Vertical Handover Supporting Pervasive Computing in Future Wireless Networks

S. Balasubramaniam J. Indulska
School of Information Technology and Electrical Engineering
The University of Queensland
Brisbane, QLD 4072
Australia

***Abstract*- A variety of current and future wired and wireless networking technologies can be transformed into seamless communication environments through application of context-based vertical handovers. Such seamless communication environments are needed for future pervasive/ubiquitous systems. Pervasive systems are context aware and need to adapt to context changes, including network disconnections and changes in network Quality of Service. Vertical handover is one of many possible adaptation methods. It allows users to roam freely between heterogeneous networks while maintaining the continuity of their applications. This paper proposes a vertical handover mechanism suitable for multimedia applications in pervasive systems. The paper focuses on the handover decision making process which uses context information regarding user devices, user location, network environment and requested QoS.**

I. Introduction

The current trend towards achieving pervasive/ubiquitous computing environments requires integration of a variety of current and future wired and wireless networking technologies to support seamless computing and communication environments for user applications. One of the requirements of seamless computing is communication network and device independence which allows users to (i) move freely between heterogeneous networks and (ii) change computing devices, if necessary, while maintaining application continuity. Such seamless computing and communication environments need to be aware of the context (situation) of the computing application and dynamically adapt to changes in this context. The issue is particularly difficult for multimedia streaming (e.g. video streaming) which has quite stringent requirements with regard to network Quality of Service (QoS) like bandwidth, delay, jitter and loss rate. In an environment in which the context of computation may change, (e.g. due to disconnections or mobility of users) the network QoS provided to communication streams may change and may no longer meet the application requirements. This needs to be compensated for by adapting the application, or its communication streams, or the networking environment, or by using a combination of these different types of adaptations.

One adaptation method which can support seamless communication infrastructures is vertical handover – handover between heterogeneous networks. Many current computing devices from PCs to PDAs and (future) mobile phones are equipped with more than one networking interface (e.g. Ethernet, WLANs, Bluetooth, GPRS, UMTS). The role of vertical handover is to dynamically redirect communication streams of a given application to a different networking interface. In the case of multimedia streaming vertical handovers are quite complex due to the stringent QoS requirements of these applications and have to depend on a rich set of context information. The system first needs to recognize that a handover is needed, then has to evaluate which of the available networks will provide the required QoS. If none of the networks fully meets the application QoS requirements some additional adaptation may be required. Finally, the whole vertical handover operation has to be designed to minimise QoS violation during the handover. Moreover, handover operations have to take into account various QoS provision schemes used in particular networking technologies.

In addition to wired networking technologies, there exist a number of wireless communication technologies which can support multimedia streaming. However, they all differ in QoS provided to such applications. GPRS (General Packet Radio Service), which is a 2.5G network, provides users with moderate data transmission rates suitable for most Internet applications but does not support video transmission well. The upcoming 3G UMTS (Universal Mobile Telecommunication System) network is an enhancement to the GPRS network and provides users with multimedia streaming capabilities (e.g. video on demand, video conferencing, etc.). Future 4G networks integrate all the different networking technologies under the IP protocol and will provide high bandwidth capabilities. Other wireless networking technologies like WLAN and Bluetooth also need to be a part of a seamless communication infrastructure, as they can be used either as a wireless extension of wired networks or peer-to-peer networks. The Bluetooth technology allows devices to form ad-hoc networks over a short range, allowing access to services without the aid of a central infrastructure [12, 16]. Vertical handover operations have a potential to change these disparate technologies into a seamless communication environment that dynamically selects the appropriate networking technology to meet application QoS requirements when the context of the application changes.

There already exist solutions for vertical handovers as indicated in the Related Work section. However, these solutions are based only on recognition of disconnections (signal strength), and this can take a reasonably long time to discover. Moreover, vertical handovers are applicable to a wider set of context changes, including (i) users moving out of network coverage (disconnection), or (ii) network QoS change to a level unacceptable for applications, or (iii) users changing devices while continuing their applications (support for user independence from devices), or (iv) users entering preferred networks (support for user preference for communication networks).

In this paper we describe a context-aware vertical handover for multimedia applications. The solution is based on the assumption that pervasive systems have to be context aware in order to support mobile users, devices and applications under varying computing environment conditions, i.e. these systems manage context information and evaluate context changes to select appropriate adaptation methods. Vertical handover is one of such adaptation methods.

This paper focuses on the handover decision making process. This process has to evaluate context information (i.e. user devices and their capabilities, user personal context information, application QoS requirements and user perceptibility of application QoS, user location, network coverage and network QoS) to decide whether handover is necessary, to which network, and whether any additional adaptations should be applied. The provided solution tries to predict disconnections based on user location and network coverage in order to minimise QoS violation during handover, and at the same time, can react to disconnections if it happens. The handover decision making process also reacts to QoS changes, device changes, and takes user preferences into account when mobile users enter the coverage of new networks.

As it is difficult for users to describe communication QoS required by their applications using typical network terms (delay, jitter, packet loss) [17], our approach uses the user perceived QoS and provides a mapping from the user perceived QoS to network QoS indices. The indices are used to make the handover decision. This allows users to specify QoS requirements for multimedia applications in an understandable, user friendly form.

The paper also briefly describes the QoS mapping between networks which is required for handovers and a mechanism employed to minimise QoS violations during handovers, however the detailed description of these two issues can be found in [4].

The structure of the paper is as follows. Section II describes related work on vertical handovers and provides a brief description of QoS issues in GPRS and UMTS networks. Section III describes a context model used in our solution. Section IV provides an overview of the whole vertical handover architecture. Section V describes the handover mechanism, followed by two examples of vertical handovers in Section VI. Section VII describes both our prototype for vertical handovers of video streams and experimental results which demonstrate the decision making process in relation to user mobility and varying network QoS. Finally, Section VIII concludes the paper.

II. RELATED WORK

The related work presented in this section addresses two issues: the existing research on vertical handovers and QoS support in GPRS and UMTS. The QoS support has an impact on QoS mapping between networks when communication streams are redirected during handovers.

A. Vertical Handover

Helal et. al [1], at the University of Florida, developed the Full Stack Adaptation (FSA) concept to allow both horizontal and vertical handovers between Ethernet, wireless LAN (WLAN), and wireless WAN. In order to allow seamless network interchange, Mobile IP was integrated into the FSA architecture. The FSA architecture employed Mobile IP in WLAN through the Subnet Architecture (SA). The subnet architecture initiates a Mobile IP handoff immediately following completion of the MAC-level handoff during migration between networks. The architecture also allows the application to participate fully in the handover process by providing recommendations in the event of QoS changes. Information about the effects of vertical handover are delivered to the Application Adaptation layer in order to perform any necessary adaptation. The vertical LAN/WAN handoff layer in the FSA architecture monitors network characteristics, available power, and application requirements in order to perform vertical handoffs between network interfaces.

The FSA architecture is limited to packet transmission that requires TCP/IP. The disadvantage of employing Mobile IP in the architecture is the added latency involved in triangular routing. whereby all packets sent to the mobile device must be sent to the mobile's home network and forwarded to the mobile's current location [8]. The FSA architecture can support user mobility and data traffic, but is not able to support real time multimedia traffic nor can it be used in the event of network QoS degradation. Furthermore, the FSA architecture requires the evaluation of signal strength before a vertical handover operation can be performed and the latency involved in discovering disconnections, finding new access points and performing handover will lead to packet losses while changing between network interfaces.

Stemm and Katz developed a vertical handover scheme for wireless networks that provides coverage over a range of geographical areas [2]. The goal of the system is to allow a mobile user to roam among multiple wireless networks in a manner that is completely transparent to applications and that disrupts connectivity as little as possible. The wireless overlay networks used in the architecture include an infrared room network, a WaveLAN network (in buildings), and a Ricochet Wide Area Network. Initial tests revealed that the time needed for vertical handovers is dominated by the time to discover that the mobile has moved in/out of coverage and therefore needs to change the network. This time tremendously affects applications that require low handover latency. Enhancements were later made to minimize this handover latency by either broadcasting beacons at higher frequency or performing packet doublecasting. The project relied heavily on the mobile device making handover decisions based on packet loss thresholds and therefore the handover time is quite long. This approach does not provide QoS support for applications during vertical handover. Similarly to the FSA architecture, this approach only supports vertical handovers due to disconnections and does not support network changes due to QoS degradation.

B. QoS support for GPRS/UMTSs

GPRS and UMTS vary in their approach to provision of communication QoS and these differences can have impact on the handover process to these networks. The GPRS network is an extension to the GSM mobile telecommunication network designed to provide packet data transmission [13]. In GPRS networks, a specific QoS profile, which is part of the PDP context profile, is allocated to every subscriber upon attachment to the network [14, 15]. The PDP context, created by the GPRS network for each session [13], includes PDP address (IP address for the mobile host – the GGSN uses this to enable packet transmission between the PDN and the mobile device), PDP type (eg. IPv4), the requested QoS, and the address of the GGSN that serves as the access point to the PDN (converting IP packets to GTP packets with the mobile's IMSI number) [14].

The user QoS profile in GPRS includes service precedence class (high, normal, low), reliability class (3 levels), delay class (4 levels), and peak and mean throughput class (8Kb/s – 2Mb/s). The service precedence is the priority of service in relation to other services. The reliability indicates the transmission characteristics required by an application. The delay parameters define the maximum values of the mean delay associated with each class. The GPRS network does not adapt communication streams if QoS changes. Therefore, support is not provided to adapt streams to suit the network conditions or device profile.

The UMTS approach to QoS is different. A significant change made by the UMTS network is the addition of OSA (Open System Architecture). The purpose of OSA is to allow third party service providers to access information in the UMTS network.. OSA consists of the Service Capability Servers (SCS), which include the Mobile execution environment (MExE) server, the Customized Application for Mobile Network Enhanced Logic (CAMEL) server, the Call Control Services server, Home

Location Register server, and the SIM application toolkit server [9]. The UMTS network provides QoS support to suit user subscribed profile, network QoS, as well as device capability. The MExE concept is developed to ensure that the content of the communication stream matches the device capability and is within the limits of the network capability and user profile. This is achieved through the MExE Service Environment (MSE) [10]. The MSE are nodes that provide MExE services to UMTS devices, and one of these services provides QoS mapping for content and capability for each request made by a user. The content negotiation ability allows adaptations of content in order to suit the device, and bi-directional capability negotiation supports the transfer of capabilities between the MSE and the MExE device [11]. To accommodate the varying types of hardware and its capabilities, a classification system known as the MExE class mark is setup to classify the devices into categories that best suit the capability of the terminal. The MExE classmarks define the category of content types for devices: classmark 1: WAP environment, classmark 2: Personal Java Environment, classmark 3: J2ME (Java 2 Micro Edition) CLDC (Connected Limited Device Configuration), classmark 4: CLI compact environment [10].

Unlike the GPRS system, where QoS is requested by the user or a default profile is assigned, UMTS defines QoS profiles for corresponding applications. The UMTS network organizes application QoS requirements into four classes: Conversational class, Streaming class, Interactive class, and Background class.

III. CONTEXT MODEL

Seamless computing requires a rich set of context information. In this paper we limit the description of context to context information which is needed to support vertical handovers. We divide this information into static and dynamically changing information. The entire context model is shown in Fig, 1.

A. Static Profile

The **static profile** holds information that does not change very often and includes description of devices, networks, and applications and their QoS requirements.

The **device** profile includes characteristics of various user devices. Each device has associated device **capability** information, device **software**, user **personal setting** for the device, and device **network interfaces**.

Device **capability** includes CPU, memory, screen size, operating system, device communication type (e.g. wireless or wireline), and content capability (e.g. text, images, sound, and video capability). Device **software** context information describes software applications for each device. The software applications are grouped based on their communication QoS requirements (Application Traffic Class Requirements - ATCR) into the following classes: Conventional Internet Services class, Playback Streaming class, and Conversational Streaming class. The Conventional Internet Services class includes applications for web-browsing, ftp, e-mails, etc. Playback Streaming class is for stored multimedia applications (eg. radio, audio). The Conversational Streaming class represents real-time multimedia applications such as live radio, video transmission, video conferencing, etc. Each class contains four attributes: delay, jitter, data rate, and packet loss. For video and audio streaming in Conversational and Playback Streaming class, context information is also provided with regard to the encoding standard used for multimedia streams (eg. Encoding scheme: Mpeg-1, Mpeg-2; Encoding method: VBR, CBR). This context information is used when QoS mapping is performed during vertical handovers. The **Personal Setting** context allows users to specify their device and network interface priorities. The priority shows the order of devices and the preferred network interfaces for each device. This information is used to redirect communication streams to the most preferred device or network interface if several options are available.

Associated with each device is context information about **network interfaces** in order to determine which different networks a particular device can access. Certain devices may have multiple network interfaces allowing access to a variety of networks, e.g. laptops which may have access to Ethernet, WLAN, Bluetooth, or GPRS networks.

The network context information includes potential network QoS and the **coverage** of the network. There are many ways in which network coverage can be modeled. For simplicity, we model network coverage as a two dimensional geographical grid map. The map provides information about the geographic coverage of each network. The different networks accessible by users currently located in a particular grid are also depicted. The grid map also describes the transition zone of each network.

In addition, the static profile describes the **user perceived QoS requirements** for multimedia applications, to allow users to express their perception of QoS for each ATCR. The user perceived QoS includes the video classification table [3] (e.g. Temporality - slow, fast; detailing - high, low), type of video content, and the tolerable level of QoS (e.g. overall video quality, disturbances, and bandwidth fluctuations permitted).

The model also includes **User Cellular Network Profiles** for GPRS and UMTS networks, which contain information of each user's subscribed profile such as the PDP context mentioned in Section II. To allow for a systematic GPRS QoS adaptation, the context description sets the QoS profile attributes for each ATCR class. The purpose of setting the QoS profile is to provide the quality the user has requested based on billing of the service. Therefore for each ATCR class, the user defines the required QoS class, e.g. video (Playback Streaming – ATCR-2): service precedence – high, packet loss – 10^{-9}, delay – 0.1ms, mean throughput – 64Kb/s (only base layer of scalable flow video). Information about the UMTS application classes is also a part of the User Cellular Network Profiles.

B. Dynamic Profile

The **dynamic profile** provides current information about users and networks such as current user **location**, current **QoS network parameters** (bandwidth, loss rate, delay, and jitter). User location is based on the grid map coordinates. The dynamic profile also includes **Impending Network Profile (INP)** which indicates to which network handover is most likely. This information changes dynamically when user location or network QoS changes.

IV. ARCHITECTURE

The proposed handover redirects communication stream between different network interfaces on one device or between network interfaces on different devices for a variety of network technologies, including telecommunication networks.

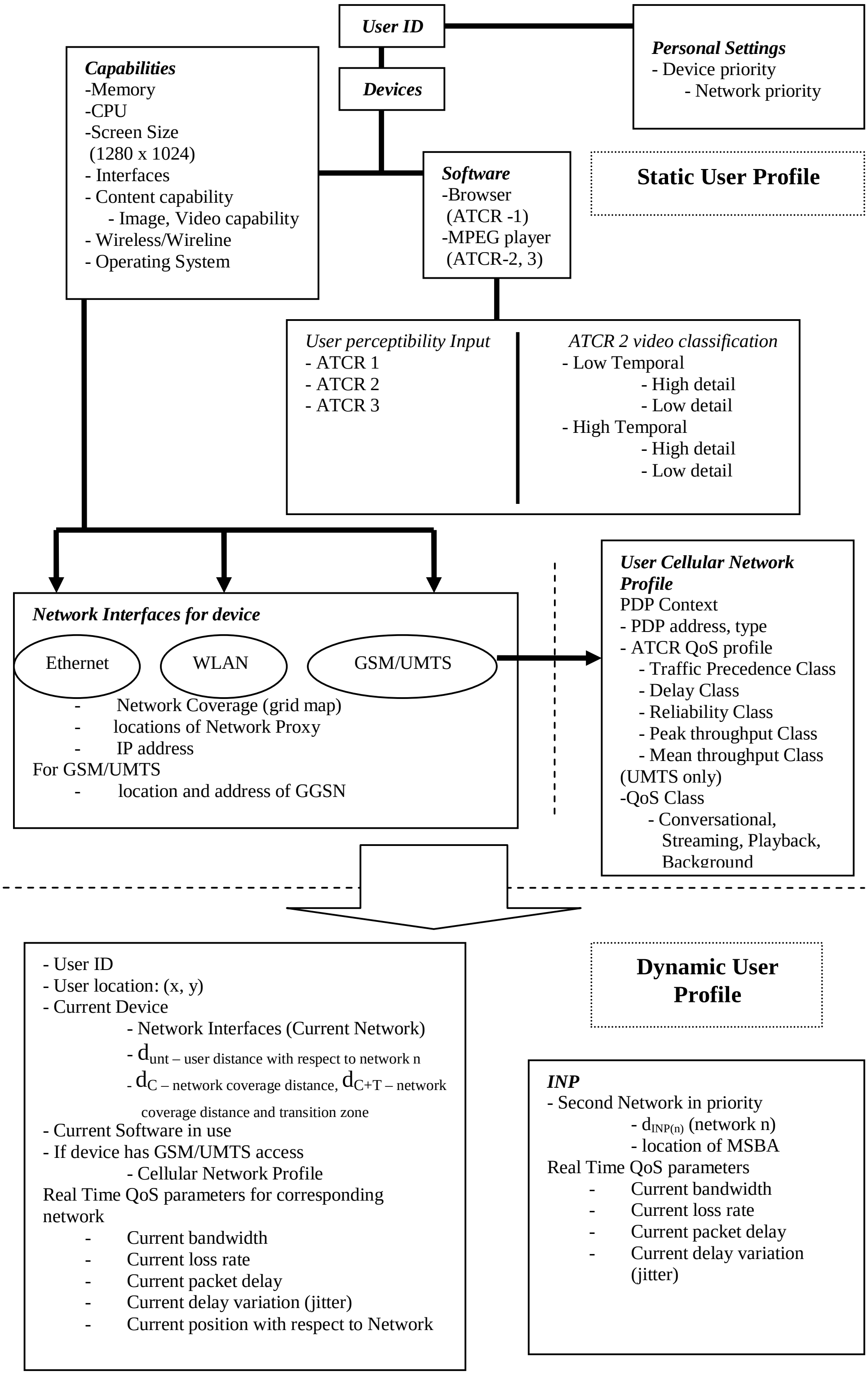

Fig. 1. Context Model used in vertical handovers

The handover mechanism operates above the transport layer to allow migrations between networks with different protocol stacks.

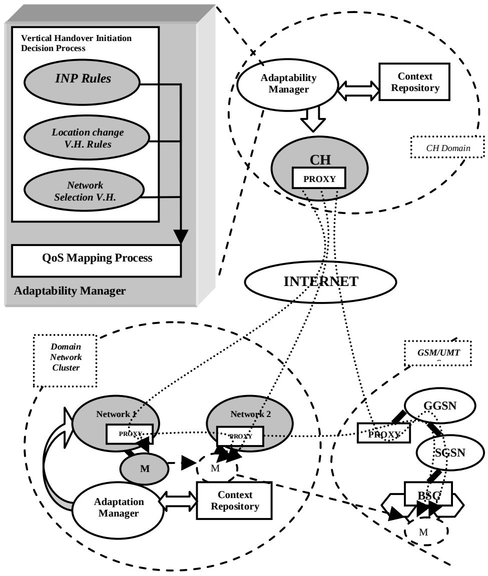

Fig. 2. Vertical handover architecture

The architecture supporting vertical handovers includes the Context Repository which gathers, manages and evaluates context information, the Adaptability Manager which makes decisions about adaptations to context changes (including decisions about handovers between networks), and Proxies which are responsible for executing handovers (Fig. 2). Each network has its own proxy. Multiple heterogeneous networks within a domain (e.g. departmental, enterprise, etc.) are clustered together into a Domain Network Cluster (DNC). Each DNC is supported by the Adaptability Manager and the Context Repository. The Adaptability Manager can subscribe to the Context Manager for notifications about particular context changes. The Adaptability Manager presented in this paper makes decisions about handovers and selects adaptations for the communication stream when necessary. The functionality of the Adaptability Manager can be divided, in the case of vertical handovers, into two main processes: ***Vertical Handover Decision Process*** and ***QoS Mapping Process***. The Components of the Adaptability Manager which provide the functionality described in the next section are shown in Fig. 2.

The proxies residing in each network are used to redirect communication streams between networks during vertical handovers. The proxies receive notification of requested handover operations from the Adaptability Manager. Since our architecture provides vertical handovers to GPRS/UMTS networks, a proxy is also placed at the interface of the GPRS gateway node (as shown in Fig. 2). The proxies provide QoS support during the handover through two operations: doublecasting the stream during handovers and dynamic packet buffering. The doublecasting operation sends the stream to the mobile host and also to the new proxy the mobile host is migrating to. This operation is shown in Fig. 2. Initially the mobile host is connected to Network 1. A packet stream is transmitted through the proxy of the Correspondent Host and streamed through the Network 1 proxy to the mobile host. When a vertical handover is triggered the stream is also sent from the Network 1 proxy to the Network 2 proxy. As the mobile migrates to Network 2 the packets are continuously streamed to the mobile host through the new proxy. During this operation, the stream is redirected from the Correspondent Host to Network 2. As soon as the redirected packets arrive, the doublecasting operation is terminated. The purpose of this operation is to minimise QoS violation during vertical handovers, i.e. eliminate packet losses and minimise delay and jitter. It is also supported by the dynamic buffering mechanism which buffers packets during the handover. The detailed description of this mechanism and its justification is presented in [4].

V. HANDOVER MECHANISM

A. Vertical Handover Decision Process

This process decides when to invoke a vertical handover operation. The decision process evaluates (i) user location changes (as users may leave or enter a particular network coverage) and (ii) QoS of the current and alternative networks. The evaluation of user location changes is carried out based on the grid map of network coverage and considers location of users as well as device and network priorities. The vertical handover process is rule based and the rules are (informally) described further in this section. The rules decide whether handover is necessary and to which network. The latter is decided by the QoS based network selection process invoked when the QoS of a particular network is below perceived acceptance quality, or a user enters the transition zone of a new network, or when determining the INP network. It has to satisfy multiple objectives including satisfying user's device preferences, achieving the highest level of bandwidth for applications while minimising packet loss, delay, jitter, and avoiding bandwidth fluctuations which may affect the applications. The decision rules and the network selection process are described below.

Decision rules:

Rule 1 (application initialisation or device change): For application intialisation, invoke the Locality based network selection process to determine the current device, current network and its INP network. In the event of device change, perform the vertical handover to new device (its INP network) then invoke the Locality based Network selection process to determine the new INP.

Rule 2 (moving out of networks): If the user approaches the transition zone of the current network (moving out of the network coverage), perform vertical handover to the INP network. Apply Locality based Network selection process to determine the new INP.

Rule 3 (entering new networks)**:** If the user enters a transition zone of a new network apply *Locality based Network selection process* to determine the network which best matches the application QoS requirements. If the network score is higher than the current network, perform vertical handover and invoke *Locality based Network selection process* to determine new INP.

Rule 4 (network QoS changes): If network QoS changes, determine if QoS for the current network (expressed as the AHP score, see below**)** drops below the score of the INP network. If yes perform vertical handover to the INP network. Invoke the *Locality based Network selection process* to determine the new INP.

Locality based network selection process

Select a set of networks which the user can currently use (overlapping coverage), and from this set select a subset of networks for which there are user devices within the user proximity. On this subset of networks, perform the *QoS based network selection process* algorithm to determine the INP network.

QoS based network selection process

Selecting a network which meets user QoS requirements, requires satisfying the following four objectives:

Objective 1: Maximizing user device preferences.
Objective 2: Maximizing application bandwidth.
Objective 3: Minimizing jitter, delay, and loss.
Objective 4: Minimizing bandwidth fluctuations.

Since a number of objectives must be satisfied, the Analytic Hierarchy Process (AHP) [5] method is employed. The AHP is a decision support tool which uses multi-level hierarchical structure of objectives and criteria. The AHP method was chosen due to its ability to vary its weighting between each objective. This fits well with our requirements that the decision making process applies user perceived QoS. The result of the process is different for each individual specification of user perceived application QoS.

The AHP calculation is a three step process:

Step 1-Calculate the objective weights from the objective pairwise comparison matrix (equation 1) based on *user QoS perceptibility* and the *personal setting* context (through relative values (RV)).

$$A = \begin{array}{c} \\ obj_1 \\ obj_2 \\ obj_3 \\ obj_4 \end{array} \begin{array}{c} \begin{array}{cccc} obj_1 & obj_2 & obj_3 & obj_4 \end{array} \\ \begin{bmatrix} 1 & RV_{12} & RV_{13} & RV_{14} \\ \frac{1}{RV_{12}} & 1 & RV_{23} & RV_{24} \\ \frac{1}{RV_{13}} & \frac{1}{RV_{23}} & 1 & RV_{34} \\ \frac{1}{RV_{14}} & \frac{1}{RV_{24}} & \frac{1}{RV_{34}} & 1 \end{bmatrix} \end{array} \quad (1) \qquad A_{norm} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \quad (2)$$

Static Profile
*Device Preference and Ranking
I. PC(Ethernet) - Input Bit Rate – 1MB/s
II. Laptop (Ethernet, WLAN)-Input Bit Rate –1MB/s
III. PC (Home) -Input Bit Rate – 1MB/s
IV. GPRS/WAP Phone-Input Bit Rate – 150Kb/s
*User Perceptibility Input
A. Current Application : ATCR 2: Slow moving, not fine, documentary
   2. User devices preference - 1
B. Video Quality – 5
C. Fluctuation of video quality – 1
D. Overall Disturbance – 3

Dynamic Profile
Current device – Laptop
Current Network – Ethernet; INP Network - WLAN

Fig. 3.Context scenario

The RV value determines the relative weight involved in each criterion, where an element R_{ij} indicates how much more important objective i is than objective j [5]. This process systematically applies weights on each objective relative to the importance of other objectives. The AHP scoring requires the RV values to be linearly scaled between 1-9 [5]. For example, RV_{12} expresses the relative ratio of user device preference (objective 1) and video quality (objective 2). Based on the example context information provided in Fig.3, where a score of 1 is assigned for user device preference, and 5 for video quality, the ratio is 1/5=0.2. Calculating the linear score, RV_{12}=(1-0.2) x 10 = 8 (a score of 8 shows a large separation gap this user has chosen between objective 1 and 2). Similarly, RV_{23} which expresses relative ratio of video quality (objective 2) and overall disturbances (objective 3), is calculated based on a score of 5 for video quality and 3 for overall disturbances, giving the ratio of 3/5=0.6. The linear score, RV_{23}=(1-0.6)x10=4 shows a fairly small gap between the two objectives. Matrix A is then normalized (equation 2), where the average values (a_{ij}) of each row for objective i is calculated (equation 3)

$$w_i = \frac{a_{i1} + a_{i2} + a_{i3} + a_{i4}}{4} \quad (3)$$

to give the weights for each objective (w_1, w_2,,w_3,w_4).

Step 2-Calculate the network scoring with respect to each objective through the network pairwise comparison matrix (equation 4).

$$\begin{array}{c} \\ Network_1 \\ Network_2 \end{array} \begin{array}{c} \begin{array}{cc} Network_1 & Network_2 \end{array} \\ \begin{bmatrix} 1 & RV_1 \\ \frac{1}{RV_1} & 1 \end{bmatrix} \end{array} \quad (4) \qquad \begin{bmatrix} 1 & 6 \\ \frac{1}{6} & 1 \end{bmatrix} \quad (5)$$

The purpose of this step is to determine the importance of each objective towards each corresponding network, unlike step 1 which concentrated on the importance of each objective towards the user. For each objective, a different scoring technique is applied to assist in efficient network scoring. For objective 3, a QoS space is employed to determine in which region, the QoS parameters (jitter, delay, and loss) lie.

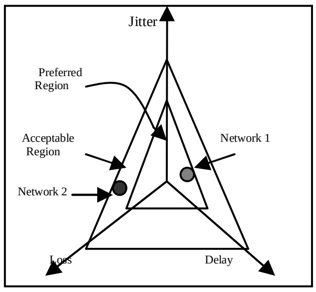

Fig. 4. QoS space scoring criteria for objective 1

$$\text{Re}\,lative\;\; ratio_{Bandwidth\;\;Fluc} = \frac{Autocorrel\;\;ation_{Network\;\;1}}{Autocorrel\;\;ation_{Network\;\;2}} \quad (6)$$

$$\text{Re}\,lative\;\; ratio_{Network\;\;Bandwidth} = \frac{Network\;\;Bandwidth_{Network\;\;1}}{Network\;\;Bandwidth_{Network\;\;2}} \quad (7)$$

As shown in Fig. 4, the effective QoS space is separated into two regions, the preferred region and the acceptable region. The optimum QoS is the origin vertex of the graph. The acceptable and the preferred regions are defined by the bounded limits for each parameter. The relative score is only apparent if two QoS network parameters are in different regions (as shown in Fig. 4). Since all scores are linearly scaled between 1-9, a score of 4.5 is allocated to the less preferable network. Objective 4 (equation 6) and objective 2 (equation 7) rely on the relative ratio difference, where the ratio is again linearly scaled between 1-9 to obtain the score. Autocorrelation of the bandwidth samples is used to determine the degree of fluctuation for objective 4.

For device preferences (objective 1), the relative score is calculated from the differences in priority and scaled between 1-9. For example, to calculate the score value for a user who has devices as described in Fig. 3 including PC device (priority 1) and Phone (priority 4), the difference in priority=2 and scaling between 1-9 gives a score of 6 (again a fairly large gap between the network preferences), which results in a pairwise matrix shown in equation 5. This calculation is performed for all objectives to determine the score for each network. Similarly to step 1, after normalizing equation 5, the average value of each row is calculated to obtain a score for each network of the corresponding objective – $S_{i1}, S_{i2},\ldots S_{ij}$ (objective j, network i). In this case $\text{Network}_{1obj1} = 0.84$ and $\text{Network}_{2obj1}=0.16$.

Step 3-Determine the sum of products of weights and network score for each network obtained from step 1 and 2 (equation 8), and <u>select the network with the highest sum</u>.

$$\text{Overall score}_{\text{Network}_i} = \sum_{1}^{n} S_{ij}(w_j) \qquad (8)$$

For the AHP calculation, **Step 1** which is based on user perception QoS, is performed only once, whereas **Steps 2** and **3**, which evaluates the current network QoS, are performed every time the *QoS based network selection process* is applied.

B. The QoS mapping process

When a decision is made that a handover should be performed, the QoS mapping process is required to select the changes to the communication stream to suit both the new networking environment and the new device capability (the latter only in the event of the device change). Our QoS Mapping process evaluates the availability of network resources (bandwidth) and, if necessary, it selects an appropriate stream conversion and filtration level to suit bandwidth availability in the new network. Using context information about the device input frame rate and network bandwidth, the QoS mapping process determines the degree of filtering required on the communication stream. The filters are applied at the proxies residing in each of the networks. The QoS mapping is extended if the user moves on to a GPRS network, to incorporate the QoS profile context chosen by the user. However, in the event the user moves to the UMTS network, the QoS mapping is performed between the device and correspondent host through the MSE, as mentioned in Section II.

VI. EXAMPLE HANDOVER

Two examples of vertical handovers are presented, which are vertical handovers between WLAN and Ethernet, and WLAN and UMTS. For both examples vertical handover is performed between network interfaces of the stream receiving device.

A. WLAN to Ethernet handover

The scenario of the Ethernet-WLAN handover is illustrated in Fig. 5 and the vertical handover protocol is shown in Fig. 6. The procedure for this vertical handover is as follows:

1. Initially packets are transmitted through PROXY-CH located at the Correspondent Host (CH) to PROXY1, which forwards the packets to the device.

2. After notification from the Context Manager about context changes, the Adaptability Manager determines that vertical handover is required, and sends a notification to PROXY1. PROXY1 triggers doublecasting to PROXY2 and the mobile device. The Adaptability Manager also asks the PROXY-CH to form a new stream from the Correspodent Host to PROXY2.

3. Once the new stream arrives, and the duplicate packets have been eliminated at the PROXY2, the next packets are streamed to the device.

4. Adaptability Manager terminates the transmission between PROXY-CH and PROXY1.

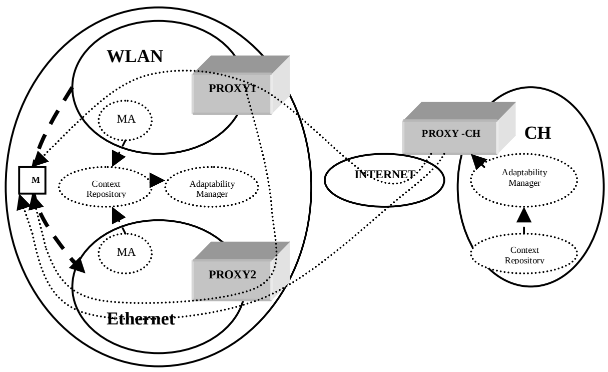

Fig. 5. Scenario for WLAN-Ethernet handover

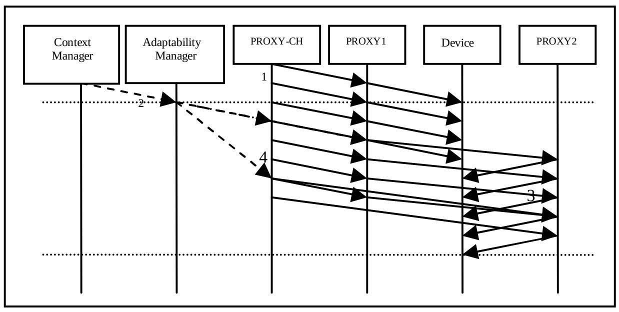

Fig. 6. WLAN-Ethernet vertical handover protocol

B. WLAN to UMTS handover

Fig.7 Illustrates vertical handover between WLAN and UMTS. The protocol for the handover is illustrated in Fig.8. The handover mechanism uses the MExE QoS support of the UMTS to provide QoS mapping between the device and packet stream during the vertical handover process. This is accomplished by pre-emptively triggering the UMTS device to negotiate with the MSE before the vertical handover takes place in order to minimise the delay during the change. This delay is mainly due to the MExE negotiation procedures as mentioned in Section II.

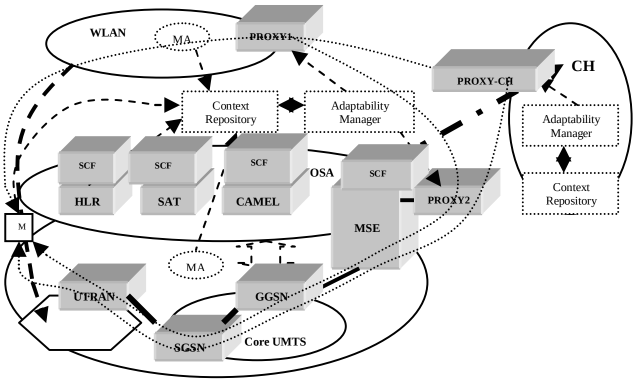

Fig. 7. Scenario for WLAN-UMTS handover

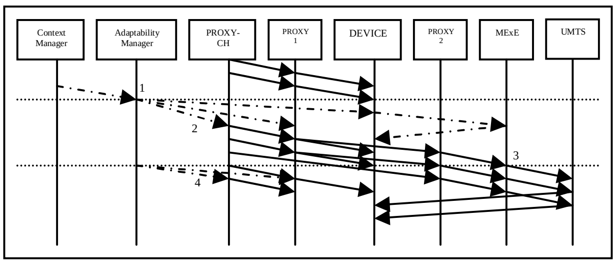

Fig. 8. WLAN-UMTS vertical handover protocol

The procedure for this vertical handover is as follows:

1. If a vertical handover is needed, the Adaptability Manager triggers the UMST device to perform content and capability negotiation with the MSE. The application service provider is asked by the MSE to determine the stream content modifications and filtering required before transmitting to the device.

2. The Adaptability Manager notifies PROXY1 that the handover is necessary and this notification triggers doublecasting from PROXY1 to PROXY2 and the mobile device. Concurrently, the Adaptability Manager triggers PROXY-CH to form a new stream to PROXY2.

3. As soon as the negotiation process between the MSE and the device is completed, PROXY2 transmits the stream to the MSE. This performs any necessary filtering and content transformation before the stream is transmitted through the UMTS network to the device.

4. The Adaptability Manager terminates the transmission between PROXY-CH and PROXY1.

VII. PROTOTYPE AND EXPERIMENTS

A prototype has been built which demonstrates vertical handover for a streaming JPEG RTP video application using the Java Media Framework (JMF). The proxies situated in each network were built from the JMF components. One of the prototype scenarios is shown in Fig. 3 (static and dynamic context information) and Fig. 9 (grid map). The static context profile shows user devices and their ranking. Only part of the dynamic context profile is presented. This profile shows the current device, the current network and the INP network.

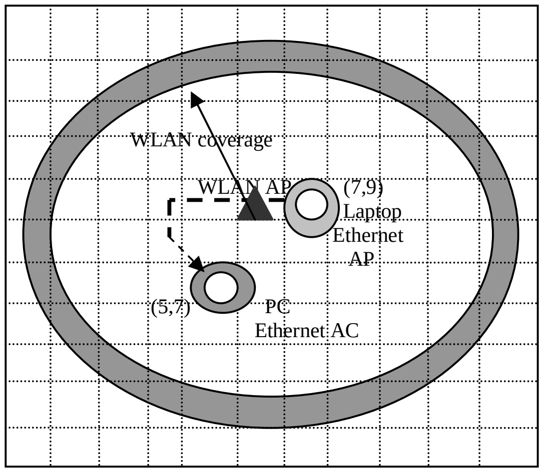

Fig.9. Grid map for prototype scenario

For this particular application, the objective weights were calculated from the user perceived QoS (step 1 AHP calculation) which resulted in the following values: w_{obj1}=0.445, w_{obj2}=0.05, w_{obj3}=0.1, and w_{obj4}=0.445. According to these objective weights, large weights were applied to objective 1 & 4, followed by objective 3 and lastly, objective 2.

Location changes are illustrated in Fig.9, and the results of the network QoS monitoring are presented in Fig. 10, 11, 12, and 13 (bandwidth, delay, jitter, and loss respectively). Notifications about users entering/leaving an area of network coverage and about QoS changes are delivered to the Context Repository by the location/QoS monitoring agents. The Adaptability Manager evaluates these changes (based on Rules 1-4) to make a decision about vertical handovers.

The grid map illustrates the transition zones (grey area) for the WLAN and Ethernet networks for the PC and Laptop. In this scenario, the user uses the laptop which is connected via Ethernet at position (7,9). Based on this position and the currently used device, the INP network was set to WLAN (Rule 1). According to Fig. 14, the average bandwidth availability on Ethernet network was 494.3 Kb/s**,** therefore the level of filtering for the stream was set to 0.1 (level of filtering was between 0.1-0.9 for bandwidth between 413Kb/s to 1.3Mb/s to provide a JPEG RTP stream at an average of 413 Kb/s as shown in Fig. 14). Before time 13, the bandwidth fluctuations (where the highest objective weight is applied – objective 4) on the Laptop-Ethernet was slightly higher than for the WLAN network. At the same time, monitored jitter, delay, and loss showed a relatively linear decline. The calculated score for the user current network and the INP network (equation 8) for each notification to the Adaptability Manager is shown in Table 1.

At approximately time 12, the overall disturbance (jitter, delay, loss) for the current network dropped below the acceptable level. Therefore, a notification was sent from the Context Manager to the Adaptability Manager which triggered the AHP calculation. The current network score (score – 0.413) was below the score of the INP network (score – 0.586), prompting the Adaptability Manager to request vertical handover from Ethernet to WLAN (Rule 4). The proxy of the Ethernet network started doublecasting to the WLAN proxy. Concurrently a new stream was created from the correspondent host to the WLAN proxy. The time for the vertical handover operation was determined by the amount of time it took for the packets of the new stream to arrive at the mobile device. The time for the stream to reach the mobile device when redirected from the Ethernet proxy to the WLAN proxy amounted to *20ms*. The new stream from the correspondent host arrived at the mobile device after an additional *30ms*. A QoS mapping operation was performed to determine the new stream bit rate. According to Fig.10, the average bit rate on the WLAN network was 965Kb/s. The Adaptability Manager therefore upgrades the filtering level to 0.6 resulting in a higher quality stream of 777.67 Kb/s, which can be seen in Fig. 14 where the average arrival rate increased.

Immediately following vertical handover, a new INP network was calculated based on the *Locality based network selection process* and is set to Ethernet. However, at time 13 the user moved away from the Ethernet access point and the INP network changed to GPRS. This change of INP was evident from table 1 at time 13, where a large variation gap is shown between the WLAN and GPRS networks. This gap was due mainly to QoS characteristics of the GPRS network which showed the average delay, loss, and jitter were relatively high compared to WLAN and also had a lower bandwidth availability from the GPRS network.

The user changed location at time 16 and moved along the path shown in the grid map. At location (5,7) a notification arrived informing the user she entered the transition zone of a new network. At this location, the user entered a new network zone with a new device, which for this test was a PC connected to the Ethernet network. The Adaptability Manager performed the AHP calculation (Rule 3) to determine if the new network could offer better resources with respect to the user's perceived QoS. This is shown in table 1, at time 22, where the current network is set to PC-Ethernet (score – 0.53) and the INP network is set to Laptop-WLAN (score – 0.375) with the network score favoring the PC Ethernet. Fig. 10 shows the bandwidth of the Ethernet network was reasonably close to the WLAN bandwidth, therefore the communication stream did not require any further adaptation. The vertical handover time from the Laptop-WLAN to the PC-Ethernet was negligent, due to the fact that double casting was performed for the two devices during the vertical handover. During the doublecasting the user received the streaming on both devices. This ceased when the Adaptability Manager terminated the stream to the Laptop-WLAN following the successful transition between the networks.

VIII. Conclusion

In this paper we have presented a context-aware vertical handover designed for future pervasive environments. The proposed handover can integrate a variety of wired and wireless technologies (2.5G, 3G, 4G, WLAN, Bluetooth) into a seamless

Time	Current Network	INP Network
12	0.413	0.586
13	0.811	0.187
22	0.53	0.375

Table 1. Current and INP network AHP scores of notification times

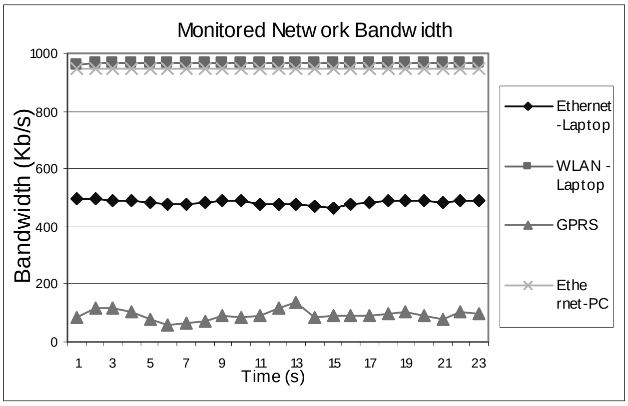

Fig. 10. Monitored bandwidth of networks

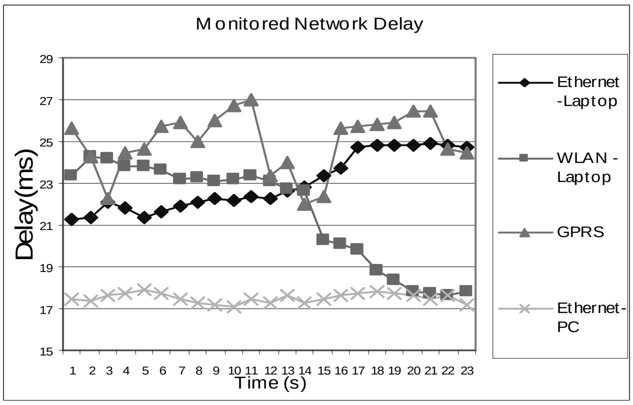

Fig. 11. Monitored packet

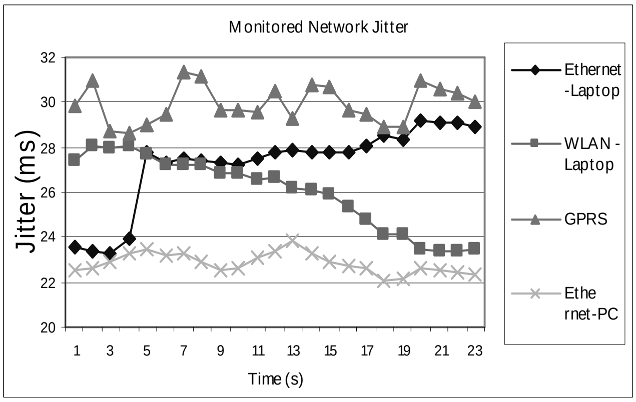

Fig. 12 Monitored packet jitter

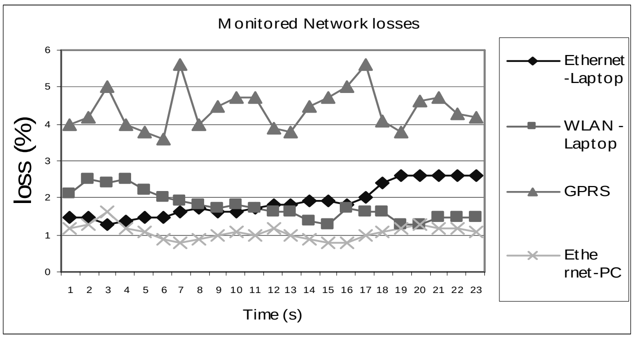

Fig. 13. Monitored packet losses

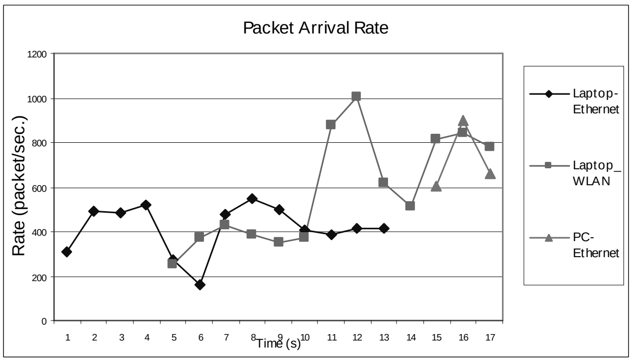

Fig. 14. Packet arrival rate

communication environment. It uses a wide range of context information about networks, users, user devices and user applications, and provides adaptations to a variety of context changes which are applicable to static or mobile users: disconnections and predicted disconnections, network QoS changes, device changes, and user preferences with regard to networks and computing devices when entering new networks. It is assumed in this solution that the application QoS is specified as user perceived QoS, as ordinary users are not able to express communication QoS requirements in terms of network indices.

A prototype handover architecture has been developed for JPEG RTP video transmission. Experimental results of the evaluation of context changes and the selection of a new network have been presented. The proposed vertical handover mechanism is currently being integrated with our infrastructure for pervasive computing [6, 7].

REFERENCES

[1]S. Helal, C. Lee, Y. Zhang, G.G. Richard III, "An Architecture for Wireless LAN/WAN Integration," *IEEE Wireless Communication and Networking Conference (WCNC) 2000*, Chicago, September 2000.

[2]M. Stemm, R.H. Katz, "Vertical handover in wireless overlay networks," *Mobile Network Application*, Vol.3, 1998, pp.335-350.

[3]R. T. Apteker, J. A. Fisher, V. S. Kisimov, "Video Acceptability and Frame Rate", *IEEE Multimedia*, vol. 2, no. 3, Fall 1995, pp. 32 – 40.

[4]J. Indulska, S. Balasubramaniam, Vertical Handover based Adaptation for Multimedia Applications in Pervasive Systems. *Joint Int. Workshop on Interactive Distributed Multimedia Systems and Protocols for Multimedia Systems (IDMS/PROMS 2002), 2002*. Lecture Notes in Computer Science, LNCS 2515, pp. 61-72.

[5]T.L. Saaty, "How to make a decision: The Analytic Hierarchy Process", *European Journal of Operational Research*, Vol. 48, 1990. pp 9 – 26.

[6]K. Henricksen, J. Indulska, A. Rakotonirainy, "Modeling Context Information in Pervasive Computing Systems", *Proc. of Pervasive'2002, Lecture Notes in Computer Science. Springer Verlag*. LNCS 2414, pp. 167-180.

[7]A. Rakotonirainy, J. Indulska, S.W. Loke, A. Zaslavsky, "Middleware for Reactive Components - An Integrated Use of Context, Roles, and Event Based Coordination", *Proc. of the IFIP/ACM Intern. Conf. on Distributed Systems Platforms. 2001*. Lecture Notes in Computer Science, Springer Verlag. LNCS 2218, pp. 77-98.

[8]C.E. Perkins, "Mobile IP", *IEEE Communication Magazine*, vol.35, May 1997, pp84-99.

[9]L. Bos, S. Leroy, "Toward an All-IP-Based UMTS System Architecture", *IEEE Network*, January/February 2001, pp36-45.

[10]3GPP TS 22.057 version 5.4.0 Release 5, "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); Mobile Execution Environment (MExE); Service Description; Stage 2", *ETSI TS 122 057*.

[11]3GPP TS 23.057 version 5.1.0 Release 5, "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); Mobile Execution Environment (MExE); Functional Description; Stage 2", *ETSI TS 123 057*.

[12] M. Frodigh, P. Johansson, P. Larsson, "Wireless ad hoc networking – The art of networking without a network", *Ericsson Review*, no.4, 2000.

[13] C. Bettstetter, H. Vogel, J. Eberspacher, "GSM Phase 2+ General Packet Radio Service GPRS: Architecture, Protocols, and Air Interface," *IEEE Communications Surveys*, Third Quarter 1999, vol.2, no.3.

[14]G. Priggouris, S. Hadjiethymiades, L. Merakos, "Supporting IP QoS in the General Packet Radio Service," *IEEE Network*, September/October 2000.

[15]O. T. W. Yu, "End-to-end Adpative QoS Provisioning over GPRS Wireless Mobile Network", *Mobile Networks and Applications*, 8, 225-267, 2003.

[16]R. Kraemer, P. Schwander, "Bluetooth based wireless Internet Applications for indoor hot spots: experience of a successful experiment during CeBIT 2001", *Computer Networks*, 41, 2003, pp. 303-312.

[17]G. Ghinea, J.P. Thomas, "An Approach Towards Mapping Quality of Perception to Quality of Service in Multimedia Communications", *Proceedings of the IEEE Multimedia Signal Processing Workshop*, Copenhagen, Denmark, 1999, pp.497-502.