Optimal Policy for Initial Selection Radio Access Technology in Next Generation Wireless Networks

Rodolfo W. L. Coutinho Pos-Graduation Program in Electrical Engineering Federal University of Para Belém, Pará, Brazil

Email: rwlc@ufpa.br

Vitor L. Coelho Faculty of Computer Science Federal University of Pará Belém, Pará, Brazil Email: vitorlimac2@hotmail.com Joao C. W. A. Costa Pos-Graduation Program in Electrical Engineering Federal University of Pará Belém, Pará, Brazil Email: jweyl@ufpa.br Glaucio H. S. Carvalho Faculty of Computer Science Federal University of Pará Belém, Pará, Brazil Email: ghsc@ufpa.br

Abstract—In this paper, we propose an optimal Joint Call Admission Control (JCAC) in next generation wireless networks, where different radio access technologies (RAT) coexist in a colocated way, and we study the impact of the different RAT's radius coverage area in the system performance. By assuming that each incoming service request may be admitted in its native RAT or in an alternative RAT, we design an optimal JCAC that has to decide in which type of RAT the incoming call has to be admitted. This decision is evaluated under the consideration that the co-located RATs have different coverage areas. Numerical results show that variation in the proportionality of the radius of RAT coverage area impact on system performance.

I. INTRODUCTION

In the next generation of wireless networks, different radio access technologies (RAT) will coexist and provide ubiquitous access with high rates for mobile users, that through multimodal terminals, will be able to connect on most appropriate RAT. Even in this new generation, the radio resources will be scarce, so that Call Admission Control (CAC) schemes will play a crucial role by determining how radio resources should be shared. However, traditional CAC schemes, which are designed for homogeneous networks, will not cope with the network heterogeneously since they do not have a whole vision of the system.

Thus, new CAC solutions, which are named Joint CAC, must be designed to perform a joint resource management that consist of: (*i*) deciding whether an incoming call should be accepted or blocked, like traditional CAC; (*ii*) selecting in which of the available RAT, the incoming service request has to be accommodated. This selection is based on criteria as signal strength, size coverage area, service cost, service class, transmission rate, network load, etc., and should improve the global system utilization, guarantee QoS, user satisfaction, and system stability. [2, 9].

Several schemes for Common Resource Management (CRM) in next generation wireless networks have been proposed in literature. For instance, in [2] is proposed an Adaptive Bandwidth Management and a Joint Call Admission Control. The objectives of the proposed adaptive JCAC scheme are to enhance average system utilization, guarantee QoS requirements of all accepted calls, and reduce new call blocking probability and handoff call dropping probability in

heterogeneous wireless networks. [8] study the dynamics of network selection in a heterogeneous wireless network using the theory of evolutionary games. The authors present two algorithms, namely, population evolution and reinforcementlearning algorithms for network selection, which consider user-driven load balancing in a heterogeneous network. In [9], two schemes of resource management are proposed. The first, named traffic-based resource management scheme (TRMS), allocates the resource based on traffic type, call type (new or handoff call), bandwidth availability, and bandwidth provision. The second scheme, named Q-learning based resource management scheme (QRMS), is formulated as a Markov decision process and the Q-learning approach is applied to conduct the resource allocation. [4] propose a Joint Radio Resource Management (JRRM) for the initial selection of RAT, considering a heterogeneous networks composed by two RATs co-located (WiMAX/UMTS). For initial selection of the most appropriate RAT, the scheme proposed considers the load in each RAT, the spatial distribution of already accepted users, the location of the newly admitted user, and its influence on global performance.

Herein, we propose an optimal Joint CAC (JCAC), where the different RATs are co-located. The contributions of the proposed JCAC are twofold. First, we consider an environment where an incoming service request may be served by its native RAT (Service Provider) or by one RAT among those, which are available. These available RATs are called alternatives and can server a mobile user with a given cost. Second, we study optimal JCAC policies based on the ratio between the radius coverage area of the co-located RATs.

We model and solve the optimal control problem by using the Semi-Markov Decision Process (SMDP) framework and compute the optimal JCAC policy by using the value iteration algorithm.

II. TRAFFIC MODEL

The incoming service requests are dichotomized into two broadly traffic classes: real time connection and non real time connection. We consider the existence i (i = 1, ..., L) types of real time service traffic classes and that all the non real time traffic are aggregated in an unique service class. For the sake of Markov modelling, the i^{th} real time service class arrives to the j^{th} RAT according to a Poisson process with parameter λ_i^j . A non real time service connection arrives in the j^{th} RAT according to a Poisson process with mean rate λ_{nr}^j . As one will be explained, we consider a unique non real time service class.

The call duration time has been assumed exponentially distributed with mean value equal to $1/\mu_{d_i}$ for real time service class connection and *ideally* $1/\mu_{d_{nr}}$ for non real time service class connection. The RAT residence time represents the time which a mobile user stay in the j^{th} RAT and follows a negative exponential distribution with mean given by [3]:

$$\frac{1}{\mu_{r_i}} = 1/(0.7182\frac{V}{R_j}).$$
(1)

where V is the mobile user average speed and R is the RAT radius. Since we work with co-located RATs, we assume that the average speed of a mobile user remains unchanged along the RAT's coverage areas. Therefore, the stay time in each RAT is *de facto* differentiated by the size of RAT radius. Let $R_j = kR_w$ $(j, w \in \{1, \ldots, K\})$ be the relation between the radius of the j^{th} RAT and the w^{th} RAT. Given Eq.(1) and the aforementioned relation, we have:

$$\frac{1}{\mu_{r_j}} = 1/(0.7182 \frac{V}{kR_w}).$$
(2)

Since $R_w = \frac{0.7182V}{\mu_{r_w}}$, we have after some mathematical manipulations:

$$\frac{1}{\mu_{r_i}} = k \frac{1}{\mu_{r_w}},\tag{3}$$

which relates the residence time between both RATs. The channel holding time is defined as the time elapsed between the instant that a channel is assigned to serve a call in a RAT and the instant it is released by either call completion or a cell boundary crossing by the mobile user. The mean value of channel holding time in the j^{th} RAT is given by Eq.(4) and Eq.(5) for real time service connection and non real time service connection, respectively.

$$\mu_{h_{i}^{j}} = \mu_{d_{i}} + \mu_{r_{j}}.$$
(4)

$$\mu_{h_{nr}^j} = \mu_{d_{nr}} + \mu_{r_j}.$$
(5)

Finally, define $\rho_i^j = \lambda_i^j / \mu_{h_i^j}$ as the i^{th} real time service class connection intensity and $\rho_{nr}^j = \lambda_{nr}^j / \mu_{h_{nr}^j}$ as the non real time service class connection intensity in the j^{th} RAT, respectively.

III. OPTIMIZATION CONTROL

A. System and Traffic Assumptions

The system under consideration consists of K co-located RATs. The j^{th} RAT (j = 1, ..., K) consists of a wireless link with B_j radio resources, which are shared by the incoming service requests. The physical meaning of a unit

of radio resource is dependent on the specific technological implementation of the radio interface. However, no matter which multiple access technology (FDMA, TDMA, CDMA, or OFDM) is used, we could interpret system capacity in terms of effective or equivalent bandwidth [5, 7].

Each real time connection request demands b_i resources. To model the non real time service, we use the degradation and compensation mechanism that allows a non real time connection to adapt its mean rate accordingly the network load [1], [10]. Thus, each non real time connection request can adjust its bandwidth in the range of values $[b_m, b_M]$ radio resources.

When an incoming service requests an access to the network, the optimal JCAC has to decide if it will be accommodated in its native RAT or any one of the alternatives RATs. A native RAT belongs to the Service Provider with which the mobile user has its service provider agreement. An alternative RAT is that one in which the mobile user's Service Provider has a Service Level Agreement, which ensures its roam or initial access in the case of overload or another JCAC decision.

B. State Space

We define in Eq.(6) the set Φ of all feasible states in which m_i^j and m_{nr}^j are the number of ongoing real time service class i connections and non-real time connections, being served in j^{th} RAT, respectively. Since a real time service class i demands b_i resources to fulfill its QoS profile, its maximum number of connections in the j^{th} RAT is $\lfloor \frac{B_j}{b_i} \rfloor$, where $\lfloor g \rfloor$ is the largest integer not greater than g. The maximum number of non real time connections is given by $\lfloor \frac{B_j}{b_m} \rfloor$.

$$\Phi = (m_i^j, m_{nr}^j, e : i = 1, \dots, L, j = 1, \dots, K) / \sum_{i=1}^{L} m_i^j b_i + m_{nr}^j b_{nr}^j(x) \le B_j \ \forall j.$$
(6)

To model the non real time service class traffic elasticity, it is used the concept of ideal departure rate,[1],[10], in which the real instantaneous departure rate of data connections is proportional to the actual bandwidth of each connection. So, with L real time service classes into the j^{th} RAT, each non real time connection will receive the bandwidth of

$$b_{nr}^{j}(x) = \min(b_{M}, \max(b_{m}, \frac{B_{j} - \sum_{i=1}^{L} m_{i}^{j} b_{i}}{m_{nr}^{j}}))/$$

$$0 \le \sum_{i=1}^{L} m_{i}^{j} b_{i} \le B_{j}, m_{nr}^{j} > 0, x \in \Phi,$$
(7)

and will be served with service rate of

$$\mu_{d_{nr}^{j}}(x) = \frac{b_{nr}^{j}(x)}{b_{M}} \mu_{d_{nr}}, x \in \Phi.$$
 (8)

It is worthy to note that inside the concept of *ideal departure* rate when a non real time connection receives the maximum bandwidth, b_M , its mean service rate will also be maximized and equal to $\mu_{d_{nr}}^j(x) = \mu_{d_{nr}}$. For each $x \in \Phi$, accordingly the concept of ideal departure, the mean value channel holding time to sessions of non real time service class is

$$\mu_{h_{nr}^{j}}(x) = \mu_{d_{nr}}(x) + \mu_{r_{j}}.$$
(9)

The random variable e, in Eq.(6), is the last event occurred. This information is introduced in the state space in order to define the set of possible actions in each state. Accordingly the system dynamics, the values of e may be

$$e = \begin{cases} 0, & j \in \{1, \dots, K\} \\ ij, & j \in \{1, \dots, K\} \land i \in \{1, \dots, L\}. \end{cases}$$
(10)

where the value e = 0 represents a departure of an ongoing call; e = ij means an arrival the i^{th} real time service class connection destined to j^{th} RAT; and e = j means an arrival of a non real time connection destined to j^{th} RAT.

C. Decision Epochs and Actions

The decision epochs are those time points when a call arrives to or leaves from the system. We assume that each state means the system's configuration just after an event occurrence and just before a decision making. The "real" decision epochs are the arrivals of real time and non real time connections, *i.e.*, $e = j, 11, 21, \ldots, ij, \ldots, LK$ while the service completion epochs are defined as "fictitious" decision epochs, e = 0. In each state $x \in \Phi$, the controller can choose one out of the possible actions:

$$A(x) = \begin{cases} 0, j \le e \le LK; \\ 1j, j \le e \le LK/b_e + \sum_{i=1}^{L} m_i^j b_i + m_{nr}^j b_m \le B_j; \\ 2w, j \le e \le LK/b_e + \sum_{i=1}^{L} m_i^w b_i + m_{nr}^w b_m \le B_w \ \forall w \ne j, \end{cases}$$
(11)

where b_e is b_m (b_i) if e = j (e = ij). It is noteworthy that accordingly to the degradation and compensation mechanism applied to non real time service class, a incoming non real time service request is admitted with b_M bandwidth whenever there are sufficient resources. Throughout the occupation the system dynamics, non real time service class connections can still be admitted with bandwidth up to b_m .

In the set of actions $a \in A(x)$, $x \in \Phi$, the action a = 0 denotes the rejection, a = 1j denotes admission in the native RAT, and a = 2w denotes admission in the alternative w^{th} RAT.

D. Expected Time Until the Next Decision Epoch

If the system is in the state $x \in \Phi$ and the action $a \in A(x)$ is chosen, then the expected time until the next decision epoch, $\tau_x(a)$, is given by Eq.(12).

$$\tau_x(a) = \frac{1}{\sum_{j=1}^{L} \sum_{i=1}^{K} \lambda_i^j + \sum_{j=1}^{L} \lambda_{nr}^j + \sum_{i=1}^{L} \sum_{j=1}^{K} m_i^j \mu_{h_i^j} + \sum_{j=1}^{K} m_{nr}^j \mu_{h_{nr}^j}(x)}$$
(12)

E. State Dynamics

The state dynamic is completely specified by stating the transition probabilities among the system states. Thus, let $p_{xy}(a)$ be the probability that in the next decision epoch the state will be $y \in \Phi$ if the present state is $x \in \Phi$ and the action $a \in A(x)$ is chosen. For $x \in \Phi$, given in Eq.(13), and $y \in \Phi$, we have the cases presented below:



• Case 1: e = 0 in Eq.(13) has:



• Case 2: e = j in Eq.(13) has:

$$p_{xy}(a) = \begin{cases} \lambda_{nr}^{j} \tau_{x}(a), a = 0, y = x; \\ \lambda_{nr}^{j} \tau_{x}(a), & \begin{bmatrix} m_{1}^{1}, \dots, m_{1}^{1}, \dots, m_{L}^{1}, m_{nr}^{1} \\ \vdots & \vdots & \vdots & \vdots \\ m_{1}^{j}, \dots, m_{i}^{j}, \dots, m_{L}^{j}, m_{nr}^{j} + 1 \\ \vdots & \vdots & \vdots & \vdots \\ m_{1}^{m}, \dots, m_{i}^{m}, \dots, m_{L}^{m}, m_{nr}^{m} \end{bmatrix}, e); \\ \lambda_{nr}^{j} \tau_{x}(a), & \begin{bmatrix} m_{1}^{1}, \dots, m_{i}^{1}, \dots, m_{L}^{1}, m_{nr}^{1} \\ \vdots & \vdots & \vdots & \vdots \\ m_{1}^{j}, \dots, m_{i}^{m}, \dots, m_{L}^{m}, m_{nr}^{m} \end{bmatrix}, e) \forall w \neq j. \\ \lambda_{nr}^{i} \tau_{x}(a), & \begin{bmatrix} m_{1}^{1}, \dots, m_{i}^{1}, \dots, m_{L}^{1}, m_{nr}^{1} \\ \vdots & \vdots & \vdots & \vdots \\ m_{1}^{m}, \dots, m_{i}^{m}, \dots, m_{L}^{m}, m_{nr}^{m} \end{bmatrix}, e) \forall w \neq j. \end{cases}$$

$$(15)$$

• Case 3: e = ij in Eq.(13) has:



• 0 for otherwise.

F. Cost Function

If the system is in the state $x \in \Phi$ and the action $a \in A(x)$ is chosen, the admission control incurs in the following cost

$$C_x(a) = C_B(x, a) + C_A(x, a),$$
 (17)

where $C_B(x, a)$ and $C_A(x, a)$ are the block cost and the alternative accept cost for a incoming service request, respectively. The former is given by

$$C_B(x,a) = c_{b_i}(c_{b_{nr}}), x \in \Phi, ij \le e \le LK(e=j), a = 0 \in A(x),$$
(18)

for real time service class (non real time service class). When the JCAC decides that the incoming service request will be accept in the w^{th} alternative RAT, the cost incurred relative this operation is given by Eq.(19) for real time incoming service request (non real time incoming service request).

$$C_A(x,a) = c_{a_i^w}(c_{a_{nr}^w}), x \in \Phi, ij \le e \le LK(e=0), a = 2w \in A(x).$$
(19)

With $\tau_x(a)$, $p_{xy}(a)$ and $C_x(a)$, using the value iteration algorithm and the uniformization method [11], we can obtain the optimal CAC stationary policy. A stationary policy R, defined by the decision rule $f : \Phi \to A$, prescribes the action $f(x) \in A(x)$ each time the system is observed in the state $x \in \Phi$.

G. Performance Measurement

The mean carried service class connection traffic is computed as Eq.(20), where $\pi_x(\forall x \in \Phi)$ is the continuous time Markov chain steady state probability distribution under the optimal policy. Given O_e^a , we can derive the real time connection blocking probability in the native RAT (e = ij or e = j and a = 1j) and alternative RAT (e = ij or e = j and a = 2w) by Eq.(21).

TABLE I System Configuration.

Parameter	Value	Parameter	Value	Parameter	Value
$B_1 = B_2$	20 channels	b_1	2 channels	$c_{b_{nr}^1} = c_{b_{nr}^2}$	0.5
μ_{d_1}	$0.0055S^{-1}$	$[b_m, b_M]$	[1,2] channels	$c_{a_1^1} = c_{a_1^2}$	0.5
μ_{dnr}	$0.0016S^{-1}$	$c_{b_1^1} = c_{b_1^2}$	1	$c_{a_{nr}^1} = c_{a_{nr}^2}$	0.1

$$O_{e}^{a} = \sum_{\substack{x \in \Phi; \\ j \le e \le LK; \\ a=1j, 2w \in A(x)}} \left(\sum_{j=1}^{K} \sum_{i=1}^{L} \lambda_{i}^{j} + \sum_{j=1}^{K} \lambda_{nr}^{j} + \sum_{i=1}^{L} \sum_{j=1}^{K} m_{i}^{j} \mu_{h_{i}^{j}} + \sum_{j=1}^{K} m_{nr}^{j} \mu_{h_{nr}^{j}}(x) \right) \pi_{i}$$

$$(20)$$

$$P_{b_i^j} = 1 - \frac{O_e^a}{\lambda_i^j}.$$
 (21)

The utilization of system is computed by Eq.(22).

$$U = \frac{1}{\sum_{j=1}^{K} B_j} \sum_{\substack{x \in \Phi; \\ a \in A(x); \\ m_i^j > 0; m_{nr}^j > 0}} (\sum_{j=1}^{K} \sum_{i=1}^{L} m_i^j b_i + \sum_{j=1}^{K} m_{nr}^j b_{nr}^j(x)) \pi_x.$$
(22)

IV. NUMERICAL RESULTS

Here, it is considered two RATs with two service classes: a real time service class and a non real time service class. We have considered k = 1 and k = 5, for performance evaluation of the model proposed. For this we consider the RAT-1 residence rate $r_1 = \mu_{d_i}/8$, [6], and the RAT-2 residence rate I shows the set of the remainder parameters used in the experiments.

Fig.1a shows that the real time service class blocking probability, in the native RAT-1, does not vary significantly with an increase in the k values. However, the same pattern is not observed in Fig.1b. The reason for this resides in the fact that the greater the k value, the shorter the RAT-2 mean residence time, and the shorter the mean channel holding time. Consequently, radio channels are quickly released and may be allocated for incoming service requests.

This same characteristic is seen in Fig.2a, but in this case the native RAT is one of number 2. Fig.2b reveals a similar behavior.

The Fig.3a shows that the system utilization decreases as k increases. This happens because the larger the RAT, more connections it holds due to the mobility model. As shown in Fig.3b the optimal cost decrease as k increases.

V. OPTIMAL POLICY

In this Section, we show the optimal policy structure for the experiment present previously. Particularly, we analyze the case where the traffic intensity is 14.

The Fig.4 shows the optimal decisions for the incoming real time service class connection destined natively for the RAT-1, for k = 1. The notation '+' represents those states at which the system would admit natively the incoming service request, notation '-' and 'o' represents those states in which actions depend on the state in the alternative RAT how show



Fig. 1. Performance Metrics versus $\rho_1 = \rho_{nr}$: Real time connection blocking probability in (a) native RAT-1. (b) alternative RAT-2



Fig. 2. Performance Metrics versus $\rho_1 = \rho_{nr}$: Real time connection blocking probability in (a) native RAT-2. (b) alternative RAT-1



Fig. 3. Performance Metrics versus $\rho_1 = \rho_{nr}$: (a) Utilization of system. (b) Optimal cost

Fig.(5).a and Fig.(5).b, respectively. Due to difference in the cost of blocking of the services class connections and due to the alternative accept cost, these connections, e = 11, are admitted in the native RAT whenever there are sufficient resources. For the other cases, the action a = 22 is selected always the alternative RAT is busy with up to half capacity, as shown Fig.(5).a and Fig.(5).b.

VI. CONCLUSION

In this paper, we analyzed an optimal Joint CAC in a heterogeneous wireless networks environment. We propose a JCAC scheme that consider the alternative accept cost, corresponding to the accept the incoming service request in a RAT of the other Service Provider. Results show that variations in proportionality of size radius of coverage area in co-located networks impact directly on the performance measurements.

ACKNOWLEDGMENT

This work is supported by the Programa de Apoio ao Doutor Recém-Contratado of the Federal University of Pará



Fig. 4. Optimal Policy for k = 1



Fig. 5. Decisions according RAT-2's occupation: (a) '-' in the Fig.4 (b) 'o' in the Fig.4

(PADRC/CAPES/FAPESPA). This is also partially supported by CNPq.

REFERENCES

- B. L. a. K. S. a. X.-R. C. B. Li and, L. Li and. Call admission control for voice/data integrated cellular networks: performance analysis and comparative study. *IEEE Journal on Selected Areas in Communications*, 22(4):706 – 718, may 2004.
- [2] O. E. Falowo and H. A. Chan. Adaptive bandwidth management and joint call admission control to enhance system utilization and qos in heterogeneous wireless networks. *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, 2007.
- [3] R. Fantacci. Performance evaluation of prioritized handoff schemes in mobile cellular networks. *IEEE Transactions on Vehicular Technology*, 49(2):485 –493, mar 2000.
- [4] K. K. Ibrahim, M. and and S. Tohme. Network-centric joint radio resource policy in heterogeneous wimax-umts networks for streaming and elastic traffic. pages 1 –6, april 2009.
- [5] G. Kesidis, J. Walrand, and C.-S. Chang. Effective bandwidths for multiclass markov fluids and other atm sources. *IEEE/ACM Transactions* on Networking, 1(4):424 –428, aug 1993.
- [6] M. Meo and M. A. Marsan. Resource management policies in gprs systems. *Performance Evaluation*, 56(1-4):73 – 92, 2004.
- [7] N. Nasser and H. Hassanein. Dynamic threshold-based call admission framework for prioritized multimedia traffic in wireless cellular networks. In *IEEE Global Telecommunications Conference*, 2004. *GLOBECOM '04.*, volume 2, pages 644 – 649 Vol.2, 29 2004.
- [8] D. Niyato and E. Hossain. Dynamics of network selection in heterogeneous wireless networks: An evolutionary game approach. *IEEE Transactions on Vehicular Technology*, 58(4):2008 –2017, may 2009.
- [9] W. Shen and Q.-A. Zeng. Resource management schemes for multiple traffic in integrated heterogeneous wireless and mobile networks. pages 1–6, August 2008.
- [10] S. Tang and W. Li. An adaptive bandwidth allocation scheme with preemptive priority for integrated voice/data mobile networks. *IEEE Transactions on Wireless Communications*, 2006.
- [11] H. C. Tijms. A first course in Stochastic models. John Wiley & Sons Ltd, 2003.