# Resource Management in Mobile Heterogeneous Networks: State of the Art and Challenges

Kandaraj Piamrat, César Viho, Adlen Ksentini, Jean-Marie Bonnin

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

# *Resource Management in Mobile Heterogeneous Networks: State of the Art and Challenges*

Kandaraj Piamrat — César Viho — Adlen Ksentini — Jean-Marie Bonnin

**N° 6459**

February 2008

Thème COM

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--6459--FR+ENG

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
RENNES - BRETAGNE ATLANTIQUE

# Resource Management in Mobile Heterogeneous Networks: State of the Art and Challenges

Kandaraj Piamrat*, César Viho†, Adlen Ksentini‡, Jean-Marie Bonnin§

Thème COM — Systèmes communicants
Équipe-Projet Dionysos

Rapport de recherche n 6459 — February 2008 — 29 pages

**Abstract:** Deployment of next-generation networks (i.e. 4G) begins to spread throughout the world. Today's emerging multimedia application has many requirements in terms of quality of service and users always want to be best connected anywhere, anytime, and anyhow. To satisfy these demands, a variety of access technologies has become available: WiFi (Wireless Fidelity), WiMAX (Worldwide Interoperability for Microwave Access), and Cellular networks. This has made it difficult for service provider to select the best network for requesting services and to control the quality level of ongoing connections. Thus, the use of resources management to prevent overloaded or underutilized networks as well as to best satisfy users is indispensable. This report addresses the state of the art on radio resource management in next-generation networks. Recent schemes in network selection and bandwidth allocation are discussed in several aspects, namely decision making, QoS, mobility, and architectural design.

**Key-words:** 4G networks, radio resource management, network selection, bandwidth allocation, mobility, quality of service, architectural design

Note: This report has also been published as IRISA report no. PI-1889

* kandaraj.piamrat@irisa.fr
† cesar.viho@irisa.fr
‡ adlen.ksentini@irisa.fr
§ jm.bonnin@telecom-bretagne.eu

Centre de recherche INRIA Rennes – Bretagne Atlantique
IRISA, Campus universitaire de Beaulieu, 35042 Rennes Cedex
Téléphone : +33 2 99 84 71 00 — Télécopie : +33 2 99 84 71 71

# Gestion des ressources dans les réseaux mobiles hétérogènes: état de l'art et défis

**Résumé :** Les réseaux de nouvelle génération (4G) commencent à se déployer dans le monde entier. Aujourd'hui, les nouvelles applications multimédia ont de nombreuses exigences en termes de qualité de service et les utilisateurs veulent être toujours mieux connectés n'importe où et à n'importe quel moment. Pour satisfaire ces exigences, une variété de technologies d'accès sont d'ores et déjà disponibles : le WiFi (Wireless Fidelity), le WiMAX (Worldwide Interoperability for Microwave Access), et les réseaux cellulaires. Le problème qui se pose alors aux fournisseurs d'accès à l'Internet mobile est celui de la sélection du meilleur réseau pour un service demandé et le contrôle de la qualité des connexions en cours. Ainsi, la gestion des ressources est indispensable afin d'éviter de surcharger ou de sous-utiliser tel ou tel réseau d'accès tout en satisfaisant au mieux les utilisateurs. Ce rapport décrit l'état de l'art sur la gestion des ressources radios dans les réseaux mobiles de nouvelle génération. Les mécanismes récents de sélection et d'allocation de bande passante sont examinés sous plusieurs aspects,à savoir ceux de la prise de décision, de la qualité de service, de la mobilité, et de l'architecture de contrôle.

**Mots-clés :** réseaux mobiles, gestion des ressources, sélection de réseaux, allocation de bande passante, mobilité, qualité de service, système architecture

# 1 Introduction

Currently, enormous progress has been done in wireless technologies. This event confirms that next generation network will become mobile heterogeneous. This kind of network integrates different radio access technologies together in order to provide services to user. With variety of available technologies, quality of service will become a crucial concern. Service providers will need to guarantee mobile users being always best connected. To achieve this goal, they will need to profit from heterogeneity of wireless networks in intelligent manner. Heterogeneous network keeps the best features of the individual networks: the global coverage of satellite networks, the wide mobility support of 3G systems, and the high speed and low cost of WLANs. Therefore, resource management in mobile heterogeneous is absolutely indispensable.

Generally, resource management covers several issues. In this report, we privileged the more important ones to be discussed. We give an overview of radio resource management (RRM) in 3 steps: information harvesting, decision making, and decision enforcement. First, in information harvesting step, important factors for making decision are given. Then we discuss decision making, the most important issue in RRM, with mechanisms used and criteria. Decision enforcement is usually included in decision schemes; hence we talk about it in the same section. We survey the literature over the period of 2002 to 2007 on RRM in wireless heterogeneous network in several aspects. We extract common challenges namely decision mechanism, QoS support, and mobility support as well as related issues such as media adaptation and architectural approach to be discussed in details.

As of our knowledge, there has not yet been a recent survey of resource management in heterogeneous network. Related works are the comparison of four IST architectures from Annoni et al. [1] and a survey on common radio resource management from Wu and Sandrasegaran [2] that focuses only on GERAN/UTRAN and WWAN/WLAN. With recent advances on resource management in wireless heterogeneous network, we decided to carry out a new investigation. For this report, we focus more on the techniques deployed for radio resource management (RRM) and we discuss architectural aspect as one of the concerning issue.

We discuss the features that would enable us to design a good mechanism to manage radio resource in mobile heterogeneous network, with more emphasis on decision making mechanism. The remainder of this report is organized as follows. We begin by giving an overview of resource management in mobile heterogeneous network in section 2. Then we follow by section 3 with decision making as well as other related issues in section 4. Finally we conclude this report in section 5.

# 2 Overviews

In this section, we give a brief overview of mobile heterogeneous network and summarize the ones considered in surveyed papers; then, we give brief overview of resource management in this type of network and explain decision factors.

| Network | Standard | Bandwidth | Coverage |
|---|---|---|---|
| Cellular | GSM | 6.9 kbps | 100m - 30 000 m |
| | UMTS | 2 Mbps | 300 m |
| WLAN | 802.11a | 54 Mbps | 100 m |
| | 802.11b | 11 Mbps | 100 m |
| | 802.11g | 54 Mbps | 100 m |
| WPAN(Bluetooth) | 802.15.1 | 2 Mbps | 30 m |
| Zigbee | 802.15.4 | 250 kbps | 100 m |
| WMAN(WiMAX) | 802.16 | 70 Mbps | 10 000 m |

Table 1: Wireless technologies

## 2.1 Mobile heterogeneous Network

In wireless networking, a heterogeneous system is composed of several wireless technologies similarly to Figure 1; they constitute together a network that connects users to the Internet. Core network, sometimes called backbone network, joints all access networks together. The technologies utilized in core and access networks may be different; resulting in different characteristics such as bandwidths illustrated in Table 1. Communication protocol mostly used in heterogeneous network is the Internet protocol (IP) either IPv4 or IPv6. We observed that some of the recent architectures such as [3] and [4] have been proposed for working with IPv6 in order to facilitate the migration from IPv4 to IPv6. However backward compatible to IPv4 is necessary since a large number of network equipments and end systems are still using this version of Internet protocol.

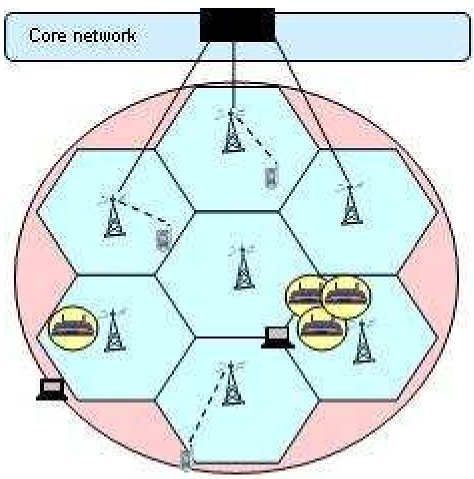


Figure 1: Heterogeneous System

In wireless heterogeneous network, stations are mobile and can move freely from one place to another while being always connected. The transition of user from one point of attachment to another is called handover or handoff. Handover or handoff refers to the process of transferring an ongoing session from one attachment point to another. Mobile node movement may cause two types of handover: intra-technology or inter-technology handovers. Intra-technology handover sometimes called horizontal handover happens when mobile node changes its point of attachment within the same technology. In contrast, inter-technology handover or vertical handover happens when mobile node move to another technology.

| **Access Techonologies** | |
|---|---|
| UMTS | WLAN(802.11x) DVB-T(Digital Video Broadcasting - Terresterial) [5]<br>WLAN(802.11e) [4]<br>WLAN [6, 5, 4] |
| GPRS | WLAN(802.11) [7] |
| WiMAX | WLAN [8]<br>CDMA WLAN(802.11) [9, 10]<br>WLAN(802.11e) WPAN(802.15.1) TD-CDMA [3] |
| Satellite | WLAN 3G(UMTS and CDMA2000) [11] |

Table 2: Surveyed heterogeneous architectures

In our survey, we investigate resource management in heterogeneous systems figured in Table 2. It can be seen that the most popular technologies deployed are UMTS and WiMAX. This is because of the great coverage areas and the high bandwidths these two technologies can provide. In contrast, WLAN is usually deployed in small area because of its small coverage; however its high-bandwidth and low-cost properties is desirable for Internet traffic such as multimedia or real time applications which are increasing significantly these days. Moreover, the deployment of WLAN is easier and cheaper than UMTS and WiMAX. As proof of it, we can see its installations everywhere, for example, in office buildings, airports, or shopping malls, we can also notice that today all laptops are integrated with Wi-Fi adapter which is not the case for UMTS and much less for WiMAX.

Although lots of benefits of WLAN are obvious, it cannot support high-speed mobile users such as mobile terminal in vehicles driving in highway, thus the supplement from cellular networks. To sum up, the attractive advantage of heterogeneous network is indeed heterogeneity: different network types are linked together and they can provide wider ranges and higher quality of service than in homogeneous network. For example, with overlapping of several coverage areas, multi-mode users have possibility to connect to the best points of attachment and to profit from best quality of service offering by heterogeneous system. With this kind of mixed architecture, good resource management is needed in order to get expected QoS while precisely minimizing the necessary resource though the appropriate access network.

## 2.2 Resource Management

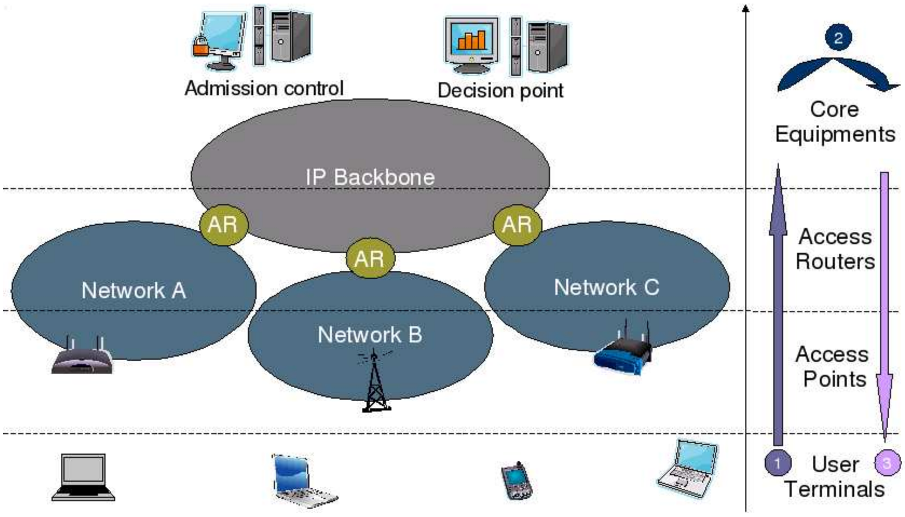


Figure 2: Resource management in mobile heterogeneous network

From management point of view, mobile heterogeneous network can be illustrated as in Figure 2. Generally, radio resource management in mobile heterogeneous network consists of 3 steps: information harvesting, decision making, and decision enforcement. Figure 2 illustrates the levels where these steps are taking place. We explain each step as follow:

1. Information harvesting: at this stage, information about user and network are gathered, this information is important factor for making decision. First users' information should be collected at user terminals level then it should be propagated up through access points (AP) and access routers (AR) levels for having more information about cell condition and network conditions respectively. Examples of information are given in Table 3.

2. Decision making: at this stage, decisions are made. It can be noticed that later in this report we also show another place than core network where decision can also be made. For example, decision can be made at user terminals in case they have possibility to choose their point of attachment or at access routers for controlling local networks. Therefore, the decision point depends essentially on where the control has been placed. We will discuss about decision making location again in section 4 as architectural aspect. Furthermore, since decision making is the most important part in resource management, we will discuss about it in more detail in section 3.

3. Decision enforcement: at this stage, decisions are enforced. Several mechanisms can be used to ensure that decisions made in step 2 are respected. Admission control is one of the enforcement mechanisms; it can be used to filter access according to the decision. In some case, it can be adopted in decision making (step 2) to screen candidate networks by comparing required service and availability on the present networks. An example of

> RRM architecture is policy-based architecture; it has been implemented by many schemes. The architecture typically has central architecture similar to Figure 2. The system consists of policy repository (PR) where all policies and network information are stored, policy decision point (PDP) where decisions are made, and policy enforcement point (PEP) where decisions are executed.

Previously, we explain steps in resource management according to locations where those steps take place in the network. Here in Figure 3 we summarize how decision mechanisms work in terms of input, processing, and output.

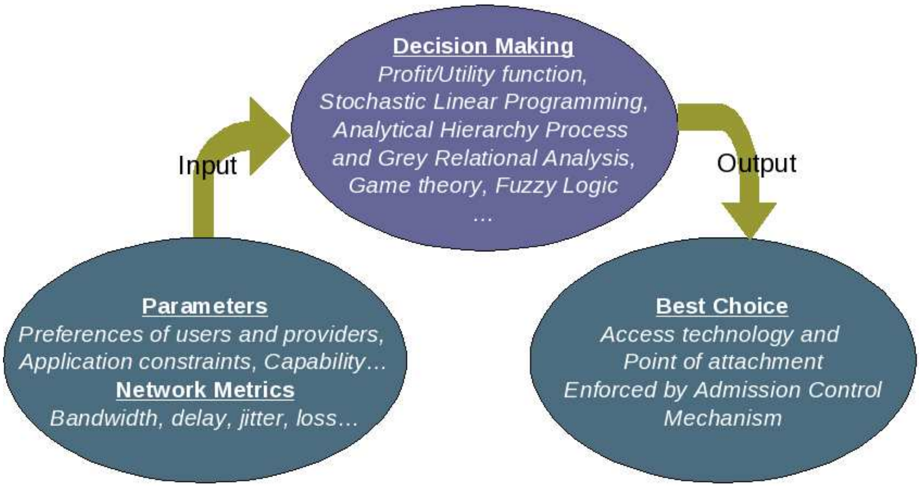


Figure 3: Decision making: input, processing, and output

Every decision mechanism in resource management needs to get information of relevant factors as inputs. These inputs are essential in order to make a good decision; referring to their nature, they can be separated into 2 categories: predetermined and time-varying. The predetermined factors are pre-defined a priori and stay for a long period of time whereas time-varying ones change in time. Predetermined factors are taken into consideration as initial policy or preference; they also include constraints of application and capabilities of technology and equipments; on the other hand, time-varying factors are monitored continuously. We give examples of both of them in Table 3.

After having information about factors, management schemes deploy various decision making techniques that give best solutions for service providers. Usually they determine best access technology and point of attachment for new and ongoing connections. Decisions are enforced afterward by mechanisms such as admission control.

| | |
|---|---|
| **Pre-Determined** | **Users preference**: cost, security, power, visual quality...<br>**Providers preference**: cost, trust, security, load balancing, latency, throughput, drop probability, user priority, topology...<br>**Application constraints**: QoS constraints, service requirement, application requirement, application context, variety of services, adaptation ability, minimum required bandwidth, maximum loss rate and latency allowed, delay bounds, traffic specification...<br>**Capabilities**: network capabilities, network equipments capability, access technologies capability, Access Point bandwidth and queue, up/downlink bandwidth, modulation scheme, terminal capability: CPU, memory size, display I/O, transmitted power, battery, network interface, built-in application, software platform ... |
| **Time-Varying** | **Availability**: network load, available radio coverage, visible AP, traffic characteristic, maximum saturation throughput of AP, transmission bandwidth, cell diameter, bandwidth per user,delay, throughput, response time, jitter, bit error rate, burst error, loss, radio condition (path loss), available service, network connection, terminal conditions ...<br>**Quality related**: SINR(signal to interference plus noise ratio), traffic intensity/connection arrival process, SNR(signal to noise ratio),connection holding time, average number of connection, bandwidth utilization, RSS(received signal strength), SIR(signal to interference ratio), SER(symbol error rate), image resolution, data rate, BER(bit error rate), MSE(mean square error), latency, jitter, PSNR(peak signal to noise ratio), user activity history, suitable application, handover latency, CIR(carrier to interference ratio), loss, dropping rate... |

Table 3: Information categories: Pre-determined and Time-varying Factors

# 3 Decision Making

In this section, we first discuss decision mechanisms. Several techniques have been deployed to achieve the best resource management; we cover some of them in section 3.1. In section 3.2, we discuss two important criteria concerning decision: QoS and mobility supports. We explain the main issues and give solution trends.

## 3.1 Decision Mechanism

We consider resource management in terms of mechanisms used to manage resources for new and ongoing connections; however, some of them may present bandwidth allocation algorithm for the whole network. From the surveyed papers, many schemes have been proposed to manage resources in wireless heterogeneous network; we can categorize them, by method used for decision making, into two groups: function-based and mathematical-based as illustrated in Figure 4.

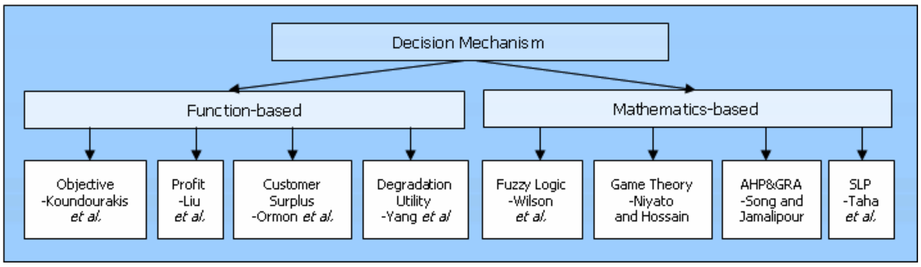


Figure 4: Decision mechanism

### 3.1.1 Function-based mechanisms

In this category, the decision mechanisms are based simply on output of the functions. Four functions have been summarized in Table 4 with their required input and desired goal, and they are described below.

**Objective function** - In objective function [5], inputs are derived from 3 different sources: user data, network data, and policy information. First, users are asked for list of visible AP with corresponding signal quality, list of requested services with corresponding nominal bit rate, and delay tolerance. Second, network data such as the AP bandwidth at the wireless interface and the delay at the queue between access router and the backbone are collected. Third, policy such as cost, compatibility, trust, preference and capability along with their weights are defined. The weights can be dynamically changed according to the network condition. Finally, with all factors and their weights, the algorithm iterates and computes the best allocation that maximizes the objective function for overall network. For more detail of this algorithm, please refer to the annex.

**Consumer surplus** - With customer surplus [12], the authors propose a user-centric solution meaning that decisions are made at users' side. The scheme has been designed for non real-time traffic with the following strategy. First, the users survey the radio interface and determine a list of available access networks. Next, they predict transfer rates of networks on the list by taking its average of the last five data transfers and then derive completion times. After that, the users compute predicted utility which is the relationship between the budget and the user's flexibility in the transfer completion time. Finally, the users compute consumer surplus, which is the difference between utility and cost charged by the network, for each candidate network; and they choose the best one to request for connection. For more detail of this function, please refer to the annex.

**Profit function -** In this function [6], the authors associate each handoff with a profit that is decided by a target function with two parameters: the bandwidth gain and the handoff cost. For the functioning of the profit function, the authors classify handoffs into reactive handoff and proactive handoff. A reactive handoff is initiated whenever a mobile node is going to roam out of the current cell, while proactive handoff can only be initiated at periodical discrete epoch when connection experience can be improved. Parameters used in the calculation of the gain include access networks along with their maximum bandwidth provided to a single user and percentage of capacity that has been used, application's maximum requirement on bandwidth, and the bandwidths of access networks used by a mobile node for handoff. The authors define handoff cost as data volume lost due to handoff delay; it corresponds to the volume of data which could have been transmitted during the handoff delay. Thus, the profit is difference between gain and cost. At each handoff epoch, mobile node will compare profit from different networks and choose the one that give maximum profit. For more detail of this function, please refer to the annex.

**Degradation utility -** The authors propose an idea of degradation utility [11] to deal with different user priorities. By degrading lower priority traffic, more bandwidth can be released for higher priority users. First, service providers need to specify levels of service in terms of bandwidth offered, classified as excellent, good basic, and rejected, for each application type such as voice, video, and data. This bandwidth specification will be used to compute released bandwidth (difference of bandwidths before and after degradation). After that, table of rewards for each user priority class are defined: there are three quality of service (excellent, good, and basic) and three kinds of disconnection (forced disconnection, handover drop, and rejected), each of them associated with reward for each type of application. This table will be used to compute lost reward points (difference of reward points before and after degradation). Finally, degradation utility is the division of released bandwidth by lost reward points. When a new connection is requested, service provider finds all potential degradable connections, computes their degradation utilities and begins to degrade the connection that give highest utility. Example of this function is in the annex.

**Discussion**

As can be seen in Table 4, objective function [5] and profit function [6] derive inputs from many sources including users in order to make decision. The inputs of these functions cover all necessary factors to make a good decision; however, decision mechanisms are no longer transparent to users since they are asked for data. Many other schemes also ask for user participation. Since user information is so important but implementing schemes will lose transparency, the challenge is finding compromise between utility and cost of involving users.

| Function | Input | Goal |
|---|---|---|
| Objective function | Quality Indicator (bandwidth, delay, signal quality) and Policy Indicator (cost and compatibility, network provider, terminal type) | Maximize sum of the input timed their weights |
| Consumer Surplus | Utility (available network, predicted completion time) and Cost charged by the network | Maximize difference between utility and cost |
| Profit function | Bandwidth Gain (available bandwidth, percentage of used capacity, application's requirement) and Handoff Cost (data lost due to handoff delay) | Maximize difference between gain and cost |
| Degradation Utility | Released bandwidth and lost reward point (point lost during degradation, according to traffic class/quality,user priority) | Maximize ratio of released bandwidth and lost reward points |

Table 4: Functions: inputs and goals

The scheme using consumer surplus [12] has been designed for non real-time application. It is not appropriate with real-time multimedia traffics that have more constraints, not only in terms of completion time. This scheme also proposes user-centric solution which may not be good for load-balancing of the whole network. It may result in congestion since each user only consider its own criteria and does not care about network load distribution.

With emerging of multimedia traffic, releasing bandwidth of low-priority traffic to give better quality for high-priority traffic become interesting strategy for service providers. Degradation utility function [11] has been designed to perform this strategy but the trade-off between satisfying degrading connections and new connection has to be well studied. Moreover, this fine-grained strategy is suitable only if service level agreement (SLA) has been a priori signed between user and service provider to specify their individual responsibilities.

For performance evaluation, no common evaluation has been done in this report because different metrics are used by the schemes depending on their goal. The scheme deploying objective function [5], with the goal to provide good bandwidth allocation, demonstrates results in terms of network loads, number of execution, and QoS handover. In addition to system throughput, profit function scheme [6] presents its results in terms of connection blocking and connection dropping rate which are essential in point of view of users' satisfaction. The scheme using degradation utility [11], with the goal to differentiate treatment among different user priorities, shows the result on degradation ratio which is lower for high priority users and it increases significantly for lower priority users. Finally, [12] with the aim to economy cost has shown the results in terms of utility (cents) by completion time.

### 3.1.2 Mathematical-base mechanism

Besides the previously described management functions, various mathematical techniques have been deployed for managing network resources recently. Four of the techniques are summarized in Table 5 with the input parameters, processing steps, and output; and their deploying schemes are described below.

**Stochastic Programming** - The authors of [4] deploy stochastic programming (SP), a mathematical programming technique used in decision making under uncertainty, to design a proactive allocation mechanism. The scheme actually uses a subset of SP called stochastic linear programming (SLP) to handle probabilistic nature of demands in wireless heterogeneous network. In the exemplary scenario, a single data service of fixed bandwidth requirement is provided by cellular network and WLAN. The idea is to associate probabilistic demands with predetermined significant probability, then formulate given scenario with allocation, underutilization, and rejection along with the predetermined probability. Thus, the goal is to obtain maximum allocation in both networks while minimizing cost of resource underutilization and demand rejection. For more information, please refer to the annex.

**Fuzzy Logic Controller** - the authors of [13] use an algorithm based on fuzzy logic controller (FLC) to evaluate fitness ranking of candidate networks. They first differentiate decision making into three phases: pre-selection, discovery, and make decision. Pre-selection phase takes criteria from user, application, and network to eliminate unsuitable access networks from further selection. If present networks are not corresponding to user's requirement, system returns to ask user for reducing their criteria. Discovery stage deals with two kinds of state: power-up users when no current connections exist, and connected user when a connection is already established but QoS is not meeting the criteria at the same time other potential networks become available. The authors implemented discovery phase based on fuzzy logic control, they fuzzify crisp values of the variables (network data rate, SNR, and application requirement data rate) into grade of membership in fuzzy set. Then they are used as input to the pre-defined logic rule base. Finally, overall ranking is obtained through defuzzification with weighted average method. For more information, please refer to the annex.

**AHP and GRA** - The authors of [14] propose network selection scheme using analytical hierarchy process (AHP) to weigh QoS factors and grey relational analysis (GRA) to rank the networks. With QoS factors, the authors construct an AHP hierarchy based on their relationships. QoS is placed in the topmost level as the objective; main QoS factors describing network condition such as availability, throughput, timeliness, reliability, security, and cost are defined in the second level. Moreover, availability, timeliness and reliability have been decomposed into sub factors and they have been arranged in the third level. Finally, available solutions are arranged in the bottommost level. The mechanism is divided into three main logical function blocks: collecting data, processing data and making decision. QoS parameters are separated into two types: user's preference and network conditions. User-based data is collected and processed by AHP in order to get global weights of second-level factors and local weights of third-level factors, and then the final weights are computed. Network-based data are normalized by GRA, then ideal network performance is defined following by calculation of the grey relational coefficient (GRC) which give grey relationship between ideal network and the other. The previously computed weights have been taken into account for the calculation of GRC; finally the network with the largest GRC is more desirable. For more information, please refer to the annex.

**Game Theory -** The authors propose bandwidth allocation and admission control algorithms based on bankruptcy game [10]. With this special type of N-person cooperative game, each network cooperates to provide the requested bandwidth to a new connection using coalition form and characteristic function. The amount of allocated bandwidth to a connection in each network is obtained using Shapley value and the stability of the allocation is analyzed using the concept of the core. User initiating new connection is analogous to bankrupt company and the requested bandwidth is the money that has to be distributed among different networks (creditors). The objective of each network is offering maximum bandwidth as possible to gain revenue from new connection, similar to creditors trying to get the most payment. Here is a scenario in bandwidth allocation: when user initiates a new connection, the information on the required bandwidth is sent to the central controller who computes the offered bandwidths by each network, then, the Shapley value is obtained. For admission control, central controller accepts new connection if the sum of allocated bandwidths is at least equal to the requested bandwidth and allocated bandwidths is in the core, meaning that everybody is satisfied. For more information, please refer to the annex.

| **Techniques** | **Parameters** | **Processing** | **Output** |
|---|---|---|---|
| Stochastic Programming | Allocation, demand, underutilisation, rejection | 1-association of predetermined probability to demands<br>2-variable formulation<br>3-SLP statement | Allocation in each network |
| Fuzzy Logic Controller | Network data rate, SNR, application - required data rate | 1-fuzzification<br>2-fuzzy inference<br>3-defuzzification | Fitness rank of each network |
| Analytical Hierarchy Process and Grey Relational Analysis | User's requirements, network conditions | 1-AHP of user's requirements<br>2-GRA of network conditions<br>3-calculation of GRC | Network rank by GRC |
| Game Theory | Available bandwidths in each network | 1-determine offered bandwidths<br>2-compute Shapley value<br>3-verify core | Bandwidth allocation |

Table 5: Mathematical techniques: parameters, processings, and outputs

**Discussion**

Table 5 shows the parameters, processing steps and output of four mathematical-based schemes. First mathematical attempt that directly address joint resource management in wireless heterogeneous networks is stochastic programming approach [4]. The authors deploy mathematical programming technique that deal with uncertainty to handle probabilistic demand nature of user requests in wireless heterogeneous network. However, the scheme is designed for supporting single common service with fixed required bandwidth which is not appropriate to variety of services along with various bandwidth requirements today.

In [13], the authors use pre-selection to filter unsuitable network for further selection according to criteria from user, application, and network. Since user and application criteria do not change during decision, pre-selection is a useful preliminary step only if several networks with different characteristics are available at the same time. However, pre-selection can be costly if several present networks have similar characteristics. In this case, the scheme will choose either all or any since if one network corresponds to criteria the other do as well and vice versa for rejection. It can also be noticed that many schemes have integrated pre-selection into decision making as only one phase.

Fuzzy logic control has been deployed in [13] for network selection. It gives good result in this case of few metrics but when number of metric increases, the system may become very complex and may give erroneous results. The critical issue in this approach is the definition of fuzzy set and rules which needs to be carefully specified. These specifications are very important in order to get a good approximation and they are very delicate to define. With similar objective as [13], the authors of [14] adopt different strategy. Instead of using FLC, they propose network selection algorithm using AHP and GRA. Many QoS parameters have been taken in AHP comparing to only three parameters in [13]. Pair-wise comparison in AHP is finer than fuzzification. Thus this scheme gives more precise solution.

Game theory is another mathematical technique used in resource management; the authors of [10] deploy cooperative game called bankruptcy game to model the bandwidth allocation problem. With this model, coalition form and respective characteristic function have to be defined appropriately. Moreover, the solution is stable (i.e. every body is satisfied) only when it belongs to the core that is not always the case. In case of unstable solution, the most preferable distribution has to be determined, thus this strategy maybe more expensive in such a case.

As already mentioned earlier, different measurements have been used to evaluate the performance. Here, in [10] with the objective to reach satisfied bandwidth allocation, the authors illustrate performance on bandwidth utilization of each network, average number of connections, and connection blocking probability. In [4], SLP has shown a better result than deterministic programming approach. However, for network-selection schemes, such as [13] and [14], there is no measurement to argue their choice.

### 3.2 Decision Criteria

In a good decision mechanism, criteria need to be met. The highest-priority criterion is quality of service since user's objective is to obtain the best quality. Moreover users are mobile in wireless environment, thus the second inevitable criterion is mobility related issues. We describe supports that have been done so far for these two criteria and then we give solution trends.

### 3.2.1 QoS support

QoS support covers wide range of aspect concerning supports in both network and end system. QoS is the most important criterion in resource management because all network providers wish to guarantee mobile users being always best connected. In this section we consider only supports from the network side since it is more general, end system support which is more specific to users' purpose will be discussed later in related issues. Figure 5 summarize methodologies and their respective mechanisms.

In cellular technologies, UMTS, GSM, EDGE, WCDMA, and CDMA2000 have sophisticated network-centric resource management supporting mixed traffic types with different QoS requirements. Similar to cellular network, Wi-MAX also has dynamic service to support different kinds of service flows. Contrary to those cellular and WMAN networks, IEEE 802.11 WLAN has limited resource management supporting only best effort data without any QoS. To enhance IEEE 802.11 standard, amendments and standards have been proposed. IEEE 802.11e defines a set of QoS enhancements through modification in the MAC layer to support bandwidth-sensitive applications such as voice and video. IEEE 802.11k aims to provide client feedback to AP. It defines a series of measurement that detail client statistics to achieve fast handover and uninterrupted service. IEEE 802.11r is also designed for the same objectives. The protocol allows client to establish a security and QoS state at a new AP before it actually moves. IEEE 802.21 enables seamless handover. The standard provides common protocol that allows handover between the same or different technologies.

Beside the above standards, numbers of techniques have been deployed to support quality. Most schemes take QoS metrics and requirements into account for decision making; other schemes, as in [8] and [3], make use of resource reservation protocol to pre-reserve resource and to guarantee requested quality. It can be noticed that many resource reservation protocols have been proposed but most of them fail to be deployed because of restriction that all network equipments need to be reservation-enabled. In addition to reservation, service differentiation has also been used to distinguish treatments for applications with different priorities. IEEE 802.11e and DiffServ are good examples. These protocols are normally accompanied by priority schedulers to help dealing with requests according to their priorities.

The arrival of mobile network results in mobile users who can move from one place to another while being always connected. Liu et al. [6] apply this transition to improve QoS, this is called QoS handover, a type of handover aimed to improve quality. Service provider can use QoS handover in accordance with network condition and user priority to control QoS level. Related technique is QoS upgrade/degrade proposed by Yang et al. [11]. This mechanism will have to be carefully studied a priori due to tradeoff between degrading and upgrading connections. When QoS upgrade takes place, someone is being degraded to release necessary bandwidth. Nevertheless, this approach is interesting because it provides suitable solution for emerging problem of multimedia traffic today.

Apart from techniques previously described. New architectures have also been designed for supporting QoS. Most of them have agents called QoS broker, as in [3] and [15], to manage QoS in the network. Controlling QoS can be done periodically but control signaling may waste bandwidth. Particularly in the case of limited-bandwidth network such as GSM or GPRS, control traffic may be the cause of bottleneck. Therefore, dynamic adaptation using triggering seams to be a more adequate solution. Triggering conditions depend on the objectives of service provider, for example, triggering may be caused by new user initiating connection or ongoing service facing QoS problem as in [5]. In heterogeneous network, central server is frequently adopted to manage overall performance. QoS negotiation [16] between access controller and central controller is an alternative for load balancing among access networks. Despite the prevention if congestion still occurs, it can be handled with strategies such as soft or hard congestion controls. In soft congestion control, only excess bandwidth is removed while in hard congestion control admitted services are totally disconnected based on priority.

To cover all aspects of QoS, a framework has been proposed by Gao et al. [17] with three planes management providing both static and dynamic QoS functions. Static functions are executed during application initiation and remain constant during the session. These functions include traffic specification, QoS translation, QoS negotiation, admission control, and resource reservation. On the contrary, dynamic functions are executed on ongoing connections to improve service quality during the session. These functions include monitoring, renegotiation, adaptation, and feed back.

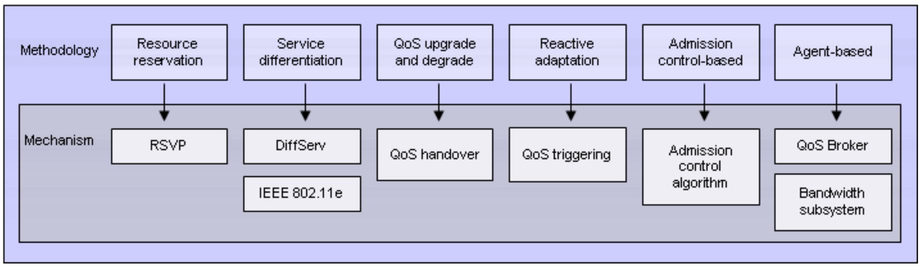


Figure 5: QoS methodology and mechanism

### 3.2.2 Mobility support

In wireless heterogeneous network, stations are mobile and can move freely from one place to another. To handle this mobility, many researchers proposed mobility management module. Most of them are based on Mobile IP and its extension such as Fast handovers for MIP or Hierarchical MIP. Some researchers focus more on handover issue with the objective to achieve seamless handover. Liu et al. [6] proposed another classification with reactive and proactive handovers to fit with wireless environment. Reactive handover is triggered when a mobile node is going to roam out of the current cell while proactive handover can take place whenever a mobile node finds that its connection can be improved through such a handover. The reactive handover is called sometimes forced handover since there is no other choice or else loosing connectivity. Akyildiz et al. [18] discuss deeper in the detail of handover process, they propose a function to determine the best handover initiation time. The goal is to find an appropriate initiation time that is neither too early nor too late. Too early initiation will result in double use of bandwidth in home and foreign network while too late initiation will result in packet loss and non-seamless handover. More recently, a technique such as multihoming has also been used to improve performance in mobile networking as in [5]. With multi-homing, it is possible to connect to multiple networks at the same time by utilization of multi-interface terminal. Advantages are decrease of handover delay and more reliable connection but, at the same time, drawback is the double bandwidth occupied by multi-homed terminal. To standardize handover, IEEE working group is developing IEEE 802.21 for media independent handover services, which will enable co-operative handover decision making of users and operators. Huge effort has been put on mobility issue because this issue will obviously result in quality of service, the final goal of both provider and user.

# 4 Related issues

As already mentioned earlier, resource management covers various aspects. In this section, we discuss media adaptation which is a hot issue dealing with multimedia transmission. Then we follow by architectural approach listing common architectures, their advantages, and drawbacks.

## 4.1 Media adaptation

In today's wireless environment, multimedia traffic such as video transmission increases considerably. With this kind of traffic and unstable condition of wireless network, media adaptation becomes essential. Media adaptation means that node adapts itself to media condition. For example, the control of encoding rate of the video stream based on the estimated available bandwidth or the error correction according to the varying wireless conditions.

Media adaptation can be performed at different locations: end systems or intermediate nodes. End systems such as sender or receiver may participate in media adaptation. The sender can adapt its parameters to be coherent with network condition and ongoing application. For example, the server adjusts its transmission rate according to congestion in the network. Stream switching is one of the techniques. The server prepares streams to be transmitted to the channel in different encoding rate and stocks them in a database. When network condition changes, the server selects stream with encoding rate accordingly. However, drawback of this technique is high consumption of disc space that cannot be possible in every case. It can be noticed that sender adaptation is appropriate in term of bandwidth since no bandwidth is wasted. Receiver can also cooperate in dynamic adaptation by sending its reception capacity to sender but this approach may be costly in terms of bandwidth. So it is not recommended in small-bandwidth networks such as GPRS. More recently, scalable video coding (SVC) is being developed. With this technique, encoding rate can change dynamically according to network condition.

Another issue in media adaptation is reliability. To deal with unreliable channel, error correction mechanisms are recommended. For example, forward error correction (FEC) and automatic repeat request (ARQ) have been deployed by Kassler et al. [7] to enforce transmission. However for real-time or delay sensitive application, ARQ is not preferred because late arrival of retransmitted packets are usually discarded. To deal with retransmission, Singh [19] has proposed selective retransmission scheme to adaptively enable retransmission according to channel condition. The retransmission should be disable when the channel is congested otherwise it should be enable with selective retransmission of I, P and B frame according to available bandwidth. Carneiro et al. [3] proposed channel adaptation module using several protocol such as H264/AVC to provide enhanced bit error resilience capability, UDP-Lite (RFC3828) to deliver erroneous packets and to deal with erroneous packet payloads, robust header compression (RoHC) to reduce IP overhead improving IP packet latency for real time services, and finally FEC to eliminate retransmission that degrade overall throughput.

## 4.2 Architectural approach

This section discusses about architectural approach and gives examples of resource management architecture. The architectures found in most of the papers can be categorized into three types described below.

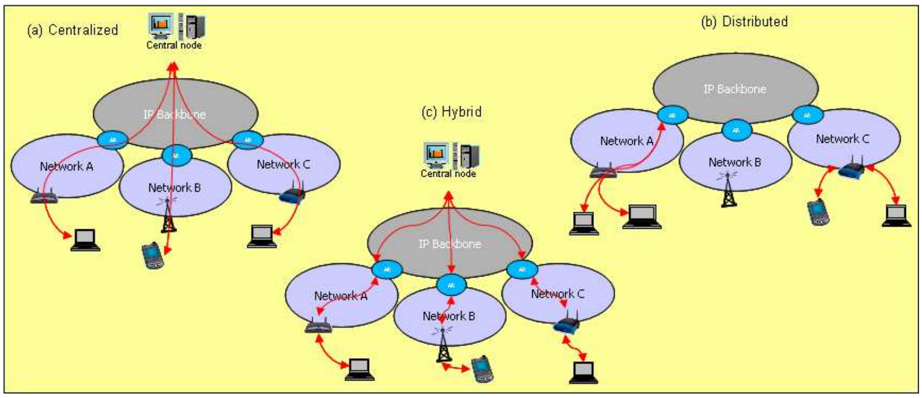


Figure 6: Different types of architecture

The first one is centralized architecture illustrated in Figure 6a used by [8, 13, 14, 4]. In this architecture, control is aggregated into one central point which is usually situated in the core network. Central node has a global view of the whole network which allows an advantageous management of overall performance. On the other hand, since management is centralized at only one point, all other nodes have to send management traffic to central point and this may wastes some bandwidth. Especially, in access network that has limited-bandwidth capacity, this approach may cause bottleneck problem. The same problem happens as well with other centralized communications such as architecture for bandwidth or QoS negotiations, where local controllers need to interact with central controller for management. Moreover, centralized architecture is not scalable and results in one point-of-failure problem.

The opposite of centralized architecture is distributed or decentralized architecture illustrated in Figure 6b. Control in this architecture is allocated into several places either on the network or eventually on the user terminal. In general, the control is placed at access router [5] if network provider wants to manage the access network. Alternately, control may also be placed at the point of attachment that represent local cell such as access points or base stations. Occasionally, distributed architecture also placed control on user terminal in order to get information from user. An extreme solution even gives user possibility to make its own decision on which network to be connected as in [6] and [12]. This solution is not recommended because it may result in congestion on particular network. For example, if terminals choose their point of attachment based on radio power detected, many users located near one AP will all connect to this AP causing overload in this AP and underutilization in the other AP. The same reasoning applies to other schemes where decisions are made at user's side. In addition to previously describe distributed approaches; the authors of [11] propose cooperative distributed system to manage the whole heterogeneous network while still being scalable.

Figure 6c illustrates the last one called hybrid architecture; this type of architecture combines the two architectures described above. It composed of central node that manage global resource and distributed nodes to manage resource locally. We also observe schemes collaborating management in distributed network node and user terminal. For example, Magnusson et al.[20] recommend the combination of distributed network and terminal management for dynamic handling of individual users and sessions. Koundourakis et al. [5] presented network-based and terminal-assisted approach to optimize resource allocation while compromising QoS constraints. Akyildiz et al. [18] develop a hybrid network selection scheme that combines terminal-based and network-based selection mechanisms. Terminal dynamically collects network condition and determines best reachable network, then network makes globally optimized selection and achieve load balancing for the whole system.

# 5 Conclusion

In this paper, we give a comprehensive survey of resource management in heterogeneous network. Research in this topic has been extensively studied in recent years, and many management schemes have been proposed. We described techniques deployed for decision mechanism; the function-based mechanisms are presented as well as mathematical-based mechanisms, and several issues have been raised.

Besides the aforementioned decision mechanisms, this paper also discusses QoS and mobility that come with the arrival of multimedia in wireless networks. These two issues influence the research and development in wide areas, and they need to be considered each time designing new scheme. Moreover, there has been an ongoing debate on architectural design in terms of performance of the system, and finally, hybrid scheme is recommended for good performance of the system because it controls globally while still being scalable

Performance evaluation of existing schemes usually based on network load (access networks and total load), processing capacity requirement, connection blocking, connection dropping, throughput, delay, number of handovers, and bandwidth utilization. In this paper, we did not give a common evaluation because each scheme has a specific goal and consequently they use different metrics for the evaluation.

# 6 Acknowledgment

This work is supported by the ANEMONE project, which is partly funded by the Sixth Framework Program of the European Commission's Information Society Technology.

# References

[1] M. Annoni, R. Hancock, T. Paila, E. Scarrone, R. Tonjes, L. Dell'Uomo, D. Wisely, and R. Mort. Radio access networks beyond 3g: a first comparison of architectures. *Personal, Indoor and Mobile Radio Communications, 2001 12th IEEE International Symposium on*, 2:G–133–G–140 vol.2, Sep/Oct 2001.

[2] Leijia Wu and Kumbesan Sandrasegaran. A survey on common radio resource management. *Wireless Broadband and Ultra Wideband Communications, 2007. AusWireless 2007. The 2nd International Conference on*, pages 66–66, 27-30 Aug. 2007.

[3] Gustavo Carneiro, Carlos Garca, Pedro Neves, Zhiqui Chen, Michelle Wetterwald, Manuel Ricardo, Pablo Serrano, Susana Sargento, and Albert Banchs. The DAIDALOS Architecture for QoS over Heterogeneous Wireless Networks. In *Proceedings of the 14th IST Mobile and Wireless Communications Summit*, June 2005.

[4] A.-E.M. Taha, H.S. Hassanein, and H.T. Mouftah. On robust allocation policies in wireless heterogeneous networks. *Quality of Service in Heterogeneous Wired/Wireless Networks, 2004. QSHINE 2004. First International Conference on*, pages 198–205, 18-20 Oct. 2004.

[5] G. Koundourakis, D.I. Axiotis, and M. Theologou. Network-based access selection in composite radio environments. *Wireless Communications and Networking Conference, 2007.WCNC 2007. IEEE*, pages 3877–3883, 11-15 March 2007.

[6] Xiaoshan Liu, Victor O. K. Li, and Ping Zhang. Nxg04-4: Joint radio resource management through vertical handoffs in 4g networks. *Global Telecommunications Conference, 2006. GLOBECOM '06. IEEE*, pages 1–5, Nov. 2006.

[7] A. Kassler, A. Schorr, L. Chen, C. Niedermeier, C. Meyer, M. Helbing, and M. Talanda. Multimedia communication in policy based heterogeneous wireless networks. *Vehicular Technology Conference, 2004. VTC 2004-Spring. 2004 IEEE 59th*, 5:2987–2992 Vol.5, 17-19 May 2004.

[8] Yen-Wen Chen, I-Hsuan Peng, and Shiu-Tsuen Guan. Dynamic bandwidth management for handoffs with rsvp in 802.16/wlan environment. *Advanced Information Networking and Applications Workshops, 2007, AINAW '07. 21st International Conference on*, 2:243–248, 21-23 May 2007.

[9] D. Niyato and E. Hossain. A cooperative game framework for bandwidth allocation in 4g heterogeneous wireless networks. *Communications, 2006. ICC '06. IEEE International Conference on*, 9:4357–4362, June 2006.

[10] Dusit Niyato and Ekram Hossain. Wlc04-5: Bandwidth allocation in 4g heterogeneous wireless access networks: A noncooperative game theoretical approach. *Global Telecommunications Conference, 2006. GLOBECOM '06. IEEE*, pages 1–5, Nov. 2006.

[11] Xu Yang, J. Bigham, and L. Cuthbert. Resource management for service providers in heterogeneous wireless networks. *Wireless Communications and Networking Conference, 2005 IEEE*, 3:1305–1310 Vol. 3, 13-17 March 2005.

[12] O. Ormond, P. Perry, and J. Murphy. Network selection decision in wireless heterogeneous networks. *Personal, Indoor and Mobile Radio Communications, 2005. PIMRC 2005. IEEE 16th International Symposium on*, 4:2680–2684 Vol. 4, 11-14 Sept. 2005.

[13] R. Malyan A.L. Wilson, A. Lenaghan. Optimising wireless access network selection to maintain qos in heterogeneous wireless environments. *Wireless Personal Multimedia Communications, 2005*, 18-22 Sep. 2005.

[14] Qingyang Song and A. Jamalipour. A network selection mechanism for next generation networks. *Communications, 2005. ICC 2005. 2005 IEEE International Conference on*, 2:1418–1422 Vol. 2, 16-20 May 2005.

[15] G.J. Hoekstra, O. sterb, R. Schwendener, J. Schneider, F.J.M. Panken, and J. van Bemmel. Quality of service solution for open wireless access networks. *European Signal Processing Conference, EUSIPCO 2005*, 4-8 Sep. 2005.

[16] R. Ferrus, A. Gelonch, O. Sallent, J. Perez-Romero, N. Nafisi, and M. Dohler. A feasible approach for qos management in coordinated heterogeneous radio access networks. *Performance, Computing, and Communications Conference, 2005. IPCCC 2005. 24th IEEE International*, pages 607–614, 7-9 April 2005.

[17] Xia Gao, Gang Wu, and Toshio Miki. Qos framework for mobile heterogeneous networks. *Communications, 2003. ICC '03. IEEE International Conference on*, 2:933–937 vol.2, 11-15 May 2003.

[18] I.F. Akyildiz, S. Mohanty, and Jiang Xie. A ubiquitous mobile communication architecture for next-generation heterogeneous wireless systems. *Communications Magazine, IEEE*, 43(6):S29–S36, June 2005.

[19] K. D. Singh. *Improving Quality of Service and Resource Utilisation for Multimedia Streaming over Third Generation Mobile Networks*. PhD thesis, University of Rennes 1, 2007.

[20] P. Magnusson, J. Lundsjo, J. Sachs, and P. Wallentin. Radio resource management distribution in a beyond 3g multi-radio access architecture. *Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE*, 6:3472–3477 Vol.6, 29 Nov.-3 Dec. 2004.

# 7 Annex

## 7.1 Function-based Approach

**Access and interface selection algorithm**

The authors denote requested service as s, belonging to the total of services S, and ap is access point belonging to the total of access points AP, the objective function OF is given by (1).

$$OF(\forall s \in S, \forall ap \in AP) = F(s, ap) + OF(\forall s' \in S, s' \neq s, \forall ap \in AP) \quad (1)$$

The value of the OF for s' represents the allocation of the rest of services. The sequence by which the OF is calculated affects the overall result, because the allocation of an application to an AP decreases its available bandwidth. Thus, all possible permutations must be considered. Function F consists of the quality part Q and the part of policies PT, with their corresponding weights ($w_q + w_{pt} = 1$).

$$F = w_q Q + w_{pt} PT \quad (2)$$

Furthermore, functions Q and PT are analyzed as:

$$Q = w_{bi} BI + w_{di} DI + w_{sqi} SQI \quad (3)$$
$$PT = w_{cci} CCI + w_{npi} NPI + w_{tti} TTI \quad (4)$$

Each term in summations of (3) and (4) represents a specific factor that is calculated as a product of an indicator with its corresponding weight. Note that $w_{bi} + w_{di} + w_{sqi} = 100$ and $w_{cci} + w_{npi} + w_{tti} = 100$. BI = bandwidth indicator, DI = delay indicator, SQI = signal quality indicator, CCI = cost and compatibility indicator, NPI = network provider indicator, and TTI = terminal type indicator.

**Profit function**

The authors associate each handoff with a profit (P), which is decided by a target function f with two parameters: the bandwidth Gain (G) and the handoff Cost (C).

$$P = f(G, C) \quad (1)$$

1. Bandwidth Gain

   Parameters used in the calculation of the gain include:

   $N_i$: the ith access network, In the model, author use two networks $i = 1, 2$;

   $n_i$: the maximum bandwidth that can be provided to a single user by $Ni$;

   $\eta_i$: the percentage of transmission capacity that has been used in $N_i$;

   $r_M$: an application's maximum bandwidth requirement; lower QoS levels with $r_{M-1}, r_{M-2}, ..., r_1$ may be tolerated;

   $m(i, t_k)$: the bandwidth of $N_i$ used by mobile node during 2 handoff decision epochs $[t_k, t_{k+1})$

Definition of the bandwidth gain $G$ of a handoff decision at $t_k$ as

$$G_i(t_k) = \begin{cases} m(i, t_k) - m(j, t_{k-1}) & k \geq 1, \\ m(i, t_k) & k = 0 \end{cases} \quad (2)$$

In (2), $k = 0$ means a mobile node is initiating its connection for the first time. $G_i(t_k)$ gives the difference in bandwidth between the next period and this period.

2. Handoff Cost
   In the model, the authors define the handoff cost as data volume lost due to handoff delay.
   $$C(t_k) = m(i, t_{k-1})d(x, y) \quad (3)$$
   Where $d(x, y)$ is the handoff delay when an MN makes a handoff from base station $x$ to $y$.

3. Profit Function
   $$Pi = (t_{k+1} - t_k)G_i(t_k) - m(i, t_{k-1})d(x, y) \quad (4)$$

**Degradation Utility**

| **Application** | **Excellent** (kbit/s) | **Good** (kbit/s) | **Basic** (kbit/s) | **Rejected** |
|---|---|---|---|---|
| Voice | 30 | 30 | 30 | 0 |
| Video | 2000 | 384 | 256 | 0 |
| Data | 100 | 50 | 10 | 0 |

Table 6: Bandwidth for different quality of service

| **Quality Level** | **Voice** | **Video** | **Data** |
|---|---|---|---|
| Excellent | 300 | 700 | 1000 |
| Good | 300 | 600 | 800 |
| Basic | 300 | 500 | 400 |
| Forced Disconnection | -5000 | -5000 | -5000 |
| Handover Drop | -5000 | -5000 | -5000 |
| Reject | -2500 | -2500 | -2500 |

Table 7: Setting rewards for user priority class 1

Referring to Table 6 and 7. Let's take an example, consider a connection: User priority 1; application type: video; quality level: excellent. When the connection is degraded to good quality level:

Released bandwidth = 2000-384 kbit/s = 1616 kbit/s
Lost reward points = 700-600 = 100
Degradation utility = 1616/100 = 16.16

**Consumer Surplus**

Figure 7 outlines the consumer surplus based decision strategy. The authors first compute the predicted completion time ($T_c$) and thus predicted utility ($U_i$) and customer surplus ($CS$) for each candidate network. $T_{c-ideal}$ denotes user ideal transfer completion time and $T_{c-max}$ the maximum transfer completion time that a user is willing to wait.

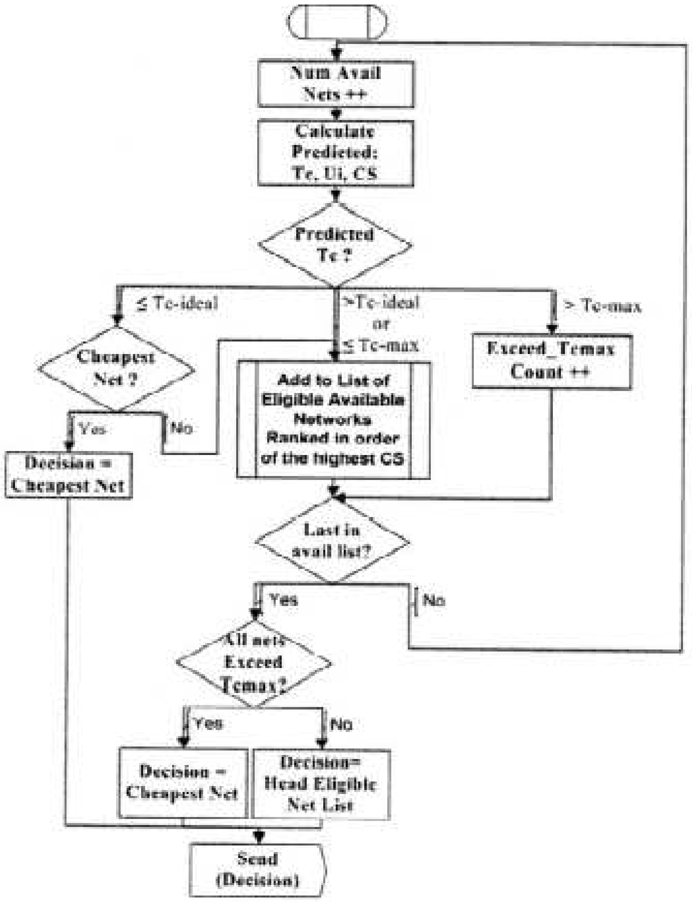


Figure 7: Dicision strategy based on predicted $T_c$

The authors define the transfer completion time (in seconds) as Tc. It is related to size of the file and depends on the rate according to the equation below.

$T_c = F_i/r$

Where $F_i$ is size of file $i$ in bits and $r$ is average rate for total transfer in bps.

The user aims to maximize the $CS$, subject to user constraints of time deadline for completed file transfer, it is calculated as:

$CS = U_i(T_c) - C_i$

subject to $T_c \leq T_{c-max}$

Where CS is consumer surplus in cent, $U_i(T_c)$ is the monetary value (in cent) that the user places on the transfer of file $i$ in the given the transfer completion time $(T_c)$ and $C_i$ is the cost charged by the network, also in cent, for the competed file transfer.

## 7.2 Mathematical-based Approach

**Fuzzy Logic Control**

1. Fuzzification
   Fuzzification is the process of transforming crisp values into fuzzy set; we give example of the output of this step (SNR and Data rate) in Figure 8 and 9 repectively.

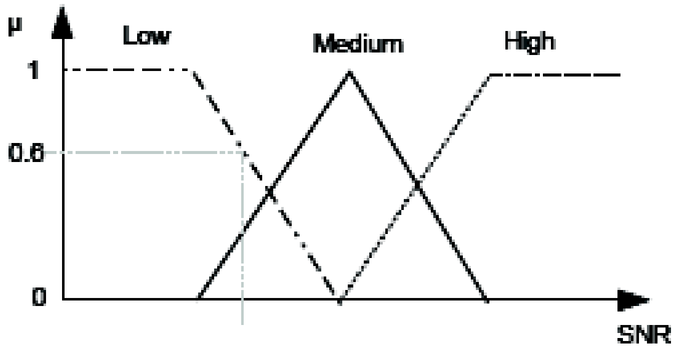


Figure 8: Membership functions of SNR

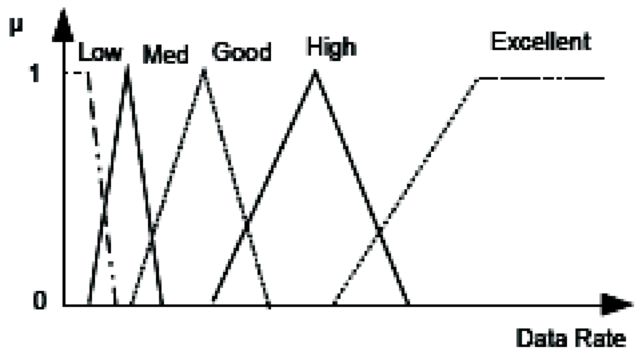


Figure 9: Membership functions of data rate

2. Rulebase
   Once crisp values are fuzzified, membership of a fuzzy set is then used as input to the logic rulebase, which are collection of linguistic IF-THEN rules similar to Figure 10.

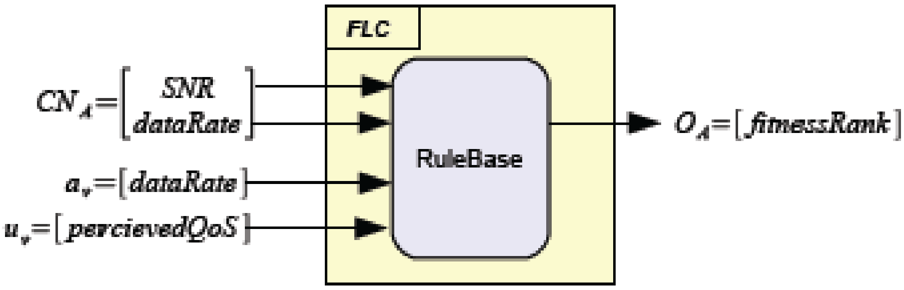


Figure 10: Schematic rulebase of FLC

3. Defuzzification The overall ranking is obtained through defuzzification using the weighted average method trigger rules to degree, which is added to a sum of all triggered rule weights. The output is illusted with Figure 11 below.

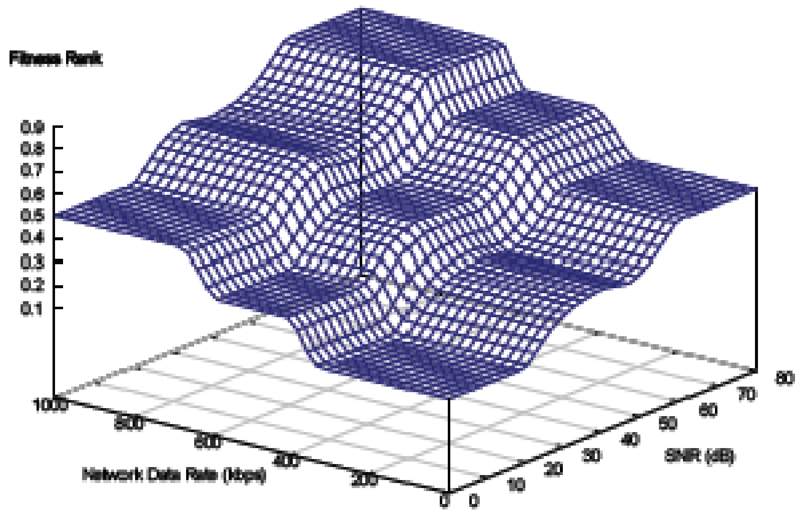


Figure 11: Fitness ranking as a surface for values of SNR and data rate

**Stochastic Linear Programming**
Here is the formulation for single common service (SCS) with probabilistic demands. Let $S$ be the set of all possible scenarios. In every scenario $s \in S$ the demand $D_{ij}(s)$ takes on specific values with a predetermined probability. The probability that the current $D_{ij}(s)$ is a specific value, i.e. $P(D_{ij}(s) = D) = p_{ij}(s)$. The demand uncertainty can be imposed on Program SCS-DD through the allocation-rejection-demand constraints, where the penalty can be applied to the rejection. In this manner, the penalty (cost) of unit rejection is $z_{ij}(s)$. As such, the return function to be maximized becomes

$$\prod_{SCS-PD} = \sum_{\forall i,j} x_{ij} A_{ij} - \sum_{c=j,\nu} y_c U_c - \sum_{\forall i,j} \sum_{s \in S} p_{ij}(s) z_{ij}(s) R_{ij}(s)$$

With this, the SLP can be stated as follows Program SCS-PD Max $\prod_{SCS-PD}$
Subject to

$$A_j + U_j = \frac{B_j}{Q} \qquad \forall j$$

$$\sum_{i \neq j} A_{ij} + U_\nu = B_\nu$$

$$A_{ij} + R_{ij}(s) = D_{ij}(s) \qquad \forall i, j, s \in S$$

All variables are positive integer.

### AHP and GRA

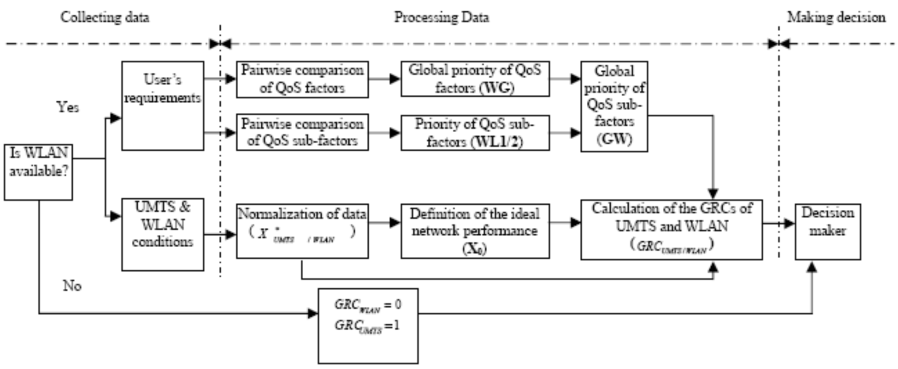


Figure 12: AHP and GRA based network selection model

### Game theory

– Bandwidth allocation algorithm:

  Based on a standard bankruptcy game, the authors propose a bandwidth allocation algorithm for a new connection which can be served simultaneously by three different wireless access networks (WLAN, cellular network, and WMAN). So the total number of agents is $N = 3$ and the set of agents is defined as $A = wl, ce, wm$ for WLAN, cellular network, and WMAN respectively.

  When a new connection requests for bandwidth M, a central controller determines the amount of offered bandwidth from each network using the following equation:

$$d_i = \begin{cases} \tilde{b}_{k,i}, & \tilde{b}_{k,i} < (B_i^{(a)})^r \\ (B_i^{(a)})^r + \aleph(B_i^{(a)} - (B_i^{(a)})^r) & \tilde{b}_{k,i} \geq (B_i^{(a)})^r \end{cases}$$

  Where $\tilde{b}_{k,i}$ is the predefined offered bandwidth by network $i$ to a new connection with subscription $k$, $(B_i^{(a)})$ is the available bandwidth in network $i$, $b_k^{(req)}$ is the amount of requested bandwidth in class

$k$, $\aleph$ is a uniform random number between zero and one, and $r$ is a control parameter which will be referred to as the bandwidth shaping parameter ($0 < r \leq 1$).

Note that, with the above definition of offered bandwidth, network $i$ offers bandwidth $\tilde{b}_{k,i}$ to a new connection under normal traffic load situation. However, when the network becomes congested (i.e., defined by the condition $\tilde{b}_{k,i} > (B_i^{(a)})^r$) the offered bandwidth is gradually shaped by the random number $\aleph$ and the shaping parameter $r$ to ensure that the network does not offer too much bandwidth to the new connection. In the proposed bandwidth allocation algorithm, the Shapley value becomes the amount of allocated bandwidth in each network $i$, i.e., $x_i = \phi_i(\nu), \forall i \in A$. The notations and the descriptions of the variables for the bankruptcy game and the bandwidth allocation algorithm are shown in Table 8.

| **Variable** | **Bankruptcy Game** | **Bandwidth Allcoation** |
|---|---|---|
| $n$ | total number of agents | total number of network |
| $M$ | money (estate) | requested bandwidth |
| $A$ | set of agents | set of networks |
| $d_i$ | claims of agent i | offered bandwidth by network i |
| $x_i$ | solution of money distributed to agent i | Bandwidth allocated to new connection in network i |

Table 8: Notation and description of the variables for bankruptcy game and proposed bandwidth allocation algorithm

– Admission control algorithm:

  The admission control algorithm ensures the requested bandwidth can be satisfied. When a mobile initiates a new connection, the information on the required bandwidth is sent to the central controller, which computes the offered bandwidth by each network. Then, the Shapley value is obtained. The new connection is accepted if $\sum_{i \in A} x_i \geq b_k^{(req)}$ and $x_i \in C, \forall i \in A$ (i.e., the Shapley value is in the core, namely, the solution is stable), and rejected otherwise.

Centre de recherche INRIA Rennes – Bretagne Atlantique
IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Saclay – Île-de-France : Parc Orsay Université - ZAC des Vignes : 4, rue Jacques Monod - 91893 Orsay Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)

ISSN 0249-6399