Loose and Tight Interworking between Vertical and Horizontal Handovers in Multi-RAT Scenarios

J. Pérez-Romero, O. Sallent
Universitat Politècnica de Catalunya (UPC)
E-mail: [jorperez, sallent] @ tsc.upc.edu

Abstract.- **Common Radio Resource Management (CRRM) refers to the functions devoted to efficiently manage the common pool of resources available in a heterogeneous network where different radio access technologies coexist. In that sense, this paper focuses on the vertical handover strategy, and analyses the interworking between the vertical handover and the horizontal handover. The impact of two interworking approaches, denoted as tight and loose, in terms of QoS performance and signalling overhead is analysed in a detailed UTRAN/GERAN scenario considering a service-based policy. Results reveal that the tight approach achieves a better fulfilment of the policy.**

I. Introduction

The heterogeneous network concept is intended to propose a flexible and open architecture for a large variety of wireless access technologies, applications and services with different QoS demands, as well as different protocol stacks. These new scenarios where different Radio Access Technologies (RATs) will coexist and will operate in a coordinated way are often referred to as beyond 3G systems. They assume the existence of multi-mode terminals, able to provide connectivity to multiple access networks.

The complementary characteristics of these networks make possible to exploit the trunking gain resulting from the joint consideration of the different networks as a whole, thus leading to a better overall performance than the accumulated performances of the stand-alone systems. Clearly, a proper management of the available radio resources is a key issue for a successful realization of beyond 3G systems, since potential gains need to turn into reality [1]. To this end, the capability to switch connections from one Radio Access Network (RAN) to another, usually known as inter-system or vertical handover (VHO), is a key enabler to properly manage the heterogeneous radio access network scenario. Research community has already identified the importance of this mechanism in future mobile scenarios and, consequently, a number of contributions have appeared in the recent years [2]-[5]. Specifically, in [2] the benefits of Common Radio Resource Management (CRRM) by carrying out load-based vertical handovers in order to balance the load of the different cells and RATs are analysed. Similarly, in [3] the advantages of distributing the load among different networks to increase flexibility and reduce network equipment costs are addressed, and in [4] different policies to overflow sessions that arrive to saturated RANs are discussed. Finally, in [5] different procedures for making measurements of different RATs are discussed as a means to provide support for vertical handover decisions.

In this context, this paper, after providing an insight on the framework for developing CRRM techniques, the vertical handover is identified as one of these functions and the rest of the paper focuses on this particular CRRM strategy. Specifically, it discusses the interworking between the vertical handover and the horizontal handover (HHO), i.e. the functionality to change the base station that the mobile is connected to in a given RAN. Two approaches are discussed, denoted as tight and loose interworking, depending on whether the vertical handover algorithm is invoked before every horizontal handover or not. Both approaches are evaluated by means of system level simulations in a scenario with UMTS Terrestrial RAN (UTRAN) and GSM/EDGE RAN (GERAN).

The rest of the paper is organised as follows. Section II describes the Common Radio Resource Management concept and the corresponding functionalities. Section III focuses on the vertical handover function, and the possibilities for interworking between horizontal and vertical handover are presented. The simulation model is described in section IV and performance results are presented in section V. Finally, conclusions are summarised in Section VI.

II. Common Radio Resource Management

Common Radio Resource Management (CRRM) refers to the set of functions devoted to ensure an efficient use of the radio resources in heterogeneous networks scenarios by means of a proper coordination among the access networks [6]-[8]. The functional model in 3GPP for CRRM operation considers the total amount of resources available divided into radio resource pools. Each pool consists of the resources in a set of cells, typically under the control of a RNC (Radio Network Controller) in UTRAN or a BSC (Base Station Controller) in GERAN. Two types of entities are considered for the management of these pools: the RRM entity, which manages the resources in one pool of a certain access network, and the CRRM entity, which is involved in the coordinated management of the pools from several RRM entities.

The RRM functionalities arising in the context of a single RAT are the admission and congestion control, the horizontal (intra-system) handover, the packet scheduling and the power control [1]. It is worth mentioning that, when these functionalities are coordinated between different RATs in a heterogeneous scenario, they can be denoted as "common" (i.e. thus having the common admission control, common congestion control, etc.). In turn, when an heterogeneous scenario is considered, two specific additional functionalities arise, namely the Initial RAT selection (i.e. the functionality devoted to decide to which RAT a given service request should be allocated) and the Vertical (inter-system) handover (i.e. the functionality devoted to decide a seamless RAT switching for an on-going service). The different possibilities envisaged for the operation between RRM and CRRM entities are further described in [8].

1-4244-0088-0/06/$20.00 ©2006 IEEE

III. Vertical Handover

After the initial RAT selection decision at session initiation, vertical handover is the first step to carry out interoperation among access networks in heterogeneous systems. The successful execution of a seamless and fast vertical handover is essential for hiding to the user the underlying enabling infrastructure. Issues related to vertical handover comprise scanning procedures for the terminal to discover available RATs, measurement mechanisms to capture the status of the air interface in the different RATs, vertical handover triggers (i.e. the events that require the system to consider whether a vertical handover is actually required or not), vertical handover algorithm (i.e. the criteria used to decide whether a vertical handover is to be performed or not) and protocol and architectural aspects to support handover execution.

Vertical handover procedures may be useful to support a variety of objectives, such as avoiding disconnections due to lack of coverage in the current RAT, blocking due to overload in the current RAT, possible improvement of QoS by changing the RAT, support of user's and operator's preferences in terms of RATs usage or load balancing among RATs. Thus, the vertical handover procedure enables another dimension into the CRRM problem and provides an additional degree of freedom in rearranging traffic. In this respect, there would be many possibilities to follow both in the decision of changing the initially selected RAT and in the return policy (i.e. deciding the suitable instant in which sessions that are not in the preferred RAT must return to it). The trade-off arising here is between flexibility in CRRM and signalling overhead.

In the following two different degrees of interworking between the vertical and the horizontal handover algorithms are presented. They are referred to as loose and tight interworking, denoted in the rest of the paper as L-VHO and T-VHO, respectively. These terms intend to describe how often the suitability to carry out a vertical handover is considered for each connection. In particular, L-VHO stands for the case that the vertical handover algorithm is executed only when a horizontal handover fails or when a call dropping is about to occur due to bad propagation conditions, so that it is seen simply as a mechanism to avoid call droppings. In this case, after having executed a vertical handover the session will remain in the new RAT as long as the propagation conditions are satisfactory and there are available resources in the subsequent horizontal handovers. In turn, T-VHO stands for the case that the vertical handover decision algorithm is executed at every instant that the horizontal handover algorithm is considered, so that both possibilities are considered prior to taking a decision. In the VHO decision algorithm the RAT selection policy is evaluated for the specific connection. Fig. 1 reflects the interactions between vertical and horizontal handover functions, indicating that horizontal handover algorithm is an inherent part of the RRM entity while vertical handover algorithm belongs to the CRRM entity. Similarly, Fig. 2 summarises in the form of a flow diagram the L-VHO and T-VHO alternatives. We note that e.g. a periodic triggering of the vertical handover algorithm would also be possible, though not considered in this paper.

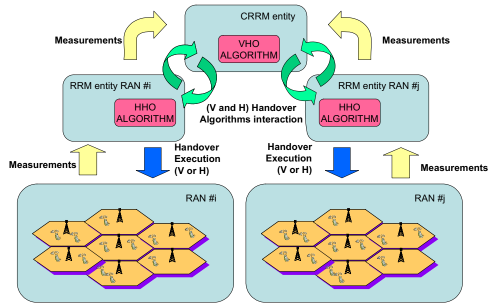

Fig. 1 Interactions between VHO and HHO.

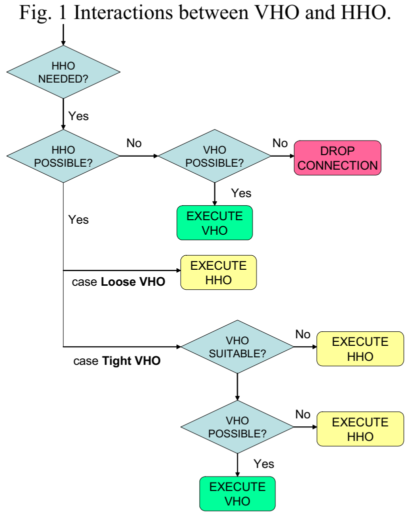

Fig. 2 Flow diagram of T-VHO and L-VHO

There are two main aspects that need to be considered when comparing T-VHO with L-VHO. Firstly, CRRM entity implementation: if the CRRM entity is implemented in every existing RNC/BSC, then HHO and VHO are tightly coupled in a natural way and the interaction is simply an internal matter of the RNC/BSC. On the contrary, if the CRRM entity is implemented only in some RNC/BSC or in an external server [7], then delay in taking decisions plays a role and tends to impact more on T-VHO because of the signalling exchange required between the nodes where the CRRM and RRM entities reside. Secondly, the achieved performance will be a result of the RAT selection policy. Then, as long as T-VHO and L-VHO imply different rates at which a vertical handover is considered, there will also be different chances to be able to follow the established policy and, consequently, different performances can be expected.

With respect to the RAT selection criterion used either at session initiation or in the VHO decision algorithm, a service policy is considered as an example in this paper [9], although the L-VHO and T-VHO approaches could also be used with other policies taking into account e.g. load balancing principles, radio propagation conditions, etc. Specifically, and assuming the existence of voice and interactive traffic, voice traffic is directed to GERAN if capacity is available while interactive traffic is directed to UTRAN. If capacity is not available in the preferred RAT (i.e. if the admission control in this RAT rejects the connection), the other RAT is selected instead. If no capacity is neither available in the alternative

RAT, the call is blocked. The rationale for this service-based policy relays in the better ability of UTRAN to handle interactive users by allocating them on dedicated channels than in case of GERAN, where interactive traffic shares the available packet data channels.

IV. Simulation Model

The considered interworking approaches have been evaluated by means of system level simulations including a detailed characterisation of UTRAN and GERAN. The scenario considers 7 omnidirectional cells for GERAN and 7 for UTRAN. The cells of both RANs are collocated with cell radius 500m. The main parameters of the User Equipment (UE) and the Base Station (BS) are summarised in Table I. It is assumed that all terminals have multi-mode capabilities, i.e. they can be connected either to UTRAN or to GERAN. The urban macrocell propagation model in [10] is considered for both systems with a shadowing with standard deviation 10 dB. The mobility model in [11] is considered with speed 3 km/h.

A mix of voice and interactive users is considered. Voice calls are generated according to a Poisson process with an average call rate of 10 calls/h/user and exponentially distributed call duration with an average of 180 s. In UTRAN, the Radio Access Bearer (RAB) for voice users is the 12.2 kb/s speech defined in [12]. The BLER target is 1%. In turn, in GERAN, voice users are allocated to one time slot in each frame. Interactive users follow the www browsing model in [11], with 5 pages per session, reading time between pages 30s, an average of 25 packets per page, and interarrival packet time 0.125s for the uplink and 0.0228s for the downlink. The average packet size is 366 bytes. A session rate of 24 sessions/h/user is assumed. This service is provided in UTRAN by means of dedicated channels (DCH) using the transport channel type switching procedure. The considered RAB assumes a maximum bit rate of 64 kb/s and 128 kb/s for UL and DL [12]. The BLER target is 10%. In GERAN, the www service is provided through a Packet Data Channel with a round robin scheduling algorithm to allocate transmissions to users sharing the same time slot. A link adaptation mechanism ois used to select the highest modulation and coding scheme (MCS) that ensures the specific sensitivity requirements.

TABLE I
UTRAN BS and UE parameters

BS parameters	**UTRAN**	**GERAN**
Maximum transmitted power	43 dBm	43 dBm
Thermal noise	-104 dBm	-117 dBm
Common Control Channels Power	33 dBm	43 dBm
Maximum DL power per user	41 dBm	N/A
Number of carriers	1	3
UE parameters	**UTRAN**	**GERAN**
Maximum transmitted power	21 dBm	33 dBm
Minimum transmitted power	-44 dBm	0 dBm
Thermal noise	-100 dBm	-113 dBm
DL Orthogonality factor	0.4	N/A
Multislot class	N/A	2 UL, 3 DL, 4 UL+DL

In the admission control in UTRAN, three conditions are checked [1], namely the uplink load factor including the new user should be below 1, the downlink transmitted power below 42 dBm and there must be code sequences available. In GERAN, voice users are accepted provided that there are available time slots, while interactive users are always accepted at session initiation in idle state. Voice users have precedence over www users, so that slots occupied by www users are allocated to incoming voice users when there are not other free slots. All slots are reversible except the slot 0 of the carrier transmitting the broadcast channel. Users are dropped after being 5 dB below sensitivity for 20s.

V. Results

This section analyses the performance obtained with the two vertical handover approaches under different conditions regarding load, traffic mix and scenario parameters. Vertical handover clearly extends the degrees of freedom in the management of the radio resources, so that it is a suitable mechanism to avoid call droppings that would occur within a single RAT due to a mobile arriving to a blocked cell. As a result of these call droppings, the total aggregated throughput is reduced when there is no VHO mechanism, as reflected in Fig. 3. In turn, when the VHO strategy is considered the dropping is reduced and the total throughput increases. It is worth mentioning that the throughput results are presented for the uplink but similar trends are also observed in the downlink.

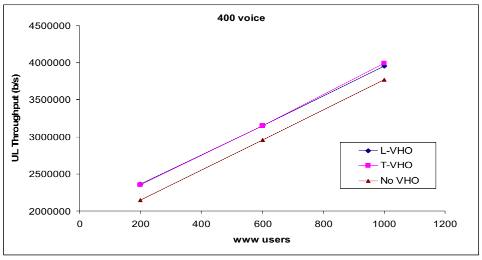

Fig. 3 Total uplink aggregated throughput

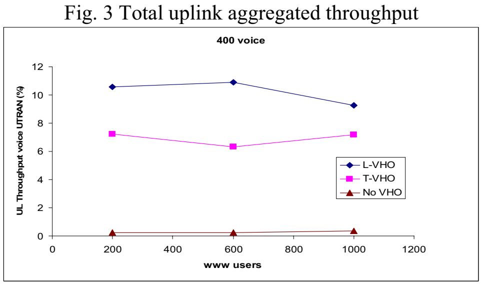

Fig. 4 Fraction of voice throughput served in UTRAN with 400 voice users

When comparing T-VHO with L-VHO, the first aspect to consider is how the traffic distribution changes in the different RATs. In particular, when 400 voice users are present in the scenario, Fig. 4 shows the percentage of voice traffic served through UTRAN as a function of the number of www users. In this case, even when no vertical handover is available, a certain portion of the voice traffic is served through UTRAN, corresponding to the new voice calls originated in blocked GERAN cells. In turn, when VHO is used, a certain number of voice users are also transferred from GERAN to UTRAN when they reach blocked cells during horizontal handovers, thus increasing the fraction of voice traffic served through UTRAN. Notice that with T-VHO there are less voice users in UTRAN than with L-VHO, because in the latter users that are handed over to UTRAN will tend to remain there, while with T-VHO a voice user in UTRAN will try to return to GERAN in the first horizontal handover.

When increasing the load up to 800 voice users, not only GERAN is most of the time saturated but even UTRAN perceives a high load. In this case, when no VHO is used, due to the lower number of voice calls that are served through UTRAN, the load is smaller than with the two vertical handover approaches, in which the uplink load factor level reaches high values, causing that several www users cannot be admitted in UTRAN and therefore are directed to GERAN at session initiation. This is reflected in Fig. 5, where the percentage of www throughput served through UTRAN is shown. Clearly, when no VHO is used, almost all the traffic is served through UTRAN. In turn, with the tight approach, also the amount of traffic served through UTRAN is more than 98%, because www users that have initiated session in GERAN will tend to return to UTRAN. On the contrary, with the loose approach www users in GERAN will tend to remain there, thus reducing the amount of www traffic in UTRAN.

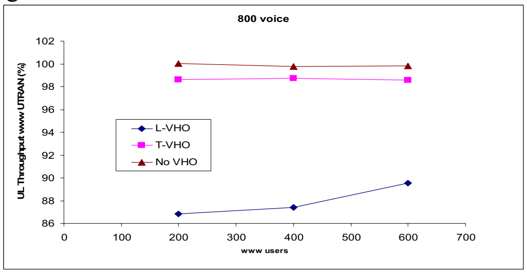

Fig. 5 Fraction of www throughput served in UTRAN with 800 voice users

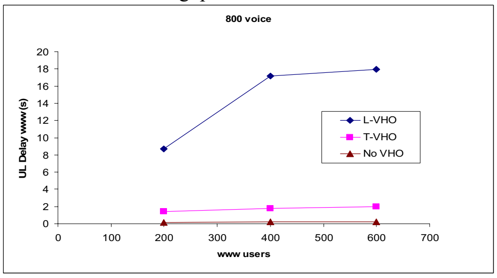

Fig. 6 Uplink packet delay for www traffic with 800 voice users

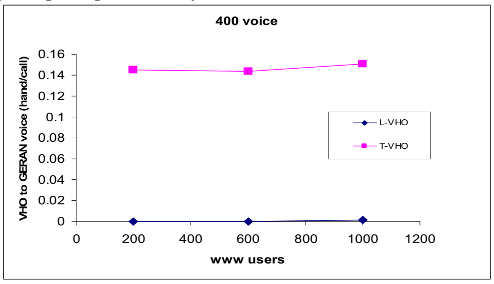

Fig. 7 Average number of VHOs from UTRAN to GERAN per voice call

As reflected in Fig. 4 and Fig. 5, T-VHO approach facilitates the fulfilment of the service RAT selection policy, consisting in serving the www traffic as much as possible through UTRAN. This has an important impact in terms of average delay for www traffic, as reflected in Fig. 6 for the 800 voice users case. The differences arise from the fact that the www delay in GERAN is significantly much higher than the delay in UTRAN, because the www users in GERAN share the few slots remaining after having allocated all the voice users. Then, in terms of the average delay considering both UTRAN and GERAN contributions, the strategies that serve most of the www traffic through UTRAN (i.e. T-VHO and no vertical handover) achieve a lower delay than L-VHO, in which more www users are served through GERAN.

On the other hand, from the point of view of signalling, T-VHO implies much more frequent vertical handover procedures, then increasing signalling load. This is shown in Fig. 7, which plots the average number of vertical handovers from UTRAN to GERAN per voice call. Clearly, T-VHO provides more chances to follow the service policy established for RAT selection, reflected in a higher number of vertical handovers to GERAN than in L-VHO, in which the voice users entering UTRAN tend to remain there until their call ends.

VI. Conclusions

This paper has focused on the vertical handover algorithms in the framework of the CRRM strategies for heterogeneous networks. The interworking between horizontal and vertical handover has been studied, with two considered approaches, denoted as the tight approach T-VHO, in which the vertical handover algorithm is executed at every time that a horizontal handover should be carried out, and the loose approach L-VHO, in which the vertical handover algorithm is executed only when a horizontal handover fails.

Acknowledgments

This work has been performed in the framework of the IST EVEREST project, which is partly funded by the European Community and by the Spanish Research Council (CICYT) under COSMOS grant TEC2004-00518 and TEC2004-0053.

References

[1] J. Pérez-Romero, O.Sallent, R.Agustí, M. Díaz-Guerra, *Radio Resource Management Strategies in UMTS*,John Wiley & Sons, 2005.

[2] A. Tölli, P. Hakalin, H. Holma, "Performance Evaluation of Common Radio Resource Management (CRRM)", ICC Conference, Vol. 5, April, 2002, pp. 3429-3433.

[3] S. Lincke-Salecker, "The Benefits of Load Sharing when Dimensioning Networks", Proceedings of the 37th Annual Simulation Symposium (ANSS'04), April, 2004.

[4] S. Lincke-Salecker, "Performance and Service Issues in Selecting Adaptive Placement as a Load Distribution Technique", IEEE 59th Vehicular Technology Conference, VTC 2004-Spring, Milan, 2004.

[5] M. Siebert, M. Schinnenburg, M. Lott, "Enhanced Measurement Procedure for Vertical Handover in Heterogeneous Wireless Systems", 14th PIMRC Conference, Beijing, 2003.

[6] 3GPP TR 25.881 v5.0.0 "Improvement of RRM across RNS and RNS/BSS"

[7] 3GPP TR 25.891 v0.3.0 "Improvement of RRM across RNS and RNS/BSS (Post Rel-5) (Release 6)"

[8] J. Pérez-Romero et al."Common Radio Resource Management: Functional Models and Implementation Requirements", 16th PIMRC Conference, Berlin, 2005.

[9] J. Pérez-Romero, O. Sallent, R. Agustí "Policy-based Initial RAT Selection algorithms in Heterogeneous Networks", 7th MWCN Conference, Marrakesh, 2005.

[10] 3GPP TR 25.942 "Radio Frequency (RF) system scenarios"

[11] UMTS 30.03 TR 101 112 "Selection procedures for the choice of radio transmission technologies of the UMTS", ETSI, April, 1998.

[12] 3GPP TS 34.108 "Common Test Environments for User Equipment (UE); conformance testing"