International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:5, No:1, 2011

A Markov Chain Model for Load-Balancing Based and Service Based RAT Selection Algorithms in Heterogeneous Networks

Abdallah Al Sabbagh

Abstract**—**Next Generation Wireless Network (NGWN) is expected to be a heterogeneous network which integrates all different Radio Access Technologies (RATs) through a common platform. A major challenge is how to allocate users to the most suitable RAT for them. An optimized solution can lead to maximize the efficient use of radio resources, achieve better performance for service providers and provide Quality of Service (QoS) with low costs to users. Currently, Radio Resource Management (RRM) is implemented efficiently for the RAT that it was developed. However, it is not suitable for a heterogeneous network. Common RRM (CRRM) was proposed to manage radio resource utilization in the heterogeneous network. This paper presents a user level Markov model for a three co-located RAT networks. The load-balancing based and service based CRRM algorithms have been studied using the presented Markov model. A comparison for the performance of load-balancing based and service based CRRM algorithms is studied in terms of traffic distribution, new call blocking probability, vertical handover (VHO) call dropping probability and throughput.

Keywords**—**Heterogeneous Wireless Network, Markov chain model, load-balancing based and service based algorithm, CRRM algorithms, Beyond 3G network.

I. Introduction

DUE to the coexistence of different RATs, NGWN is predicted to be heterogeneous in nature and to integrate different RATs such as UMTS Terrestrial Radio Access Network (UTRAN), GSM/EDGE Radio Access Network (GERAN) and Wireless Local Area Network (WLAN) through a common platform. Each RAT has particular and different characteristics in capacity, coverage area, cost of service, security and QoS levels provided to subscribers. Using a multimode User Equipments (UE) in a heterogeneous wireless network will allow the subscriber to be able to have access to the wireless network through any of the available RATs.

The motivation in NGWN comes out from the fact that no single RAT could support widespread coverage and provide continuous high QoS levels over multiple smart areas, e.g. office, cafe, public smart areas, etc. [1]. In this case, multiple access networks that come from different technologies are spread in the same geographical space. The 3rd Generation Partnership Project (3GPP) has proposed an interconnected heterogeneous wireless network Beyond 3G (B3G) architecture which interconnects GERAN, UTRAN and WLAN. More details about B3G architecture is described in Section II. In B3G network, the need for CRRM is required to provide the demanded QoS and to support the efficient use of radio resources. CRRM will support the integration over existing wireless network systems [2], [3]. A suitable CRRM algorithm can maximize system performance and QoS by allocating users to the most suitable RAT when two or more RATs are co-located in the same area. A number of RAT selection algorithms have been proposed in the literature [4]. The load-balancing based and service based CRRM algorithms are studied in this paper. In the load-balancing based CRRM algorithm, users are always allocated to the least loaded RAT [5], [6]. However, in the service based CRRM algorithm, users will be allocated based on service types and network properties [7], [8].

The user level Markov model is used to analyze the probabilities of a single user being in different states. In the User Level Markov model, it is assumed that the network capacity is sufficient to serve the user.

The rest of the paper is organized as follows. Section II presents the architecture of heterogeneous B3G network. UTRAN is integrated with GERAN and WLAN. In Section III, the Markov model is presented. The load-balancing based CRRM algorithm is studied in Section IV. In Section V, the service based CRRM algorithm is studied. A comparison for the performance of load-balancing based and service based CRRM algorithms is presented in Section VI. Load-balancing based and service based CRRM algorithms are evaluated in terms of traffic distribution, new call blocking probability, VHO call dropping probability and throughput. Finally, conclusions are drawn in Section VII.

II. Heterogeneous Beyond 3G Network Architecture

The heterogeneous B3G network is expected to propose an open and flexible architecture to support different wireless access technologies and provide services and application with

Abdallah AL Sabbagh is with the School of Computing and Communications, and Centre for Real-Time Information Networks (CRIN) at the University of Technology, Sydney (UTS), P.O. Box 123 Broadway, NSW 2007 Australia (e-mail: asabbagh@eng.uts.edu.au).

different QoS demands [2]. Various Radio Access Networks (RANs) will be interfacing the common core network. Fig. 1 presents the architecture of the heterogeneous B3G network. The core network infrastructure is composed to Circuit Switched (CS) and Packet Switched (PS) domains that are connected to the different wireless access technologies. The CS domain contains a Mobile Switching Centre (MSC) that it is interconnected with all Radio Network Controllers (RNCs) in UTRAN via Iu_CS interfaces and all Base Station Controllers (BSCs) in GERAN via A/Iu_CS interfaces. The MSC allows the interconnection with the external fixed networks, Public Switched Telephone Network/Integrated Services Digital Network (PSTN/ISDN). The PS domain contains Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN) that are interconnected via Gn interface. The SGSN handles the mobility management functions for all packet switched data. It is responsible for the delivery of data packets, from and to the UE within its geographical service area. SGSN is interconnected with all RNCs in UTRAN via Iu_PS interfaces, all BSCs in GERAN via Gb/Iu_PS interfaces and all APCs in WLAN via Wr interfaces. The GGSN allows the interconnection with the external IP networks (internet).

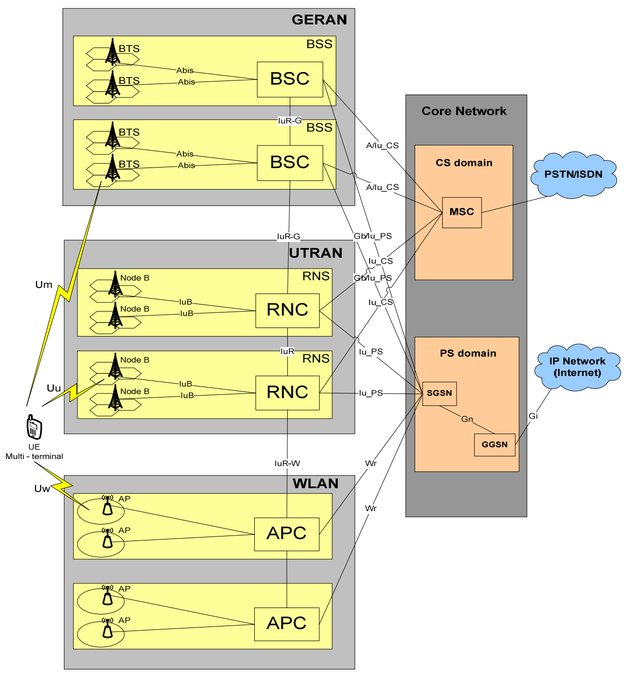

Fig. 1 Heterogeneous Beyond 3G Network Architecture.

The UTRAN infrastructure contains different entities called Radio Network Subsystems (RNSs) that allow the connection of the mobile terminal to the core network. Each RNS contains a number of Nodes B and one RNC that are interconnected via IuB interfaces. The RNC controls the available resources at Nodes B and allocates and de-allocates them depending on the service needs. It controls the handover procedures between the Nodes B that are connected to it. Node B is connected to the UE through a radio interface Uu and it handles the radio transmission procedures. It is composed of one or several cells, each cell has Cell ID.

The GERAN infrastructure contains different entities called Base Station Subsystems (BSSs) that are connected to the core network. Each BSS contains one BSC and several BTSs that are interconnected via Abis interfaces. The BTS is connected to the UE through a radio interface Um and handles the radio transmission procedures. The BSC is the node responsible for controlling the use of the radio resources in the BTSs. The BSC is interconnected with the core network. It controls the handover between the BTSs.

WLAN is composed of different entities constituted by an Access Point Controller (APC) and a set of stations denoted as Access Points (APs). The APC has the role of RNC in UTRAN and BSC in GERAN. It is responsible for controlling the use of the radio resources in the APs. The APC is interconnected with the core network. The AP is connected to the UE through a radio interface Uw and handles the radio transmission procedures.

III. User Level Markov Model

In this section, a user level Markov model is presented. A user in the system can be in one of the following six states:

- State 0: Not connected. User is idle.
- State 1: Inside the hotspot area and connected to GERAN.
- State 2: Inside the hotspot area and connected to UTRAN.
- State 3: Inside the hotspot area and connected to WLAN.
- State 4: Outside the hotspot area and connected to GERAN.
- State 5: Outside the hotspot area and connected to UTRAN.

Fig. 2 shows the user state transition diagram. Let P_0, P_1, P_2, P_3, P_4 and P_5 be the probabilities of a user being in State 0, 1, 2, 3, 4 and 5 respectively. The steady state probability transition matrix is:

$$\mathrm{P} = \begin{bmatrix} P_{00} & P_{01} & P_{02} & P_{03} & P_{04} & P_{05} \\ P_{10} & P_{11} & P_{12} & P_{13} & P_{14} & P_{15} \\ P_{20} & P_{21} & P_{22} & P_{23} & P_{24} & P_{25} \\ P_{30} & P_{31} & P_{32} & P_{33} & P_{34} & P_{35} \\ P_{40} & P_{41} & P_{42} & P_{43} & P_{44} & P_{45} \\ P_{50} & P_{51} & P_{52} & P_{53} & P_{54} & P_{55} \end{bmatrix}, \tag{1}$$

where $0 \le P_{ij} \le 1$ for i, j = 0,1,2,3,4,5.

The sum of each row of the matrix is 1:

$$\sum_{j=0}^{5} P_{ij} = 1, \text{ for i = 0,1,2,3,4,5.} \tag{2}$$

P_{00} is the probability of the system staying in State 0 where a new user doesn't arrive. P_{01}, P_{02} and P_{03} are the probabilities of the new arrival user being allocated to GERAN, UTRAN and WLAN respectively inside the hotspot area. P_{04} and P_{05} are the probabilities of the new arrival user being allocated to GERAN and UTRAN respectively outside the hotspot area. P_{12}, P_{13}, P_{14}, P_{15}, P_{21}, P_{23}, P_{24}, P_{25}, P_{31}, P_{32}, P_{34}, P_{35}, P_{41}, P_{42}, P_{43}, P_{45}, P_{51}, P_{52}, P_{53} and P_{54} are the VHO probabilities. P_{11}, P_{22}, P_{33}, P_{44} and P_{55} are the probabilities that an on going call stays in the RAT that is currently serving it.

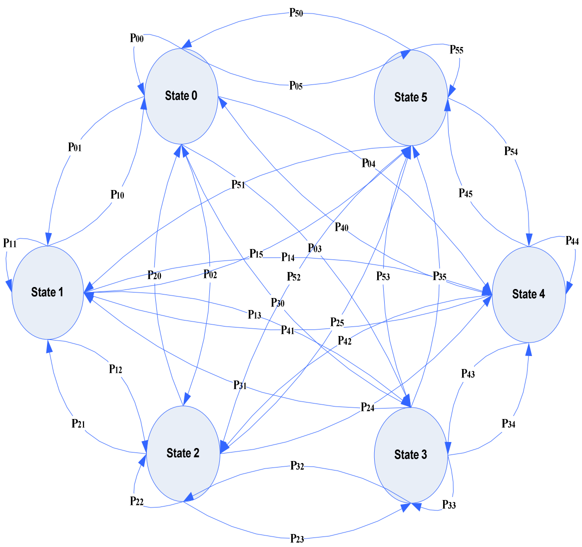

Fig. 2 User state diagram.

The steady state probabilities can be worked out by solving the following equation [9]:

$$\pi = \pi P, \tag{3}$$

where π is the state probability vector given by $\pi = [P_0, P_1, P_2, P_3, P_4, P_5]$. Since a user can only be in the six states at any point in time,

$$P_0 + P_1 + P_2 + P_3 + P_4 + P_5 = 1 \tag{4}$$

A user level Markov models is studied in Section IV and Section V for both load-balancing based and service based RAT selection algorithms.

IV. Load-Balancing Based CRRM Algorithm

Note: This section is based heavily on [13] (co-authored by the author of this paper). The results and figures presented are drawn from the earlier paper and are presented here for the purpose of allowing clearer comparison with alternative approaches, particularly those described in Section VI.

In this section, a load-balancing based CRRM algorithm is studied using the Markov model. Load-balancing based CRRM algorithm aims to distribute traffic load between all available RATs in a heterogeneous wireless network. Balancing load between all available RATs offers an efficient utilization of the radio resources [10], [11]. Traffic load could be continuously balanced [10], balanced at a specific part of time [11], or balanced while a specific load threshold is being reached [12]. Also in a heterogeneous wireless network, load-balancing could be forced or unforced. Reference [13] proposed a user level Markov model for load-balancing based CRRM algorithm in a simple and complex scenario. A multi-access cellular network is shown in Fig. 3. It is considered that users can move inside and outside the hotspot area.

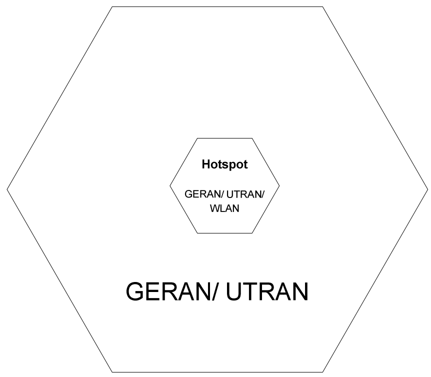

Fig. 3 Multi-Access Cellular Network.

The following definitions were made. P_{new} is the probability of a new call arrives. P_{term} is the probability of an existing call is terminated. P_{in} is the probability of a user residing in the hotspot area. P_{out} is the probability of a user residing outside the hotspot area, where $P_{out} = 1 - P_{in}$. P_{new_h} is the probability of a new call arriving in the hotspot area, where $P_{new\_h} = P_{new} \times P_{in}$. The probability of a new call arriving outside the hotspot area is then to be $P_{new} \times P_{out}$. P_{ex} is the probability of a user exiting the hotspot area during a session. P_{en} is the probability of a user entering the hotspot area during a session.

In a load-balancing based CRRM algorithm, users are allocated to the least loaded RAT. L_G, L_U and L_W were defined as the loads of GERAN, UTRAN and WLAN respectively. The following weighting parameters were introduced:

$$a = \begin{cases} 1, & if \quad L_G = \min(L_G, L_U, L_W) \\ 0, & if \quad L_G \neq \min(L_G, L_U, L_W) \end{cases} \tag{5}$$

$$b = \begin{cases} 1, & if \quad L_U = \min(L_G, L_U, L_W) \\ 0, & if \quad L_U \neq \min(L_G, L_U, L_W) \end{cases} \tag{6}$$

$$c = \begin{cases} 1, & if \quad L_W = \min(L_G, L_U, L_W) \\ 0, & if \quad L_W \neq \min(L_G, L_U, L_W) \end{cases} \tag{7}$$

$$a_2 = \begin{cases} 1, & if \quad L_G = \min(L_G, L_U) \\ 0, & if \quad L_G \neq \min(L_G, L_U) \end{cases} \tag{8}$$

$$b_2 = \begin{cases} 1, & if \quad L_U = \min(L_G, L_U) \\ 0, & if \quad L_U \neq \min(L_G, L_U) \end{cases} \tag{9}$$

It is assumed that new calls arrive according to the Poisson process with a mean arrival rate of λ. Call duration T_{call} is exponentially distributed with a mean of $1/\mu$. The call completion rate is μ. If we set the time unit of user state transition diagram to be the same as call arrival and completion rate, the call completion probability is:

$$P_{term} = \mu. \tag{10}$$

The new call arriving probability is:

$$P_{new} = \lambda. \tag{11}$$

The steady state probabilities are as follows:

$$P_0 = P_{term} \; / \; (P_{term} + P_{new}) \;, \tag{12}$$

$$P_1 = P_{new\_h} \times a \times P_0 + (1 - P_{term}) \times (1 - P_{ex}) \times a \times P_{in} + P_{en} \times (1 - P_{term}) \times P_4 \;, \tag{13}$$

$$P_2 = P_{new\_h} \times b \times P_0 + (1 - P_{term}) \times (1 - P_{ex}) \times b \times P_{in} + P_{en} \times (1 - P_{term}) \times P_5 \;, \tag{14}$$

$$P_3 = P_{new\_h} \times c \times P_0 + (1 - P_{term}) \times (1 - P_{ex}) \times c \times P_{in} \;, \tag{15}$$

$$P_4 = (c_1 \times a_2 \times P_0 + c_2 \times (P_{new\_h} \times a \times P_0 + (1 - P_{term}) \times (1 - P_{ex}) \times a \times P_{in}) + c_2 \times a_2 \times P_3 + c_3 \times a_2 \times P_{out}) \; / \; (1 - c_2 \times P_{en} \times (1 - P_{term})) \;, \tag{16}$$

$$P_5 = P_{out} - P_4 \;, \tag{17}$$

where $c_1 = 1 - P_{new\_h}$, $c_2 = P_{ex} \times (1 - P_{term})$, $c_3 = (1 - P_{term}) \times (1 - P_{en})$.

Numerical results are presented to validate the above analysis. It is assumed that the loads of the three RATs are the same. Fig. 4 shows that with the increase of the probability of a new call arrival rate in the hotspot area, the probability of a user being served inside the hotspot area increases, while the probability of a user being served outside the hotspot area decreases.

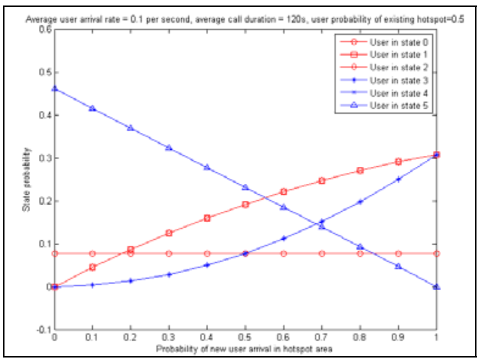

Fig. 4 Increasing the probability of new user arrival rate in the hotspot area. Reproduced from [13]

Fig. 5 shows that with the increase of a new user arrival rate, the probability of a user being idle decreases and the probability of a user being served increases.

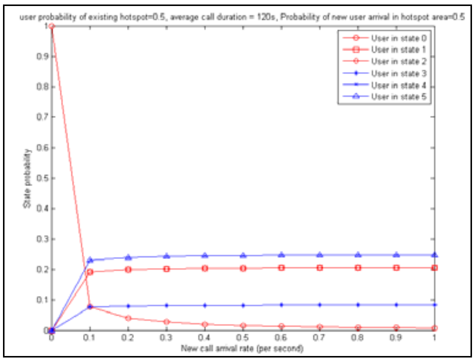

Fig. 5 Increasing new call arrival rate. Reproduced from [13]

V. SERVICE BASED CRRM ALGORITHM

Note: This section is based heavily on [15] (co-authored by the author of this paper). The results and figures presented are drawn from the earlier paper and are presented here for the purpose of allowing clearer comparison with alternative approaches, particularly those described in Section VI.

In this section, a service based CRRM algorithm is studied using the Markov model. A service based CRRM algorithm allocates calls into a specific RAT based on the class of service, such as voice, video streaming, data, etc [14]. This algorithm is based on the reality that different RATs are designed to provide different classes of service. As an example, GSM is designed for voice services however UMTS is designed for data services. A service based CRRM algorithm allocates new calls into a RAT that can better provide and support the service class of the call. Reference [15] proposed a user level Markov model for service based CRRM algorithm in a simple and complex scenario. The considered multi-access cellular network was that users can move inside and outside the hotspot area. Two types of services, voice and data are considered. Voice users are allocated to GERAN, UTRAN and WLAN in order and data users are allocated in the inverse order [7]. A user was randomly determined as voice or data. P_{voice} and P_{data} were defined as the probabilities of a call to be voice and data respectively.

$$P_{voice} + P_{data} = 1. \tag{18}$$

The following definitions were made. P_{new} is the probability of a new call arrives. P_{term} is the probability of an existing call is terminated. P_{in} is the probability of a user residing in the hotspot area. P_{out} is the probability of a user residing outside the hotspot area, where $P_{out} = 1 - P_{in}$. P_{voice} is the probability of a new call being voice call and P_{data} is the probability of a new call being data call, where $P_{data} = 1 - P_{voice}$. P_{new_h} is the probability of a new call arriving in the hotspot area, where $P_{new\_h} = P_{new} \times P_{in}$. The probability of a new call arriving outside the hotspot area is then to be $P_{new} \times P_{out}$. P_{ex} is the probability of a user exiting the hotspot area during a session. P_{en} is the probability of a user entering the hotspot area during a session.

The steady state probabilities are as follows:

$$P_0 = P_{term} /(P_{term} + P_{new}), \quad (19)$$

$$P_1 = P_{in} - P_2 - P_3, \quad (20)$$

$$P_2 = d_3 \times P_5, \quad (21)$$

$$P_3 = (P_{new\_h} \times P_{data} \times P_0 + d_2 \times P_2) / (1 - d_2), \quad (22)$$

$$P_4 = [c_1 \times P_{voice} \times P_0 + c_2 \times d_1 \times P_0 /(1-d_2)]/[1-c_2 \times d_3/(1-d_2)-c_3], \quad (23)$$

$$P_5 = P_{out} - P_4, \quad (24)$$

Where $c_1 = P_{new} \times P_{out}$, $c_2 = P_{ex} \times (1 - P_{term})$, $c_3 = (1 - P_{term}) \times (1 - P_{en})$, $d_1 = P_{new\_h} \times P_{voice}$, $d_2 = (1 - P_{term}) \times (1 - P_{ex})$, $d_3 = P_{en} \times (1 - P_{term})$.

$$P_{in} = P_1 + P_2 + P_3 = 1 - P_0 - P_{out}, \quad (25)$$

$$P_{out} = P_4 + P_5 = (c_1 \times P_0 + c_2 \times (1 - P_0))/(1 + c_2 - c_3). \quad (26)$$

Fig. 6 shows that with the increase of the probability of new users arriving in the hotspot area, P_2 depends on the number of data users served in UTRAN outside the hotspot area and the probability of users cross the border. A higher number of users outside the hotspot area will cause a lower probability of users entering the hotspot area. The maximum value of P_2 occurs at P_h=0.5.

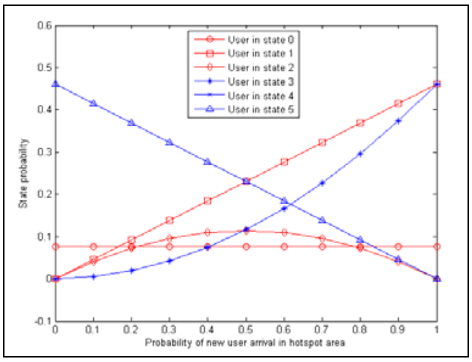

Fig. 6 Increasing the probability of new user arrival in the hotspot area. Reproduced from [15]

Fig. 7 shows that with the increase of the probability of a user being a real time (RT) user, the probabilities of a user to be served in GERAN/UTRAN increases while the probabilities of the user to be served in WLAN decreases.

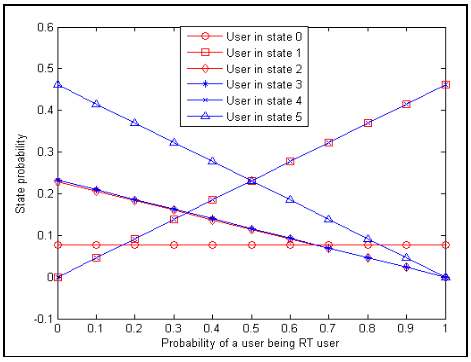

Fig. 7 Increasing the probability of a user being a real time (RT) user. Reproduced from [15]

VI. Comparison Between Load-Balancing Based and Service Based CRRM Algorithms

In this section, we extend the analysis presented in [13] and [15] to compare the performance of load-balancing based and service based CRRM algorithms in multi-access cellular networks. GERAN, UTRAN and WLAN are assumed to co-exist in the same coverage area (hotspot area). It is assumed that users arrive and move only inside the hotspot area. Voice and data are the two considered service types.

In GERAN, each voice user is allocated by one channel. The bit rate of voice user in GSM is 12.2 kbps. Each data user can occupy one channel when the GERAN capacity is sufficient. However, multiple data users are forced to share one channel when the GERAN capacity is not sufficient. The bit rate of data user is 59.2 kbps/channel. Load GERAN is calculated by the following equation:

$$L_G = (N_{VG} + N_{DG}) / N_{CG}, \quad (27)$$

where

- N_{VG} is the number of voice users served in GERAN,
- N_{DG} is the number of data users served in GERAN,
- N_{CG} is the total number of channels in GERAN.

In UTRAN, voice users are served at a bit rate of 12.2 kbps. For data users, they equally share the capacity not used by voice users. Data users can achieve a maximum bit rate of 128 kbps when the UTRAN capacity is sufficient. However, data users will reduce their bit rate towards 16 kbps when the UTRAN capacity is not sufficient. The new arrival data user will be blocked when it will cause the bit rate of data users to be lower than 16 kbps. The load factors for uplink and downlink are calculated as follows [16]:

$$\eta_{UL} = (1+i) \times \sum_{j=1}^{N} \frac{1}{1 + \dfrac{W}{(E_b/N_o)_j \times R_j \times \upsilon_j}}, \quad (28)$$

where N is the number of users per cell, υ_j is the activity factor of user j at physical layer, E_b / N_o is the signal energy per bit divided by noise spectral density, W is the chip rate,

International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:5, No:1, 2011

R_j is the bit rate of user j and i is the other cell to own cell interference ratio seen by the base station receiver. The load factor for the downlink is:

$$\eta_{DL} = \sum_{j=1}^{N} \upsilon_j \times \frac{(E_b / N_o)_j}{W / R_j} \times [(1-\alpha_j) + i_j], \qquad (29)$$

where α_j is the orthogonality of channel of user j and i_j is the ratio of other cell to own cell base station received by user j.

In WLAN, the resource consumption for each voice user is 2 × 22.8 kbps (uplink and downlink). Data users equally share the capacity not used by voice users. A data user can occupy the whole WLAN bandwidth when there are no other users served. The new arrival data user will be blocked when it will cause the bit rate of data users to be lower than 16 kbps. Load WLAN is calculated by the following equation:

$$L_W = N_{VW} \times 2 \times 22.8e3 / W_C + N_{DW} \times 2 / W_C, \qquad (30)$$

where

- N_{VW} is the number of voice users served in WLAN,
- N_{DW} is the number of data users served in WLAN,
- W_C is the available WLAN capacity.

Load-balancing based and service based CRRM algorithms are evaluated in terms of traffic distribution, new call blocking probability, VHO call dropping probability and throughput respectively.

Fig. 8 shows the distribution of traffic among the three RATs. It can be seen that for load-balancing based algorithm, voice and data users are distributed equilibrium between GERAN, UTRAN and WLAN. However, for the service based algorithm more voice users are allocated to GERAN and less voice users are served by WLAN. Evidently, in the service based algorithm the most of data users are allocated to WLAN.

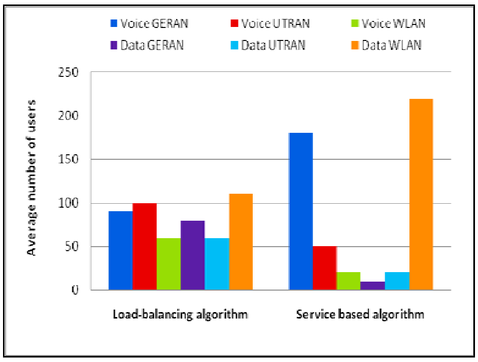

Fig. 8 Traffic distribution.

Fig. 9 and 10 show the blocking and dropping probabilities for the load-balancing based and service based CRRM algorithms. Simulation results show that new call blocking probability and VHO dropping probability in load-balancing based algorithm are lower than in service based algorithm.

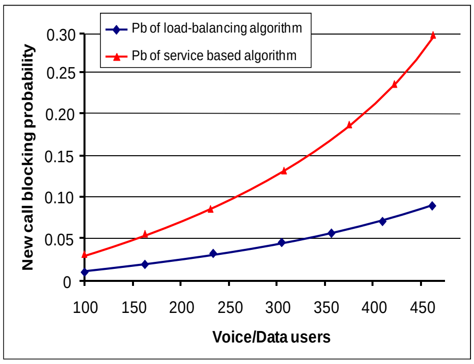

Fig. 9 New call blocking probability.

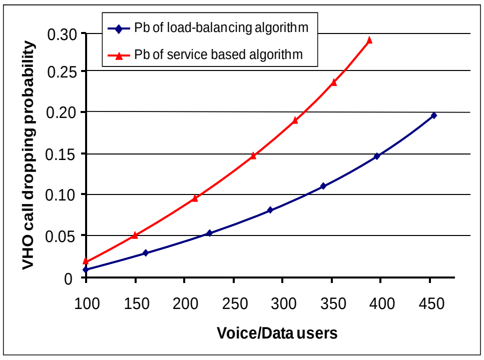

Fig. 10 VHO call dropping probability.

Fig. 11 illustrates the throughput for the load-balancing based and service based CRRM algorithms. It can be seen that load-balancing based algorithm performs better than the service based algorithm when traffic load is low. However, when the traffic load becomes high, service based algorithm outperforms the load-balancing based algorithm in terms of throughput.

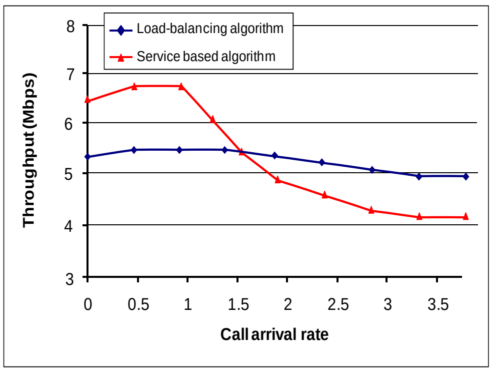

Fig. 11 Throughput.

VII. Conclusion

CRRM algorithms can bring a significant benefit to the heterogeneous network. This paper has presented a user level Markov model for a heterogeneous B3G network. The load-balancing based and the service based CRRM algorithms have been considered using the presented Markov model. A comparison for the performance of the two CRRM algorithms has been studied in terms of traffic distribution, new call blocking probability, VHO call dropping probability and throughput where GERAN, UTRAN and WLAN co-exist in the same coverage area. Simulation results show that in terms of dropping and blocking probability, the service based algorithm performs better than the load-balancing based algorithm. In terms of throughput, the load-balancing based algorithm outperforms the service based algorithm when traffic load is low. However, the service based algorithm performs better when the traffic load becomes high.

Acknowledgment

This work is sponsored by the Centre for Real-Time Information Networks (CRIN) in the Faculty of Engineering & Information Technology at the University of Technology, Sydney (UTS). The author would also like to acknowledge the contributions to earlier versions of the work described in this paper (particularly the work described in Section IV and Section V) by Wu et al.

References

[1] K. Murray, R. Mathur, and D. Pesch, "Network Access and Handover Control in Heterogeneous Wireless Networks for Smart Space Environments", *1st International Workshop on Managing Ubiquitous Communications and Services (MUCS 2003)*, Waterford, Ireland, December 11, 2003.

[2] J. Perez-Romero, O. Sallent, R. Agusti and M. A. Diaz-Guerra, *Radio Resource Management Strategies in UMTS*, John Wiley & Sons Ltd, Chichester, UK 2005.

[3] R. Agusti, O. Salient, J. Perez-Romero, and L. Giupponi, "A Fuzzy-Neural Based Approach for Joint Radio Resource Management in a Beyond 3G Framework", *First International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks (QSHINE'04)*, Dallas, Texas, USA, October 18–20, 2004, pp. 216–224.

[4] O. E. Falowo and H. A. Chana, "Joint call admission control algorithms: Requirements, approaches, and design considerations", *Computer Communications*, vol. 31, no. 6, April 2008, pp. 1200-1217.

[5] 3GPP TR 25.881 v5.0.0, "Improvement of RRM across RNS and RNS/BSS (Release 5)", 2001.

[6] A. Tolli, P. Hakalin, and H. Holma, "Performance Evaluation of Common Radio Resource Management (CRRM)", *2002 IEEE International Conference on Communications (ICC 2002)*, New York, USA, April 28 – May 2, 2002, pp. 3429- 3433.

[7] A. Baraev, L. Jorguseski, and R. Litjens, "Performance Evaluation of Radio Access Selection Procedures in Multi-Radio Access Systems", *The 8th International Symposiumon Wireless Personal Multimedia Communication (WPMC'05)*, Aalborg, Denmark, September 18 – 22, 2005.

[8] J. Perez-Romero, O. Sallent, and R. Agusti, "Policy-based Initial RAT Selection Algorithms in Heterogeneous Networks", *The 7th IFIP International Conference on Mobile and Wireless Communications Networks (MWCN 2005)*, Marrakech, Morocco, September 19-21, 2005.

[9] S. M. Ross, *Introduction to Probability Models*, Ninth Edition, Academic Press, 2007.

[10] A. Tolli and P. Hakalin, "Adaptive Load Balancing Between Multiple Cell Layers", *2002 IEEE 56th Vehicular Technology Conference (VTC 2002)*, vol. 3, Vancouver, Canada, September 24–28, 2002, pp. 1691–1695.

[11] K. H. Suleiman, H.A. Chan, and M.E. Dlodlo, "Load Balancing in the Call Admission Control of Heterogeneous Wireless Networks", *International Conference on Communications and Mobile Computing (IWCMC 2006)*, Vancouver, British Columbia, Canada, July 3–6, 2006, pp. 245–250.

[12] A. Pillekeit, F. Derakhshan, E. Jugl, and A. Mitschele-Thiel, "Force-Based Load Balancing in Co-located UMTS/GSM Networks", *2004 IEEE 60th Vehicular Technology Conference (VTC 2004)*, vol. 6, Los Angeles, California, USA, September 26-29, 2004, pp. 4402–4406.

[13] L. Wu, A. E. Al-Sabbagh, K. Sandrasegaran, and M. Elkashlan "A User Level Markov Model for Load Balancing Based RAT Selection Algorithm", *8th International Information and Telecommunication Technologies Symposium (I2TS 2009),* Florianopolis, Santa Catarina State, Brazil, December 09 - 11, 2009, pp. 1-5.

[14] W. Zhang, "Performance of Real-time and Data Traffic in Heterogeneous Overlay Wireless Networks", *19th International Teletraffic Congress (ITC 19)*, Beijing, China, August 29 – September 2, 2005.

[15] L. Wu, A. E. Al-Sabbagh, K. Sandrasegaran and M. Elkashlan, "A User Level Markov model for service based CRRM algorithm", *International Conference on Multimedia Computing and Information Technology (MCIT'2010)*, Sharjah, UAE, March 2-4, 2010, pp. 41 – 44.

[16] H. Holma and A. Toskala, *WCDMA for UMTS – HSPA Evolution and LTE*, fourth Edition, John Wiley & Sons Ltd, Chichester, UK 2007.

Abdallah AL Sabbagh received a B.Sc. in Information Technology and Computing (Hons) from the Arab Open University, Lebanon and the Open University, UK (2006) and a Master of Engineering Studies (MES) in Telecommunication Networks from the Faculty of Engineering at the University of Technology, Sydney (2010). He is currently with the School of Computing and Communications, and Centre for Real-Time Information Networks (CRIN) at the University of Technology, Sydney (UTS). His recent research is the Radio Resource Management (RRM), Admission Control, Vertical Handover (VHO) and packet scheduling for Next Generation Wireless Networks (NGWN).

Abdallah is member of the Centre for Real-time Information Networks (CRIN) and Engineers Australia.