WIRELESS WORLD
RESEARCH FORUM

Artemis Koutsorodi, Evgenia Adamopoulou, Konstantinos Demestichas,
Michael E. Theologou

Terminal-Controlled Access Network Selection in Heterogeneous Networks

***Abstract*—This paper presents a mobile terminal architecture for devices operating in heterogeneous environments, which incorporates intelligence for supporting mobility and roaming across legacy access networks. It focuses on the structure and functionality of the proposed scheme that supports terminal-initiated and terminal-controlled access network selection in heterogeneous networks. It discusses the decomposition of the proposed Terminal Management System into separate modules, responsible for retrieving link-layer measurements from the attachment points in the terminal's neighborhood, for handling the user's profile and for performing intelligent access network selection. This latter function aims at independently determining the optimal local interface and attachment point through which applications can be obtained as efficiently as possible, by taking into account network status and resource availability, user preferences and service requirements.**

***Index Terms*—access selection, always best connected, optimization process, user preferences.**

Introduction

THE next generation of mobile systems is expected to comprise heterogeneous networks consisting of diverse radio segments, able to host multimode wireless terminals, each of them capable of alternatively operating in the diverse radio segments available in the infrastructure. The different radio segments, or access technologies (e.g. WLANs, cellular and broadcast networks), will be interconnected by a backbone (e.g. an IP-based fixed network) and jointly operated in an optimized fashion that will allow for an improved overall resource management [1].

A challenging issue related to the above is the development of management frameworks, both for the terminal and the network ([2],[3]), that will enable ubiquitous service provisioning regardless of the network the end user is connected to, thus allowing the latter to benefit from being able to access his/her subscribed services anywhere and anytime. Another important aspect is the fact that the user will need to control the usage of the available networks, especially when this usage comes with a price [4]. This involves a potentially complex decision making process which may be guided by policy management tools, with support from both the user terminals and the networks, thus giving rise to the significant issue of optimally distributing intelligence between the network and the mobile terminal in order to support seamless mobility and service provisioning.

Consequently, the exploitation of the composite radio infrastructure requires innovative management schemes. The deployment of a central network and service management system represents one solution. An alternative, decentralized approach is to capitalize on the growing capabilities and computational power of today's mobile terminals to remove some of the management work-load from network equipment and to distribute it to the terminals [5].

In this paper, we argue that appropriate functionality must be in place at the mobile terminal to handle basic mobility management tasks and to support the applications in dealing with the dynamics and heterogeneity of available access networks. We discuss a management architecture for composite radio infrastructures that incorporates an innovative Terminal Management System (TMS), located at the mobile terminal and capable of dynamically and independently selecting the appropriate access network through which services can be obtained efficiently in terms of cost and QoS, in a transparent manner.

There are several arguments for the terminal to incorporate functionality for performing access network selection: First of all, the terminal is the entity that is aware of the different access technologies in its surroundings for two reasons: (a) it knows which hardware interfaces it has implemented, and (b) it can detect the availability of access networks in its physical surroundings [6]. Moreover, as the decision to initiate a handover, as well as the handover target selection process itself, may be based on user preferences, the terminal may provide the user with his options on a GUI, and the user may also be able to switch interfaces on or off manually and even dynamically alter his preferences.

Optimal system operation is to be achieved through the joint contributions of both the network and the terminals: On the network's side, the network management entity will perform tasks such as the balancing of the traffic load between the different access networks, and the general coordination thereof – the goal being optimal joint resource management. On the terminal's side, the TMS will be responsible for detecting available access networks in the mobile terminal's neighbourhood, for making the optimal selection, based on network availability, services and user preferences, of the appropriate interface(s) through which services may be provided to the user, and for interacting with the management system, both for submitting its decisions for approval and for receiving network-initiated handover decisions. This interaction will ensure that the network's and the terminal's estimations of radio conditions and QoS levels in the infrastructure are beneficially combined for making an educated selection of the appropriate access technologies through which services can be obtained as efficiently as possible. Thus, both the network and the terminals contribute intelligence towards optimal system operation.

Terminal Management Architecture

Overview

Fig. 1 illustrates the main components of the TMS architecture, namely the Network Interface Adaptation Module (NIAM), the Mobility Management Module (MMM) and the User Preferences Module (UPM).

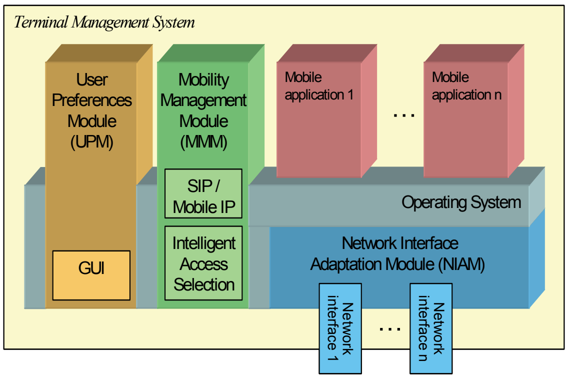

Fig. 1. TMS architecture

The NIAM is responsible for providing the terminal with a level of abstraction from the different network drivers, the UPM for accessing and processing the user profile, and the MMM for handling mobility events and for optimally computing handover targets.

Network Interface Adaptation Module (NIAM)

The NIAM is responsible for identifying the different interfaces present in the terminal, for monitoring their status, for collecting measurements from each interface and for executing the selection and de-selection of the appropriate interface. More specifically, the NIAM serves two purposes: (a) the connection and de-connection of the appropriate interface during power-up of the mobile terminal or during a handover and (b) the retrieval of layer-2 measurements in the network interface. This means that the NIAM is able to provide the terminal with measurements retrieved from the different network drivers, reflecting the signal quality or connectivity status in each of these interfaces in an abstracted way (e.g. good, average or poor signal strength). In all, the NIAM is able to provide the terminal with a list for each attachment point in the terminal's neighbourhood, each list comprising information about the attachment point's signal strength and bandwidth availability, its type of technology, its network operator and about the cost at which it is able to provide each available service at each of the service's permissible quality levels. This set of parameters, coupled with some additional pieces of information as will be explained in the following, constitutes a minimal set based on which the terminal can perform intelligent access selection in a multiple access environment.

Therefore, the NIAM should incorporate a mechanism for retrieving such measurements from attachment points in its range (such as the Candidate Access Router Discovery protocol [7]) and for processing this information in order to detect if a new attachment point has appeared in the terminal's neighbourhood, if an old one has disappeared, or if the perceived signal strength from an already selected attachment point has severely deteriorated. In all of these cases, the MMM is notified accordingly for the purpose of triggering the process of optimally distributing all running applications to appropriate interfaces.

User Preferences Module (UPM)

The UPM is responsible for storing, accessing and editing the user's profile. A graphical user interface (GUI) through which the user will be able to specify his preferences, will allow the user to give different priorities to parameters that may influence the access network selection process. This prioritization is equivalent to the specification of values for the different coefficients w_q, w_o, w_t and w_c, which correspond to parameters 'quality' (i.e. the maximum QoS that the user is willing to be charged for), preferred 'network operator', preferred 'technology type' and 'cost' (i.e. the maximum cost the user is willing to pay) respectively, and represent the measure by which each one of these parameters is weighted in the access network selection algorithm. For example, if the user chooses to specify that at a given moment 'quality' is for him the most important factor in access network selection, 'technology type' comes second, 'cost' comes third and last comes 'network operator', then the respective coefficients will be assigned values $w_q > w_t > w_c > w_o$.

Mobility Management Module (MMM)

The MMM is responsible for handling all events relating to access network selection. This functionality is carried out by the Intelligent Access Selection (IAS) function. The latter is responsible for optimally selecting the mobile terminal's local interface (technology) and the network's point of attachment (access router, access point), both in the case of an intra-technology handover (horizontal HO) and in the case of an inter-technology handover (vertical HO).

The MMM is responsible for providing the IAS function with the required input (retrieved from the UPM and the NIAM), for triggering its execution and, finally, for relaying its decisions to the NIAM for handover execution.

Intelligent Access Selection

The IAS function is triggered in the following cases: (a) when a new service request appears; (b) when the user changes his/her profile, and therefore the weights w_q, w_o, w_t, w_c attributed to the different selection criteria are altered; (c) when the NIAM issues a notification of severe signal degradation or complete signal loss; (d) when the availability of a new attachment point or the unavailability of an old one is detected by the NIAM. Whenever one of the aforementioned events occurs, the IAS algorithm is executed for the purpose of finding the optimal attachment point both for the provision of the newly requested service (in the case of trigger (a)), and for the possible handover of the already running services to newly computed optimal attachment targets. The optimization process carried out by the IAS is illustrated in Fig. 2.

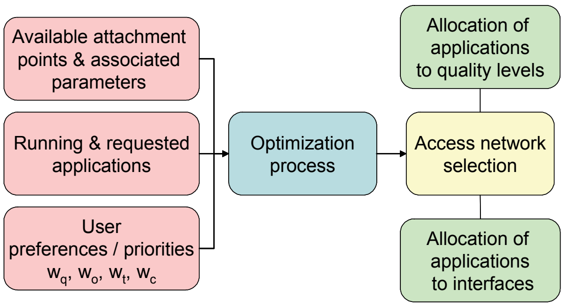

Fig. 2. Intelligent Access Selection optimization process

The optimization problem relies on the following input data: (a) a set of measurements reflecting the availability, signal quality and other parameters perceived from each of the available attachment points, as provided by the NIAM; (b) the set of applications that are already running on the mobile terminal, the corresponding quality levels at which these applications are being provided as well the set of applications that the user is requesting to use; (c) the set of user preferences, according to which the parameters 'quality', 'network operator', 'technology type' and 'cost' are prioritized.

The optimally computed access network selection is equivalent to an optimal allocation of both the requested and the already running services to appropriate quality levels, and an optimal allocation of the requested services to network interfaces. These allocations are bound to certain constraints: (a) every application should be allocated to an acceptable quality level; (b) the quality level requested by the user should be guaranteed; (c) the user's preferences should be satisfied in the specified order.

The aforementioned allocations should optimize an objective function associated with the weights attributed to the different selection criteria, and computed for each of the requested/running applications separately. Let P be the set of attachment points that the terminal perceives, $P = \{p_1, p_2, ..., p_n\}, n \in \mathbb{N}$, and $Q(p)$ the set of quality levels at which attachment point p can offer the service under consideration, $Q(p) = \{q_1, q_2, ...q_m\}, m \in \mathbb{N}$. The goal is the computation of:

$$OF(p,q) = w_q \times \text{Quality}(p,q) + w_o \times \text{Operator}(p)$$
$$+ w_t \times \text{Technology}(p) - w_c \times \text{Cost}(p,q)$$
(1)

for all $p \in \text{P}$ and $q \in \text{Q}(p)$, and the determination of: $\max\limits_{\forall p \in P} \left\{ \max\limits_{\forall q \in Q(p)} \{OF(p,q)\} \right\}$ as the optimal attachment point and quality level for each of the requested/running services.

In the proposed implementation, coefficients w_q, w_o, w_t and w_c are assigned values 0.8, 0.6, 0.4 or 0.2, according to the position of factors 'quality', 'cost', 'technology type' and 'network operator' in a prioritized list. I.e. the coefficient that corresponds to the first and most important factor will be assigned the value 0.8, the coefficients corresponding to the second, third and fourth factors will be assigned the values 0.6, 0.4 and 0.2, respectively.

Factor $\text{Quality}(p,q)$ in (1) is not an expression of the nominal quality level offered by each attachment point, but rather an expression of the combined effect of the nominal quality level and the perceived signal strength from each attachment point. Therefore:

$$\text{Quality}(p,q) = q_s \times q$$

where q_s expresses the strength of the received signal and q expresses the quality level at which attachment point p can offer the service under consideration. If 5 different levels for signal strength are assumed, coefficient q_s may be assigned the following values:

$$q_s = \begin{cases} 1.0, \text{ for signal level 5 (excellent signal)} \\ 0.95, \text{ for signal level 4} \\ 0.8, \text{ for signal level 3} \\ 0.6, \text{ for signal level 2} \\ 0.4, \text{ for signal level 1 (weak signal)} \end{cases}$$

Factor $\text{Cost}(p,q)$ in (1) represents the cost of a specific allocation decision, i.e. the cost at which attachment point p can offer the service under consideration at quality level q. In the proposed implementation, information about the cost at which services are offered is received from the network every time the mobile terminal powers up, or at regular time intervals (e.g. once a day), and is stored in the terminal in the form of an XML document. The data in this document correspond to the cost of a service being provided at a specific quality level, by a specific network operator and through a specific technology, per data volume unit (e.g. Kb) or per time unit (e.g. sec).

The cost values that are retrieved from the XML document are normalized by the optimization algorithm before being used in the computation of the objective function. The normalization process is necessary in order to convert the values of factor $\text{Cost}(p,q)$ to a scale that renders them comparable to the values of $\text{Quality}(p,q)$.

Let $\text{actualCost}(p,q)$ be the cost per unit of the requested service, as it is retrieved from the XML document, maxCost the maximum cost per unit for this specific service and noQoSLevels the number of the different QoS levels that the service can be provided at. Then:

$$\text{Cost}(p,q) = \frac{\text{noQoSLevels}}{\text{maxCost}} \times \text{actualCost}(p,q)$$

Finally, as far as the factors $\text{Operator}(p)$ and $\text{Technology}(p)$ in (1) are concerned, the user has the capability to specify a preferred network operator and a preferred technology type, through a graphical user interface. In case a candidate attachment point belongs to the preferred operator and/or supports the preferred technology, it is granted a 'bonus'. The value of this bonus is a percentage of the absolute value of the following difference:

$$w_q \times \text{Quality}(p,q) - w_c \times \text{Cost}(p,q)$$

and in the proposed implementation stands as follows:

$$\text{Provider}(p) = \begin{cases} 0.5 \times Abs[w_q \times \text{Quality}(p,q) \\ -w_c \times \text{Cost}(p,q)], \text{ if they match} \\ 0, \text{ otherwise} \end{cases}$$

$$\text{Technology}(p) = \begin{cases} 0.25 \times Abs[w_q \times \text{Quality}(p,q) \\ -w_c \times \text{Cost}(p,q)], \text{ if they match} \\ 0, \text{ otherwise} \end{cases}$$

Simulation

The IAS algorithm's functionality is tested within the framework of a scenario that simulates a typical day in the life of a university professor, Prof. Green.

Scenario description

Prof. Green is in the university campus, walking from his laboratory towards the classroom where he is due to teach a class, and stopping in the meantime by his office and the underground parking area. During his walk inside the campus, he makes use of several services via his 4G terminal, supplied at the most appropriate quality levels and through connections to the best available access networks (GSM, UMTS or WLAN).

The user profile in use during this time specifies that 'quality' is the most important factor in access network selection, 'cost' comes second, 'network operator' comes third and 'technology type' comes fourth. Prof. Green has not specified any preferred choice, neither for 'network operator' nor for 'technology type'.

In the following, every step of the scenario, along with the corresponding allocations of services to networks and quality levels, are described in more detail. In most cases, the IAS algorithm is triggered by a new service request. For every allocation that is computed, the algorithm's execution time is given within parentheses (in ms). As may be observed, in all cases this time is less than half a second, and could be reduced to even less if the execution time of the result recording process in a log file were not taken into account.

Results

Table I lists the total number of quality levels and the maximum allowed quality level for each service. This latter setting is specified by the user profile and also determines the maximum cost that the user is willing to pay for each service. The idea here is that the delivery of a service at any quality level that satisfies all other user requirements is preferable to a service request being denied.

Table II presents a list of all available attachment points, during the course of the scenario. For each attachment point, the signal level in each step of the scenario is given in a scale from 1 (weakest) to 5 (strongest).

(a) Inside the lab

Prof. Green initiates a web browsing service, in order to download a large file.

IAS is triggered by a new service request (web browsing).

IAS selects:

- WLAN #3, web browsing, QoS-4 (390ms)

Reason for this allocation: WLAN technology is preferred because it offers high bit rates at low cost. Among the 3 available WLAN attachment points, the one that offers the same quality level at the lowest cost is selected. Quality level 4 is the maximum allowed.

(b) Near the lab

Prof. Green exits the laboratory and sends an SMS to a colleague, to remind him of their meeting later in the day.

IAS is triggered by a new service request (SMS).

IAS selects:

- WLAN #3, web browsing, QoS-4
- UMTS #2, SMS, QoS-1 (234ms)

Reason for this allocation: Web browsing continues to be delivered through the same attachment point as before. Short Message Service is provided through a UMTS attachment point, which outranks all other available attachment points in signal strength (Prof. Green has specified 'quality' as his no. 1 priority), with the exception of two available GSM attachment points, compared to which though, it provides the requested service at a lower cost.

(c) On the way to the office

Prof. Green calls his assistant and informs him that he is running a bit late for class.

IAS is triggered by the unavailability of access point WLAN #3, through which web browsing is being delivered, and by a new service request (voice call).

IAS selects:

- GSM #2, web browsing, QoS-3 (94ms)
- GSM #1, voice call, QoS-3 (78ms)

Reason for this allocation: As the attachment point through which web browsing was being delivered suddenly becomes unavailable, attachment point GSM #2 now becomes the best option for this service, as it provides good signal strength and comparatively low cost. GSM #1 is selected for the voice call, as it also combines strong signal and low cost.

These two services are allocated to different attachment points because GSM #1 is cheaper for voice call delivery but more expensive for web browsing, compared to GSM #2. Quality level 3 for web browsing is the maximum that can be offered from a GSM access point, whereas for the voice call it is the maximum allowed based on the user's preferences.

TABLE I
AVAILABLE AND MAXIMUM ALLOWED QUALITY LEVELS FOR SERVICES IN SCENARIO

	voice call	SMS	MMS	e-mail	video call	Video streaming	Web browsing
Total number of quality levels	5	1	2	1	5	5	5
Maximum allowed quality level	3	1	2	1	4	3	4

TABLE II
SIGNAL LEVELS FROM AVAILABLE ATTACHMENT POINTS IN SCENARIO STEPS

Access Point	Operator	(a) Inside the lab	(b) Near the lab	(c) On the way to the office	(d) Inside the office	(e) Among trees	(f) In the garage	(g) Towards the classroom	(h) Inside the classroom
GSM #1	#1	4	5	5	4	1	1	3	2
GSM #2	#2	5	5	5	5	n/a	n/a	4	3
UMTS #1	#1	5	4	4	4	1	1	2	3
UMTS #2	#3	4	5	5	5	1	n/a	4	3
WLAN #1	#4	5	1	n/a	n/a	n/a	n/a	n/a	n/a
WLAN #2	#5	5	2	n/a	n/a	n/a	n/a	n/a	n/a
WLAN #3	#6	5	1	n/a	n/a	n/a	n/a	n/a	n/a
WLAN #4	#6	n/a	n/a	n/a	5	n/a	n/a	n/a	n/a
WLAN #5	#5	n/a	n/a	n/a	n/a	n/a	n/a	n/a	5

(d) Inside the office

As he enters his office, Prof. Green starts looking for his lecture notes but can't seem to find them anywhere. He initiates a video call to his assistant to ask him if he has any idea where the notes may be. His assistant cannot be of any help either so the call is terminated.

IAS is triggered by the availability of a new access point (WLAN #4).

IAS selects:

- WLAN #4, web browsing, QoS-4 (188ms)
- WLAN #4, video call, QoS-4 (125ms)

Reason for this allocation: WLAN technology is preferred because it offers high bit rates at low cost. Both services are delivered in the maximum allowed quality level.

(e) Among trees

The way to the underground parking area goes through a park with thick trees, so the received signal from some access points deteriorates significantly. Prof. Green decides to take advantage of the time needed to get to his car, so he initiates a video streaming service to watch the lecture of a colleague.

IAS is triggered by the unavailability of access point WLAN #4, through which web browsing is being delivered.

IAS selects:

- GSM #1, web browsing, QoS-1 (63ms)
- GSM #1, video streaming, QoS-1 (63ms)

Reason for this allocation: Both services are allocated to the same attachment point and to quality level 1, because this solution provides, given the circumstance, the best combination of quality and cost. The allocation of both services to minimal quality levels is due to the poor quality of the perceived signal.

(f) In the garage

When he reaches the parking area and picks up his briefcase, Prof. Green calls his assistant to let him know that he found his notes but will be late for class. Video streaming is automatically put on hold while the voice call is ongoing.

IAS is triggered by the request for the voice call.

IAS selects:

- GSM #1, web browsing, QoS-1
- GSM #1, video streaming, QoS-1
- GSM #1, voice call, QoS-2 (47ms)

Reason for this allocation: Web browsing and video streaming continue to be delivered through the same attachment point, and the requested voice call is allocated to the same attachment point and to quality level 2.

(g) Towards the classroom

Prof. Green leaves the parking area and heads for the classroom.

IAS is triggered by the availability of new access points (GSM #2 and UMTS #2).

IAS selects:

- GSM #2, web browsing, QoS-3 (78ms)
- UMTS #2, video streaming, QoS-3 (94ms)
- UMTS #2, voice call, QoS-3 (106ms)

Reason for this allocation: The appearance of new available attachment points leads to the reconsideration of the current allocations of services to networks and quality levels. This results in the running services being allocated to higher quality levels, as the present availability of networks offers higher levels of quality.

(h) Inside the classroom

After hanging up the voice call, Prof. Green alters his profile by changing his choice of preferred technology type from 'none' to 'WLAN'. When he finally reaches the classroom he is pleased to see that his students haven't been discouraged by his delay. He terminates the video streaming and web browsing services and sends a short e-mail with questions to his colleague, whose lecture he was previously watching through video streaming.

IAS is triggered by the availability of a new access point (WLAN #5) and a new service request (e-mail).

IAS selects:

- WLAN #5, web browsing, QoS-4 (109ms)
- WLAN #5, video streaming, QoS-3 (110ms)
- WLAN #5, e-mail, QoS-1 (93ms)

Reason for this allocation: WLAN technology is preferred because it offers high bit rates at low cost. Web browsing and video streaming are allocated to service levels 4 and 3 respectively as these are the maximum allowed. This decision is also reinforced by the fact that Prof. Green's profile at that moment specifies that WLAN is the preferred technology type.

Conclusion

The terminal management architecture discussed in this paper incorporates functionality for retrieving and processing a minimal set of parameters based on which the terminal can perform intelligent access selection in a heterogeneous environment. This latter function is able to indicate the optimal allocation of services to network interfaces and quality levels in near real-time.

An interesting extension of this work involves the replacement of the graphical user interface through which the user specifies his preferences with a system that can autonomously and dynamically predict the user's priorities, according to the usage context and the service that is being requested. Consequently, the allocation decisions that the IAS algorithm reaches based on these predictions may be submitted to the user for approval or rejection, thus allowing the system to adapt to his pattern of behaviour.

References

[1] E. Gustafsson and A. Jonsson, "Always Best Connected", IEEE Wireless Communications, February 2003.

[2] G. Koundourakis, N. Koutsouris, V. Stavroulaki, L. Papadopoulou, V. Tountopoulos, D. Kouis, P. Demestichas, N. Mitrou, "Network and Service Management System for Optimising Service Delivery and Traffic Distribution in Composite Radio Environments" In Proc. IST Mobile &Wireless Telecommunications Summit 2003, Aveiro, Portugal, June 2003.

[3] P. Demestichas, N. Koutsouris, G. Koundourakis, K. Tsagkaris, A. Oikonomou, V. Stavroulaki, L. Papadopoulou, M. Theologou, G. Vivier, K. El Khazen, "Management of Networks and Services in a Composite Radio Context", IEEE Wireless Communications Magazine, August 2003 (Evolution towards 4G Mobile Communication Systems).

[4] J. van Bemmel, H. Teunissen, D.-J. Plas, A. Peddemors, "A Reference Architecture for 4G Services", In: 7th WWRF Meeting in Eindhoven, The Netherlands, December 2002.

[5] D. Chantrain, K. Handekyn, H. Vanderstraeten, "The Soft Terminal: Extending Service Intelligence from the Network to the Terminal", Alcatel Telecommunications Review, 2nd Quarter 2000.

[6] B. Busropan, J. van Loon, F. Vervuurt, R. van Eijk, "Access Network Selection in Heterogeneous Networks and the Role of the Operator", In: 9th WWRF Meeting in Zurich, Switzerland, July 2003.

[7] M. Liebsch, A. Singh (editors), H. Chaskar, D. Funato, E. Shim, "*The Candidate Access Router Discovery Protocol*", draft-ietf-seamoby-card-protocol-07.txt, work in progress.