in higher revenue as compared to the case where the bandwidth reservation is done across LTE and WiFi.

To show the impact of varying the session durations on the revenue, we experimented with fixed arrival rates with values  $\sigma_1 = 0.028$  and  $\sigma_2 = 0.03$  and varied the average session holding times (Figure 2). In this case, similar trends are observed as in Figure 1. We notice, however, that the achieved revenue has become lower as compared to the case where the session durations are fixed. This is because, for the same traffic

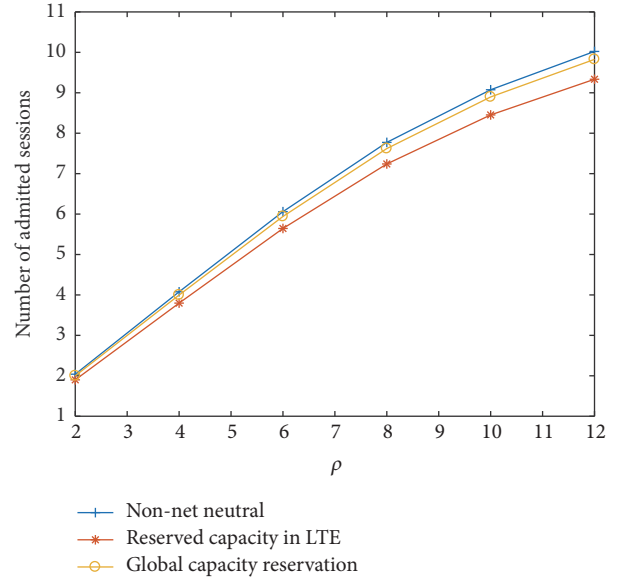


FIGURE 3: Social benefit.

intensity level, Figure 1 resulted from higher traffic rates (or indeed higher call arrival rates) with shorter session duration times, while Figure 2 is from lower traffic rates (or indeed lower call arrival rates) with longer session duration times. As a result, the wireless resources become reserved for shorter time and more calls could be admitted for the case of Figure 1, providing higher revenue, while, for the case of Figure 2, with increased session duration times, implying lower call arrival rates with the same traffic intensity level, the wireless resources become reserved for longer time, fewer arrivals will be admitted giving hence lower revenue.

**5.2. Social Benefit.** The social benefit is a metric that offers insight into the user experience and is defined as the total number of admitted sessions into the network regardless of the traffic type. Figure 3 depicts the number of admitted sessions obtained by the three considered RAT selection policies. We observe that the non-net neutral policy is the one that achieves the highest social benefit as it allows rearranging the traffic between LTE and WiFi in order to admit the highest possible number of sessions. It is also shown that when reserving capacity for SS spans the whole HWN, it becomes possible to admit more sessions (particularly IAS sessions) than in the case where the reserved bandwidth is in LTE only and therefore a better social benefit is realized.

**5.3. Blocking Probability.** The blocking probability obtained by the considered RAT selection policies is depicted in Figures 4 and 5 for SS and IAS traffic, respectively. We observe that the blocking probability for SS traffic is lower when the reserved bandwidth is in LTE only, while for IAS traffic it is the opposite, which also confirms the results for revenue and social benefit depicted in Figures 1, 2, and 3. On the other hand, while the non-net neutral policy provides the lowest blocking probability for SS traffic due to the priority granted

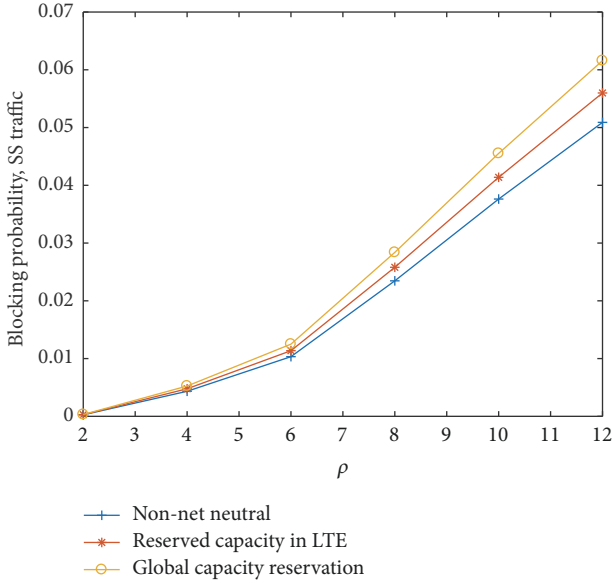


FIGURE 4: Blocking probability, specialized services traffic.

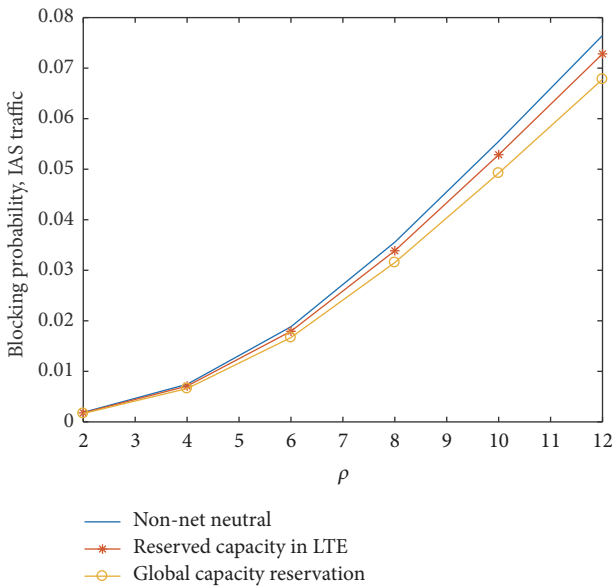


FIGURE 5: Blocking probability, Internet access traffic.

to the latter, it provides the highest blocking probability for IAS traffic.

Having compared the three performance metrics (revenue, social benefit, and blocking probability), we conclude that reserving capacity for SS traffic in the whole HWN offers more advantages despite an insignificant loss in revenue, by providing a better social benefit and lower blocking probability for IAS traffic.

**5.4. Reserved Capacity for SS: Impact on Revenue.** In this part, we investigate the impact of the ratio of reserved capacity for SS traffic on the generated revenue for both studied scenarios, namely, when the dedicated capacity for SS traffic is reserved

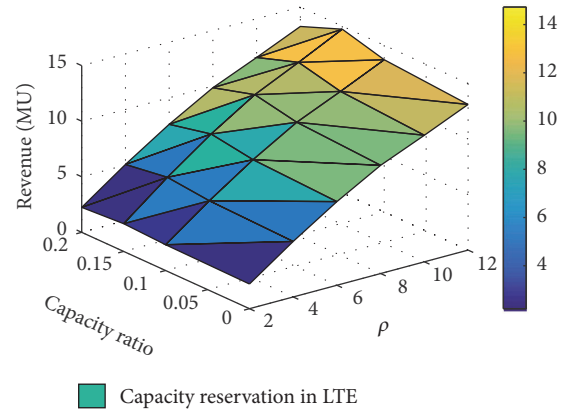


FIGURE 6: Revenue, different ratio of reserved capacity (LTE only).

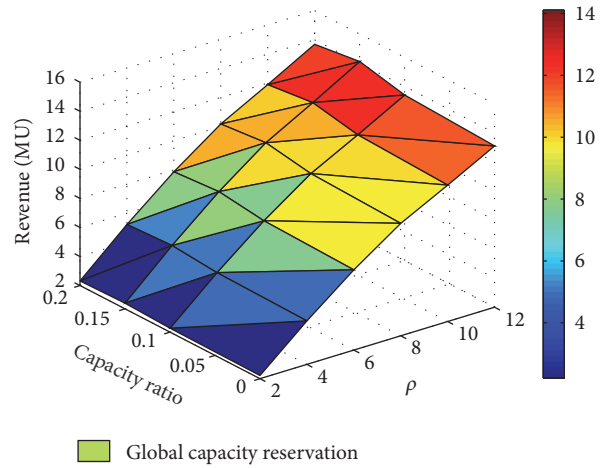


FIGURE 7: Revenue, different ratio of reserved capacity (global reservation).

in LTE only and when it is reserved across the HWN. Unlike the investigation above where the reserved capacity for SS is fixed as shown in Table 1, the values of reserved capacity ratio vary from 0% to 20% in the investigation of this subsection.

In the first case, that is, capacity reserved in LTE only, the revenue tends to increase with the increase of the reserved bandwidth up to a certain point, after which the revenue starts to decrease (Figure 6). This is due to the fact that when the ratio of reserved capacity is high, some resources will be reserved unnecessarily, while they could have been exploited to admit IAS traffic. We lose therefore some revenue that could have been achieved if more IAS traffic was admitted.

In the second case, where the bandwidth reservation for SS traffic is done across all RATs in the HWN, the behavior is similar (Figure 7). However, we notice that the graph is smoother than in the first case. This allows us to conclude that varying the ratio of reserved capacity for SS traffic in this case has less impact on the revenue than when we reserve capacity in LTE only. This is due to the possibility of admitting larger number of IAS sessions when the bandwidth reservation for SS traffic is done across the HWN.

## 6. Conclusion

In this work, we present a model for RAT selection in HWNs where net neutrality is taken into consideration and with the objective of maximizing the revenue. We study particularly the exemption granted to SS traffic that the net neutrality regulation allows. Two variants of the RAT selection strategy are proposed, namely, bandwidth reservation in LTE only and bandwidth reservation in the whole HWN. The MDP formulation for the two variants is presented, and a model for the spatial distribution of the BSs, APs, and the users is provided. Our objective has been to give insight into how the bandwidth reservation for SS traffic should be made in order to ensure a maximum revenue while remaining compliant to net neutrality and how the ratio of reserved bandwidth should be chosen. In terms of performance, we conclude that reserving resources in the whole HWN may be more beneficial as it guarantees better social benefit than the other variant, as well as lower blocking probability for IAS traffic, at the expense of a marginal loss in the generated revenue. Moreover, the impact of the ratio of reserved capacity for SS traffic on the achieved revenue was investigated with both variants of the RAT selection policy.

As a final remark, using MDP in this work, we have managed to formulate the RAT selection problem. However, we were not able to obtain closed-form expression for the solution, that is, expressing the revenue as an explicit function of the adopted net neutrality approach and the involved parameters. To address this limitation, we have resorted to using mathematical tool to numerically solve the MDP problem and get the results. The numerical results presented in the paper were obtained through implementing the mathematical model in MATLAB. By varying the inputs along different angles, mainly traffic density (due to either traffic rate changes or session service time changes) and the ratio of reserved capacity, we have tried to give an overall picture. For the use of the results in this paper, an ISP, given its traffic condition, could do similar numerical investigation (e.g., as for Figure 6) to find out how much capacity it could reserve for SS traffic to maximize the revenue.

## Notations

$ C_{\text{targ}} $ :	Coverage area of the targeted LTE base station
$\phi^{(k)}$ :	Poisson Point Process distribution of RAT $k$
$R$ :	Radius of the circular area covered by an AP
$r$ :	Distance separating a typical user from the nearest AP
$\lambda^{(k)}$ :	Number of BS/AP per unit area of RAT $k$
$\lambda^{(u)}$ :	Number of camping users per unit area
$\lambda_i$ :	The arrival rate of class $i$
$P_{c,k}$ :	Coverage probability of RAT $k$
$\sigma_i$ :	The average number of sessions per second generated by a class $i$ user
$1/\mu_i$ :	The average session holding time of class $i$
$C_k$ :	Capacity (number of channels) of RAT $k$

$n_{j,i}$ :	Number of sessions of class $i$ in RAT $j$
$\mathbf{s}_k$ :	The state vector of RAT $k$
$\mathbf{s}$ :	The state vector of the system
$S$ :	The state space of the system
$a_i$ :	The action resulting from the arrival of a class $i$ session
$\mathbf{a}$ :	Vector representing the action taken following a decision epoch
$A$ :	The action space of the MDP
$A_s$ :	The action space of state $\mathbf{s}$
$C_{1,\text{res}}$ :	Number of reserved bbu for SS traffic in LTE
$C_{\text{res}}$ :	Number of reserved bbu for SS traffic in the HWN
$\tau(\mathbf{s}, \mathbf{a})$ :	Expected sojourn time in state $\mathbf{s}$ when action $\mathbf{a}$ is chosen
$P_{ss'}(\mathbf{a})$ :	Transition probability from state $\mathbf{s}$ to state $\mathbf{s}'$ when action $\mathbf{a}$ is chosen
$P_{c,2}^*$ :	Coverage probability of the targeted AP
$\pi_s$ :	Policy chosen at state $\mathbf{s}$
$\Pi$ :	Set of admissible policies
$r(\mathbf{s}, \mathbf{a})$ :	Reward function for state $\mathbf{s}$ when action $\mathbf{a}$ is chosen
$w_{j,i}$ :	Weight associated for admitting a class $i$ session in RAT $j$
$\pi^*$ :	Optimal RAT selection policy
$P_h$ :	Price charged for SS traffic
$P_l$ :	Price charged for IAS traffic
$\rho$ :	Total traffic load: sum of traffic load of SS and IAS traffic.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported in part by the EU FP7 Marie Curie Actions under Grant Agreement no. 607584 and in part by the EU H2020 Marie Skłodowska-Curie Actions under Grant Agreement no. 699924.

## References

- [1] S.-P. Yeh, S. Talwar, G. Wu, N. Himayat, and K. Johansson, "Capacity and coverage enhancement in heterogeneous networks," *IEEE Wireless Communications Magazine*, vol. 18, no. 3, pp. 32–38, 2011.
- [2] "Berec guidelines on the implementation by national regulators of European net neutrality rules," <http://berec.europa.eu/eng/documentregister/subjectmatter/berec/regulatorybest.practices/guidelines/6160-berec-guidelines-on-the-implementation-by-national-regulators-of-european-net-neutrality-rules>.
- [3] F. Sørensen, "On the origin of specialised services," Tech. Rep., 2014, <https://eng.nkom.no/topical-issues/news/on-the-origin-of-specialised-services>.



- [4] J. Tang and R. T. B. Ma, "Regulating monopolistic ISPS without neutrality," in *Proceedings of the 22nd IEEE International Conference on Network Protocols, ICNP*, pp. 374–384, USA, October 2014.
- [5] J. P. Choi and B.-C. Kim, "Net neutrality and investment incentives," *The RAND Journal of Economics*, vol. 41, no. 3, pp. 446–471, 2010.
- [6] R. T. B. Ma, J. Wang, and D. M. Chiu, "Paid prioritization and its impact on net neutrality," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 2, pp. 367–379, 2017.
- [7] T. Wu, "Wireless carterfone," in *Proceedings of the International Journal of Communication*, pp. 389–426, 2007.
- [8] R. W. Hahn, R. E. Litan, and H. J. Singer, "The economics of "wireless net neutrality"," *Journal of Competition Law and Economics*, vol. 3, no. 3, pp. 399–451, 2007.
- [9] L. Martínez, O. Álvarez San-Jaime, and J. Markendahl, "Net neutrality principles and its impact on quality of experience based service differentiation in mobile networks," in *Proceedings of the Regional Conference of the International Telecommunications Society (ITS)*, 2015.
- [10] P. Hosein, W. Choi, and W. Seok, "Disruptive network applications and their impact on network neutrality," in *Proceedings of the 17th IEEE International Conference on Advanced Communications Technology, ICACT 2015*, pp. 663–668, Republic of Korea, July 2015.
- [11] R. T. B. Ma and V. Misra, "The public option: a nonregulatory alternative to network neutrality," *IEEE/ACM Transactions on Networking*, vol. 21, no. 6, pp. 1866–1879, 2013.
- [12] E. Altman, M. K. Hanuwa, and R. Sundaresan, "Nonneutral network and the role of bargaining power in side payments," *NET-COOP2010-4th Workshop on Network Control and Optimization*, 2010.
- [13] 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.
- [14] E. Khloussy and Y. Jiang, "The impact of net neutrality on revenue and quality of service in wireless networks," in *the IEEE Consumer Communications Networking Conference (CCNC)*, pp. 12–15, 2018.
- [15] E. Stevens-Navarro, Y. Lin, and V. W. S. Wong, "An MDP-based vertical handoff decision algorithm for heterogeneous wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 2, pp. 1243–1254, 2008.
- [16] P. C. Fishburn and A. M. Odlyzko, "Dynamic behavior of differential pricing and quality of service options for the internet," *Decision Support Systems*, vol. 28, no. 1, pp. 123–136, 2000.
- [17] J. Hou, J. Yang, and S. Papavassiliou, "Integration of pricing with call admission control to meet QoS requirements in cellular networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 13, no. 9, pp. 898–910, 2002.
- [18] H.-S. Jo, Y. J. Sang, P. Xia, and J. G. Andrews, "Heterogeneous cellular networks with flexible cell association: a comprehensive downlink SINR analysis," *IEEE Transactions on Wireless Communications*, vol. 11, no. 10, pp. 3484–3494, 2012.
- [19] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A new tractable model for cellular coverage," in *the 48th Annual Allerton Conference on Communication, Control, and Computing, Allerton 2010*, pp. 1204–1211, USA, October 2010.
- [20] O. E. Falowo and H. A. Chan, "Multiple-RAT selection for reducing call blocking/dropping probability in cooperative heterogeneous wireless networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, 2012.
- [21] Y. Wang, K. I. Pedersen, P. E. Mogensen, and T. B. Sørensen, "Carrier load balancing methods with bursty traffic for LTE-advanced systems," in *Proceedings of the IEEE 20th Personal, Indoor and Mobile Radio Communications Symposium, PIMRC 2009*, Japan, September 2009.
- [22] H. Chen, C.-C. Cheng, and H.-H. Yeh, "Guard-channel-based incremental and dynamic optimization on call admission control for next-generation QoS-Aware heterogeneous systems," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 5, pp. 3064–3082, 2008.
- [23] D. Kumar, E. Altman, and JM. Kelif, "Globally optimal user-network association in an 802.11 WLAN & 3G UMTS hybrid cell," in *Managing Traffic Performance in Converged Networks*, pp. 1173–1187, Springer, Heidelberg, Berlin, 2007.
- [24] I. Chadès, G. Chapron, M.-J. Cros, F. Garcia, and R. Sabbadin, "MDPtoolbox: a multi-platform toolbox to solve stochastic dynamic programming problems," *Ecography*, 2014.

