Optimal Radio Access Technology Selection Policy for LTE-WiFi Network

Arghyadip Roy, Abhay Karandikar
Department of Electrical Engineering, Indian Institute of Technology Bombay, India 400076
Email: {arghyadip, karandi}@ee.iitb.ac.in

***Abstract*—As a part of Fourth Generation (4G) wireless communication system, a user can associate with any radio access technology (RAT) when there are multiple RATs available and can move seamlessly among them. To handle the explosive growth of traffic in cellular network, the idea of mobile data traffic offloading to Wireless Fidelity (WiFi) has been proposed. In this paper, we focus on mobile data offload assisted optimal association problem in a heterogeneous network. The problem has been formulated under the framework of constrained Markov decision process (CMDP). Value iteration and gradient descent algorithm has been used to determine the optimal policy. We aim to maximize the expected average per-user throughput of the system subject to a constraint on blocking probability of voice users. In our model, we consider the possibility of mobile data user offload from one RAT to another during association or departure of a user. Optimal policy has been found to follow a threshold structure. Numerical results are presented to show how the average per-user throughput of the system varies under different load conditions.**

I. Introduction

A heterogeneous network (HetNet) consists of multiple radio access technologies (RATs) and network architectures to improve coverage and capacity. In a typical HetNet, a subscriber can associate with a RAT from a pool of RATs, so as to provide the RAT with the best possible performance to the subscriber. Such HetNet systems may allow a user to move seamlessly among different RATs. The mechanism behind seamless movement of users from one RAT to another is called vertical handoff. Various algorithms have been proposed for vertical handoff in the last few years.

Traditionally, there has been many related works [1], [2], [3], [4] which investigate into the optimal access technology selection and admission control in the presence of different RATs. In [1], optimal RAT selection problem has been considered in a system which includes vertical handoff. In this work, the authors have determined Markov decision process (MDP) based optimal policies taking different objective functions like blocking probability and throughput into account. Based on the optimal policies, some heuristic policies with low computation complexity have been proposed and they showed significant improvement over the existing policies. In [2], optimal association among wireless stations (STAs) and access points (APs) has been considered based on the value of the sum of utilities of an association. The limitation of this work is that user arrival and departure scenario is not taken into consideration. In [3], the authors have focused on optimal client-AP association in a wireless Local Area Networks (WLAN) network and formulated the problem using continuous time MDP. The limitation in [2] has been addressed in this work. However, in this work, departure of a user has not been considered as a decision epoch. We have addressed this limitation in our work and formulated appropriate actions in case of a departure also.

A typical HetNet consists of IEEE 802.11 based WLAN (popularly known as Wireless Fidelity (WiFi)) APs, femtocells etc. overlaid with Long term evolution (LTE) base-stations (BSs). The widespread use of smartphones has resulted in significant growth in mobile data traffic. Different methods have been proposed to mitigate the problem of ever-increasing traffic in cellular networks. One of them is to steer the mobile data users to WiFi which typically operates in unlicensed band, providing a cost-effective solution. This approach requires close co-ordination between cellular and WiFi networks. With the widespread deployment of WiFi APs and IEEE 802.11u capable Hotspot 2.0, WiFi has become a prominent candidate for offloading of data users.

There are two different approaches to perform offloading of data from LTE based cellular network to WiFi, namely user-based offloading and network based offloading. User-based offloading gives user the capability to choose the RAT and doesn't require much changes in the network side. On the other hand, in network-based offloading, network takes the offloading decision. One user-based data offloading scheme from LTE to WLAN has been proposed in [5]. In this paper, an offloading approach based on combined information on signal quality and network load has been proposed. This kind of distributed user-based offloading algorithms tend to perform greedily and fail to converge to a global optimum. This often causes network-wide performance degradation. Moreover, the issue of user mobility is also not taken into account in this work.

Based on the time of offloading, offloading can again be classified into two types: *on-the-spot* and *delayed* offloading [6], [7]. On-the-spot offloading works on the principle that mobile data traffic is steered to WiFi, if available. Whenever there is no WiFi availability, mobile data traffic is transmitted using the cellular network. On the other hand, delayed offloading waits for some fixed time, if WiFi is not available. If WiFi availability doesn't resume in this interval, data is transmitted using cellular network. Performance evaluation of on-the-spot offloading has been investigated in [6] in terms of

978-3-9018-8273-9/15/ ©2015 IFIP

expected delay as the system metric. However, on-the-spot offloading offers limited performance improvement over existing schemes. In [7], a quantitative study of on-the-spot and delayed WiFi offloading has been done.

In [8], a network assisted WiFi offloading model has been considered with an objective of maximizing per-user throughput. The authors have formulated this problem as an optimization problem that determines the optimal fraction of users to be steered to WiFi. It has been found that the proposed scheme outperforms on-the-spot offloading [6]. However, the authors have not considered the blocking probability constraint and mobility aspect of users in this work.

In this paper, we study the optimal association problem in a LTE-WiFi heterogeneous network. We consider the possibility of network-based data offload of an active user from one RAT to another at the instance of the association of a new user and the departure of a user. To study the optimal association, a constrained discrete time MDP (CMDP) based model has been formulated. Per-user throughput has been set as the objective function subject to a constraint on the voice blocking probability. Our work is partly motivated by [1], but the networks under study are completely different. Above all, in the per-user throughput maximization problem formulation, we have introduced the voice blocking probability constraint which is not considered in [1]. In this work we have considered offloading only for data users and assumed voice users to be served using LTE only.

Our contribution in this work are: 1) We formulate the optimal association in a data offload capable heterogeneous network containing a LTE BS and a WiFi AP as an *infinite horizon* discrete time CMDP problem to determine optimal policy, 2) we analyze the optimal policy to observe that it has a threshold structure for a fixed number of users in WiFi. As the number of users in WiFi changes, the structure retains but the threshold value changes and 3) numerical results are presented under different arrival and service rates for both voice and data users.

The rest of the paper is organized as follows. The system model is described in Section II. In Section III, the problem has been formulated as CMDP with a description of the optimality criteria. Section IV presents the structure of the optimal policy and numerical results under different arrival and service rates for voice and data users. In Section V, we conclude the paper.

II. System Model

In this section, we formulate an optimal association problem as a CMDP problem. The system consists of a geographical area comprising single LTE BS and single WiFi AP as shown in Fig. 1. LTE BS has a larger coverage area than that of the WiFi AP. We assume voice users to be present under the coverage area of LTE BS. On the other hand, data users are distributed uniformly in the geographical area where both LTE and WiFi coverage is present. We don't consider mobility of users here. Voice and data user arrivals follow poisson

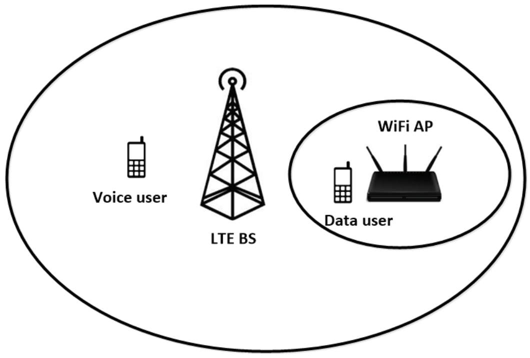


Fig. 1: LTE-WiFi Heterogeneous Network Architecture

process with mean λ_v and λ_d respectively. Service time for voice user is exponentially distributed with mean $\frac{1}{\mu_v}$. Similarly, service time for data user is exponentially distributed with mean $\frac{1}{\mu_d}$. Voice users are served only by LTE network. However, data users can be served using either LTE or WiFi. When a data user is admitted, the controller chooses an appropriate RAT based on certain criteria.

The main elements of a CMDP are decision epochs, states, actions along with corresponding transition probabilities, rewards and costs [9]. At every decision epoch, action taken based on the current state moves the system to a different state with finite probability. In this work, we have considered a discrete time MDP problem. Our objective is to maximize the expected average reward of the system over infinite horizon subject to a constraint on the expected average cost.

A. State Space

The system can be modeled as a 4 dimensional discrete time Markov chain (DTMC). The state in the state-space S is represented as a 4-tuple $s = (i, j, k, l)$ where i, j, k represent number of voice users in LTE, number of data users in LTE and number of data users in WiFi respectively. The fourth parameter l may take any of these five values as follows:

$$l = \begin{cases} 1, & \text{if voice user arrives,} \\ 2, & \text{if data user arrives,} \\ 3, & \text{if voice user departs,} \\ 4, & \text{if data user departs from LTE,} \\ 5, & \text{if data user departs from WiFi.} \end{cases}$$

For example, state $s = (2, 3, 4, 1)$ represents 2 voice users and 3 data users in LTE, and 4 data users in WiFi when a voice user arrival has occurred. State $s = (2, 3, 4, 3)$ represents a similar scenario just after the departure of a voice user.

For the sake of simplicity, we assume that a data or voice user in the LTE network consumes one unit of resource block. This makes the state-space S defined above, to be bounded by the set of constraints:

$$S_{constarint} = \begin{cases} i \le C, \\ j \le C, \\ i + j \le C, \\ k \le W, \end{cases}$$

where C is the total number of resource blocks in the LTE network and W is the maximum number of users supported in the WiFi network, so that per-user throughput in WiFi is greater than a predefined threshold value (say 0.1 Mbps).

B. System Metric

We are interested in the overall per-user throughput of the system as the system metric T. T can be expressed as:

$$T = \sum_{s \in S} Q(s) \frac{i.R_{L,V} + j.R_{L,D} + k.R_{W,D}(k)}{i+j+k}, \quad (1)$$

where $Q(s)$ denotes steady state probability that the system is in state s. $R_{L,V}$ and $R_{L,D}$ represent bit rate of voice and data user in LTE respectively. In our model, it is assumed that both of them are constant (with different values). In general, voice user is a constant bit rate (CBR) traffic source. $R_{W,D}(x)$ denotes per-user data throughput in WiFi when there are x users in WiFi. We have followed IEEE 802.11 Distributed Coordination Function model as described in [10] for computational purposes.

C. Action Space

The set of actions defines a set of possible strategies in case of an arrival or a departure of a user in the LTE-WiFi system. We denote the action space as A. It is assumed that the time interval between successive decision epochs is small enough so that the probability of more than one arrival or departure between two successive epochs can be ignored. For numerical calculations in Section IV, we have taken suitable value for the interval between successive decision epochs so that the above assumption is justified. Based on whether it is an arrival or a departure, we have the set of actions as described below.

$$A = \begin{cases} 1, & \text{Reject the arriving user or do nothing at departure,} \\ 2, & \text{Accept voice/data user in LTE,} \\ 3, & \text{Accept data user in WiFi,} \\ 4, & \text{Accept voice user in LTE with one data user offload from LTE to WiFi,} \\ 5, & \text{Move one data user to one RAT (from which departure has occurred) from another RAT after a departure.} \end{cases}$$

Note that when no arrival/departure occurs between two successive decision epochs, no decision is involved and the system remains in the same state. Clearly, in case of voice user arrival, actions $\{1, 2, 4\}$ are applicable. In case of data user arrival, $\{1, 2, 3\}$ are the feasible actions. When a departure occurs, we have two choices of actions, namely action 1 and action 5. In this paper, we consider voice blocking as a feasible action in all states of the form $(i, j, k, 1)$ except the state $(0, 0, 0, 1)$. We consider voice blocking probability as the performance metric for voice users. Voice blocking probability is defined as the fraction of incoming voice users that are blocked. For data users, we consider throughput as the performance metric. In our setting, data user blocking is possible only when capacity has been reached for both the networks. Symbolically, it is possible only in states $(i, j, k, 2)$ with $i + j = C$ and $k = W$. Action 5 moves one data user from RAT y to RAT x and is triggered at the departure of one user from RAT x. Hence, states of the form $(i, j, k, 3)$ have choice of action 1 or 5, when both actions are feasible. However, for a situation where after the departure of a user from RAT x, there is no data user in RAT y, the only possible action is 1.

D. Transition Probabilities

Let $p_{ss'}(a)$ represents transition probability from state $s = (i, j, k, l)$ to state s' under an action a. The interval between two successive decision epochs is denoted by δ. Associated with each (i, j, k), we define two quantities $v(i, j, k)$ (which is the sum of arrival and service rates of all the users) and $\hat{v}(i, j, k)$ so that

$$v(i,j,k) = \lambda_v + \lambda_d + i\mu_v + j\mu_d + k\mu_d, \quad (2)$$
$$\hat{v}(i,j,k) = 1 - v(i,j,k)\delta. \quad (3)$$

Under the assumption that at most one arrival or departure can occur between two consecutive decision epochs, transition probabilities are described below.

$$p_{ss'}(1) = \begin{cases} \lambda_v\delta, & s' = (i,j,k,1), \\ \lambda_d\delta, & s' = (i,j,k,2), \\ i\mu_v\delta, & s' = (i-1,j,k,3), \\ j\mu_d\delta, & s' = (i,j-1,k,4), \\ k\mu_d\delta, & s' = (i,j,k-1,5), \\ \hat{v}(i,j,k), & s' = (i,j,k,l). \end{cases}$$

For action 2 in state $s = (i, j, k, 1)$,

$$p_{ss'}(2) = \begin{cases} \lambda_v\delta, & s' = (i+1,j,k,1), \\ \lambda_d\delta, & s' = (i+1,j,k,2), \\ (i+1)\mu_v\delta, & s' = (i,j,k,3), \\ j\mu_d\delta, & s' = (i+1,j-1,k,4), \\ k\mu_d\delta, & s' = (i+1,j,k-1,5), \\ \hat{v}(i+1,j,k), & s' = (i,j,k,1). \end{cases}$$

For action 2 in state $s = (i, j, k, 2)$,

$$p_{ss'}(2) = \begin{cases} \lambda_v\delta, & s' = (i,j+1,k,1), \\ \lambda_d\delta, & s' = (i,j+1,k,2), \\ i\mu_v\delta, & s' = (i-1,j+1,k,3), \\ (j+1)\mu_d\delta, & s' = (i,j,k,4), \\ k\mu_d\delta, & s' = (i,j+1,k-1,5), \\ \hat{v}(i,j+1,k), & s' = (i,j,k,2). \end{cases}$$

For action 3 in state $s = (i, j, k, 2)$,

$$p_{ss'}(3) = \begin{cases} \lambda_v\delta, & s' = (i,j,k+1,1), \\ \lambda_d\delta, & s' = (i,j,k+1,2), \\ i\mu_v\delta, & s' = (i-1,j,k+1,3), \\ j\mu_d\delta, & s' = (i,j-1,k+1,4), \\ (k+1)\mu_d\delta, & s' = (i,j,k,5), \\ \hat{v}(i,j,k+1), & s' = (i,j,k,2). \end{cases}$$

For action 4 in state $s = (i, j, k, 1)$,

$$p_{ss'}(4) = \begin{cases} \lambda_v \delta, & s' = (i+1, j-1, k+1, 1), \\ \lambda_d \delta, & s' = (i+1, j-1, k+1, 2), \\ (i+1)\mu_v \delta, & s' = (i, j-1, k+1, 3), \\ (j-1)\mu_d \delta, & s' = (i+1, j-2, k+1, 4), \\ (k+1)\mu_d \delta, & s' = (i+1, j-1, k, 5), \\ \hat{v}(i+1, j-1, & \\ k+1), & s' = (i, j, k, 1). \end{cases}$$

For action 5 in state $s = (i, j, k, 3)$ and $s = (i, j, k, 4)$,

$$p_{ss'}(5) = \begin{cases} \lambda_v \delta, & s' = (i, j+1, k-1, 1), \\ \lambda_d \delta, & s' = (i, j+1, k-1, 2), \\ i\mu_v \delta, & s' = (i-1, j+1, k-1, 3), \\ (j+1)\mu_d \delta, & s' = (i, j, k-1, 4), \\ (k-1)\mu_d \delta, & s' = (i, j+1, k-2, 5), \\ \hat{v}(i, j+1, & \\ k-1), & s' = (i, j, k, l). \end{cases}$$

Similarly, For action 5 in state $s = (i, j, k, 5)$,

$$p_{ss'}(5) = \begin{cases} \lambda_v \delta, & s' = (i, j-1, k+1, 1), \\ \lambda_d \delta, & s' = (i, j-1, k+1, 2), \\ i\mu_v \delta, & s' = (i-1, j-1, k+1, 3), \\ (j-1)\mu_d \delta, & s' = (i, j-2, k+1, 4), \\ (k+1)\mu_d \delta, & s' = (i, j-1, k, 5), \\ \hat{v}(i, & \\ j-1, k+1), & s' = (i, j, k, 5). \end{cases}$$

E. Rewards and Costs

Let the reward function and cost function be represented by $r(s, a)$ and $c(s, a)$ respectively. Since we are interested in voice blocking probability, we define the cost in our model in the following way. Whenever the controller blocks a voice user, one unit cost is incurred, otherwise it will be zero. Thus,

$$c(s, a) = \begin{cases} 1, & \text{if } s = (i, k, l, 1) \text{ and } a = 1, \\ 0, & \text{else.} \end{cases}$$

The reward function is defined as the average per-user throughput of the system and can be expressed as:

$$r(s, a) = \frac{i.R_{L,V} + j.R_{L,D} + k.R_{W,D}(k)}{i + j + k}. \quad (4)$$

It is to be noted that the reward function associated with each state is independent of both l and action a.

III. CMDP Formulation And Optimality Criteria

Let $\pi^t \colon S \to A$ denotes a *decision rule* which prescribes the action to be taken at different states and decision epochs. Let *policy* R be a sequence of decision rules $(\pi^1, \pi^2, \ldots, \pi^t, \ldots)$ taken at decision epochs. A policy is said to be *stationary* if the decision rules are independent of time. If the decision rules of a stationary policy are deterministic, then it is called *pure policy*.

We take the average expected reward of the system as the performance metric to be optimized over infinite horizon. Let $\mathbb{R}$ be the set of all memoryless policies. With initial state i and following the policy R, let the average reward and cost of the system over infinite horizon be denoted by $V^R(i)$ and $B^R(i)$ respectively. Our objective can be summarized as follows:
maximize

$$V^R(i) = \liminf_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \mathbb{E}_{i,R} r(s_t, a_t), \quad (5)$$

subject to

$$B^R(i) = \limsup_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \mathbb{E}_{i,R} c(s_t, a_t) \le B_{\max}. \quad (6)$$

where $B_{\max}$ denotes the maximum tolerable average fraction of voice users blocked over infinite horizon and $\mathbb{E}_{i,R}$ denotes the usual expectation operator with initial state i and following a policy R. Note that we have taken $\liminf$ and $\limsup$ to represent long-run expected average reward and cost respectively, to capture the worst case reward and cost scenario.

In other words, our objective is to find a policy R^* in $\mathbb{R}$ which satisfies $V^{R^*}(i) \ge V^R(i)$ for all $R \in \mathbb{R}$. Such a policy is called *optimal policy* in $\mathbb{R}$.

The above constrained MDP problem can be solved using Lagrangian approach [11]. For a fixed value of β, the constrained MDP problem can be written as an equivalent unconstrained MDP as:

$$r(s, a; \beta) = r(s, a) - \beta c(s, a). \quad (7)$$

For a fixed value of β, the maximization problem can be solved using Value Iteration Algorithm (VIA) [9]. It results in maximum expected average reward under a pure policy π_β. Let β^* maximize the reward subject to the cost constraint. To determine the value of β^*, we employ gradient descent algorithm following [12]. In kth iteration, we have

$$\beta_{k+1} = \beta_k + \frac{1}{k}(B^{\pi_\beta} - B_{\max}). \quad (8)$$

After determining β^*, we need to determine the optimal policy for CMDP. It is discussed in [11] that the optimal policy for a CMDP with a single constraint is a mixture of two pure policies, obtained by small perturbation of β^* by amount ϵ in both directions. This results in two pure policies $\pi_{\beta^*-\epsilon}$ and $\pi_{\beta^*+\epsilon}$ with associated long-term expected average costs $B_{\beta^*-\epsilon}$ and $B_{\beta^*+\epsilon}$ respectively. In the next step, we determine the value of the parameter p such that

$$pB_{\beta^*-\epsilon} + (1-p)B_{\beta^*+\epsilon} = B_{\max}. \quad (9)$$

Finally, the optimal policy π^* is computed by mixing two policies $\pi_{\beta^*-\epsilon}$ and $\pi_{\beta^*+\epsilon}$ and can be expressed as

$$\pi^* = p\pi_{\beta^*-\epsilon} + (1-p)\pi_{\beta^*+\epsilon}. \quad (10)$$

The physical interpretation of Equation (10) is that at every decision epoch, first and second policy are chosen with probability p and $(1-p)$ respectively. This results in a randomized optimal policy for the considered CMDP problem.

IV. Numerical Results And Analysis

In this section, optimal policy computed by solving the CMDP is studied. Additionally, we evaluate the performance of the optimal policy which maximizes the expected average per-user throughput, subject to a constraint on the voice blocking probability. The objective is to study the impact of different load conditions on the system performance.

A. Observations on the Structure of Optimal Policy

In this section, we study the optimal policy computed by solving the CMDP and observe that it has a threshold structure for a fixed value of k (number of data users in WiFi). As the value of k changes, the threshold structure is maintained but value of the threshold changes.

For investigation, we consider parameters of the LTE and WiFi network model as described in Table I and II. We choose the parameter values so as to keep the CMDP problem computationally tractable. In Table II, we consider the saturation throughput WiFi model as described in [10]. Minimum acceptable per-user throughput in WiFi network is assumed to be equal to 100 kbps. Based on this parameter, we can derive value of the constant W as described in Section II. To be specific, the value of $W = x$ if $R_{W,D}(x) \geq 100$ kbps and $R_{W,D}(x+1) <100$ kbps. As we have taken the channel bit rate as 1 Mbps, the value of W comes out to be equal to 8 which is a reasonable value for computational purpose. Moreover, minimum acceptable per-user throughput in WiFi should be chosen to be less than the data bit rate of a single user in LTE (which is 240 kbps). Otherwise, a data user would always get associated with WiFi as it gives higher throughput for any value of k. Propagation delay in WiFi network is assumed to be negligible.

TABLE I: LTE Network Model

Parameter	Value
Maximum voice capacity	10 users
Maximum data capacity	10 users
Voice bit rate of a single user	20 kbps
Data bit rate of a single user	240 kbps

TABLE II: WiFi Network Model

Parameter	Value
Channel Bit rate	1 Mbps
MAC header	272 bits
PHY header	128 bits
Packet payload	8184 bits
Slot duration	50 μs
Short inter-frame space (SIFS)	28 μs
Acknowledgment packet size	112 bits + PHY header
Distributed Coordination Function IFS (DIFS)	128 μs
Minimum acceptable per-user throughput	100 kbps

The value of the constraint $B_{\max}$ is .05. Now, optimal policy for a CMDP is a randomized policy which is a mixture of two pure policies $\pi_{\beta^*+\epsilon}$ and $\pi_{\beta^*-\epsilon}$ with corresponding weights p and $(1-p)$. We compute the optimal policy with $\lambda_v = \lambda_d = 0.5$ users/s, $\mu_v = \mu_d = 1.0$ users/s and $B_{\max} = 0.05$. The interval between two successive decision epochs (δ) is taken as 20 ms. When the capacity has been reached (i.e., $i + j = 10$ and $k = 8$) for both LTE and WiFi, the value of $v(i, j, k)$ (see Equation (2)) is maximum and is equal to $0.5 + 0.5 + 10 * 1 + 8 * 1 = 19$. In that case, the average interval between two events (arrival or departure) is 1/19 s=52.63 ms which is significantly less than δ. This justifies the assumption that more than one arrival or departure between two successive epochs occur with negligible probability.

1) Observations on the Structure of Optimal Policy for Voice User Arrival

First we consider the optimal policy for voice user arrival. Fig. 2 illustrates the optimal policy at different states of the system where arrival of voice user occurs (states with $l = 1$). Optimal policy has been plotted for different values of k. It can be clearly seen that in most of the states, actions under $\pi_{\beta^*+\epsilon}$ and $\pi_{\beta^*-\epsilon}$ are identical. Same observation is applicable for optimal policies for arrival of data users and departure of voice/data users also. Thus, in most of the states, the optimal action to be taken is a deterministic one except for few states where it is probabilistic. When $k = 0$, i.e there is no data user in WiFi, the optimal policy $\pi^*(j, 0)$ can be described as a function of j. We have,

$$\pi^*(j,0) = \begin{cases} 2, & j = 0 \text{ and } 2 \text{ is a feasible action} \\ 4, & j > 0 \text{ and } 4 \text{ is a feasible action} \end{cases} \tag{11}$$

Note that the only possible action at states where $i+j = C$ is blocking (action 1). When $k = 0$, the optimal action

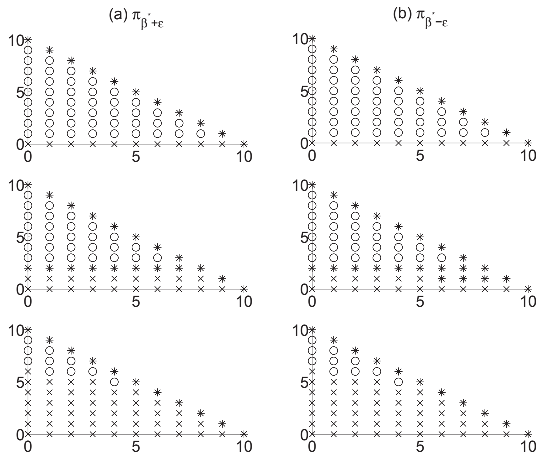


Fig. 2: Optimal policy for voice user arrival for k (number of data users in WiFi)$= 0, 2, 5$ respectively. X and Y axis represents number of voice users and data users in LTE. The stars, crosses and circles represent the optimal action $a = 1$ (blocking), $a = 2$ (accept in LTE), $a = 4$ (accept in LTE with data offload to WiFi) respectively.

in all states is always 4, wherever it is a feasible action. When there in no active data user in LTE, action 4 is not feasible as there is nothing to offload to WiFi. This forces the controller to choose $a = 2$ when $j = 0$. Thus, offloading a data user to WiFi is preferred when admitting voice user in LTE.

However, when the value of k is more (say $k = 2$), there are multiple threshold values of j after which the optimal action changes. In Fig. 2, we can clearly see that for $k = 2$, the optimal action changes from 2 to 1 when $j = 2$ and then it changes to 4 when $j = 3$. This happens because voice user has less contribution to overall per-user throughput of the system. So, voice users are getting blocked in few states to save resources for data users which contribute significantly to the system metric. At higher values of k (say $k = 5$), voice user blocking is not an optimal action anymore. As we have put constraint on the voice blocking probability, choosing action 1 as optimal is limited to few states only. In such situations, action 2 is chosen as the optimal action. This happens due to the fact that when WiFi occupancy is more, offloading one more data user to WiFi results in decreased overall per-user throughput. However, when $j > k$, action 4 is preferred to action 2. When $j > k$, choosing action 4 does some kind of load balancing of data users.

Combining all these behaviors, the generalized threshold structure of the optimal policy as a function of j and k can be expressed as:

$$\pi^*(j,k) = \begin{cases} 2, & j \leq j_{va}(k) \text{ and 2 is a feasible action} \\ 1, & j_{va}(k) < j \leq k \\ 4, & j > k \text{ and 4 is a feasible action} \end{cases} \tag{12}$$

where $j_{va}(k)$ is a threshold which depends on k.

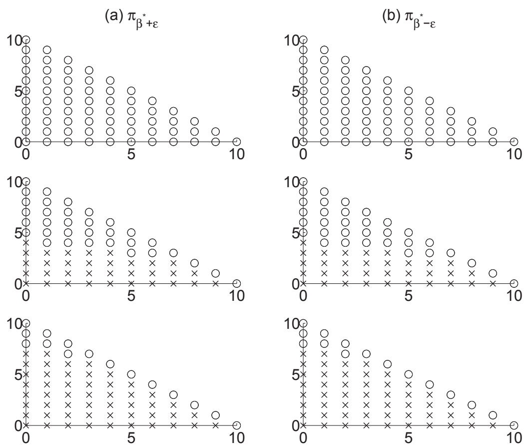


Fig. 3: Optimal policy for data user arrival for k (number of data users in WiFi)= $0, 4, 7$ respectively. X and Y axis represents number of voice users and data users in LTE. The crosses and circles represent the optimal action $a = 2$ (accept in LTE), $a = 3$ (accept in WiFi) respectively.

2) *Observations on the Structure of Optimal Policy for Data User Arrival*

Similar result is observed for data user arrival also. However, instead of multiple thresholds as we have seen in case of voice user arrival, single threshold can be determined. As seen in Fig. 3, when there is no data user in WiFi ($k = 0$), the optimal action is always 3. This is because when $k = 0$, a data user will get more throughput in WiFi. Note that, when $i+j = C$, the only possible action is 3.

As the value of k increases, the fraction of the states where action 3 is optimal, decreases. When WiFi occupancy is more, a user will get more throughput if it is associated with LTE. So, action 2 becomes optimal. However, in Fig. 3 it is seen that after j crosses certain threshold, action 3 becomes optimal again. Due to the constraint on the blocking probability of voice users, it is necessary to reserve some resources for them in LTE. Thus, for higher values of k, when j crosses certain threshold, data users are associated with WiFi instead of LTE. The optimal policy $\pi^*(j, k)$ can be described as:

$$\pi^*(j,k) = \begin{cases} 2, & j \leq j_{da}(k) \text{ and 2 is a feasible action} \\ 3, & j > j_{da}(k) \text{ and 3 is a feasible action} \end{cases} \tag{13}$$

where $j_{da}(k)$ is a threshold which depends on k.

3) *Observations on the Structure of Optimal Policy for Departure of User*

Let $j_{vd}(k)$ be a threshold for departure of voice users and it is a function of k. As can be seen from Fig. 4, when $k = 0$, optimal action is 1 as action 5 is not feasible. When $k > 0$, for lower values of $j(j < j_{vd}(k))$ action 5 is preferred and when $j \geq j_{vd}(k)$, action 1 is preferred. When $j < j_{vd}(k)$, offloading one data user from WiFi to LTE is preferred as that will lead to overall increase in per-user throughput. For example, when $k = 1$ or 2, per-user throughput in WiFi comes out to be 0.8708 or 0.4232 Mbps respectively. In LTE, one data user will get 240 kbps data rate (Details of these parameters are described in Table I and II). Now in state $(0, 0, 2, 3)$, if we choose action 1, then per-user throughput is 0.4232 Mbps. However, if we choose action 5, the per-user throughput becomes $(1 * 0.240 + 1 * 0.8708)/2$ which is more than 0.4232 Mbps. When $j \geq j_{vd}(k)$, offloading one user to LTE doesn't improve the system metric. Moreover, it consumes LTE resources which may lead to blocking of future voice users. Thus, for $j \geq j_{vd}(k)$ action 1 is preferred.

The structure of optimal policy for data user departure is shown in Fig. 5(a) and (b). It is clearly seen that action 5 is the preferred action in all the states wherever it is feasible. Whenever there is a departure of data user, choosing action 5 as the optimal action increases per-user throughput of the overall system. Additionally, it does some amount of load balancing also.

These observations on threshold structure of optimal policies can be used to propose an algorithm with low computational complexity. The observation that for data user departure, action 5 is always optimal, can be used to reduce the state space. We can eliminate the states with $l = 4, 5$ and whenever data user departure occurs, we can always offload one user from another RAT. This will lead to smaller state space and contribute to faster computation of optimal policies.

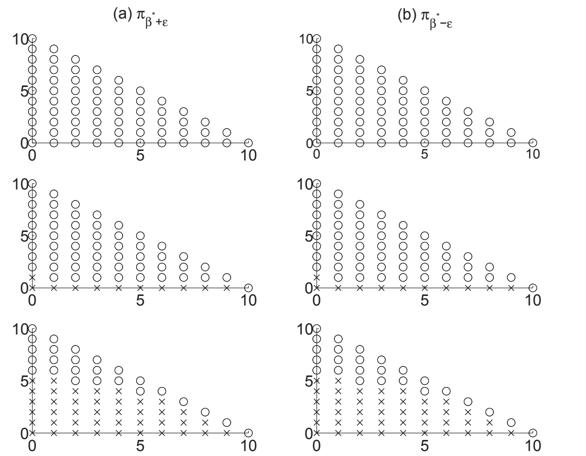
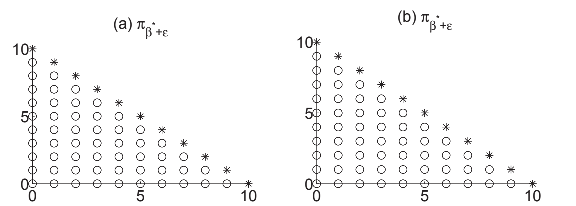
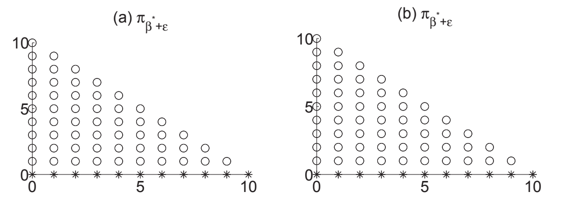


Fig. 4: Optimal policy for voice user departure for k (number of data users in WiFi)$= 0, 2, 6$ respectively. X and Y axis represents number of voice users and data users in LTE. The circles and crosses represent the optimal action $a = 1$ (do nothing during departure), $a = 5$ (Move one data user to one RAT (from which departure has occurred) from another RAT) respectively.



(a) Departure from LTE



(b) Departure from WiFi

Fig. 5: Optimal policy for data user departure from LTE or WiFi. X and Y axis represents number of voice users and data users in LTE. The stars and circles represent the optimal action $a = 1$ (do nothing during departure), $a = 5$ (Move one data user to one RAT (from which departure has occurred) from another RAT) respectively.

B. Voice and Data User Arrival Rate Variation

In this section, we describe how the system metric, the average per-user throughput, varies with variations in voice and data user arrival rate. We vary λ_v from 0.1 users/s to 1.0 users/s in steps of 0.1 users/s with $\lambda_d = 0.5$ users/s and $\mu_v = \mu_d = 1$ users/s. Fig. 6 shows variation of average per-user throughput T (as defined in Equation (1)) as a function of λ_v. As λ_v increases, T increases monotonically. As λ_v increases, the probability of no arrival/departure between two successive decision epochs decreases. As a result, the optimal policy causes the system to spend more time in states with higher values of (i, j, k) and steady state probabilities of these states increase. When λ_v is low, such states have very low steady state probabilities. As our optimal policies do significant amount of load balancing, the benefit of

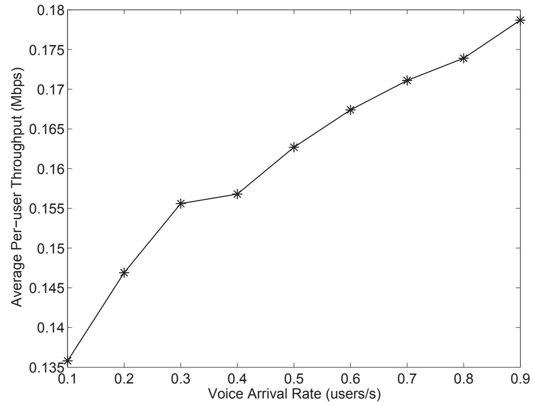


Fig. 6: Average per-user throughput under different voice user arrival rates

such policies is properly exploited when λ_v is high. This results in higher value of average per-user throughput as λ_v increases. However, as voice users contribute lesser to the overall system metric, the range of T (0.1358 to 0.1787 Mbps) in this interval is not very high.

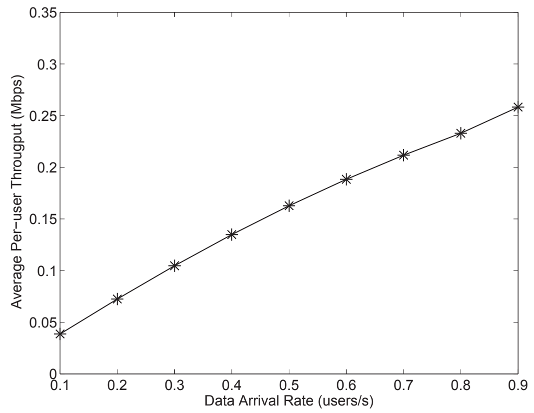
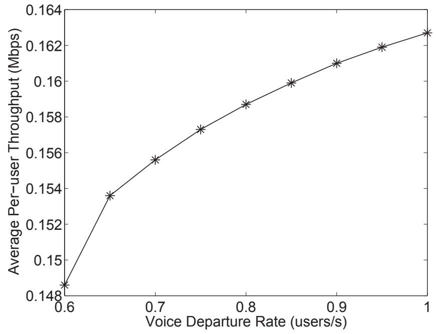
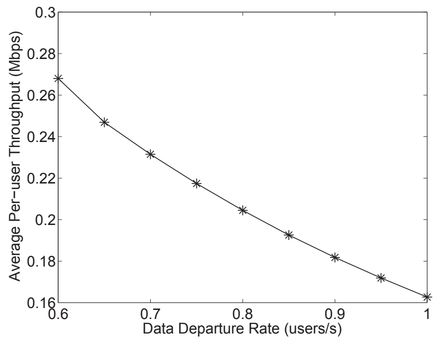


Fig. 7: Average per-user throughput under different data user arrival rates

Similar behavior is observed when we vary the data user arrival rate and observe the average per-user throughput. However, as data users have more contribution to the overall per-user throughput, the range of T (0.0386 Mbps to 0.2583 Mbps) is greater than the range in case of λ_v variation. Fig. 7 shows variation of T as we change λ_d from 0.1 users/s to 1.0 users/s in steps 0.1 users/s. Other parameters are kept constant at $\lambda_v = 0.5$ users/s, $\mu_v = \mu_d = 1$ users/s. As λ_d increases, T increases. On the other hand, the rate of increase of T decreases. If we consider the problem of maximizing average per-user throughput of the system without any constraint on the voice blocking probability, then the problem can be formulated as an unconstrained MDP. In the unconstrained problem, with increase in arrival rate, the optimal policy for maximizing average per-user throughput would result in increased blocking probability. However, in the constrained case, blocking probability of voice users is upper bounded. So, the average per-user throughput in the constrained case becomes less than the corresponding value in the unconstrained case. When λ_d is less, the difference in T in these two



(a) Voice user departure



(b) Data user departure

Fig. 8: Average per-user throughput under different departure rates

cases is less. As λ_d increases, this difference increases and results in decreasing slope of the T vs λ_d curve.

C. Voice and Data User Service Rate Variation

This section is focused on impact of voice and data user service rate (μ_v and μ_d respectively) variation on average per-user throughput of the system.

In Fig. 8(a), as μ_v increases, T also increases. Other system parameters are $\lambda_v = \lambda_d = 0.5$ users/s and $\mu_d = 1$ users/s. With increase in μ_v, fraction of voice users in the system at any point of time decreases. Voice user has lesser contribution to the overall per-user throughput of the system than data user. As μ_v increases, fraction of data user in the system increases. Thus the system metric value increases with increase in μ_v.

The variation of system metric T as a function of μ_d has been shown in Fig. 8(b). With $\lambda_v = \lambda_d = 0.5$ users/s and $\mu_v = 1$ users/s, as μ_d increases from 0.6 users/s to 1.0 users/s, T decreases. As μ_d increases, data users are served at a faster rate. As a result, there are increasing number of free resources in the system for voice users. Thus, the voice blocking probability reduces with increase in μ_d. As μ_d increases, number of voice users in the system increases. As voice users contribute less to the overall per-user throughput of the system, increase of μ_d causes decrease in T.

V. Conclusion

In this paper, we formulated the optimal association problem in a LTE-WiFi heterogeneous network as a CMDP with an objective of maximizing per-user throughput of the system, subject to a constraint on the blocking probability of voice users. We computed the optimal policy and found it to follow a threshold structure, for both arrival of users and departure of voice users. Departure of data users from a RAT has been always found to choose offloading of one data user from another RAT as the optimal action. We also presented numerical results which show variation of the average per-user throughput of the system under different arrival and service rates of users. In this paper, we have considered users to be stationary and CBR traffic for data users. Parameters have been selected suitably so as to keep the computation tractable. In future, we would like to simulate the arrival and departure of users in a LTE-WiFi HetNet with 3GPP recommended parameters and evaluate the performance of our optimal policy. In addition, we would like to propose a computationally efficient algorithm based on the structure of the optimal policy, for both stationary and moving users.

VI. Acknowledgment

This work is funded by the Department of Electronics and Information Technology (DeitY), Government of India.

References

[1] D. Pacheco-Paramo, V. Pla, V. Casares-Giner and J. Martinez-Bauset, “Optimal Radio Access Technology Selection on Heterogeneous Networks”, Physical Communication, Volume 5, Issue 3, pp. 253-271, September 2012.

[2] A. Kumar and V. Kumar, “Optimal Association of Stations and APs in an IEEE 802.11 WLAN”, National Conference on Communications (NCC), pp. 1-5, January 2005.

[3] G.S. Kasbekar, P. Nuggehalli and J. Kuri, “Online client-AP association in WLANs”, International Symposium on Modeling & Optimization in Mobile, Ad Hoc & Wireless Networks (WiOpt), pp. 1-8, April 2006.

[4] D. Kumar, E. Altman and J. Kelif, “Globally optimal user-network association in an 802.11 WLAN & 3G UMTS hybrid cell”, International Teletraffic Congress (ITC), pp. 1173-1187, June 2007.

[5] S. Ranjan, N. Akhtar, M. Mehta and A. Karandikar, “User-based integrated offloading approach for 3GPP LTE-WLAN network”, National Conference on Communications (NCC), pp. 1-6, February 2014.

[6] F. Mehmeti and T. Spyropoulos, “Performance analysis of “on-the-spot” mobile data offloading”, IEEE Global Communications Conference (Globecom), pp. 1577 - 1583, December 2013.

[7] K. Lee, J. Lee, Y. Yi, I. Rhee and S. Chong, “Mobile data offloading: how much can WiFi deliver?”, IEEE/ACM Transactions on Networking, Volume 21, No. 2, pp. 536-550, April 2013.

[8] B. H. Jung, N. Song and D.K. Sung, “A Network-assisted User-centric WiFi-Offloading Model for Maximizing Per-user Throughput in a Heterogeneous Network”, IEEE Transactions on Vehicular Technology, Volume 63, Issue 99, pp. 1940 - 1945, October 2013.

[9] M. L. Putterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley and Sons, 1994.

[10] G. Bianchi, “Performance analysis of the IEEE 802.11 distributed coordination function”, IEEE Journal on Selected Areas in Communications, Volume 18, No. 3, pp. 535-547, March 2000.

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

[12] D. V. Djonin and V. Krishnamurthy, “Q-Learning Algorithms for Constrained Markov Decision Processes With Randomized Monotone Policies: Application to MIMO Transmission Control”, IEEE Transactions on Signal Processing, Volume 55, No. 5, pp. 2170-2181, May 2007.