A Comprehensive Survey on RAT Selection Algorithms for Heterogeneous Networks

Abdallah AL Sabbagh, Robin Braun and Mehran Abolhasan

Abstract— Due to the coexistence of different Radio Access Technologies (RATs), Next Generation Wireless Networks (NGWN) are predicted to be heterogeneous in nature. The coexistence of different RATs requires a need for Common Radio Resource Management (CRRM) to support the provision of Quality of Service (QoS) and the efficient utilization of radio resources. RAT selection algorithms are part of the CRRM algorithms. Simply, their role is to verify if an incoming call will be suitable to fit into a heterogeneous wireless network, and to decide which of the available RATs is most suitable to fit the need of the incoming call and admit it. Guaranteeing the requirements of QoS for all accepted calls and at the same time being able to provide the most efficient utilization of the available radio resources is the goal of RAT selection algorithm. The normal call admission control algorithms are designed for homogeneous wireless networks and they do not provide a solution to fit a heterogeneous wireless network which represents the NGWN. Therefore, there is a need to develop RAT selection algorithm for heterogeneous wireless network. In this paper, we propose an approach for RAT selection which includes receiving different criteria, assessing and making decisions, then selecting the most suitable RAT for incoming calls. A comprehensive survey of different RAT selection algorithms for a heterogeneous wireless network is studied.

Keywords— Heterogeneous Wireless Network, RAT selection algorithms, Next Generation Wireless Network (NGWN), Beyond 3G Network, Common Radio Resource Management (CRRM).

I. INTRODUCTION

OVER the past twenty years, a number of different wireless technologies have been developed. In First Generation (1G), Narrow Band analogue wireless networks were used. Second Generation (2G) is based on digital system technology. It uses circuit switching technology. Global System for Mobile Communication (GSM) is the commercially standard of 2G. Time Division Multiple Access (TDMA) and Code Division Multiple Access (CDMA) were

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).

Robin Braun is a Professor of Telecommunications in 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: robin.braun@uts.edu.au).

Mehran Abolhasan is a Senior Lecturer at the School of Computing and Communications in the Faculty of Engineering and Information Technology at the University of technology, Sydney (UTS), P.O. Box 123 Broadway, NSW 2007 Australia (e-mail: mehran.abolhasan@uts.edu.au).

the two types of standards used in 2G technologies. 2G is used to transmit voice and it introduced a low volume digital data for mobiles such as Short Message Service (SMS) or Multimedia Message Service (MMS). The 2.5 generation adds packet switched functionality. In 2.5G, General Packet Radio Service (GPRS) was introduced as an extension for GSM to improve throughput and offers faster data rates to deliver packet-based services efficiently over GSM networks. Enhanced Data Rates for Global Evolution (EDGE) was introduced as 2.75G. EDGE improves the data transmission rates of GPRS. The third generation (3G) uses packet switching technology. 3G includes CDMA2000 and Universal Mobile Telecommunications System (UMTS) based on Wideband CDMA (W-CDMA) technologies; it allows high-speed digital data transmission. References [1]-[3] contain more technical details about the different wireless networks technologies.

NGWN is expected to be a heterogeneous network which integrates all 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. The question is how to allocate users for the most suitable RAT for them. The solution to this question can lead to maximization of radio resources and better performance for service providers and provision a QoS with low costs to users. 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 will be described in the section II.

Currently, Radio Resource Management (RRM) including admission control, congestion control, power control, scheduling, handover (HO), initial RAT selection algorithm and vertical HO (VHO) algorithm is implemented efficiently for the RAT that it was developed for. However, it is not suitable for the heterogeneous wireless network. CRRM was proposed to manage radio resource utilization in B3G network within the different RATs.

The remainder of this paper is organized as follows. Section II presents the architecture of a heterogeneous B3G network. UTRAN is integrated with GERAN and WLAN. Section III proposes a RAT selection approach for selecting the appropriate RAT for requested calls. In Section IV, a comprehensive survey of RAT selection algorithms is studied. Finally, conclusions are drawn in Section V.

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 different QoS demands [4]. 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 User Equipments (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).
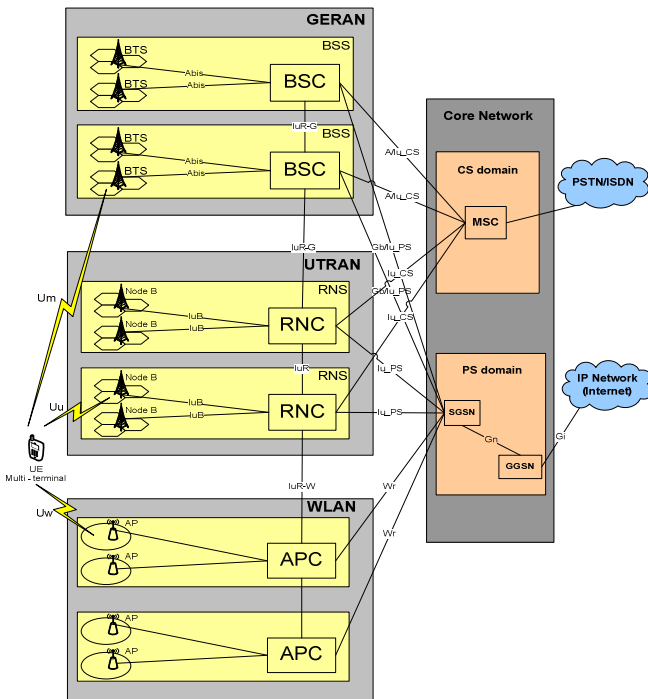

Fig. 1 Heterogeneous Beyond 3G network architecture.

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 Base Transceiver Stations (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. RAT Selection Approach

In heterogeneous wireless networks, different RATs coexist in the same coverage area. The challenge is how to decide the most suitable RAT for each call. An optimized RAT selection algorithm will guarantee the targeted QoS for the requested call, offer high capacity, maintain the planned coverage area and lead to efficient utilization of radio resources.

This section proposes a RAT selection approach for selecting the appropriate RAT for requested calls. The proposed RAT selection approach tends to focus on three main parts: considered criteria (input), RAT selection algorithm (decision making) and Admitted call (output). This is shown in Fig. 2. In the input, different criteria are considered such as feedback/measurements from local RRM of the available RATs, supported RATs for the requested service, operator policies, user preferences, terminal capability of UE, demanded QoS for the selected call, type of service (such as voice, data, video streaming, interactive service), load conditions for the available RAT, interference conditions and the cost of requested service in each RAT. In the decision making, RAT selection algorithms such as random based, load-balancing based, service-class based, service-cost based, path-loss based, policy based, layer based, non computational-intelligence, or computational-intelligence based RAT selection algorithms will consider all these criteria or some of them to make a decision. These RAT selections algorithms will be described in Section IV. In the output, after the decision has been made, the call will be allocated to the selected RAT. A Radio Recourse Unit (RRU) and a bit rate

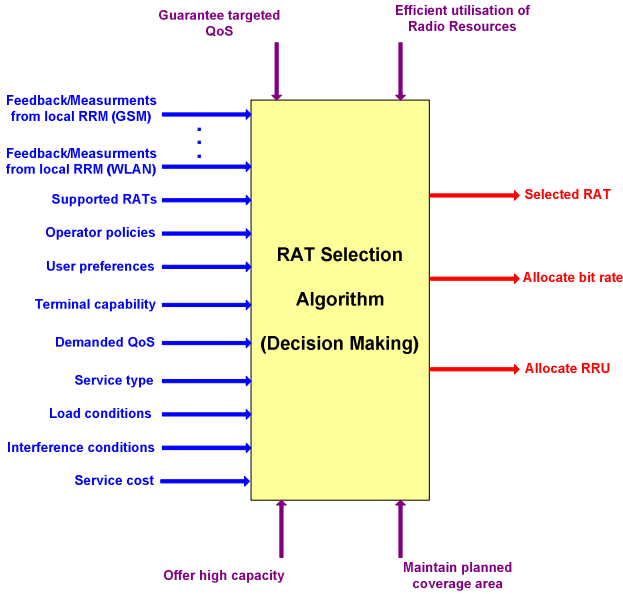will be allocated for this call from the selected RAT.

Fig. 2 RAT selection approach.

IV. Comprehensive Survey of RAT Selection Algorithms

A number of RAT selection algorithms including initial RAT selection and vertical HO have been proposed in the literature for heterogeneous wireless networks. Reference [5] presents a good revision on these algorithms. Each one has its benefits and limitations. These algorithms are described in the next subsections.

A. Random based RAT Selection Algorithm

In this algorithm, while incoming new or vertical handover call is arrived, any of the available RATs will be selected randomly for this call. When no enough radio resources are available to serve the call in the chosen RAT, the call will be dropped or blocked. By dealing with random based selection that works on CRRM, if the chosen RAT is not able to serve the call, another RAT will be randomly chosen. If none of the available RATs are able to serve this call, the incoming call will be dropped or blocked.

Random based RAT selection algorithm is basically used within performance validation of another RAT selection algorithm by differentiation. The benefit of this algorithm is that it is simple to participate. Furthermore, it has a high call dropping/blocking probability, and less utilization of radio resource efficiency.

B. Load-balancing based RAT Selection Algorithm

Load-balancing based RAT selection algorithm aims to distribute traffic load between all available RATs in a heterogeneous wireless network. Balancing load between all available RATs in a heterogeneous wireless network offers an efficient utilization of the radio resources [6], [7]. Traffic load could be continuously balanced [6], balanced at a specific part of time [7], or balanced while a specific load threshold is being reached [8].

Also in a heterogeneous wireless network, load-balancing could be forced [8] or unforced. Forced load-balancing is implemented by transferring some ongoing calls from a high loaded RAT to a less loaded RAT, which is called "call reassignment". However, unforced load-balancing is implemented only while there is an admission of a new call or a necessary vertical handover call resulting in a user terminal that is moving outside the coverage area of its current RAT. In this situation, a new or a vertical handover call is allocated to the least loaded RAT between the available RATs. A great benefit of load-balancing algorithm is high network stability due to addressing load distribution. However, this algorithm does not guarantee the required QoS for the admitted calls. It is network-centric.

C. Service-class based RAT Selection Algorithm

Service-class based RAT selection algorithm allocates calls into a specific RAT based on the class of service, such as voice, video streaming, data, etc [9]. 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-class based RAT selection algorithm allocates new calls into a RAT that can better provide and support the service class of the call. Therefore, this algorithm has the benefit of achieving high QoS, it is a user-centric. However, it may admit too many calls to a specific RAT and lead to high unbalanced network load.

Service-class based RAT selection algorithm can be sorted as rigid or flexible. Rigid service-class based algorithm attempts to allocate an ongoing call of a particular class into a specific RAT. If the chosen RAT for this call was not able to accept the call, might because of no enough radio resource available, other RATs are unacceptable. Therefore the call will be blocked. On the other hand, flexible service-class based algorithms attempt to allocate an ongoing call for a particular class into a specific RAT. If the chosen RAT for this call was not able to accept the call, other RATs are acceptable. Flexible service-class based algorithm has less call blocking probability by comparing to rigid service-class based algorithm.

D. Service-cost based RAT Selection Algorithm

Service-cost based RAT selection algorithm allocates incoming calls through the cheapest RAT cost for the requested service, so the user will be serving by the cheapest service cost in the heterogeneous wireless network. The objective of this technique is working basically on the reality that the service cost varies from a RAT to another. The service cost can be attributed by variation to the cost of hardwares and the cost of license spectrum. This algorithm has the benefit of decreasing the service cost that served by users in general. However, it might incur a serious unbalanced network load.

E. Path-loss based or Network-Controlled Cell-Breathing (NCCB) RAT Selection Algorithm

Path-loss based or Network-Controlled Cell-Breathing (NCCB) RAT selection algorithm allocates incoming calls into a specific RAT based on the measurements of path-loss taken in the cells of each RAT [10]. It is calculated by measuring the received downlink power of the best cell in UTRAN provided by the terminal in the establishment phase. If the path-loss of UTRAN is below the path loss threshold value, the call will be allocated to UTRAN; otherwise it will be admitted to GERAN. Path-loss based algorithm achieves high throughput and low bit error rate; however it might lead to high frequency of vertical handover among available RATs.

F. Policy based RAT Selection Algorithm

Policy based RAT selection algorithm allocates incoming calls into a specific RAT based on specific rules specified by the network. A simple policy has been proposed in [4], [11] which include three policies: Voice GERAN (VG), Voice UTRAN (VU) and Indoor (IN) policies. In VG policy, it allocates voice services to GERAN and other services to UTRAN. On the other hand, In VU policy it allocates voice services to UTRAN and other services to GERAN. In IN policy, it allocates indoor services to GERAN and outdoor services to UTRAN. A complex policy has proposed in [11] which include three policies: VG*VU, VG*IN and IN*VG policies. In VG*VU policy, it allocates voice users to GERAN firstly to fill the available capacity and then it will direct them to UTRAN. Furthermore, the data users will be allocated firstly to UTRAN to fill the available capacity and then it will direct them to GERAN. In VG*IN policy, it always allocates indoor voice users to GERAN and outdoor data users to UTRAN. Outdoor voice users will be allocated firstly to GERAN to fill the available capacity and then it will direct them to UTRAN. In contrast, indoor data users will be allocated firstly to UTRAN to fill the available capacity and then it will direct them to GERAN. In IN*VG policy, it always allocates indoor voice users to GERAN and outdoor data users to UTRAN. Outdoor voice users will be allocated firstly to UTRAN to fill the available capacity and then it will direct them to GERAN. Therefore, indoor data users will be allocated firstly to GERAN to fill the available capacity and then it will direct them to UTRAN. This algorithm has the benefit of reducing the average delay for data users. However, it does not consider RAT load and it only considers GERAN/UTRAN scenario.

G. Layer based RAT Selection Algorithm

Layer based RAT selection algorithm allocates calls into a specific layer. If the layer is not able to accept the call, the layer based RAT selection algorithm attempts to allocate the call through the following available layer. Layer based algorithm is simple to implement, however it might incur a serious unbalanced network load.

H. Non computational-intelligence based RAT Selection Algorithm

Non computational-intelligence based RAT selection algorithm allocates incoming calls into a specific RAT basically on certain utility or cost functions [12] or a fittingness factor [13]-[16] directed from a number of characters without referring to the computational intelligence techniques. This algorithm has high efficiency, however it is complicated to implement and it might lead to high computational overheads.

I. Computational-intelligence based RAT Selection Algorithm

Computation-intelligence based RAT selection algorithm allocates incoming calls into a selected RAT by applying the computational intelligence techniques to RAT selection Criteria. Some of the most used computational intelligence techniques are Fuzzy logic technique as in [17], Fuzzy neural technique as in [4], [18] and [19], Fuzzy MADM (Multiple Attribute Decision Making) technique as in [20]-[22], and genetic algorithm technique.

Fuzzy logic is the most used techniques for RAT selection algorithms. This is shown in Fig. 3. It includes fuzzifier, inference engine, fuzzy based rules and defuzzifier. The fuzzifier is the input; it allocates a value from 0 to 1 for each input, such as signal strength, available resource for each RAT and mobile speed, corresponding to the degree of membership of this input. In the inference engine, for each fuzzy subset defined in the fuzzifier the inference engine will use some of the fuzzy based rules to indicate if it is suitable to be selected for each of the available RATs. The output of this part will be a combination of the available number of RATs, and it will have a value that varies between Y (yes), N (no), PY (probably yes), and PN (probably no). The fuzzy based rules are a series of "If then" rules. The defuzzifier will convert the output of the inference engine to a value between 0 to 1 named Fuzzy Selected Decision (FSD) for each of the available RATs. The RAT with the highest value will be the preferred one to be selected.
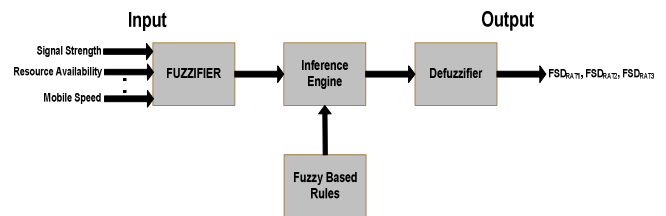

Fig. 3 Fuzzy Logic.

Computation-intelligence based RAT selection algorithm is very efficient, it increase the satisfaction of users in the heterogeneous wireless network. However it is complicated to implement it.

V. Conclusion

An optimized RAT selection algorithm is definitively necessary in heterogeneous B3G network. In this paper, we present the architecture of a heterogeneous B3G network. A RAT selection approach is proposed. Then a comprehensive survey of different RAT selection algorithms has been studied. Single criteria algorithms such as load-balancing based and service based are simple and easy to implement however they may lead to unbalanced network, high call blocking probability or do not guarantee QoS. Multiple criteria algorithms such as non computational-intelligence based and computational-intelligence based are more efficient however they are complicated to implement. The best solution could be either by developing an intelligent one that it is simple and easy to implement or by distributing the role of the RAT selection algorithm among the UE, Node B, RNC and core network.

References

[1] V. K. Garg, *Wireless Communications and Networking*, Morgan Kaufmann Publishers, San Francisco, CA, USA 2007.

[2] P. Nicopolitidis, M. S. Obaidat, G. I. Papadimitriou, and A. S. Pomportsis, *Wireless Networks*, John Wiley & Sons Ltd, Chichester, UK 2003.

[3] J. Schiller, *Mobile Communications*, 2nd Edition, Addison Wesley, Harlow UK 2003.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] A. Umbert, L. Budzisz, N.Vucevic and F. Bernardo, "An all-IP heterogeneous wireless testbed for RAT selection and e2e QoS evaluation", The *2007 International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST 2007)*, Cardiff, Wales, UK, September 12-14, 2007, pp. 310–315.

[11] 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.

[12] O. Ormond, J. Murphy, and G-M. Muntean, "Utility-Based Intelligent Network Selection in Beyond 3G Systems", *2006 IEEE International Conference on Communications (ICC 2006)*, Istanbul, Turkey, June 11–15, 2006, pp. 1831–1836.

[13] J. Perez-Romero, O. Sallent, and R. Agusti, "A Novel Metric for Context-Aware RAT Selection in Wireless Multi-Access Systems", *2007 IEEE International Conference on Communications*, Glasgow, Scotland, June 24–28, 2007, pp. 5622-5627.

[14] O. Sallent, J. Perez-Romero, R. Ljung, P. Karlsson, and A. Barbaresi, "Operator's RAT Selection Policies Based on the Fittingness Factor Concept", *16th IST Mobile and Wireless Communications Summit*, Budapest, Hungary, July 1-5, 2007, pp. 1-5.

[15] J. Perez-Romero, O. Sallent, A. Umbert, A. Barbaresi, R. Ljung, and R. Azevedo, "RAT Selection in Wireless Multi-Access Systems", *First Ambient Networks Workshop on Mobility, Multiaccess, and Network Management (M2NM 2007)*, Sydney, Australia, October 16-19, 2007.

[16] J. Perez-Romero, O. Salient, and R. Agusti, "A Generalized Framework for Multi-RAT Scenarios Characterisation", *The 65th IEEE Vehicular Technology Conference (VTC2007-Spring)*, Dublin, Ireland, April 22-25, 2007, pp. 980 - 984

[17] P. M. L. Chan, R. E. Sheriff, Y. F. Hu, P. Conforto, and C. Tocci, "Mobility Management Incorporating Fuzzy Logic for a Heterogeneous IP Environment", *IEEE Communications Magazine*, vol. 39, no.12, December 2001, pp. 42-51.

[18] 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.

[19] L. Giupponi, R. Agusti, J. Perez-Romero, and O. Sallent, "A novel joint radio resource management approach with reinforcement learning mechanisms", *24th IEEE International Performance, Computing, and Communications Conference (IPCCC 2005)*, Phoenix, Arizona, USA, Appril 7-9, 2005, pp. 621- 626.

[20] W. Zhang, "Handover Decision Using Fuzzy MADM in Heterogeneous Networks", *2004 IEEE Wireless Communications and Networking Conference (WCNC 2004)*, vol. 2, Atlanta, GA USA, March 21-25, 2004, pp. 653 - 658.

[21] A. L. Wilson, A. Lenaghan, and R. Malyan, "Optimising Wireless Access Network Selection to Maintain QoS in Heterogeneous Wireless Environments", *International Symposium on Wireless Personal Multimedia Communications 2005 (WPMC 2005)*, Aalborg Denmark, September 18-22, 2005, pp. 1236-1240.

[22] O.E. Falowo and H.A. Chan, "Fuzzy Logic Based Call Admission Control for Next Generation Wireless Networks", *3rd International Symposium on Wireless Communication Systems (ISWCS'06)*, Valencia, Spain, September 5–8, 2006.