Optimal Radio Access Technology Selection in an SDN based LTE-WiFi Network

Arghyadip Roy*, Prasanna Chaporkar*, Abhay Karandikar*†, Pranav Jha*
*Department of Electrical Engineering, Indian Institute of Technology Bombay, India 400076
Email: {arghyadip, chaporkar, karandi, pranavjha}@ee.iitb.ac.in
†Director and Professor, Indian Institute Technology Kanpur, India 208016
Email: karandi@iitk.ac.in

***Abstract*—Today's wireless networks consist of a multitude of Radio Access Technologies (RATs), each being controlled individually, leading to suboptimal utilization of network resources. However, the unprecedented growth of data traffic is creating the need for an efficient inter-working of various RATs to circumvent the problem of suboptimal utilization of resources. Application of Software Defined Networking (SDN) principles enables the control and management of various RATs in a unified way. In this paper, we specifically focus on the inter-working between Long Term Evolution (LTE) and Wireless Fidelity (WiFi). We propose an SDN based architecture for a network comprising LTE Base Stations (BSs) and WiFi Access Points (APs). Users can be offloaded from one RAT to another based on different criteria, viz., user priority and channel state of users. We consider the problem of optimal RAT selection to maximize the total system throughput subject to constraints on blocking probability of high priority users and offloading probability of high priority users and formulate it as a Constrained Markov Decision Process (CMDP). We propose a low-complexity RAT selection algorithm which does not require the knowledge of the statistics of system dynamics. To conduct experiments, we develop a Network Simulator-3 (ns-3) based evaluation platform in accordance with the SDN principles. Experimental results demonstrate that the proposed algorithm provides a near-optimal performance.**

I. Introduction

With the advent of Fourth Generation (4G) cellular networks, data-hungry applications such as video, social networking are becoming popular. Simultaneously, the number of mobile subscribers is also increasing. To cater to the increasing data traffic consumption and data rate demand, network operators are deploying low cost IEEE 802.11 based Wireless Local Area Network (WLAN) (popularly known as Wireless Fidelity (WiFi)) Access Points (APs) in hotspot areas. These kind of networks is known as Heterogeneous Networks (HetNets). While 4G Long Term Evolution (LTE) Base Stations (BSs) are deployed aiming at providing ubiquitous coverage, WiFi APs target to provide high data rate in hotspot regions. In regions where both LTE BS and WiFi AP coverages are present, a user can be associated with either of them and steered from one to another. This steering mechanism introduced in 3GPP Release 12 specifications [1] is known as mobile data offloading.

With the future Fifth Generation (5G) [2] standardization in progress, it is expected that future wireless networks will be a mixture of a large number of RATs. In existing network, every RAT is controlled by RAT-specific elements. For example, LTE is controlled by control elements such as Mobility Management Entity (MME) and Evolved NodeB (eNodeB), and WLAN is controlled by WLAN controllers. Even in the upcoming 5G network [2] which supports multiple RATs, radio access decisions are taken by RAT-specific elements. Therefore, while choosing control and management decisions, a global view of different RATs is not present in today's networks. This results in a suboptimal utilization of network-wide resources. To achieve the optimal network performance, common functionalities supported by different RATs such as admission control, flow control, mobility management need to be controlled and managed in a unified manner. Recent developments in Software Defined Networking (SDN) [3] may enable us to achieve unified control of various RATs.

SDN enables the split of control and data plane elements and functionalities in a network. Using SDN, the control plane functionalities of different RATs can be decoupled from network elements of various RATs and aggregated in the control plane. While the resulting control plane consists of control and management protocols and elements, the data plane consists of protocols and elements for data transfer. Since the control plane has a global view of the entire network, this approach facilitates the optimal utilization of network resources contrary to distributed control in today's network.

In this paper, we focus on the interworking between LTE and WiFi networks. We propose an SDN based network architecture which unifies the control and management functionalities of LTE and WiFi RATs using an SDN controller. The LTE BS and the WiFi AP forward the radio resource management messages to the SDN controller which takes the control and management decisions. We focus on the optimal RAT selection problem. We consider that users of different priorities are present in the network. The controller takes admission control decisions based on the user priority. We assume that high priority users are those users which require Guaranteed Bit Rate (GBR) (such as Voice Over Internet Protocol (VoIP), live streaming). High priority users are always served using LTE since WiFi may not provide the required Quality of Service (QoS). Low priority users are best effort class of users which may be served using LTE or WiFi. We assume that the available resource blocks in LTE, after a fixed number of resource block is allocated to every GBR high priority user, are equally distributed among the low priority users. However, the data rate obtained by an individual low priority user depends on the channel condition of the user. A high priority user may be blocked if it is not possible to provide the required QoS using LTE. We assume that the arrival of a high priority user and the departure of a high/low priority user from LTE (WiFi) can trigger the offloading of a low priority user to LTE (WiFi). Our target is to maximize the total system throughput. Generally, WiFi provides better throughput to users compared to LTE when WiFi load is less. However, depending on the channel conditions of the users, under high WiFi load, association with LTE may be preferable since the total throughput in WiFi decreases [4] with load. However, maximizing the total system throughput may result in excessive blocking of high priority users since their contribution towards the total system throughput is usually less than that of low priority users. Therefore, we consider a constraint on the blocking probability of high priority users. This problem has been addressed in our earlier works [5], [6]. However, the channel states of users are not considered in these works. Maximizing the total system throughput subject to a blocking probability constraint [5], [6], may lead to excessive offloading of low priority users. For example, upon the admission of a new high priority user in LTE, an existing low priority user may be offloaded to WiFi. However, if another high priority user departs from LTE, it may be optimal to offload one existing WiFi user to LTE. As a result, it may happen that within a short time interval, one user moves from LTE to WiFi and back to LTE again, leading to 'ping-pong' kind of behavior. Similar instances can occur in case of departures followed by arrivals also. This may generate additional control signaling in the backhaul. To address this, along with the high priority user blocking probability constraint, we also take into account the offloading probability of low priority users (i.e., fraction of offloaded low priority users) as a constraint. This problem can be modeled as a Constrained Markov Decision Process (CMDP) problem.

The conventional Dynamic Programming (DP) methods to solve the CMDP problem is computationally expensive in the face of large state and action spaces. Moreover, the computation of the optimal policy using DP methods requires the knowledge of transition probabilities of the underlying model which are governed by the statistics of the system dynamics, viz., the arrival rates of high and low priority users. This is difficult to obtain in reality. To address these issues, we propose an algorithm based on which RAT selection and offloading decisions can be taken in the SDN controller. Unlike DP based algorithms, the proposed algorithm has low computational and storage complexities. Furthermore, the proposed algorithm does not require the knowledge of the statistics of system dynamics and hence, is suitable for practical implementation. We develop an SDN based evaluation platform in Network Simulator-3 (ns-3) (a discrete event network simulator) to conduct the experiments in an integrated LTE-WiFi network. Building of this platform requires a significant restructuring of existing ns-3 modules. Experimental results demonstrate that the proposed algorithm provides a near-optimal performance.

A. Related Work

RAT selection and offloading solutions proposed in the literature can be mainly classified into two categories, viz., user-initiated [7]–[9] and network-initiated [5], [6], [10], [11]. In [9], "on-the-spot offloading", which always steers a data user to WiFi inside the WiFi coverage, is proposed. In [8], the problem where each user attempts to maximize its own utility is formulated as a non-cooperative game. However, due to the emergence of SDN as a part of future 5G networks, network-initiated RAT selection and offloading solutions are gaining popularity. Among the SDN based network-initiated RAT selection and offloading approaches [12]–[15], authors in [13] consider an SDN-enabled dynamic path selection problem in a multi-RAT system and propose an algorithm which chooses the path based on the rate obtained. The rate obtained takes into account factors like radio conditions, performance requirement of different flows and load conditions. In [14], authors propose a QoS-aware RAT selection algorithm based on a metric which takes into account bit rate requirements of users and capabilities of different RATs. A user association heuristic which considers multiple traffic classes and scales well with the LTE/WiFi HetNet system, is proposed in [15].

Unlike [5], [6], we consider channel states of users and a constraint on the offloading probability of low priority users for RAT selection. The proposed algorithm in this paper does not require the knowledge of the model. Although model-free learning techniques are adopted in literature [11], contrary to our approach, they still suffer from the curse of dimensionality. The development of SDN based evaluation platform in ns-3 enables us to characterize the performance of our proposed algorithm in a practical LTE-WiFi network.

The rest of the paper is organized as follows. In Section II, we present the system architecture. Section III and IV describe the system model and the problem formulation, respectively. We describe the proposed algorithm in Section V along with an analysis of computational and storage complexities. Performance of the proposed algorithm in the SDN based evaluation platform is described in Section VI. Section VII concludes the paper.

II. Proposed System Architecture

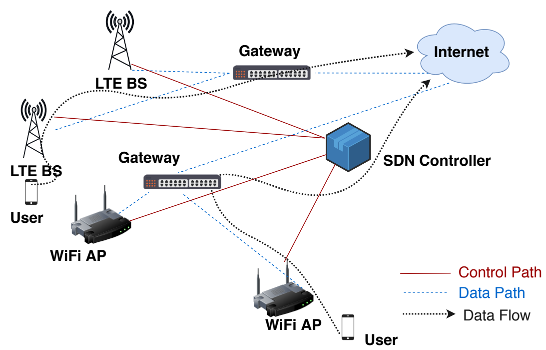


Figure 1: SDN based LTE-WiFi architecture.

In this section, we propose an overlay architecture that allows us to handle LTE and WiFi networks together using SDN. The proposed architecture consists of an SDN controller. As demonstrated in Fig.1, the SDN controller handles all control and management related functionalities. To this end, the Radio Resource Management (RRM) unit of the LTE BS is moved to the SDN controller. In effect, decision making related functionalities are implemented in the SDN controller which has a unified view of the entire network. RRM related control messages sent by users in LTE are forwarded by the LTE BS to the SDN controller. For example, the Radio Resource Control (RRC) connection request message reaches the controller via the LTE BS. Remaining functionalities of RRC after the removal of RRM, remains in the LTE BS. Similarly, the association request message is forwarded by the WiFi AP to the SDN controller. In spite of the fact that we have a single controller, scalability issues do not arise since only a small fraction of control signals (which are RRM related) is handled by the SDN controller. Note that channel condition information of users is needed for taking RAT selection and offloading decisions. For this purpose, channel condition information of users are forwarded to the SDN controller at the time of association of users.

III. System Model

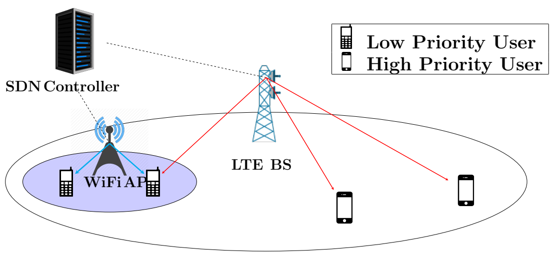


Figure 2: SDN based LTE-WiFi network.

The system model described in Fig. 2 consists of an LTE BS and a WiFi AP inside the coverage area of the LTE BS. The LTE BS and the WiFi AP are connected to the SDN controller using lossless links. We assume that high and low priority users are present at any geographical point inside the coverage area of the LTE BS. Low priority users which are present outside the dual coverage area of LTE BS and WiFi AP, get associated with the LTE BS always. Without loss of generality, we consider only those low priority users which are present in the common coverage area of the WiFi AP and the LTE BS. Low priority users can be associated with either the LTE BS or the WiFi AP. We assume that high and low priority users are allocated resources in LTE from a common resource pool. We assume that in LTE, users can be of either "good" or "bad" channel state. We assume that based on the location of users, the coverage area of the LTE BS is divided into two regions, viz., cell center and cell edge regions. Since cell edge users are present in the vicinity of cell boundary, usually they receive weaker signal strength than the cell center users. Therefore, it is assumed that users present in the cell center region have good channels, whereas cell edge users have bad channels. Selection of cell center/ cell edge region can be done based on the average Channel Quality Indicator (CQI) experienced by the users in LTE. If the average CQI of a user exceeds a certain threshold, then the user is called a cell center user, a cell edge user otherwise. We assume that the users are stationary, and the channel states do not change with time once the user is admitted. Channel states of incoming users are assumed to be known at the controller, however channel states in LTE are either good or bad with finite probabilities. Since the coverage area of the WiFi AP is small, we assume that channel states of users in WiFi are always good.

We assume that high and low priority user arrivals are Poisson processes with means λ_H and λ_L, respectively. The service times for high and low priority users are exponentially distributed with means $\frac{1}{\mu_H}$ and $\frac{1}{\mu_L}$, respectively. Assumptions on service times are in accordance with [16].

A. State Space

We model the system as a controlled continuous time stochastic process $\{X(t)\}_{t\geq 0}$. We represent a state s in the state space $\mathcal{S}$ as $s = (i_G, i_B, j_G, j_B, k_G, k_B)$, where i_G, i_B denote the number of high priority users associated with the LTE BS with good and bad channels in LTE, j_G, j_B denote the number of low priority users associated with the LTE BS with good and bad channels in LTE, and k_G, k_B denote the number of low priority users associated with the WiFi AP with good and bad channels in LTE, respectively. Note that we do not explicitly mention the channel states of users in WiFi since the channel states of users in WiFi are always good. The arrival and departure of high and low priority users with good and bad channel states in LTE are taken as decision epochs. It is easy to see that the system changes state only at these decision epochs. Also, due to Markovian nature of the system, it is sufficient to observe the system state at these decision epochs and not at other points in time.

Whenever there is an arrival or a departure of user, we refer to it as an event. The system changes state whenever an event occurs. Let the set of all events be denoted by $\mathcal{E}$. $\mathcal{E}$ consists of arrival and departure of high and low priority users. Let the arrival of a high and low priority user with good (bad) channel be denoted by $E_1(E_3)$ and $E_2(E_4)$, respectively. We assume that the departure of a high and low priority user with good (bad) channel are denoted by $E_5(E_6)$ and $E_7(E_8)$, respectively. We denote the departure of a low priority user from WiFi with good and bad channel in LTE by E_9 and E_{10}, respectively. Note that, the channel states of users in WiFi does not appear in the event space because the channel states of users in WiFi are always good. At every decision epoch, the SDN controller chooses a decision based on the current system state and the event. Based on the decision, the system makes a transition to different states with finite probabilities.

Let the LTE system is composed of C_L resource blocks. We assume that $s = (i_G, i_B, j_G, j_B, k_G, k_B) \in \mathcal{S}$ if $(i_G + 2i_B) \leq C_L$, $(j_G + 2j_B) \leq N$ and $(k_G + k_B) < W$, where N is a sufficiently large positive integer ($N \gg C_L$). The first two conditions are based on the assumption that a users with bad channel requires twice as many resource blocks as required by a user with good channel. The first condition also signifies that the admitted high priority user is provided the required number of resource blocks, whenever resources are available. The quantity W signifies the maximum number of users that can be supported in WiFi with a specified minimum per-user throughput guarantee. Note that the per-user throughput of WiFi decreases monotonically with the number of WiFi users [4]. Since high priority users require GBR ($R_{L,H}$, say), a fixed number of resource blocks are allocated to high priority users based on the channel condition of the user. However, since low priority users are best-effort in nature, the remaining resources in LTE are allocated uniformly among low priority users. Therefore, the bit rates obtained by low priority users (which is a function of the channel state) depend on the number of high priority users in the system. We assume that the bit rate obtained by a low priority user with bad channel is $\frac{1}{d}(d > 1)$ times that of a low priority user with good channel, where d is a constant.

B. Action Space

Let the action space (set of all possible association decisions in case of arrivals and departures) be denoted by $\mathcal{A}$. Action A_1 corresponds to blocking of an arriving user or doing nothing during a departure. Actions A_2 and A_3 refers to association with LTE and WiFi, respectively. Action A_4 accepts a high priority user in LTE and offloads a low priority user with bad channel to WiFi. Action A_5 offloads a low priority user with bad (good) channel from LTE (WiFi) to WiFi (LTE) upon the departure of a user from WiFi (LTE). Action A_6 accepts a high priority user in LTE and offloads a low priority user with good channel to WiFi. Action A_7 offloads a low priority user with good (bad) channel from LTE (WiFi) to WiFi (LTE) upon the departure of a user from WiFi (LTE). In case of high priority user arrivals, the feasible action set is $\{A_1, A_2, A_4, A_6\}$. In case of low priority user arrivals, the feasible action set is $\{A_2, A_3\}$. In case of departures, the feasible action set comprises A_1, A_5 and A_7, respectively. Note that blocking is a feasible action for high priority users only when the system is non-empty. On the contrary, low priority users are blocked only when $(j_G + 2j_B)$ becomes N.

C. Transition Probabilities

From each state $s \in \mathcal{S}$ and under each feasible action $a \in \mathcal{S}$, the system moves to a different state $s' \in \mathcal{S}$ with a positive probability $p_{ss'}(a)$. Let the sum of arrival and service rates of users in state $s = (i_G, i_B, j_G, j_B, k_G, k_B)$ be denoted by $v(i_G, i_B, j_G, j_B, k_G, k_B)$. Therefore,

$$\begin{aligned} v(i_G, i_B, j_G, j_B, k_G, k_B) &= \lambda_H + \lambda_L + (i_G + i_B)\mu_H \\ &+ (j_G + j_B + k_G + k_B)\mu_L. \end{aligned}$$

Let $\hat{s} = (i'_G, i'_B, j'_G, j'_B, k'_G, k'_B)$ and $e_{\{i:1 \leq i \leq 6\}}$ be a set of 6 dimensional vectors with all elements being zero except the i^{th} element being '1'. Then,

$$p_{ss'}(a) = \begin{cases} \frac{\lambda_H p_g}{v(i'_G, i'_B, j'_G, j'_B, k'_G, k'_B)}, & s' = \hat{s}, \\ \frac{\lambda_H (1-p_g)}{v(i'_G, i'_B, j'_G, j'_B, k'_G, k'_B)}, & s' = \hat{s}, \\ \frac{\lambda_L p_g}{v(i'_G, i'_B, j'_G, j'_B, k'_G, k'_B)}, & s' = \hat{s}, \\ \frac{\lambda_L (1-p_g)}{v(i'_G, i'_B, j'_G, j'_B, k'_G, k'_B)}, & s' = \hat{s}, \\ \frac{i'_G \mu_H}{v(i'_G, i'_B, j'_G, j'_B, k'_G, k'_B)}, & s' = \hat{s} - e_1, \\ \frac{i'_B \mu_H}{v(i'_G, i'_B, j'_G, j'_B, k'_G, k'_B)}, & s' = \hat{s} - e_2, \\ \frac{j'_G \mu_L}{v(i'_G, i'_B, j'_G, j'_B, k'_G, k'_B)}, & s' = \hat{s} - e_3, \\ \frac{j'_B \mu_L}{v(i'_G, i'_B, j'_G, j'_B, k'_G, k'_B)}, & s' = \hat{s} - e_4, \\ \frac{k'_G \mu_L}{v(i'_G, i'_B, j'_G, j'_B, k'_G, k'_B)}, & s' = \hat{s} - e_5, \\ \frac{k'_B \mu_L}{v(i'_G, i'_B, j'_G, j'_B, k'_G, k'_B)}, & s' = \hat{s} - e_6, \end{cases}$$

where p_g denotes the probability that the channel state of the arriving user in LTE is good. Values of $i'_G, i'_B, j'_G, j'_B, k'_G, k'_B$ as a function of different actions a (conditioned on events E_l) are described in Table I.

Table I: Transition Probability Table.

$\boldsymbol{a \vert E_l}$	$\boldsymbol{(i'_G, i'_B, j'_G, j'_B, k'_G, k'_B)}$
$A_1 \vert \mathcal{E} \cap (E_2 \cup E_4)^{\complement}$	$(i_G, i_B, j_G, j_B, k_G, k_B)$
$A_2 \vert E_1$	$(i_G + 1, i_B, j_G, j_B, k_G, k_B)$
$A_2 \vert E_2$	$(i_G, i_B, j_G + 1, j_B, k_G, k_B)$
$A_2 \vert E_3$	$(i_G, i_B + 1, j_G, j_B, k_G, k_B)$
$A_2 \vert E_4$	$(i_G, i_B, j_G, j_B + 1, k_G, k_B)$
$A_3 \vert E_2$	$(i_G, i_B, j_G, j_B, k_G + 1, k_B)$
$A_3 \vert E_4$	$(i_G, i_B, j_G, j_B, k_G, k_B + 1)$
$A_4 \vert E_1$	$(i_G + 1, i_B, j_G, j_B - 1, k_G, k_B + 1)$
$A_4 \vert E_3$	$(i_G, i_B + 1, j_G, j_B - 1, k_G, k_B + 1)$
$A_5 \vert (E_5 \cup \ldots \cup E_8)$	$(i_G, i_B, j_G + 1, j_B, k_G - 1, k_B)$
$A_5 \vert (E_9 \cup E_{10})$	$(i_G, i_B, j_G, j_B - 1, k_G, k_B + 1)$
$A_6 \vert E_1$	$(i_G + 1, i_B, j_G - 1, j_B, k_G + 1, k_B)$
$A_6 \vert E_3$	$(i_G, i_B + 1, j_G - 1, j_B, k_G + 1, k_B)$
$A_7 \vert (E_5 \cup \ldots \cup E_8)$	$(i_G, i_B, j_G, j_B + 1, k_G, k_B - 1)$
$A_7 \vert (E_9 \cup E_{10})$	$(i_G, i_B, j_G - 1, j_B, k_G + 1, k_B)$

D. Rewards and Costs

Depending on the system state and the action chosen, a finite amount of reward is obtained. In WiFi, the total throughput depends on the total load of WiFi comprising low priority users both with good and bad channels in LTE. Let $R_{W,D}(k)$ denote the per-user throughput of k users in WiFi under full buffer traffic model [4]. $R_{W,D}(k)$ is a function of success and collision probabilities which arise due to the contention-based medium access of WiFi users and slot times for idle, busy (due to collision) and successful transmissions. Let the reward rate in state s and under action a be denoted by $r(s, a)$. The reward rate under a state-action pair is the sum of throughput of users in LTE and WiFi under the action. Let us define

$$\begin{aligned} R(i_G, i_B, j_G, j_B, k_G, k_B) &= (i_G + i_B)R_{L,H} \\ &+ \frac{(C_L - i_G - 2i_B)}{(j_G + j_B)} R_{L,L}(j_G + \frac{j_B}{d}) \\ &+ (k_G + k_B)R_{W,D}(k_G + k_B), \end{aligned} \tag{1}$$

where $R_{L,L}$ is the data rate corresponding to a single resource block in LTE for low priority data users with good channel state. The exhaustive description of reward rates in state s under different event-action pair is provided in Table II.

Table II: Reward Rate Table.

$\boldsymbol{(a \vert E_l)}$	$\boldsymbol{r(s, a)}$
$(A_1 \vert \cup_{E_i \in \mathcal{E}} E_i)$	$R(i_G, i_B, j_G, j_B, k_G, k_B)$
$(A_2 \vert E_1)$	$R(i_G + 1, i_B, j_G, j_B, k_G, k_B)$
$(A_2 \vert E_2)$	$R(i_G, i_B, j_G + 1, j_B, k_G, k_B)$
$(A_2 \vert E_3)$	$R(i_G, i_B + 1, j_G, j_B, k_G, k_B)$
$(A_2 \vert E_4)$	$R(i_G, i_B, j_G, j_B + 1, k_G, k_B)$
$(A_3 \vert E_2)$	$R(i_G, i_B, j_G, j_B, k_G + 1, k_B)$
$(A_3 \vert E_4)$	$R(i_G, i_B, j_G, j_B, k_G, k_B + 1)$
$(A_4 \vert E_1)$	$R(i_G + 1, i_B, j_G, j_B - 1, k_G, k_B + 1)$
$(A_4 \vert E_3)$	$R(i_G, i_B + 1, j_G, j_B - 1, k_G, k_B + 1)$
$(A_5 \vert E_5 \cup \ldots \cup E_8)$	$R(i_G, i_B, j_G + 1, j_B, k_G - 1, k_B)$
$(A_5 \vert E_9 \cup E_{10})$	$R(i_G, i_B, j_G, j_B - 1, k_G, k_B + 1)$
$(A_6 \vert E_1)$	$R(i_G + 1, i_B, j_G - 1, j_B, k_G + 1, k_B)$
$(A_6 \vert E_2)$	$R(i_G, i_B + 1, j_G - 1, j_B, k_G + 1, k_B)$
$(A_7 \vert E_5 \cup \ldots \cup E_8)$	$R(i_G, i_B, j_G, j_B + 1, k_G, k_B - 1)$
$(A_7 \vert E_9 \cup E_{10})$	$R(i_G, i_B, j_G - 1, j_B, k_G + 1, k_B)$

We consider two types of cost functions, due to blocking and offloading, respectively. Let the cost rate for blocking and offloading in state s under action a be denoted by $c_b(s, a)$ and $c_o(s, a)$, respectively. Whenever the SDN controller blocks one high priority user, $c_b(s, a)$ is unity, else it is zero. Therefore,

$$c_b(s, a) = \begin{cases} 1, & \text{if high priority users are blocked,} \\ 0, & \text{otherwise.} \end{cases}$$

Whenever the SDN controller offloads one low priority user from one RAT to another, $c_o(s,a)$ is unity, else it is zero.

$$c_o(s,a) = \begin{cases} 1, & \text{if } a = (A_4||\ldots||A_7), \\ 0, & \text{otherwise.} \end{cases}$$

IV. Problem Formulation & Solution Techniques

We aim to determine a *policy* for the association of high and low priority users which maximizes the total system throughput subject to constraints on the blocking probability of high priority users and the offloading probability of low priority users. A policy is a mapping from a state to an action specifying which action is to be chosen in a state. The problem can be formulated as a CMDP problem. Since arrivals and departures of high and low priority users can occur at any arbitrary time, the considered problem is continuous time in nature. In this case, a *stationary randomized* optimal policy, i.e., a mixture of pure policies with corresponding probabilities, exists [17].

A. Problem Formulation

Let the set of memoryless policies be denoted by $\mathcal{M}$. We assume that the Markov chains induced by memoryless policies are unichain to guarantee a unique stationary distribution. Let the average reward, the cost due to blocking of high priority users and the cost due to offloading of low priority users over infinite horizon under policy $M \in \mathcal{M}$ be denoted by V^M, $C^{B,M}$ and $C^{O,M}$, respectively. Let the total reward, the cost due to blocking and the cost due to offloading till time t be denoted by $R(t)$, $C_B(t)$ and $C_O(t)$, respectively. The CMDP problem can be described as follows,

$$\begin{aligned} &\text{Maximize:} \quad V^M = \lim_{t\to\infty} \frac{1}{t}\mathbb{E}_M[R(t)], \\ &\text{subject to:} \quad C^{B,M} = \lim_{t\to\infty} \frac{1}{t}\mathbb{E}_M[C_B(t)] \le B_{\max} \text{ and} \quad (2) \\ &\quad C^{O,M} = \lim_{t\to\infty} \frac{1}{t}\mathbb{E}_M[C_O(t)] \le O_{\max}, \end{aligned}$$

where $\mathbb{E}_M$ is the expectation operator under policy M and $B_{\max}, O_{\max}$ denote the constraints on the blocking probability of high priority users and offloading probability of low priority users, respectively. Since the optimal policy is stationary, the limits in Equation (2) exist.

B. Conversion to Discrete-Time MDP and Lagrangian Approach

Optimal policy can be obtained using Relative Value Iteration Algorithm (RVIA) [18]. However, before that, we need to adopt Lagrangian approach [17]. For fixed values of Lagrange Multiplier (LM) β_b and β_o, the equivalent unconstrained reward function is given by

$$r(s,a;\beta_b;\beta_o) = r(s,a) - \beta_b c_b(s,a) - \beta_o c_o(s,a).$$

Using DP, the optimality equation for the considered Semi-Markov Decision Process (SMDP) $\forall s, s' \in \mathcal{S}$ is

$$V(s) = \max_a [r(s,a;\beta_b;\beta_o) + \sum_{s'} p_{ss'}(a)V(s') - \rho\bar{t}(s,a)],$$

where $V(s), \rho, \bar{t}(s,a)$ represent the value function of state $s \in \mathcal{S}$, the optimal average reward of the system and the mean transition time for state s and action a, respectively. Since the sojourn times are known to be exponential, this becomes a special case of continuous time controlled Markov chain, and therefore, the following equation holds.

$$0 = \max_a [r(s,a;\beta_b;\beta_o) - \rho + \sum_{s'} q(s'|s,a)V(s')], \quad (3)$$

where $q(s'|s,a)$ are controlled transition rates which satisfy $q(s'|s,a) \ge 0$, for $s' \ne s$ and $\sum_{s'} q(s'|s,a) = 0$. Scaling the transition rates by a positive scalar quantity is equivalent to time scaling. This scales the average reward for every policy without changing the optimal policy. Therefore, we assume (without loss of generality) that $-q(s|s,a) \in (0,1) \forall a$. This implies that $q(s'|s,a) \in [0,1]$ for $s' \ne s$. We add $V(s)$ to both sides of Equation (3) to obtain the following equation for an equivalent discrete-time MDP ($\{X_n\}$ say) with controlled transition probabilities $p_{ss'}(a)$.

$$V(s) = \max_a [r(s,a;\beta_b;\beta_o) - \rho + \sum_{s'} p_{ss'}(a)V(s')], \quad (4)$$

where $p_{ss'}(a) = q(s'|s,a)$ for $s' \ne s$ and $p_{ss'}(a) = 1 + q(s'|s,a)$ for $s' = s$. For the rest of the paper, instead of the original continuous-time MDP, we focus on the obtained equivalent discrete-time MDP in Equation (4).

For fixed β_b and β_o, we can use RVIA to solve the unconstrained maximization problem in Equation (4) using the following equation.

$$V_{n+1}(s) = \max_a [r(s,a;\beta_b;\beta_o) + \sum_{s'} p_{ss'}(a)V_n(s') - V_n(s^*)], \quad (5)$$

where $V_n(s)$ is the value function estimate of state s after n iterations, and s^* is a fixed state. We aim to obtain the the optimal values for β_b and β_o, viz., ${\beta_b}^*$ and ${\beta_o}^*$, which maximize the average reward subject to cost constraints. The following equations describe gradient descent routines to update the values of β_b and β_o in k^{th} iteration.

$$\beta_{b,k+1} = \beta_{b,k} + \frac{1}{k}(B^{\pi_{\beta_{b,k}}} - B_{\max}),$$

$$\beta_{o,k+1} = \beta_{o,k} + \frac{1}{k}(O^{\pi_{\beta_{o,k}}} - O_{\max}),$$

where $\beta_{b,k}, \beta_{o,k}$ are the values of β_b and β_o in k^{th} iteration, and $B^{\pi_{\beta_{b,k}}}, O^{\pi_{\beta_{o,k}}}$ denote the high priority user blocking probability and low priority user offloading probability in k^{th} iteration, respectively. Note that the optimal policy for the considered CMDP is a randomized policy with randomizations in at most two states [19].

V. Proposed RAT Selection Algorithm

In Section IV, the maximization of total system throughput subject to constraints on high priority user blocking probability and low priority user offloading probability is formulated as a CMDP problem which can be solved using DP techniques. However, DP based methods suffer from the curse of dimensionality. For example, in traditional policy iteration

[18], the computational complexity is $O(|\mathcal{A}|^{|\mathcal{S}|})$ which is exponential in the cardinality of the state space. Furthermore, computation of the optimal policy requires the knowledge of the state transition probabilities which are governed by the statistics of arrival processes of high and low priority users. In practice, statistics of arrival processes may be unknown. Although learning based approaches [11] which do not require the knowledge of statistics of arrival processes may be adopted, usually their convergence rate is very slow. To address these issues, we propose a low-complexity algorithm which is practically implementable. Moreover, it does not need the knowledge of statistics of arrival processes.

A. Myopic with Constraint Satisfaction Algorithm

In this subsection, we propose an algorithm which is myopic in the sense that it only optimizes based on the current reward and does not look into the future utility. However, the proposed algorithm, called Myopic with Constraint Satisfaction Algorithm (MCSA), satisfies the associated constraints on blocking probability of high priority users and offloading probability of low priority users. The complete description of MCSA is provided in Algorithm 1.

We first determine the event in the current decision epoch. Then, we determine the best action (denoted by a^*) based on the current reward (Line 4). If the current event is low priority user arrival (event E_2 and E_4), then irrespective of the channel condition of the incoming user, we always choose the action a^*. Since the feasible actions (A_2 and A_3) in case of low priority user arrivals neither affect the blocking probability of high priority users nor affect the offloading probability of low priority users, we always act in a myopic manner. However, if the current event is high priority user arrival (event E_1 and E_3), then we initially increment the counter corresponding to the number of high priority user arrivals (denoted by A_H). If the current value of blocking probability (denoted by B_H) is less than the specified constraint $B_{\max}$, then we block the arriving high priority user (Line 16). Note that, we keep a small margin ϵ_B on $B_{\max}$ to ensure that in the long run the system operates below $B_{\max}$. However, if B_H is more than $B_{\max} - \epsilon_B$ and the current value of offloading probability of low priority users (denoted by O_L) is less than the specified constraint $O_{\max}$, then we always choose the action a^* (Line 11). If the current value of O_L violates the constraint, then A_2 is selected (Line 12) since choosing A_4 or A_6 may further increase the value of O_L. Similar to $B_{\max}$, we keep a small margin ϵ_O on $O_{\max}$. Based on whether action involving blocking (A_1) or offloading (A_4 and A_6) is chosen (denoted by F_B and F_O, respectively), we update the current value of B_H and O_L (Line 20 and 21). Similar procedures are followed in case of departures, where the corresponding counter (denoted by D) is updated, and depending on the value of O_L, actions are selected (Line 25-27). Based on whether A_5 or A_7 is chosen, O_L is updated (Line 28-29). Note that unlike DP methods, MCSA does not require the knowledge of transition probabilities of the underlying model.

B. Complexity Analysis

In this subsection, we analyze the computational and storage complexities of the optimal policy and the proposed MCSA.

Algorithm 1 Myopic with Constraint Satisfaction Association Algorithm.

Input: $R_{L,H}, R_{L,L}, R_W(.), B_{\max}, O_{\max}$.
1: Initialize $A_H \leftarrow 0$, $D \leftarrow 0$, $B_H \leftarrow 0$, $O_L \leftarrow 0$, $F_B \leftarrow 0$ and $F_O \leftarrow 0$.
2: **while** TRUE **do**
3: Determine the event E in the current decision epoch.
4: Set $a^* \leftarrow \arg\max_{a \in \mathcal{A}} r(s,a)$.
5: **if** $(E = E_2 || E_4)$ **then**
6: Choose action $a = a^*$.
7: **else if** $(E = E_1 || E_3)$ **then**
8: $A_H \leftarrow A_H + 1$.
9: **if** $B_H > (B_{\max} - \epsilon_B)$ **then**
10: **procedure** HP–CONSTRAINT–VIOLATION
11: If $O_L < (O_{\max} - \epsilon_O)$ choose $a = a^*$.
12: Else choose $a = A_2$.
13: $F_0 \leftarrow I_{\{a=A_4||A_6\}}$.
14: **end procedure**
15: **else**
16: Choose action $a = A_1$.
17: **end if**
18: **procedure** UPDATE–BP–OP
19: $F_B \leftarrow I_{\{a=A_1\}}$.
20: $B_H \leftarrow \frac{B_H A_H + F_B}{(A_H+1)}$.
21: $O_L \leftarrow \frac{O_L(A_H+D)+F_O}{(A_H+D+1)}$.
22: **end procedure**
23: **else**
24: **procedure** DEPARTURE–POLICY
25: $D \leftarrow D + 1$.
26: If $O_L < (O_{\max} - \epsilon_O)$, choose $a = a^*$.
27: Else choose $a = A_1$.
28: $F_0 \leftarrow I_{\{a=A_5||A_7\}}$.
29: $O_L \leftarrow \frac{O_L(A_H+D)+F_O}{(A_H+D+1)}$.
30: **end procedure**
31: **end if**
32: **end while**

The optimal policy needs to store the optimal action corresponding to every state, resulting in a storage complexity of $O(|\mathcal{S}|)$. Also, the computation of the optimal policy using traditional policy iteration [18] involves a worst case complexity of $O(|\mathcal{A}|^{|\mathcal{S}|})$ since the total number of feasible policies is $|\mathcal{A}|^{|\mathcal{S}|}$. Therefore, it is very cumbersome to compute the optimal policy using traditional DP methods.

In the case of MCSA, whenever an event occurs, we need to compute the best action a^*. Therefore, the per-iteration computational complexity of the MCSA is $O(|\mathcal{A}|)$. MCSA requires to store the running value of A_H, D, B_H and O_L. However, it does not need to store any information regarding the state space. Therefore, the resulting storage complexity is $O(1)$, which is significantly better than the optimal policy.

VI. SIMULATION RESULTS

In this section, we implement the proposed algorithm in an SDN based evaluation platform built by us using ns-3. We observe the performance of MCSA in terms of blocking probability of high priority users, offloading probability of low priority users and total system throughput and compare with those of the optimal policy.

A. Simulation Setup and Methodology

We setup an evaluation platform (based on ns-3) which provides a framework to simulate the SDN controller based integrated control of LTE and WiFi networks. To this end, we restructure few existing modules present in ns-3. We create an SDN controller node which has two interfaces towards the LTE BS and the WiFi AP, respectively, over Internet Protocol (IP) connections. RRM functionalities present in the LTE BS are moved to the SDN controller. RRM related control signals in LTE and control signals in WiFi are forwarded to the SDN controller. However, data plane traffic is routed directly from the LTE BS to the gateway. To enable the communication between the LTE BS and the SDN controller, an application is developed. The application sends the control messages encapsulated in another control message with suitable headers. Whenever an event (arrival/departure of users) occurs, a control packet is sent from the user to the BS/AP which forwards the packet to the SDN controller, The SDN controller then chooses an action according to the implemented algorithm.

The considered network model comprises a 3GPP LTE BS and an IEEE 802.11g WiFi AP inside the coverage area of the LTE BS. Users are assumed to be stationary. We consider that the radius of the coverage area of the WiFi AP is approximately 30 m. The distance between the LTE BS and the WiFi AP is approximately 50 m. We assume that the WiFi AP is deployed by the cellular operators, and hence, the interworking is trusted in nature. LTE and WiFi network parameters summarized in Tables III and IV, are chosen based on 3GPP [20]- [21] models and saturation throughput [4] IEEE 802.11g WiFi [22] model. In simulations, we assume that the maximum data rate which a low priority user can obtain is 10 Mbps due to bottleneck in the access network. We set $B_{\max} = O_{\max} = 0.05$, $\epsilon_B = \epsilon_O = 0.01$.

Table III: LTE Network Model.

Parameter	**Value**
High priority user capacity	4 users
Bit rate of a high priority user	20 kbps
Voice packet payload	50 bits
Data packet payload	600 bits
Tx power for BS and MS	46 dBm and 23 dBm
Noise figure for BS and MS	5 dB and 9 dB
Antenna height for BS and MS	32 m and 1.5 m
Antenna type for BS and MS	Isotropic Antenna
Path loss (R in kms)	$128.1 + 37.6 \log(R)$
Multi-path fading	Extended Pedestrian A model [23]

B. High Priority Arrival Rate Variation

Fig. 3a describes the high priority user blocking probability of the proposed algorithm and the optimal policy. As λ_H increases, the blocking probability of the optimal policy increases. Since MCSA blocks high priority users based on the value of $B_{\max}$ without considering λ_H, the blocking probabilities are nearly same for all λ_Hs. In Fig. 3b, we plot the low priority user offloading probabilities for the considered algorithms. The low priority user offloading probability of MCSA is a constant for all values of λ_H due to similar reasons as that of Fig. 3a. However, in case of optimal policy, it gradually rises with λ_H because actions involving offloading (A_4, A_5, A_6, A_7) are selected more frequently. However, the total system throughput of MCSA is very close to that of the optimal policy (see Fig. 3c).

Table IV: WiFi Network Model.

Parameter	**Value**
Channel bit rate	54 Mbps
UDP header	224 bits
Packet payload	1500 bytes
Slot duration	20 μs
Short inter-frame space (SIFS)	10μs
Distributed Coordination Function IFS (DIFS)	50μs
Minimum acceptable per-user throughput	4.5 Mbps
Tx power for AP	23dBm
Noise figure for AP	4 dB
Antenna height for AP	2.5 m
Antenna parameter	Isotropic antenna
Path loss (R in kms)	$140.3 + 36.7 \log(R)$
Fading	Rayleigh fading

C. Low Priority Arrival Rate Variation

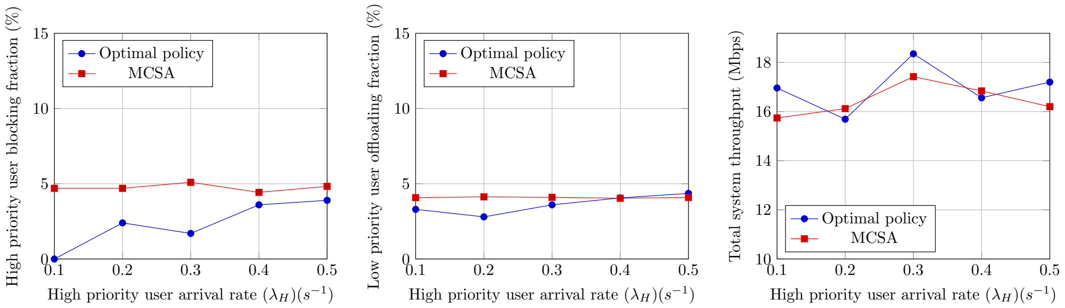
Fig. 4a describes the high priority user blocking probability of MCSA and optimal policy. Similar to Fig. 3a, MCSA exhibits blocking probability which is close to the given constraint for every value of λ_L. In Fig. 4b, we plot the offloading probability of low priority users as a function of λ_L. The offloading probability of the optimal policy grows with λ_L since more frequently actions involving offloading are selected. MCSA provides offloading probability which is close to the given constraint for all values of λ_L. In Fig. 4c, we observe that the performance of MCSA is close to optimal in terms of total system throughput. This happens because of the load balancing mechanism (similar to that of optimal policy) facilitated due to the centralized nature of MCSA.

VII. Conclusion

In this paper, we propose an SDN based architecture for an LTE-WiFi network and consider the optimal RAT selection problem in a system where users of multiple priorities are present. We aim to maximize the total system throughput subject to constraints on high priority user blocking probability and low priority user offloading probability. This problem is formulated as a CMDP problem. We then propose an algorithm for RAT selection which has low computational and storage complexities. Moreover, the proposed algorithm does not require the knowledge of the underlying transition probabilities of the model and hence, is suitable for practical implementation. Although myopic in nature, the proposed algorithm satisfies the associated constraints. To measure the performance of the proposed algorithm, we develop an SDN based evaluation platform which is implemented in ns-3. Experimental results exhibit that the proposed algorithm provides a near-optimal performance.

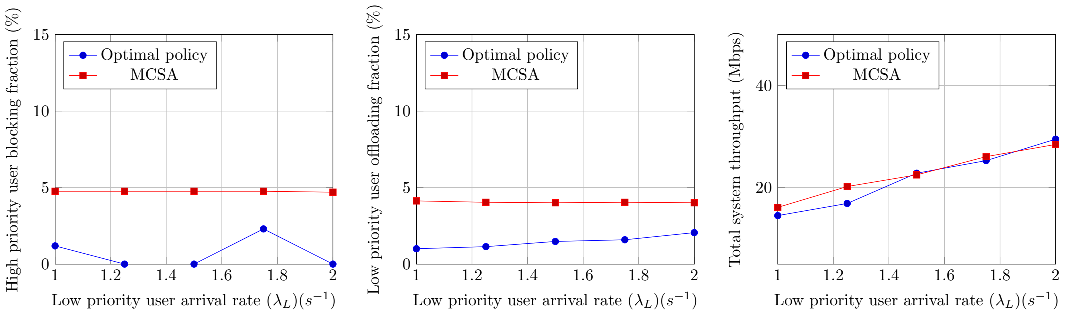
Acknowledgment

This work has been supported by the Department of Telecommunications, Ministry of Communications, India as part of the Indigenous 5G Test Bed project. The authors thank Ashish Sharma, Rohan Kharade and Abhishek Dandekar for their contributions towards development in the SDN based evaluation platform in ns-3.



(a) High priority user blocking percentage vs. λ_H.
(b) Low priority user offloading percentage vs. λ_H.
(c) Total system throughput vs. λ_H.

Figure 3: Plot of different system parameters for different algorithms under varying λ_H ($\lambda_L = 1, \mu_H = 1$ and $\mu_L = 1$).



(a) High priority user blocking percentage vs. λ_L.
(b) Low priority user offloading percentage vs. λ_L.
(c) Total system throughput vs. λ_L.

Figure 4: Plot of different system parameters for different algorithms under varying λ_L ($\lambda_H = 0.2, \mu_H = 1$ and $\mu_L = 1$).

References

[1] 3GPP TR 37.834 v0.3.0, "Study on WLAN/3GPP Radio Interworking," 2013.

[2] 3GPP TS 23.501, "System Architecture for the 5G System," 2017, [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144.

[3] A. M. Nayak, P. Jha, and A. Karandikar, "A centralized SDN architecture for the 5G cellular network," in *IEEE 5G WF*, 2018, pp. 147–152.

[4] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, 2000.

[5] A. Roy, P. Chaporkar, and A. Karandikar, "Optimal radio access technology selection algorithm for lte-wifi network," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 6446–6460, 2018.

[6] A. Roy and A. Karandikar, "Optimal radio access technology selection policy for LTE-WiFi network," in *IEEE WiOpt*, 2015, pp. 291–298.

[7] A. Whittier, P. Kulkarni, F. Cao, and S. Armour, "Mobile data offloading addressing the service quality vs. resource utilisation dilemma," in *IEEE PIMRC*, 2016, pp. 1–6.

[8] E. Aryafar, A. Keshavarz-Haddad, M. Wang, and M. Chiang, "RAT selection games in hetnets," in *IEEE INFOCOM*, 2013, pp. 998–1006.

[9] K. Lee, J. Lee, Y. Yi, I. Rhee, and S. Chong, "Mobile data offloading: How much can WiFi deliver?" *IEEE/ACM Transactions on Networking*, vol. 21, no. 2, pp. 536–550, 2013.

[10] S. Barmpounakis, A. Kaloxylos, P. Spapis, and N. Alonistioti, "Context-aware, user-driven, network-controlled RAT selection for 5G networks," *Computer Networks*, vol. 113, pp. 124–147, 2017.

[11] A. Roy, P. Chaporkar, and A. Karandikar, "An on-line radio access technology selection algorithm in an LTE-WiFi network," in *IEEE WCNC*, 2017, pp. 1–6.

[12] K. Chen, J. Liu, J. Martin, K.-C. Wang, and H. Hu, "Improving integrated LTE-WiFi network performance with SDN based flow scheduling," in *IEEE ICCCN*, 2018, pp. 1–9.

[13] S. Borst, A. Ö. Kaya, D. Calin, and H. Viswanathan, "Dynamic path selection in 5G multi-RAT wireless networks," in *IEEE INFOCOM*, 2017, pp. 1–9.

[14] A. Raschellà, F. Bouhafs, G. Deepak, and M. Mackay, "QoS aware radio access technology selection framework in heterogeneous networks using SDN," *Journal of Communications and Networks*, vol. 19, no. 6, pp. 577–586, 2017.

[15] D. Harutyunyan, S. Herle, D. Maradin, G. Agapiu, and R. Riggio, "Traffic-aware user association in heterogeneous LTE/WiFi radio access networks," in *IEEE/IFIP NOMS*, 2018, pp. 1–8.

[16] T. Bonald and J. W. Roberts, "Internet and the Erlang formula," *ACM SIGCOMM Computer Communication Review*, vol. 42, no. 1, pp. 23–30, 2012.

[17] E. Altman, *Constrained Markov decision processes*. CRC Press, 1999.

[18] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[19] F. J. Beutler and K. W. Ross, "Optimal policies for controlled Markov chains with a constraint," *Journal of mathematical analysis and applications*, vol. 112, no. 1, pp. 236–252, 1985.

[20] 3GPP TR 36.814 v9.0.0, "Further Advancements for E-UTRA Physical Layer Aspects," 2010.

[21] 3GPP TR 36.839 v11.1.0, "Mobility Enhancements in Heterogeneous Networks," 2012.

[22] IEEE 802.11-2012, Part 11, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," 2012.

[23] 3GPP TS 36.104 V10.2.0, "Base Station (BS) Radio Transmission and Reception," 2011.