Handover Criteria Considerations in Future Convergent Networks

Mehdi Mani
GET-INT, France
mehdi.mani@int-evry.fr

Noël Crespi
GET-INT, France
noel.crespi@int-evry.fr

***Abstract*—Design of Handover mechanisms in converged wireless and wire-line networks where different access technologies overlap are challenging. As the first step, defining the handover criteria for selecting the best candidate among available accesses plays an important role for efficient use of network resources. Different parameters should be considered such as: QoS, load balancing factor and inter technology handover cost. For load balancing the goal is selecting an access technology for a client according to its active sessions in a manner that the set of overlapping accesses can admit the most number of clients. It does not always mean selecting the access with the most available resources. We define the handover criteria according to access characteristics, network load condition and end user preferences. We also consider distributing of different sessions belonging to one user over different access technologies as a new feature. We analyze the effect of our approach in efficient use of resources in different test scenarios.**

I.INTODUCTION

Convergence of wire-line and different wireless network technologies can bring out new services and capabilities such as unique numbering and running bandwidth consuming applications proposed in mobile domain via wire-line or WLAN accesses with better QoS and cheaper cost. In 3GPP Release 6, the work on interconnection of WLAN and 3G technologies started and the standardization is being continued in Release 7. In B3G, the network is supposed to be more heterogeneous supporting 2G, 3G, WLAN, WiMAX and Bluetooth. Moreover, for Next Generation Networks (NGN), ITU-T seeks for convergence of wireless and wire-line technologies. In all of 3G, B3G and NGN service-related functions are independent from underlying *transport-related technologies by using 3GPP IP Multimedia Subsystem*.

IMS as an overlay on transport infrastructure provide convergent services for the users. However, Seamless mobility between hybrid technologies is a key feature.

Different mobility management protocols developed in IETF like MIP, HMIP, Fast MIP and NETLMM approaches can be considered to support roaming and handover locally and globally [1-5]. However, regardless of the mobility management protocol, defining the criteria to select the best connection in handover or roaming time is critical:

IMS is based on a model where a network operator and service provider control access to the network and services for which customers are billed. Therefore the handover/roaming criteria shouldn't focus only on selecting the connection with best QoS condition and should also consider the cost. Moreover conventional handover criteria used in homogenous networks, based on Received Signal Strength-RSS, firstly, are not able to choose the best access according to other QoS requirements of active sessions (bandwidth, delay …) of a user. And secondly with such simple criteria the resources of the hybrid network won't be used efficiently.

Hence, four main goals should be referenced by mobility management systems to cope with the challenges of handover and roaming in heterogeneous networks: 1) Selecting the best access compatible to user constraints (ie cost constraints) according to the needed resources for the active services. 2) Minimizing the signalling load pushed by handover mechanism. 3) Load balancing among different network technologies. 4) Having the capability of distributing different ongoing sessions belonging to one user over different accesses (if user is accessible simultaneously via different connections). We call this *distribution of sessions*.

The majority of the research done for inter technology handover consider handover between WLAN and mobile networks trying to minimize the handover delay and defining the required signalling and security functionalities [7,8,9]. However, almost all of the works for defining handover criteria are either accomplished for 2G [10,11].

The focus of this paper is on defining the proper criteria for selecting the best access technology among the other availabilities by following these four goals.

In the rest of the paper, at first we define the require handover criteria for convergent networks. In section IV we will introduce our approach for defining the criteria. The simulation results will be presented in section VIII and finally the paper will be concluded.

II. HANDOVER TRIGGERING MOTIVATION

Diversity in access technology provides more alternative for clients to receive their requested service via the most suitable access technology. Hence, the Handover in the future telecommunication networks (B3G, NGN and 4G) may happen with different motivation:

1) QoS Problems: The service received by user is not satisfactory. It can be because of:
   a. Received Signal Degradation.
   b. Congestion and Traffic Load (Load Balancing).
2) Requested service: The service requested can't be supported in the current network because either this service is not available in the current network or it conflicts with the network admission control policies.

3) Availability of better price and QoS for the active applications: An access of other technology can provide the running service cheaper and/or with better QoS.
4) Distribution of Sessions.

All of these possible handover scenarios necessitate well defined handover mechanism capable of handling different situation. Defining Handover criteria which can select the best access among diverse technologies with different characteristics according to the active session of the user play an important role in handover mechanism.

III- HOW TO DEFINE HANDOVER CRITERIA

In 2G cellular systems, Handover is managed in data-link layer according to the physical layer information (RSS, BER) [9]. The main supported service is voice and network is homogenous, therefore: firstly, the handover criteria are very simple and secondly the needed amount of resource for a session is fixed and therefore predictable.

However, in convergent networks, according to the diverse capability of different access technologies beside of various requirements of different services (Bandwidth, delay jitter, E2E delay, security, …), different criteria should be considered for handover mechanisms.

These criteria can be divided into four main categories:

Access network information:

Access parameters are, Downstream available bandwidth, upstream available bandwidth, Link quality condition (SNR, PER, retransmission rate), security level and access cost.

User Preferences:

User preferences include, expected QoS level for the service, preferred access point, the priorities that he/she defines for each service etc.

Terminal Capabilities:

User terminal capabilities indicate if it is a multi-mode terminal or not. If yes, which access technologies it supports. The size of the screen, battery life, etc. can be considered as terminal capabilities too.

Service Type:

Each service including VoIP, Video Phoney, Conferencing, web, email, File transferring applications…needs different QoS and Security support. Therefore, according to the active sessions of the user, the best candidate access should be selected in Handover time.

In [6], a cost function is suggested to select the best access candidate for handover. In the time of handover (when received signal level degrade), this cost function is calculated for available candidate accesses. Finally, the access with the least cost function will be considered as new access. The cost function is formulated as below:

$$f = \sum_a [(w_{c,a}.\ln C_{n,a} + w_{b,a}.\ln \frac{1}{B_n} + w_{p,a}.\ln P_n).P_a]$$

$$\sum_i w_{i,a} = 1$$

Cn is the charging rate of access n. Bn indicates the bandwidth and Pn stands for power consumption in the access n.

The weights ($w_{i,a}$) for each parameters will be defined per network per application according to the importance of the parameter. For example if charging rate is very important to the user, he/she gives the highest value to the w_c.

For each active application in the decision time the cost is calculated and then multiplied by the application priority (P_a); then a sum across all applications, will give the cost function.

However, the load balancing, distribution of sessions and vertical handover cost (because of huge amount of signalling exchange) is not considered in this function.

These issues are addressed in continue in our approach.

IV- PROPOSED HANDOVER CRITERIA FOR CONVERGENT NETWORKS

In our approach, in the first step, similar to [6] a cost function is calculated based on enhanced parameters of access network *n* with considering user preferences and user terminal capabilities for each active application ($UACF_{n,a}$). Then the total cost function for all applications will be calculated ($UACF_n$):

$$UACF_{n,a} = \sum_j w_{j,a} f_{j,n} \quad (1)$$

$$UACF_n = \sum_a UACF_{n,a}.P_a \quad (2)$$

Access parameters, $f_{j,n}$ include: Downstream Bandwidth (DBw), Upstream Bandwidth (UBw), Power consumption and Charging Rate.

Different operators may define different billing strategies for their clients including: Flat rate (ie 20$/month), per minute charging or per Kbyte charging. Hence, different billing strategies should be translated to a comparable charging rate factor. Moreover, the total amount that a client will be charged is not completely independent of QoS. It means that it is likely that in a more expensive access with better QoS, the session can be finished faster and then leads to cheaper cost.

In Table 1, we have shown how to calculate the charging rate factor in different billing scenario for each service. We have considered Retransmission Factor (RTF) in these calculations. RTF shows the retransmission rate of packets because of packet error in the network. Then it is a QoS factor.

In the next phase, we will give priority to intra technology handover by adding the additional cost as follow:

$$VCF_n = UACF_n + h_{VHO,n} \quad (3)$$

If network technology of candidate access is different from the current access, $h_{VHO,n}$ adds extra cost.

$$h_{VHO,n} = \begin{cases} c_v \ln(N) & If\ Vertical\ HO \\ 0 & else \end{cases} \quad (4)$$

N= Number of required signalling exchange in Vertical handover.

Finally, Load Balancing Factor (LBF) is added in order to distribute the traffic over different accesses in a manner that the total resources of the overlaying accesses be used the most efficiently. According to its importance, we explain our approach for calculation this factor in a dedicated section.

$$LCF_n = UACF_n + LBF_n + h_{VHO,n} \quad (5)$$

Another aspect we are looking for is the effect of distribution of sessions on overall load of the network. Distribution of sessions is the capability of allocating the most matched access technology to each active application of a user. Hence, the resources of the network will be utilized more efficiently. In this approach, the decision to select the best access candidate will be made for each active application of a user separately. Then the cost function will be as follow:

$$LCF_{n,a} = UACF_{n,a} + LBF_n + h_{VHO,n} \quad (6)$$

V. LOAD BALANCING

The simplest criterion for load balancing is the "*available bandwidth*" in each network. This criterion may be helpful in homogenous networks with the same capabilities but it is not efficient in hybrid technologies. In fact, with considering this criterion, the access with the most idle resources and bandwidth will be selected [12]. However the efficient way to use all resources of different network is to classify the access networks for different applications. For example, allocating accesses with lower bandwidth to the voice and considering accesses with higher bandwidth to the bandwidth consuming applications.

To reach this end, we define *resource space* for each access network. In this space, each axis represents allocated resources to a service class. Then the total amount of used resources in this access is represented as a point in this space: $X=(x_1,\ldots,x_a,\ldots)$. For instance, Figure 1 shows a space of allocated resources with three service classes which consume respectively: 16, 32 and 128 Kbps for an access with maximum bandwidth of 2Mbps.

Then we define a Goal Point (GP) in each space.

Defining the $GP= (gp_1,\ldots,gp_a,\ldots)$ is an optimization problem and is out of the scope of this paper. However, very simply we can say that two main parameters indicate the GP:

1) The minimum guarantied resources for each service class in the access.
2) The overall available capacity for each service class in the set of overlapping accesses.

So the best choice for GP is a combination of different services to maximize the overall capacity of all overlapping accesses and satisfying the minimum guarantied services in each access. With this definition, in the time of handover the distance between the point X, indicating the consumed resources (if handover request be accepted), and the GP can be a good factor for load balancing (formula 7): The more this distance, the more worthy the access network. However, in some situations it is not possible to reach the GP. For example when the maximum number of a service class indicated in GP has already completed in an access; and again a request of admission for this service arrives. If there is enough bandwidth (predicted for other service classes) and the user has no other access possibility, then the request will be accepted. In this condition a new GP should be defined according to the new condition. We call that Shifted GP (SGP). The distance between SGP and ideal GP can be also a criterion for load balancing.

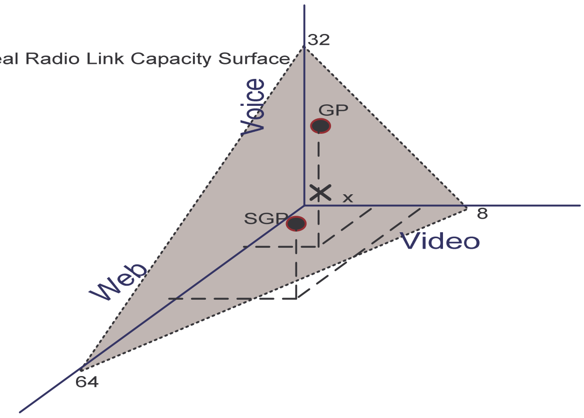

Figure 1. Ideal Radio Link Capacity Surface

According to these consideration the Load Balancing Factor (LBF) for access n, is formulated as below:

$$LBF_n = \begin{cases} c1_L\left(\dfrac{d(X_n, GP_n)}{\sqrt{\sum_a gp_a^2}} - 1\right) & \text{if } GP \text{ is not shifted} \\ c2_L(d(SGP_n, GP_n)) & \text{else} \end{cases} \quad (7)$$

$$a \in \{user \quad active \quad services\}$$

In this formula, when GP is not shifted, LBF will be negative and reduce the cost function. However, when GP is shifted the cost function amount will be increased.

VI. HANDOVER PROCEDURE

Without distribution of sessions:

After triggering the handover procedure because of the four motivations declared in section II, the handover procedure will be as follow:

1. Creating the set of the candidate accesses:
   Choosing the accesses whose RSS is more than predefined threshold.
2. Classifying the candidate accesses by the LCF_n.
3. If there are some LCF_n< Cost Threshold:
   The best choice is the one with the least LCF_n.
4. If there is no LCF_n< Cost Threshold:
   If: ∃ *an active application* ∈ *{Platinum Priority} Then*
   i. Choose the access with the least LCF_n for the session(s) with Platinum Priority.
   ii. Terminate all other sessions
   Else:
   Terminate all the sessions.

In the case of distribution of the sessions, the decision is made for each application of a user separately. The handover procedure will be exactly the same with replacing LCF_n with $LCF_{n,a}$.

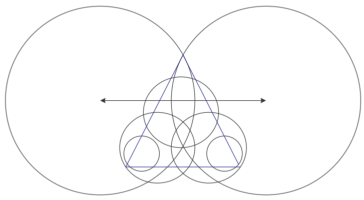

Figure 2 . Test-bed Overlapping Accesses

VII. TEST-BED AND SIMULATION SCENARIO

Our simulation is taken place in a three-overlapping-access test bed as depicted in figure 2. Three kinds of services including voice, video and text_chat are considered. The services (S_i) arrive respecting Poisson random process with the rate of λ_i. Duration of each service is modelled by exponential distribution with the mean of μ_i. Table 2 and 3 represent the parameters of the model. Arriving calls in each cell type will be distributed uniformly in that cell type. In triangular region of figure 2, we consider arrival of handover calls with the rate of $\lambda_{handover}$. Handover arrivals with the probability of .5, .25 and .25 will be respectively voice, video and text_chat. Moreover, RSS is simulated by the distance.

VIII. SIMULATION RESULTS

Figures 3-6 represent the simulation results. In the figure 3, Resource Utilization Factor (RUF) is demonstrated versus network requested load and the handover rate.

RUF is equal to the allocated resources over requested resources. As depicted in this figure when considering Distribution of the sessions with $LCF_{n,a}$ as the handover criteria ($LCF_{n,a}$ DS), the RUF remains always upper than 90%. It means that the resources are allocated efficiently. However for the case that only handover criterion is RSS, RUF falls down to 40%.

Figure 4 shows the 1-CDF of vertical handover ratio for different handover criteria. Vertical Handover Ratio is the proportion of vertical handovers to the total handovers.

In handover with only RSS criterion, access candidate will be firstly sought among the same technology. If there were no satisfactory candidate, other technologies will be considered too. Therefore, the Vertical handover ratio in this approach remains the least. However considering ($LCF_{n,a}$ DS) increase the Vertical Handover ratio; the probability that vertical handover ratio is more than 15% is 0.5 in comparison to the .18 of RSS criteria.

However, the rate of blocked handover with RSS criteria is the highest because of non-efficient traffic distribution (Figure 5) despite of ($LCF_{n,a}$ DS) which has least blocked handover ratio.

To calculate the satisfaction level of users with different handover criteria, a metric named Satisfaction Ratio (SR) is defined as follow:

$$SR = \frac{1}{A}\sum_{a}\left(\frac{UCF_{c,a}}{UCF_{\min,a}} P_a\right) \quad (8)$$

A= Number of active applications of user
a= active service
c= chosen access
min= access with minimum $UCF_{n,a}$ among access candidates.

TABLE 1. UNIFYING CHARGING FACTORS

Billing Method	T1 $/min	T2 $/kb	T3 Flat Rate ($/month)
Uniformed Charging Factor	T1(1+RTF)	T1(1+RTF)*60*application bit rate	T3*μi/30/24/60

TABLE 2. SIMULATION PARAPETERS

M	20min	3min	30min
Packet Size	200 Bytes	80 Bytes	64 Bytes

TABLE 3. SIMULATION PARAPETERS

	WLAN (PicoCell)	WLAN (MicroCell)	MacroCell
Coverage Diameter	30	100	800
Channel Bw	4Mbs	4Mbps	2Mbps
GP (Video,Voice,Text Chat)	(10,12,28)	(6,30,20)	(2,20,8)
PER	10^{-2}	$4*10^{-2}$	10^{-1}
Voice Charging Rate	9 Cents/min	12 Cents/min	10Cents/Min
Video Charging Rate	0.008 Cents/Kb	40Cents/Min	45Cents/Min
Text Chat Charging Rate	0.001 Cents/Kb	0.001 Cents/Kb	0.0015 Cents/Kb
λ_{video} calls/Second	.5	.75	1.5
λ_{voice}	4	5	7
λ_{Text_Chat}	3	4	3

Figure 6-a, shows the mean SR for the admitted users during the simulation time. SR doesn't change considerably versus load with RSS criteria and it is always around 60%. VCF has the best Mean SR. It is because the accesses are chosen according to the cost function regardless of load balancing. However when considering the load balancing (LCF), the performance degrade specially in high load. Distribution of sessions improves the performance but it remains still behind VCF.

Mean SR doesn't consider the number of admitted users. In other words, it doesn't consider how efficiently the overlaying resources are used. To cope with this shortage, in figure 6-b, we use the multiplication of Resource Utilization Factor (RUF) and Mean SR. Considering this metric, the performance of mobility management mechanism with RSS and VCF decrease dramatically with augmentation of load. However LCF and ($LCF_{n,a}$ DS) show steady performance.

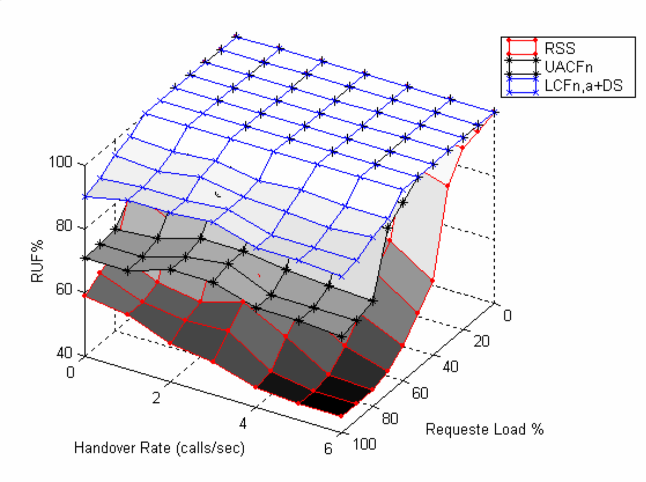

Figure 3. Resource Utilization Factor

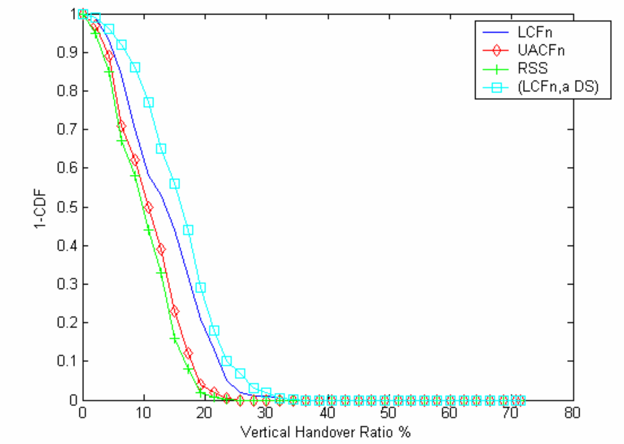

Figure 4. 1-CDF of Vertical Handover Ratio

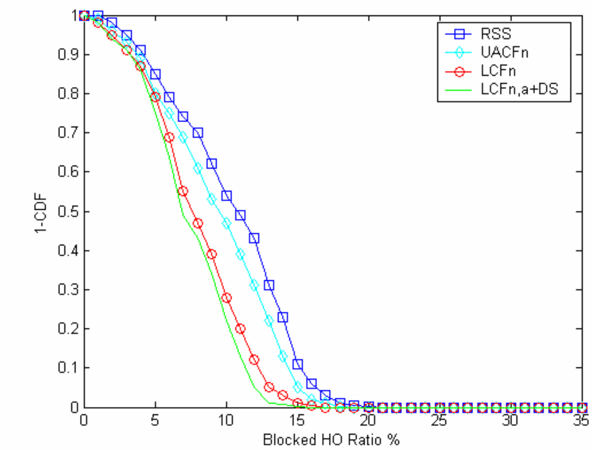

Figure 5. 1-CDF of Blocked HO Ratio

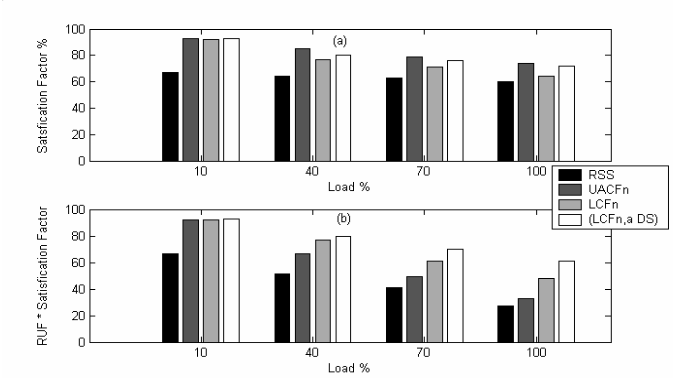

Figure 6. Mean Satisfaction Factor

IX. CONCLUSION

In this paper, we have introduced new handover criteria considering user preferences, service characteristics and radio link characteristics (Cost, Bandwidth, PER). We have defined a cost function for each access candidate. Load Balancing and Vertical Handover extra signalling load is also considered in our approach as decision criteria. Moreover, as a new feature we have considered distribution of sessions belonging to one user on different access technologies in the time of handover. The performance of our cost function in improving handover decision is evaluated by resource utilization factor, Satisfaction factor, Blocked Handover Ratio and Vertical Handover Ratio. Distribution of Sessions can lead to 40% improvement in resource utilization factor.

REFRENCES

[1]IETF, RFC 3344, IP Mobility Support for IPv4.
[2]IETF RFC 3775, Mobility Support in IPv6, June 2004.
[3]IETF, RFC 4140, Hierarchical Mobile IPv6 Mobility Management (HMIPv6), August 2005.
[4]IETF, RFC 4068, Fast Handovers for Mobile IPv6, July 2005.
[5] http://vesuvio.ipv6.cselt.it/giaretta/netlmm/
[6] Helen J. Wang, Randy H. Katz, and Jochen Giese, "Policy-Enabled Handoffs Across Heterogeneous Wireless Networks", Proc. IEEE Workshop, Mobile Computation Systems and Applications, Feb 1999.
[7] Hyeyeon Kwon, Kwang-Ryul Jung, Aesoon Park, Jae-Cheol Ryou; "Consideration of UMTS-WLAN Seamless Handover", _Seventh IEEE International Symposium on Multimedia, Dec. 2005
[8] HyangDuck Cho, Wooshik Kim; "seamless handover between cdma2000 and WLAN for 3G-WLAN interworking service continuity", Advanced Industrial Conference on Telecommunications/Service Assurance with Partial and Intermittent Resources Conference/ELearning on Telecommunications Workshop, July 2005
[9] Li, M., Kumbesan Sandrasegaran, Huang, X.; "Identity management in vertical handovers for UMTS-WLAN networks", International Conference on Mobile Business, ICMB 2005.
[10] Ivanov, K.; Spring, G.; "Mobile speed sensitive handover in a mixed cell environment", IEEE 45th Vehicular Technology Conference, 1995
[11] Murase, A.; Symington, I.; Green, E.; "Handover criterion for macro and microcellular systems", 41st IEEE Vehicular Technology Conference, 1991.
[12] Yanmaz, E.; Tonguz, O.K.; "Handover performance of dynamic load balancing schemes in cellular networks", 10th IEEE Symposium on Computers and Communications, ISCC 2005.