SEAMLESS MOBILITY IN HETEROGENEOUS WIRELESS NETWORKS

Faouzi Zarai[1], Ikram Smaoui[1], jean-Marie Bonnin[2], Lotfi Kamoun[1]

[1] LETI Laboratory, University of Sfax, Tunisia
[2] Telecom Bretagne CS 17607 35576, Cesson Sevigne, France
E-mail addresses: faouzi.zarai@isecs.rnu.tn

ABSTRACT:

In next generation of wireless networks, different technologies belonging to one or more operators should be integrated to form a heterogeneous environment based on an IP core network infrastructure. This ensures user mobility and service continuity by maintaining connections when switching between various technologies and it introduces new resources and possibilities for applications. In this context, an automatic interface selection based on instantaneous and practical constraints and user preferences (Quality of Service (QoS) parameters, available resources, security, power consumption, etc) is therefore required. The different network selection and handover schemes proposed in the literature can be classified into three approaches according to who is responsible for making the handover decision: the terminal, the network or thanks to a cooperation between both of them. However, these approaches keep presenting some drawbacks; namely the problem of resources management and network load balancing whenever the selection is controlled by the mobile terminal (MT) and the problem of scalability and unknown operator's management policy whenever the selection is rather controlled by the network.
In this article, first we propose a new architecture and new network selection scheme that explicitly take into account the current resource usage and the user preferences. Furthermore, our solution ensures the selection of the most suitable network for each flow while taking into consideration its expectations in terms of QoS. A feasibility study of our architecture is then triggered on a single MT by using typical scenarios and using various algorithms to evaluate their performances.

INDEX TERMS: *Heterogeneous wireless networks, 3GPP LTE, Best network selection, Qos Parameters, CanuMobiSim.*

I. INTRODUCTION

In the recent years, the democratization of wireless networks such as WCDMA/HSPA, LTE, WiMAX, and WLAN has encouraged the emergency of many applications (VoIP, video on demand, web applications…) that take advantages of the mobility. However, a single technology can hardly satisfy all the applications expectations (required delay, bandwidth, Bit Error Rate, security level…). Therefore, the integration of different wireless technologies in a heterogeneous environment has offered best opportunity for applications to be well delivered. However, in several situations, mobile terminals tend to associate with networks guaranteeing the best performances to stay "Always Best Connected" which leads to overload the most attractive technology while keeping the others technologies underutilized. Then, in order to overcome this

problem and to benefit from technology diversity, an automatic interface selection based on instantaneous and practical constraints and user preferences as well as operator and network resources management constraints has become an inevitable trend.

In reality, several actors (such as the user, the terminal, the applications,…) affect the best network selection decision [1, 2]; each one tends to influence this decision to satisfy its benefits. Moreover, the networks heterogeneity with different characteristics and mobility management policies constitutes a technical constraint that makes the selection decision more complex. In this context, several architectures and schemes using multi criteria decision making approaches have been proposed in the literature [1, 3-7]. These approaches have treated principally three approaches: a decision taken by the MT [5], by the access network [6] or a cooperative solution taken by both of them [7].

Mechanism proposed in [1] is based on a unique decision process that uses non-compensatory and compensatory multiattribute decision making jointly on the network side to assist the terminal to select the top candidate network. The mechanism proposed in [3] optimizes the network selection process through minimizing the average power consumption cost and the average user dissatisfaction and models the problem of multi constraint access selection as a variant of bin packing problem which is a too complex solution. Moreover, the minimization of the power consumption factor and the user dissatisfaction alone are not able to present the whole network performance.

In [5], the authors propose an intelligent access selection mechanism which considers the user's preferences, the network conditions and the applications requirement. However, there are still some limitations to these works due to the fact that these mechanisms don't take into consideration the terminal performances such as the battery life time, the memory capacity and the CPU.

The MT Selection decision is a much solicited alternative due to the fact that the selected network is the best one that satisfies the selection criteria from the MT's point of view without operator intervention that aims fulfilling its own needs. However, if we aim to ensure load balancing between the different networks, the MT should have a global vision about the network resources management which it is generally not possible due to operators' security concerns that inhibit the diffusion of such operational information. In the second approach, which it is based on network or operator decision making, the problem of network resources management is resolved thanks to the availability of network information such as network current load, operator policies, network conditions and capacities, etc. On the other hand, we will be faced the problem of operator's profit [8] that tends always to select the most beneficial network for him from monetary cost criteria or resources control by applying his proper policy independently of the user or applications expectations. Moreover, we find the constraint of MT context transfer to take into considerations the battery status and the terminal memory capacity that it is no scalable and leads to excessive overhead exchanges in the network. In the last strategy, MT and networks tend to cooperate to find the best network satisfying user and applications requirements and resolve the limits of previous approaches. However, some complexity and scalability concerns still persist.

In this article, we present a network selection architecture and scheme that provide a resource-efficient mobility management that aim at selecting the most suitable network interface for each application. Our proposed architecture intends to resolve the limitations of previous schemes by satisfying user preferences while guaranteeing a best network resources management. We focus on the case of an operator having several access network technologies and we propose an architecture able to share the load among the different available technologies in order to satisfy as much as possible applications requirements and user expectations. Then, we evaluate the feasibility of our approach by studying mobile terminals behaviours with various multi criteria decision making algorithms.

This article is organized as follows. In section II, we describe the details of our proposed network selection architecture. The selection scheme is presented in the section III. Finally, we develop various simulations that highlight the contribution developed in the previous sections. A conclusion is provided in section IV, where some perspectives are depicted.

II. THE PROPOSED ARCHITECTURE FOR BEST NETWORK SELECTION

The wireless network landscape is changing gradually from homogeneous to heterogeneous and future generation networks will be characterized the coexistence of a large variety of Radio Access Networks (RAN) (like GPRS, WCDMA/HSPA, LTE, WiMAX, and WLAN), with different protocol stacks and supporting a number of applications and services with different QoS requirements to be provisioned to terminals with different degree of multi-mode capabilities to access the available networks [9, 11]. Each mobile station and radio access network characterized by the specific air interface technology, cell size, multiple access scheme, coverage, mobility type, etc [12].
This section presents the proposed next generation wireless network architecture, especially focuses on the best network selection architecture and the QoS mapping mechanism in heterogeneous wireless network environments.

A. System Architecture

Our proposed architecture, illustrated in Figure 3, supposes that MT is responsible for network selection and handoff decisions. The RAN influences these decisions by performing data collection and analysis.The core of the next generation infrastructure is expected to be the IP based multi-service network that provides connectivity and transport thorough RAN, including legacy 3G, 3GPP LTE, WiMax, WLAN, and emerging technology. The multi-access infrastructure supports services and users having a wide variety of multi-access capable terminals. The proposed architecture introduces new nodes that may be responsible for coordinating and managing the radio resources between access networks. These nodes are described later in this section, as well as it also adds some new functionalities to the mobile terminal. The proposed architecture as shown in the figure does not include all the logical functional nodes that would be present in a heterogeneous network environment. Only those nodes that are related to the network

selection process have been presented. Other existing functional nodes, e.g., security, are outside the scope of this article.
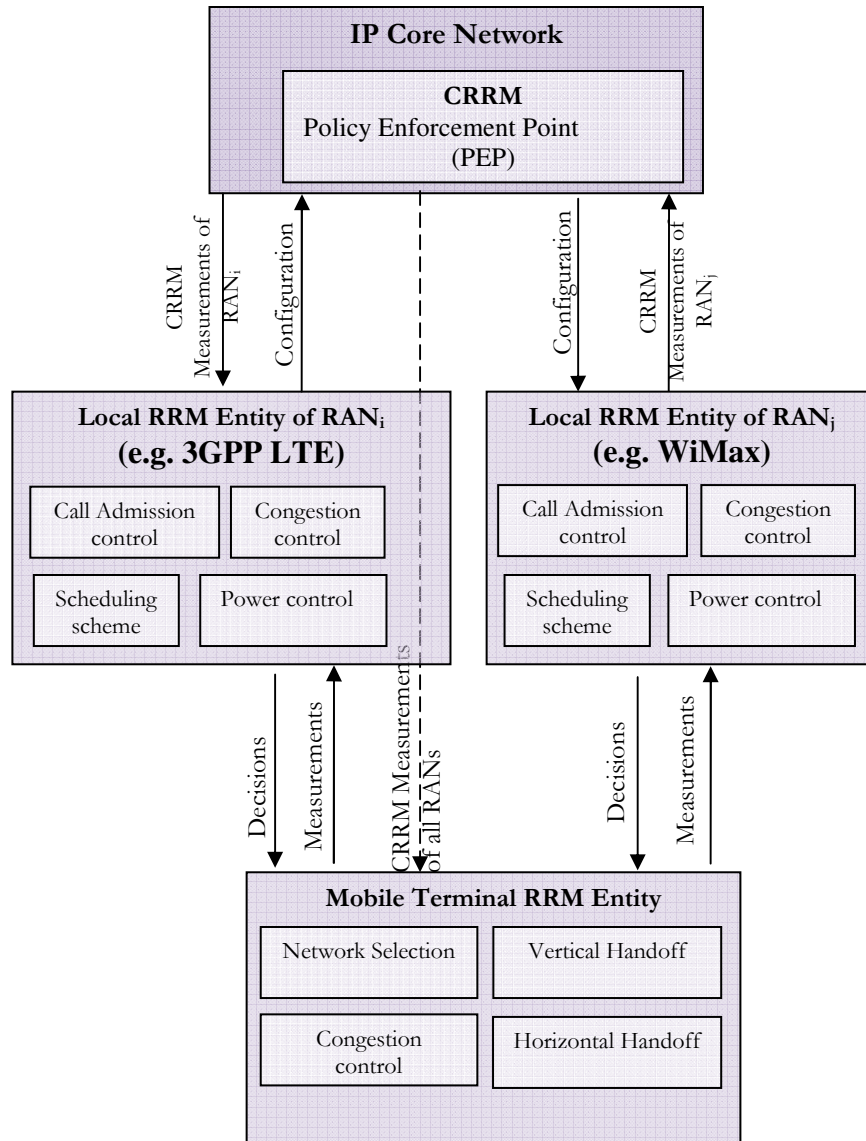

Figure 1: The proposed architecture

In our approach, the Common Radio Resources Management (CRRM) is considered simply as a Policy Enforcement Point (PEP) that translates the specific policies into an adequate configuration of the RRM algorithms. Notice that almost all functionalities reside in the local RRM entity, which is responsible of the call admission at the beginning of a session and the exchange of candidates RANs capabilities information to execute an vertical handoff between

different RANs in the middle of a session, taking into account the intra and intersystem measurements provided by the mobile terminals as well as the cell measurements from other RRM entities. The CRRM estimates available resources in different RANs in the near future time and sends it to the mobile terminal entity through local RRM entity of its RAN.

The mobile terminal is responsible of the decision to execute an intersystem or vertical handover between different RATs in the middle of a session.

The local RRM entity is the main responsible of the undertaken decisions. As an example, a policy could be defined as "provide better QoS to vertical handoff than to new call". The PEP at the CRRM entity would, in this case, configure some parameters at the local RRM, whose algorithms would be fully responsible for the management of the air interface. For example, the scheduling algorithm in LTE-RAN (E-UTRAN) would be configured in such a way that higher priority would be given to vertical handoff.

B. Common Radio Resources Management (CRRM)

The basic idea behind proposed architecture is based on resources management to accommodate more calls while satisfying at the same time applications expectations and load balancing between the different networks. Thus, our proposed resource management approach has been designed on the basis of following foundations:

- First of all, it is assumed that reliable historical information about different types of session duration (call-duration) is available. Therefore, by knowing the call time instant initiation, an approximation of its time instant termination can be done and an estimation of the released resources becomes possible,

- Secondly, we assume that not all of the resources reserved by an ongoing or incoming session are fully utilized,

- Thirdly, a slight degradation of agreed QoS parameters is acceptable for most of the applications, in case expected resources are not freed.

On the basis of above assumptions, our CRRM design is based on following network entities:

- Call Duration Statistical Module (CDSM): is responsible for controlling and managing call/session duration and elapsed time. For example, in voice session it has a common value of 120 seconds mean, but for other type of non-real time services like streaming services, this can be estimated through expected data size. In case of real time stream, a time unit may be defined for the session. Therefore, in our proposed model, the CDSM is responsible for dynamically updating values of call duration based on the historical and current data.

- Resource Estimation Module (REM) is designed to estimate available resources in the near future time (within milliseconds). This is an important component of proposed scheme as its efficiency can ensure better call admission in local RRM entity depending on the accurate future resource availability.

-   Resource Collection Module (RCM) is responsible for collecting information of network capabilities (e.g. total bandwidth, bit error rate, etc) and network conditions (e.g. network utilization and traffic load) using a standard set of parameters so that these parameters have the same interpretation across different networks using different access technologies. Moreover, the RCM is responsible for monitoring fast recovery of resources on their release from allocated session. This operation generally has a few seconds latency in most cellular networks and consequently fast recovery considerably enhances chances to accommodate incoming calls. Since our scheme uses anticipatory resources collection, it reduces such latencies greatly.

-   Resources Reservation Module (RRM) is responsible for reserving resources to new incoming and handed-over calls. This module interacts with other modules to determine maximum available resources at a given time and accordingly allocate resources to the incoming calls with a purpose of accommodating as most as possible maximum number of calls.

The main reason for using above mentioned approach is to guide the MT in making decision about the network which it is better to select on the basis of current information set.

Therefore, once the user terminal has a service flow for which it needs to choose the best network for its delivery, it requests and gets information about current network conditions and available resources from the RCM located within the access network which interacts with the REM and the CDSM, as it is depicted in Figure 2, to provide reliable information. Afterwards, if the user is interested on this network, it asks the RRM to pass a resources reservation request. The RRM interacts with the REM every time it receives a new resources reservation request to have an overview on the available resources in the network at a given time or whenever some reserved resources are released, the RRM asks the REM to update resources status.
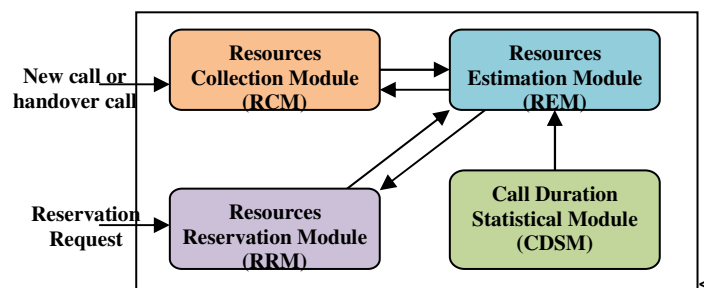

Figure 2 : Resources Reservation and Management Model

Mobile Terminal architecture and Behaviour

In our proposed architecture, similarly to some related works in the literature [5, 13], we have supposed that each MT includes a Network Selection Decision Mechanism (NSDM) that involves three modules:

Profile Database: it maintains all the required information to assist the selection decision algorithm when it makes the best network selection decisions. The following data are stored in the database:

- Data related to user preferences and operator constraints such as preferred and forbidden access networks, the weights affected to different parameters participating in the selection decisions as policies. Therefore, we specify for each application which is the more important goal by providing the suitable values for the selection goal parameters.

- Data related to applications QoS requirements. It contains mainly the QoS level required by each application. For example, the useful parameters from the application QoS requirements could be: Minimum necessary bit rate (kb/s), supported bit error rate, required security level and maximum tolerated delay.

- Data related to the available networks performances such as the mean bit rate, the maximum packet size, the packet error rate, the bit error rate, and the average latency to send a packet.

Profile Manager (PM): it supervises all the entities influencing the best network selection decision (network, user, application, and terminal) and stores the necessary information in the profile database. The PM also determines when it is necessary to trigger the Selection Decision Algorithm and assists the SDA when it makes the choice of the "best" access.

Indeed, the PM triggers the SDA in the following cases:

A modification of network interface status,

An application flow has been created or deleted,

Flow monitored parameters values modification,

User preferences or operator constraints change,

Network performances modification.

Moreover, the PM can make the automatic selection of an access network by maintaining all the necessary information for proper interface configuration.

Selection Decision Algorithm (SDA): The operating principle of this scheme is described in section IV.

III. PROPOSED SELECTION DECISION ALGORITHM (SDA)

In heterogeneous environments, criteria to selection the best network is one of the main challenges for seamless mobility as there does not exist a single factor than can provide a clear idea of which to select. Some of the most important decision factors are:
- User satisfaction degree,
- Offered Bandwidth,
- Velocity,

- Signal Strength,
- Interference (Load balancing),
- Power Requirements.

Each network selection strategies treated in the literature have their own advantages; they are all designed to meet individual mobile user's needs in bandwidth, reliability, cost or power conservation. They did not pay much attention on the system performance, such as the blocking probability of originating calls and the dropping probabilities of horizontal and vertical handoff calls. In this section, we propose a multicriteria dynamic access selection strategy which can be stated as follows. We consider a set of access networks available and incoming traffic flows from mobile terminals that need to be assigned to access networks while satisfying the following objectives:

- Maximizing the offered bandwidth from the selected network N_j,
- Accommodate the user's preferences such as velocity support, the connections requirements such as the RSS level and the system performances in terms of reducing new blocking and handoff dropping probabilities,
- Establishing a priority mechanism for handoff calls over new calls for each class of traffic,
- Achieving a load balancing among available Radio Access Technologies (RAT) to improve average system utilization,
- Minimizing the power consumption on MT.

A. System Model

In this section, we model the best network selection problem in order to find the optimal selected network that satisfies the concept of Always-Best-Connected.

1. Application Description

Let $A = (A_1, A_2, A_3 ... A_n)$ be the set of new applications for which access networks have to be determined.

A QoS Request (QR) should be generated by the MT before detection a new access network. The QoS request contains the call type, service class, and bandwidth and delay requirements.

The QoS Request of the application A_i, $QR(A_i)$ is described using a 6-tuple:

$(P_i, B_i^{min}, B_i^{req}, B_i^{max}, D_i^{req}, D_i^{max})$ where :

B_i^{req} : is the required bandwidth capacity of the application A_i,

B_i^{min} : is the minimum bandwidth capacity of the application A_i,

B_i^{max} : is the minimum bandwidth capacity of the application A_i,

P_i : is the service class of the application A_i,

D_i^{req} is the delay requirements of the application A_i,

D_i^{max} : is the delay maximum of the application A_i.

When the system is underutilized, all arriving new and handoff calls are admitted with the highest bandwidth level B_i^{max} for the calls. This approach increases bandwidth utilization for the heterogeneous wireless network. However, when the system is fully utilized, bandwidth adaptation controller is invoked.

It is assumed that each call is assigned a priority from one of $(P_1, P_2, P_3 ... P_K)$, where $P_i < P_{i+1}$ and the larger index number indicates the higher priority. This relative priority of each call is mapping with 3GPP QoS definition consistently. For example, P_4 indicates the conversation service class, P_3 indicates the steaming service class, P_2 indicates the interactive service class, and P_1 indicates the background service class. The conversation service class will have higher priority than the steaming service class, the streaming service class will have higher priority than the interactive service class, and the interactive service class will have a higher priority than the background service class.

2. Access Network Description

Let:

- $N = (N_1, N_2, N_3 ... N_m)$ be the set of access networks available.

- $V = (V_1, V_2, V_3 ... V_m)$ be the set of threshold values of velocities for a mobile station for the respective networks.

- $RSS = (RSS_1, RSS_2, RSS_3 ... RSS_m)$ be the set of threshold values of received signal strengths of respective networks.

- $C = (C_1, C_2, C_3 ... C_m)$ be the communication costs of one unit time of respective networks.

- B_j is the total bandwidth capacity of access network N_j,

- D_j is the maximum communication delay of A_i

- PB_j, PT_j, PR_j are power consumption cost parameters defined below.

- $O_{(j,t)}$ denotes the number of ongoing call in a cell of access network N_j at time t.

- $X_{(j,t)}$ denotes the number of calls in a cell of access network N_j at time t.

- $n_{(j,t)}$ and $h_{(j,t)}$ denote respectively, the number of ongoing new calls and handoff calls, in N_j.

- $B_{(i,j,t)}^{asig}$ denotes the assigned bandwidth by network N_j for call i at time t.

- $D_{(i,j,t)}^{off}$ denotes the offered delay to call i by access network N_j at time t.

- I (state) denote the condition expression as below.

$$\begin{cases} I(\text{condition expression}) = 1, if \ \text{expression} = true \\ I(\text{condition expression}) = 0, if \ \text{expression} = false \end{cases}$$

An access network N_j $(1 \leq J \leq m)$ is described using a 5-tuple $(B_j, D_j, PB_j, PT_j, PR_j)$.

The assignment function $A\text{sgn\_}Net$ maps a application A_i to access network N_j, i.e. $A\text{sgn\_}Net(A_i) = N_j$ if A_i is assigned to N_j.

A call i is degraded if $B_{(i,j,t)}^{assg} < B_i^{req}$ whereas the call is upgraded if $B_{(i,j,t)}^{assg} > B_i^{req}$.

The sum of occupied bandwidth of the cell N_j at time t.

$$B_{(j,t)}^{occp} = \sum_{i=1}^{O_{(j,t)}} B_{(i,j,t)}^{assg} \tag{1}$$

$B_{(j,t)}^{avai}$: The available bandwidths in network N_j at time t.

$$B_{(j,t)}^{avai} = B_j - B_{(j,t)}^{occp} \tag{2}$$

3. Handoff decision metric calculation

The handoff decision metric calculation is performed among candidate's networks, in order to avoid measures at the mobile terminal side, each N_j computes the network quality value, then this value is sent to the mobile terminal.

Bandwidth degraded call ratio (BDCR)

We define the bandwidth degraded call ratio of ongoing calls in access network N_j as:

$$BDCR_{(j,t)}^o = \sum_{i=1}^{O_{(j,t)}} \frac{(B_i^{req} - B_{(i,j,t)}^{assg})}{B_i^{req}} I(B_{(i,j,t)}^{assg} < B_i^{req}) \tag{3}$$

We define the bandwidth degraded call ratio of new calls in access network N_j as:

$$BDCR_{(j,t)}^n = \sum_{i=1}^{n_{(j,t)}} \frac{(B_i^{req} - B_{(i,j,t)}^{assg})}{B_i^{req}} I(B_{(i,j,t)}^{assg} < B_i^{req}) \tag{4}$$

We define the bandwidth degraded call ratio of handoff call j as:

$$BDCR_{(j,t)}^{h} = \sum_{i=1}^{h_{(j,t)}} \frac{(B_i^{req} - B_{(i,j,t)}^{assg})}{B_i^{req}} I(B_{(i,j,t)}^{assg} < B_i^{req}) \qquad (5)$$

Finally, we define the weighted bandwidth degraded call ratio as:

$$BDCR_{(j,t)} = \alpha_1 BDCR_{(j,t)}^{o} + \alpha_2 BDCR_{(j,t)}^{h} + \alpha_3 BDCR_{(j,t)}^{n} \qquad (6)$$

The network operator assigns weights to the bandwidth degradation parameters in order to determine the parameters' levels of importance.

The sum of the weights $\alpha_1, \alpha_2 \text{ and} \alpha_3$ must be equal to 1. The parameters with high weight value are more important to the user than those with low weight value and vice versa.

Bandwidth upgraded call ratio (BUCR)

We define the bandwidth upgraded call ratio of ongoing calls in access network N_j as:

$$BUCR_{t,t}^{o} = \sum_{i=1}^{o_{(j,t)}} \frac{Max\left((B_{(i,j,t)}^{assg} - B_i^{req}), B_i^{max}\right)}{B_i^{req}} I(B_i^{req} < B_{(i,j,t)}^{assg}) \qquad (7)$$

We define the bandwidth upgraded call ratio of new calls in access network N_j as:

$$BUCR_{j,t}^{n} = \sum_{i=1}^{n_{(j,t)}} \frac{(B_{(i,j,t)}^{assg} - B_i^{req})}{B_i^{req}} I(B_i^{req} < B_{(i,j,t)}^{assg}) \qquad (8)$$

We define the bandwidth upgraded call ratio of handoff call j as:

$$BUCR_{j,t}^{h} = \sum_{i=1}^{h_{(j,t)}} \frac{(B_{(i,j,t)}^{assg} - B_i^{req})}{B_i^{req}} I(B_i^{req} < B_{(i,j,t)}^{assg}) \qquad (9)$$

Finally, we define the bandwidth upgraded incoming call ratio as:

$$BUCR_{(j,t)} = \alpha_1 BUCR_{(j,t)}^{o} + \alpha_2 BUCR_{(j,t)}^{h} + \alpha_3 BUCR_{(j,t)}^{n} \qquad (10)$$

Delay degraded call ratio (DDCR)

Similarly to BDCR and BUCR, we define:
- The delay degraded call ratio of ongoing calls in access network N_j as the delay between required delay and offered delay of existing calls at time t:

$$DDCR_{(j,t)}^{o} = \sum_{i=1}^{o_{(j,t)}} \frac{(D_{(i,j,t)}^{off} - D_i^{req})}{D_i^{req}} I(D_i^{req} < D_{(i,j,t)}^{off}) \qquad (11)$$

- The delay degraded call ratio of new calls in access network N_j as:

$$DDCR_{(j,t)}^{n} = \sum_{i=1}^{n_{(j,t)}} \frac{(D_{(i,j,t)}^{off} - D_i^{req})}{D_i^{req}} I(D_i^{req} < D_{(i,j,t)}^{off}) \tag{12}$$

- The delay degraded call ratio of handoff calls in access network N_j as:

$$DDCR_{(j,t)}^{h} = \sum_{i=1}^{h_{(j,t)}} \frac{(D_{(i,j,t)}^{off} - D_i^{req})}{D_i^{req}} I(D_i^{req} < D_{(i,j,t)}^{off}) \tag{13}$$

We define the d DDCR as:

$$DDCR_{(j,t)} = \alpha_1 DDCR_{(j,t)}^{o} + \alpha_2 DDCR_{(j,t)}^{h} + \alpha_3 DDCR_{(j,t)}^{n} \tag{14}$$

Delay upgraded call ratio DUCR

Similarly to BDCR and BUCR, we define:

- The delay upgraded call ratio of ongoing calls in access network N_j as:

$$DUCR_{(j,t)}^{O} = \sum_{i=1}^{o_{(j,t)}} \frac{(D_{(i,j,t)}^{off} - D_i^{req})}{D_i^{req}} I(D_{(i,j,t)}^{off} \le D_i^{req}) \tag{15}$$

- The delay degraded call ratio of new calls in access network N_j as:

$$DUCR_{(j,t)}^{n} = \sum_{i=1}^{n_{(j,t)}} \frac{(D_{(i,j,t)}^{off} - D_i^{req})}{D_i^{req}} I(D_{(i,j,t)}^{off} \le D_i^{req}) \tag{16}$$

- The delay degraded call ratio of handoff calls in access network N_j as:

$$DUCR_{(j,t)}^{h} = \sum_{i=1}^{h_{(j,t)}} \frac{(D_{(i,j,t)}^{off} - D_i^{req})}{D_i^{req}} I(D_{(i,j,t)}^{off} \le D_i^{req}) \tag{17}$$

We define the DDCR as:

$$DUCR_{(j,t)} = \alpha_1 DUCR_{(j,t)}^{h} + \alpha_2 DUCR_{(i,t)}^{n} \tag{18}$$

B. Proposed Selection Decision Algorithm (SDA)

The SDA does not need to know how the parameters are collected within profiles or how the selection decisions are enforced. The PM must restrain unnecessary triggers, otherwise the SDA could be activated too often and it will exhaust the CPU or the battery. Based on the triggers received from the Profile Manager, the SDA interrogates the Profile Database and starts its computational procedure. Thus, various active profiles are used as input by the SDA in order to

select the "best" interface for each application flow. The procedure used by the SDA is to define access network score functions to solve this multiple-goal problem. It is worth to note that the SDA could implement diverse selection decision process, but few of those are able to take into account such a large panel of parameters (mean bit rate, average delay, security level…) as well as to work on a flow-per-flow basis.

In order to resolve the problem of resources control and management, we have opted to place new entities in the network side devoted to monitoring, analysing and managing the network resources which guarantee networks load balancing while at the same time making the decision taken by the MT without operator exigencies that aims generally to take benefits. The MT deploys flexible selection decision algorithms that use the concept of weights which will be assigned to different parameters according to user preferences and applications requirements. Therefore, our approach allows selecting the best network responding to user and applications expectations as well as best resources management and network utilization optimization.

To evaluate the behaviour of MT, we propose to use various selection decision algorithms designed to solve multi criteria selection analysis. First of all, we use the Ubique Algorithm (UA) [14] that does not adopt the concept of weights to observe the behaviour of MTs when there are no policies.

Then, we use more flexible selection decision algorithms that use weights to demonstrate the influence of various parameters participating in handover decisions making:

- Simple Additive Weighting (SAW) [4], [15]: it calculates the score of each alternative by simple addition of all weighted attribute values.

- Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) [4], [16] it considers the best alternative as the nearest alternative to the ideal solution and the furthest from the non-ideal solution. The ideal and the non-ideal alternatives are obtained by considering respectively the best and the worst values for each parameter.

- Grey Relational Analysis (GRA) [4], [17]: it is used to build a grey relationship with the ideal network by calculating grey relational coefficient (GRC) that defines the level of similarity and variability with the ideal solution.

- Multiplicative Exponent Weighting (MEW) [4]: it calculates the score of each alternative by considering the product of weighted attribute values.

IV. PERFORMANCE EVALUATION

The goal of this section is to study the feasibility of our approach and evaluating its performances by running simulation of several typical scenarios. For this purpose, we have used the framework UBIQUE [13] as a Selection Decision Mechanism (SDM) considering the different actors participating in the decision-making process, while modifying each time the SDA module.

In the different simulated scenarios that we have made, we have used the different emulators and generators already implemented in the framework UBIQUE such as the mobility emulator that uses a movement generator to create realistic movements within an inserted real map using the CanuMobiSim simulator [18] in order to obtain realistic measurements. Moreover, we have used

the mobile context changes generator which considers all information retrieved from the mobile that may be used to define the behavior of the system, such as the battery level changes. The applications launching/stopping events generator that models the applications flow initiation and termination on the MT and their corresponding QoS requirements such as the required throughput, the maximum loss rate and the maximum tolerated delay. As well as the emulator of the available networks and their characteristics (e.g. throughput, bit error rate, delay) and user preferences that can prohibit the selection of some access networks and express their perceived QoS by assigning a suitable set of weights obtained by means of some procedures as the Analytic Hierarchy Process (AHP) [7] or the entropy method [19] (see figure 3).

In the other hand, in order to include the role of the resources reservation mechanism, we have designed and implemented a resources reservation module that aims interacting with all the existing modules, managing and sharing the available resources while at the same time guaranteeing resources that are expected and reserved by each mobile.

To compare our approach to some related works done in the same context [8, 20], we have decided to simulate similar scenarios using the same parameters (Table I, II, III and IV) and we have been interested in controlling the ability of our solution to achieve network utilization balancing and fulfilling user preferences.
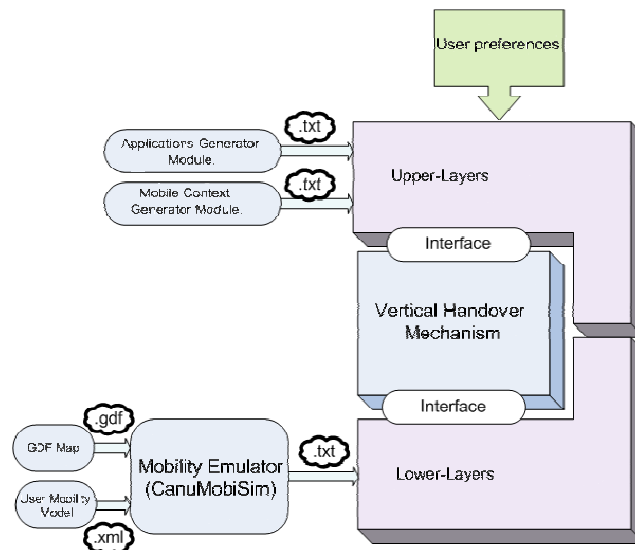

Figure 3: UBIQUE Framework

A. Network utilization balancing

In this section we are interested in evaluating the performances of our solution concerning its ability to achieve network utilization balancing. For this purpose, we have considered a multi-interfaces mobile terminal moving in a heterogeneous environment formed by six access

networks (see Table I) and running four applications (see Table II) belonging to the four traffic classes defined by 3GPP in [21].

For each class of traffic, we have assigned adequate weights values to the different parameters as shown in Table III, taking into consideration user preferences and applications requirements. For example, we remark that delay is the most important parameter for the conversational flow, while interactive flows require minimum error rate and a high security level.

As a first study, we have tried to observe the behavior of the Ubique algorithm as it does not use the concept of weights to study the case where there is no policy. Then, we have changed the SDA by the SAW algorithm and TOPSIS algorithm, respectively, to observe the behavior of multi decision making algorithms in our approach.

Figure 4 shows that the Ubique Algorithm considers the best network as the one that takes into consideration all the parameters participating in the selection decision process such as cost, delay, bandwidth, etc. in order to guarantee to the user to be always best connected, which leads generally to select the same interface for all the flows independently of service class primordial requirements and networks resources management issues.

Indeed, the Ubique Algorithm does not adopt the concept of weights in its interface selection decision. For this reason, we remark that the network *N4* offering the lowest cost is overloaded while the other networks are kept under-utilized.

To resolve this problem, the resources reservation module proposed in our solution can share load between all the available networks by preventing mobiles terminals reserving resources in the same network leading to overloaded it while others under-utilized networks can satisfy its applications requirements expressed by the defined weights. Thus, Figures 5 and 6 show that our approach by adopting the resources reservation module and applying algorithms using the concept of weight can share traffics between the different access networks according to the applications requirements and the networks conditions and performances. Therefore, network utilization balancing is successfully achieved like in [8] and [20] without operator's intervention that tends each time to change the parameters weights which leads to block the mobile control in the network selection decision.

TABLE I: SIMULATED NETWORKS CHARACTERISTICS

NETWORK	TECHNOLOGY	DELAY (ms)	PER	COST (€/MB)	SECURITY LEVEL
N1	WiFi	55	0.02	0.2	2
N2	WiMax	80	0.02	0.1	3
N3	WiFi	110	0.01	0.1	5
N4	WiFi	130	0.01	0	2
N5	GPRS	150	0.02	0.3	3
N6	WiFi	90	0.015	0.1	2

TABLE II: SIMULATED APPLICATIONS CHARACTERISTICS

APPLICATION	TYPE	SERVICE CLASS	BIT - RATE (Kb/sec)	ERROR RATE	DELAY (ms)
VoIP	Real Time	Conversational	20	10^{-3}	150
VIDEO	Real Time	Streaming	64	10^{-2}	300
SSH	SERVICE	Interactive	5	0	1000
FTP	DATA	Background	5	0	10000

TABLE III: WEIGHTS ASSIGNED TO THE DIFFERENT CLASS OF SERVICE

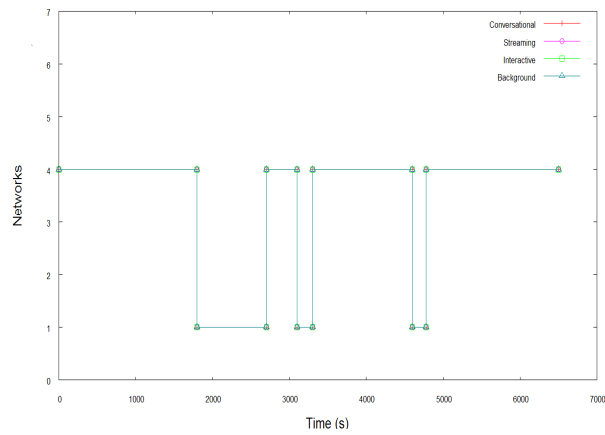| SERVICE CLASS	B	E	D	S	C
Conversational	0.065	0.065	0.613	0.128	0.128
Streaming	0.545	0.035	0.178	0.121	0.121
Interactive	0.04	0.365	0.04	0.365	0.19
Background	0.13	0.4	0.038	0.038	0.4

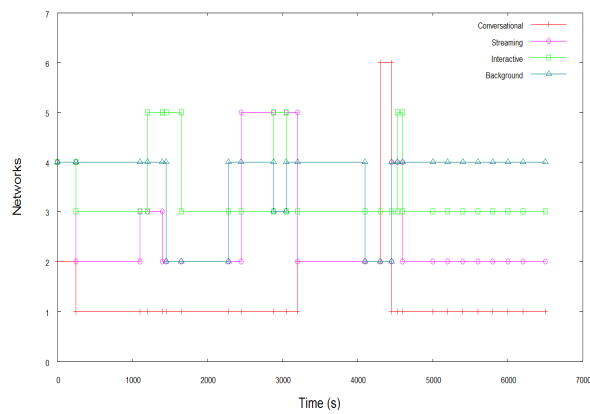Figure 4: Flows transmission when using Ubique Algorithm

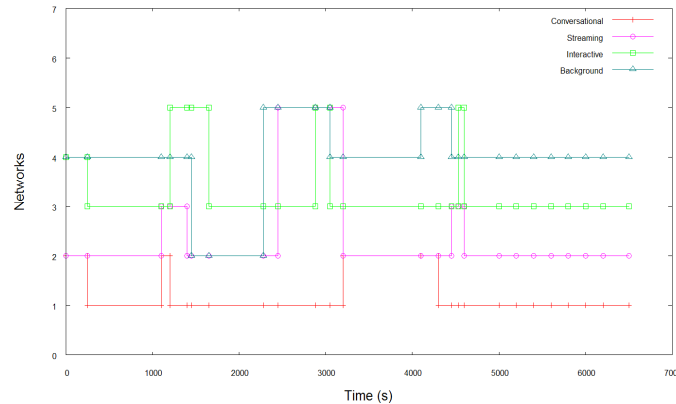Figure 5: Flows transmission when using SAW Algorithm

Figure 8: Flows transmission when using TOPSIS Algorithm

B. Impact of user preferences in network selection decision

In this scenario, we study the ability of our approach to satisfy users' preferences reflected by increasing the weight assigned to the transmission time delay when making the network selection. For this reason, we have considered a conversational application launched on a multi-interfaced MT moving in heterogeneous access networks as depicted in Table IV. Thanks to the different emulators and generators, the Ubique middleware is provided with instantaneous measurements and performances of the available networks, the mobile context, the user preferences, and the applications requirements.

In each simulation, we increased the weight assigned to the delay parameter and we measured the average transmission time offered by the selected networks. The simulation results presented in Figure 7 show that the increase in the value of the weight assigned to the delay influences the network selection decisions that aim always to make the most suitable combination of selected networks which can guarantee the minimum average delay. However, except to all the multi criteria decision making algorithms such as SAW, GRA, TOPSIS and MEW, the Ubique Algorithm didn't try to decrease the average transmission time while making the network selection due to the fact that the UA didn't adopt a specific weight for the delay parameter.

TABLE IV: CHARACTERISTICS OF ACCESS NETWORKS

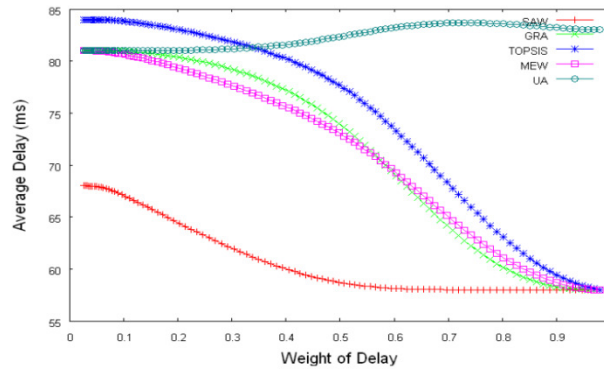| NETWORK	TECHNOLOGY	DELAY (ms)	PER	COST (€/MB)	SECURITY LEVEL
N1	Wi-Fi	70	0.01	0.2	2
N2	Wi-Fi	65	0.01	0.2	1
N3	WiMax	85	0.01	0.3	3
N4	Wi-Fi	75	0.01	0.2	3
N5	Wi-Fi	55	0.01	0.2	3
N6	UMTS	80	0.03	0.2	5

Figure 7: Impact of the variation of the delay's weight on network selection

V. CONCLUSION

The future-generation wireless systems will combine heterogeneous wireless access technologies to provide mobile users with seamless access to a diverse set of applications and services. Selection of the most optimal access network is an important issue for service delivery in a heterogeneous wireless system. The decision is influenced by several actors where each one tends to extract profit. In this article we have proposed a novel architecture and scheme for best network selection and resources management in heterogeneous environment based on mobile terminal decision with network assistance for information collection, monitoring and resources reservation. Then, we have run typical scenarios in order to study the feasibility of our solution. We proved in this article that thanks to the different network nodes interaction and the resources reservation module, in addition to the use of weights concept, network utilization balancing and user and applications expectations are successfully assured without operator intervention.

In this article, we did not consider the power consumption of vertical handoff caused by substitution as well as the handling overhead incurred by partitioning. Actually substitution and partitioning may not always be helpful in reducing the total energy consumed, however they definitely helps meet user preferences and applications requirements.

Also, it would be interesting to evaluate the behavior of a large fleet of terminals and to propose a new solution for reducing handover delay by optimizing authentication procedure in heterogeneous wireless networks.

Reference

[1] F. Bari and V. C. M. Leung, "*Automated Network Selection in a Heterogeneous Wireless Network Environment*", *IEEE Network*, vol. 21, no. 1, Jan/Feb. 2007, pp. 34-40.

[2] I. Smaoui, F. Zarai, R. Bouallegue and L. Kamoun, "*Multi-Criteria Dynamic Access Selection in Heterogeneous Wireless Networks*", IEEE ISWCS'09, Sienna-Italy.

[3] B. Xing and N. Venkatasubramanian, "*Multi-Constraint Dynamic Access Selection in Always Best Connected Networks*", IEEE MobiQuitous'05, pp. 56-64, San Diego, CA, July 2005.

[4] E. Stevens-Navarro and V.W.S. Wong, "*Comparison between Vertical Handoff Decision Algorithms for Heterogeneous Wireless Networks*", VTC 2006-Spring. IEEE 63rd Vehicular Technology Conference, vol.2, pp. 947-951.

[5] L. Suciu, J.-M. Bonnin, K. Guillouard, B. Stevant., "*Towards a Highly Adaptable User-Centric Terminal Architecture*", 7th International Symposium on Wireless Personal Multimedia Communications (WPMC'04), Abano Terme, Italy, September 2004.

[6] S.K. Lee, et al., "*Vertical Handoff Decision Algorithms for Providing Optimized Performance in Heterogeneous Wireless Networks*", IEEE GLOBECOM 2007, Washington D.C.

[7] S. Qingyang, A. Jamalipour, *"A network selection mechanism for next generation networks"*, Communications, 2005. ICC IEEE International Conference on, vol. 2, pp 1418-1422.

[8] I. Lassoued, J.-M. Bonnin, and A. Belghith. "*Towards an architecture for mobility management and resource control*". In WCNC'08 (IEEE wireless communications and networking conference), page 6 pp., Las Vegas, Nevada, USA, March 31st 2008.

[9] Q-T. Nguyen-Vuong, N. Agoulmine, Y. Ghamri-Doudane, "Novel Approach for Load Balancing in Heterogeneous Wireless Packet Networks", in IEEE NOMSWorkshops2008, pp.26-31.

[10] 3GPP Technical Report 25.814, version 7.1.0, *Physical Layer Aspect for Evolved Universal Terrestrial Radio Access (UTRA)*, September 2006

[11] J. Perez-Romero, O. Sallent, R. Agusti, "*A Novel Metric Context-Aware RAT Selection in Wireless Multi-Access System",* IEEE, ICC, 2007.

[12] TA. Weiss, FK. Jondral, "*Efficient signaling of spectral resource spectrum pooling system*", IEEE Comm. Mag., vol.42, Nov., 2003.

[13] L. Suciu, J.-M. Bonnin, K. Guillouard, B. Stevant, *"Achieving "Always Best Connected" Through Extensive Profile Management"*, SPRINGERVERLAG Germany 2004, ISSU 3260.

[14] L. Suciu, K. Guillouard and JM. Bonnin, *"A methodology for assessing the vertical handover algorithms in heterogeneous wireless networks"*, ACM International Conference Proceeding Series, 2006.

[15] J. Geldermann, and O. Rentz, "Bridging the gap between American and European MADM-approaches", Presented at the 51st meeting of the European working group "Multicriteria Aid for Decisions" in Madrid, 2000.

[16] Lee-Ing Tong, Chung-Ho Wang and Hung-Cheng Chen, *"Optimization of multiple responses using principal component analysis and technique for order preference by similarity to ideal solution"*, The International Journal of Advanced Manufacturing Technology, Springer London 2005.

[17] Qinbao Song and M. Shepperd and C. Mair, *"Using grey relational analysis to predict software effort with small data sets"*, 2005. 11th IEEE International Symposium Software Metrics, sep, 2005.*

[18] I. Stepanov and J. Hahner, C. Becke and Jing Tian and K. Rotherme, *"A Meta-Model and Framework for User Mobility in MobileNetworks"*, ICON2003. The 11th IEEE International Conference, 2003, pp 231-238.

[19] B. Bakmaz, Z. Bojkovic, M. Bakmaz, "*Network selection algorithm for heterogeneous wireless environment*", in the proceedings of the 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'07).

[20] I. Lassoued, J.-M. Bonnin, Z. Ben Hamouda and A. Belghith. *"A Methodology for Evaluating Vertical Handoff Decision Mechanism"*. In ICN'08 (IEEE Seventh International Conference on Networking), Mexico, April 2008.

[21] 3GPP, *"QoS concepts and architecture"*, TS 22.107(v6.3.0), June 2005.