

Active Node supporting Context-aware Vertical Handover in Pervasive Computing Environment with Redundant Positioning

Sasitharan Balasubramaniam, Tom Pfeifer
Telecommunication Software and Systems Group
Waterford Institute of Technology
Cork Rd., Waterford, Ireland
sasib@tssg.org, t.pfeifer@computer.org

Jadwiga Indulska
School of Information Technology and Electrical
Engineering
The University of Queensland
Brisbane, Queensland 4072, Australia
jaga@itee.uq.edu.au

Abstract—A major requirement for pervasive systems is to integrate context-awareness to support heterogeneous networks and device technologies and at the same time support application adaptations to suit user activities. However, current infrastructures for pervasive systems are based on centralized architectures which are focused on context support for service adaptations in response to changes in the computing environment or user mobility. In this paper, we propose a hierarchical architecture based on active nodes, which maximizes the computational capabilities of various nodes within the pervasive computing environment, while efficiently gathering and evaluating context information from the user's working environment. The migratable active node architecture employs various decision making processes for evaluating a rich set of context information in order to dynamically allocate active nodes in the working environment, perform application adaptations and predict user mobility. The active node also utilizes the Redundant Positioning System to accurately manage user's mobility. This paper demonstrates the active node capabilities through context-aware vertical handover applications.

I. INTRODUCTION

The emerging pervasive computing technology aims to support user activities and perform tasks on behalf of the users, while adapting the application to changes in the computing environment. This vision is beginning to take shape with the advancement in various wireless technology infrastructures (e.g. WLAN, Bluetooth, 3G), proliferation of new devices (e.g. wearable computers), sensors and actuators embedded into users living environment, as well as software models for integrating these disparate technologies. A key requirement towards achieving full-scale pervasive computing systems is the integration of context-awareness and the capability of gathering and evaluating context information from the computing environment. However, most infrastructures for pervasive computing are based on centralized architectures which can lead to bottleneck effects when the number of users increases and are usually prone to system failures. Since the pervasive computing environment will incorporate various heterogeneous networks and devices, the context gathering and evaluation mechanisms should be distributed to various nodes within the infrastructure to alleviate the computational stress on

centralized systems and at the same time increase system reliability.

In order to counter the drawbacks found in centralized architecture, we propose a hierarchical architecture that consists of an intelligent *active node* context gathering and evaluation system based on active network technologies. The proposed solution requires active nodes to be distributed through various regions of the user's working environment, also known as Personal Computing Environment (PCE), and provides active node hopping mechanism as users move between PCEs. Each node subscribes and evaluates the context information with regards to the user's activity and computing environment within the PCE. The active node provides an efficient infrastructure for pervasive computing applications and employs the *Redundant Positioning System (RDS)* [4] to accurately provide user location in order to support efficient handover decisions and operations.

In this paper, we particularly focus on the application of active nodes for *vertical handover* operations. Since pervasive computing environments are built from heterogeneous networks and devices, one of the goals of pervasive systems is to intelligently support user tasks and provide seamless computing capabilities within this environment. Vertical handover provides an essential adaptation mechanism for pervasive systems, by providing transparent redirection of packets between heterogeneous networks or between different computing devices while maintaining application continuity. Vertical handover operation may be due to disconnections and predicted disconnections, users performing device changes, degradation in network QoS with respect to application QoS or user preferences with regards to networks and computing devices when entering new environments.

The structure of the paper is organized as follows. Section II discusses the related work on active environments for pervasive systems and vertical handovers and context transfers mechanisms using active network technologies. Section III introduces the architecture of the proposed solution and descriptions of the various components. Section IV briefly describes the active node. Section V discusses the vertical handover operation and context transfer mechanisms using the

active nodes. Section VI presents the prototype and experimental results, while section VII concludes.

II. RELATE WORK

The related work is separated into two sections. Section A provides background on the current active environment infrastructures, while section B discusses the mobility support mechanisms for current vertical handover architectures.

A. Active Environment

The meta-operating system Gaia [12] is built as a distributed middleware infrastructure that coordinates software entities and heterogeneous devices contained within a specific geographical region. Gaia provides the interaction of services where the interactions allow user and developers to abstract ubiquitous computing environments as a single reactive and programmable entity instead of a collection of heterogeneous individual devices. Gaia defines an active space as a physical space with physical objects and devices, coordinated by a context-based software infrastructure that supports user's activities. Context services use a context provider registry in order to find providers of the context they desire. By employing context-awareness, the user's sessions are dynamically mapped to the active space resources when the user enters an active space (e.g. such as execution nodes, display or speakers that could be used to support the current type of application). Gaia has three major components, which includes the kernel, the application framework, and the application. The event manager notifies applications when the current state of the environment changes, and performs application adaptations to suit the changes. The active space stores all information about the software and hardware abilities in the space and lets applications browse and retrieve entities on the basis of specific attributes.

However, Gaia does not provide a scalable mechanism to use the infrastructure to support services in the active space. The middleware only addresses how the resources within the active space can be efficiently used to support the application, but does not address the mechanism to efficiently process context information and manage context information transfer between components of multiple active spaces.

The Project Aura [7] creates a pervasive system that supports user task and activity that integrates wireless communication, wearable computers, handheld devices, and smart spaces. The Aura infrastructure provides a number of key requirements in supporting pervasive computing environment. This includes cyber foraging, where the capabilities of resource limited mobile devices are amplified through surrogate servers that are located near the mobile device. Aura also employs a wireless bandwidth adviser which monitors the resources of the wireless access points and predicts the available resources in the near future. The wireless bandwidth adviser provides a mechanism to support application adaptations in the event of bandwidth changes. The Aura also provides a user location service which is based on signal strength and access point information from WLAN networks. The location services employs two types of algorithms for location sensing which includes pattern matching and triangular based remapped interpolated algorithms.

However, the algorithms are required to run on the user device which leads to unnecessary power consumptions especially for resource constrained mobile devices. Since the location information are derived from the signal strengths of the WLAN card, certain environmental effects can result in inaccurate location sensing, such as rotating the cards from its original position or obstructions on antennas. Since the location service only relies on one source, which is the WLAN card, it does not provide a flexible mechanism to support user location.

The Prism component of the Aura architecture provides a layer which explicitly represents the user's intent to provide certain reconfiguration or adaptation of services in the event of environmental changes. Each user's working environment contains a task manager which coordinates with the Context Observer to monitor the user's environment changes. In the event of changes in the computing environment, the Context Observer coordinates with the Environmental Manager to perform application adaptation. However, in the event of a user moving from one environment to the next, the infrastructure requires the task to pause until the user enters a new environment. This is not suitable for seamless computing, where minimal application disruption is required as users migrate between different environments. The migration of tasks between two static environments only supports static context gathering and evaluation that are deployed to specific regions (e.g. home task manager or office manager).

B. Mobility Support for Vertical Handover

Traditionally, Mobile IP has been used in a variety of vertical handover models and employs a home agent and foreign agent architecture [1] [3] [9] [10] [11]. However, Mobile IPv4 is not appropriate for pervasive systems since the protocol requires triangular routing when the user visits the foreign network. The triangular routing leads to extra latency and is not appropriate for multimedia applications that have end-to-end delay requirements. Mobile IPv6 protocols overcomes a number of shortcomings found in Mobile IPv4 by providing a larger address space, route optimization (does not require triangular routing), and improved security [2]. The MIPv6 protocol includes a location management mechanism using Binding Updates (BU), where each time the mobile host changes location, a BU is transmitted to the home network as well as to the Correspondent Host (CH) to allow packets to be transmitted directly from the CH to the mobile node's new location. The detection time is the amount of time from when the mobile enters the coverage of a new network to the time it receives the first Router Advertisement (RA) from the access router in the network. Although Mobile IPv6 eliminates triangular routing when the mobile host migrates to the foreign network, the protocol requires that the mobile host receive the RA from the new network before a handover is performed. This could increase the handover latency in the event that the mobile device is performing a handover from a high bandwidth network to a lower bandwidth network (e.g. WLAN to GPRS). Therefore, the handover performance will depend greatly on the bandwidth performance between the heterogeneous networks. Application continuity will also be disrupted in the event the mobile device moved out of the current network coverage before a handover is performed to the new network (e.g. WLAN to GPRS). Both Mobile IPv4 and v6 require the

packets to be transmitted to either the home agent or foreign agent before transmitting the packets to the mobile device. This will lead to a bottleneck effect in the event the number of users increases, and since the architecture relies on a centralized proxy, it has a single point of failure.

The Context-Aware handover solution developed at DoCoMo labs [6] aims to perform handover decision to the right access point by not only considering signal strength but also context information. The architecture integrates a context management framework, an active platform and a service deployment scheme to provide functionalities needed for context-aware handover. The context management framework is responsible for collecting context information from the Location Information Server (LIS), Network Traffic Monitor (NTM), and the user profile repository. Since the context information changes and evolves faster than network layer functions, the solution employs software agents and active networks to cope with the changing nature of context information. A number of handover decision mechanisms have been proposed which includes a context-exchange protocol approach, dynamic download of context-aware decision agent, and flexible framework that combines context-exchange with a dynamic agent. In the context-exchange protocol approach, context information is exchanged between the centralized context collection point and the handover decision point, where the context is evaluated by the context interpretation algorithm and a decision is made on the handover operation. In the dynamic download of a context-aware decision agent, a software agent that encapsulates the context information and the algorithms for interpretation of the data is downloaded to the mobile device, before the handover is performed. The input to the software agent includes reachable access points, applications requests and sessions and the output is the handover decision.

However, there are a number of drawbacks to the context-aware handover solution developed by DoCoMo. Firstly, the handover decision making process is performed at the mobile device, which is not appropriate for devices that have limited processing power. Secondly, although the context information is updated each time the device performs a handover, the context collection point remains at the same location and must process context information for all users in the computing environment, which in turn can affect the processing capacity when the number of users or amount of context information increases. Lastly, the mobility management of the devices is based on Mobile IP, where the mobile device listens to the RA from the new access points before selecting the right access point for connection. Before the mobile device performs the handover to the new access point, the new context profile for the new access point must be transmitted to the device. However, in the event of the mobile device migrating from a low performance network to a high performance network, this could result in the mobile device disconnecting from the current network coverage and missing the new context information required for the new network. Therefore, this will result in a failure in the handover operation.

In conclusion to this section, the current infrastructures do not provide scalable and efficient context gathering and

evaluation mechanism for seamless computing yet and are not suitable for multimedia applications.

III. ARCHITECTURE

A. Overview

Our aim is to provide a scalable mechanism to acquire context information and to perform optimized context evaluation. The architecture for the proposed solution applied on various heterogeneous networks is illustrated in Fig. 1. It consists of a Central Context Repository (CCR), Active Node repository, Redundant Positioning System, QoS monitoring agents, and a filter repository. Each active node is responsible for context management, evaluation, and decision making process for each user's PCE. Based on the user's location, the active node is dispatched to the most capable node that is within close proximity of the user's location and device (e.g. base station, router, desktop, etc.), where the location is decided by the CCR. The active nodes are not only deployable to local area networks, but may also be dispatched to other types of networks including GPRS/UMTS and ad hoc networks, as shown in Fig. 1. The interactions between various components of the architecture are illustrated in Fig. 2. The CCR uses a context model developed for vertical handover. The model defines three types of context information: *Static*, *Dynamic*, and *Infrastructure* based context information. The CCR also holds the Context Transfer module, which is responsible for context information transfer to the active nodes, and the Active Node Placement Decision Process (ANPDP). The Active Node Repository holds the active nodes and dispatches the nodes dynamically.

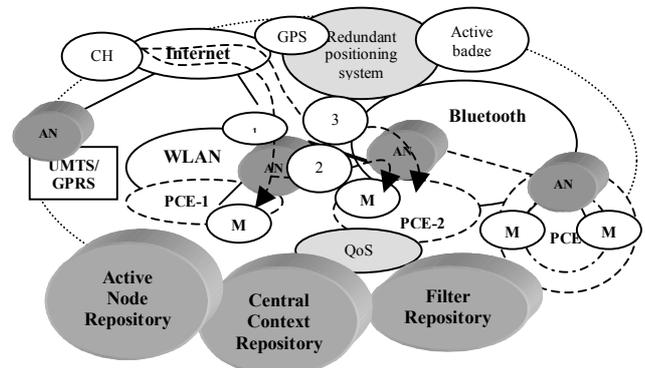


Figure 1. Architecture of active node support for Pervasive computing environment

B. Central Context Repository

The CCR contains the context information within the computing environment that is essential towards evaluation of handover operation, application adaptations, and placements of active node with respect to the environment.

Static Context: The static context information is information that does not change too often, and contains information regarding the **user's profiles**. The user profile context includes the types of devices the user possesses (e.g. Laptop, PDAs) and the different **networks** that each device has access to. Associated with each user device are the types of **applications** and their QoS requirements. The application

context information also contains information on **user perceived QoS satisfaction** context information, which provides information on the effects of different filtering mechanisms of a video stream with respect to the user's satisfaction requirements. This information is used by the decision process to determine the type of stream adaptation that is required to suit the user and network QoS, since various filtering mechanisms will have varying QoS effects on the user's satisfaction. The **user QoS perceptibility requirement** allows users to provide a description of their QoS perceptibility for each application. The user profile also contains information on the types of **location positioning systems** that the user has access to, which may include the location positioning systems on devices (e.g. GSM/GPRS location system on mobile phones) or access to specific positioning systems (e.g. active badge, RFID). The user profile also contains information on the **user mobility history**, which is associated with particular activities of the user. As shown in Fig. 2, the feedback from the active node also provides the latest user mobility record that gets updated with the current mobility history context, where the updates are performed by the Context update algorithm.

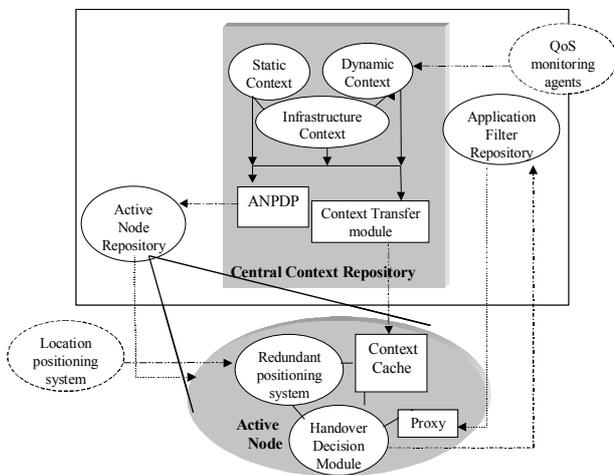


Figure 2. Interactions between components of architecture

Dynamic Context: The dynamic context information is current information that is associated to the user and computing environment. Such information includes the **user and device location**, and **network QoS** (e.g. delay, loss rate, bandwidth, and jitter). The location context information is collected from various location positioning systems. Since each location information is unique and associated to a particular user, the information is subscribed by the active node and transmitted directly to the RDS. However, the dynamic network QoS context, which is common to all active nodes in the same network, is transmitted directly to CCR and each active node subscribes for this information from the CCR.

Infrastructure Context: The infrastructure context information contains the **network coverage** information of the various networks under the domain, the type of nodes in the network (e.g. routers, base station), as well as the capacity of each node. The network coverage is modeled as two-dimensional grid maps.

When an active node is formed, the CCR transmits the network coverage that is only related to the PCE of the active node. For more information on network coverage map, please refer to [8].

C. Active Node Placement Decision Process

The ANPDP algorithm is a decision process in the CCR and is responsible for deciding where the active node will be placed in the new computing environment. The decision process takes into consideration the capacity of each node and the amount of processing power at the node with respect to the number of users. The decision process uses optimization techniques by evaluating the traffic load at access points and delay impacts on the significance of dynamic context information from the monitoring agents, before selecting the most appropriate node.

IV. ACTIVE NODE

The active nodes are dynamically placed in the computing environment and the location is determined by the ANPDP in the CCR.

Context Cache: The context cache holds the relevant information of the user for the specific PCE that may lead to possible vertical handovers. Such information maybe related to specific activities associated to the location of the PCE.

Redundant positioning system: The RDS provides an accurate mechanism to locate user in different computing environments across multiple administrative domains by combining various multiple individual positioning technologies to obtain wider coverage and a more precise location [4]. The mechanism used to combine various heterogeneous positioning technologies to obtain an accurate location is through data fusion. Considerations such as imperfect or misleading information (e.g. uncertainties in positioning measurement technologies used, relationship changes between target user and information obtained, etc.) are also taken into account in determination of location accuracy. The RDS is an essential requirement due to the fact that the placements of active nodes rely largely on accurate user mobility tracking and a prediction of the path that the user may take. Since the RDS require multiple location positioning technologies to transmit user location to the CCR, this could add further stress on the CCR. Therefore, to reduce the stress on the CCR, the active nodes will subscribe to the specific location positioning system directly for current location of the user (e.g. user has mobile phone and active badge, therefore subscribe to GPRS/GSM location system and active badge location system).

Handover Decision Process: The Handover Decision Process (HDP) algorithm is a decision process that determines when a handover operation is performed. The decision process is based on the user mobility, user performing a device change or degradation in network QoS, where further information on handover algorithms can be found in [8]. Handover operation due to user mobility must take into consideration the user's current location with respect to the network coverage grid map and the history of user's mobility based on user's activity. By evaluating multiple context information, the accuracy of handover operations can be improved. For example if the user

the active node are transmitted to the mobile device. The active node also subscribes for filters from the filter repository in the event that filtration is required for the doublecast stream when a handover operation is performed. The active node monitors the user's location context information from the redundant positioning system and the network QoS context from the network QoS monitoring agent and determines when a vertical handover is required. The user then changes location, where position changes have been transmitted to the active node from the positioning systems. The active node notifies the CCR of the handover operation and notifies the CCR of the predicted location of the user in the new network (WLAN) based on the mobility history. The ANPDP determines the node with the highest capacity with respect to the current network traffic to place the new active node (AN-new – Fig. 3). In this particular case, the WLAN base station was selected. When the new active node (AN-new) has been successfully configured, the old active node begins transmitting the user profile context information to the new active node. Upon receiving the context information, the new active node subscribes to the location positioning devices for the RDS. Before the active node begins the doublecasting to the new active node, the old active node configures the filters for the new one and transmits the configuration to the CH. The active node once again refers to the user perceived QoS satisfaction context and configures the stream to 320x240 resolution, quantization scale of 11, and at 30 fps in order to suit the 2Mbps bandwidth of WLAN. The old active network then performs the doublecasting process to the new active node (stream 2 – Fig. 3), where at the same time the new active node makes a new request for a new stream from the CH (stream 3 – Fig. 3).

VI. PROTOTYPE

A prototype has been developed to demonstrate the vertical handover between heterogeneous networks for streaming JPEG RTP video application using Java Media Framework, where the results are presented in Table 1. Although series of tests have been performed for different networks, this paper will only present selected results for WLAN to Bluetooth handover.

TABLE I. HANDOVER TIMES VS. NOTIFICATION DELAY SCENARIOS FOR WLAN TO BLUETOOTH HANDOVERS.

	Handover between stream 1 and 2 (ms)	Handover between stream 2 & 3 (ms)	Total Handover time with doublecast (ms)	No doublecast (ms) – Handover between stream 1 & 3
Delay 1	0.0	28.2 (1 to 3)	28.2	54.0
Delay 2	30.0	18.0	48.0	466.6
Delay 3	28.2	10.0	38.2	1620.2
Delay 4	25.0	20.0	45.0	2987.5

The results presented in this paper are only average handover times during the vertical handover operation. The reason is, firstly, the WLAN stream to the mobile device (stream 1 in Fig.1) is replaced by the stream traveling through

both the WLAN and the Bluetooth active nodes (stream 2). Secondly, stream 2 is replaced by a direct stream from the CH (stream 3). The increase in each delay scenario represents an increase in distance between the CH and the active node. As shown in Table 1, the impact of doublecasting has demonstrated low handover latency as opposed to handover with no doublecasting for all delay scenarios.

VII. SUMMARY

This paper has presented a migratable active node architecture for context gathering and evaluation for location dependent pervasive computing applications, where the active node has the capability to hop between PCE and adapt to the new user computing environment. We have focused on applying the active node for vertical handover operations between heterogeneous networks in pervasive computing environments, where handover decisions are performed in the active nodes.

REFERENCES

- [1] C. E. Perkins. Mobile IP. *IEEE Communications Magazine*, 35, 5 (May 1997), 84-99.
- [2] R. Chakravorty, P. Vidales, K. Subramaniam, I. Pratt, J. Crowcroft. Performance Issues with Vertical Handovers – Experiences from GPRS Cellular and WLAN Hot-spots Integration. In *Proceedings of 2nd IEEE International Conference on Pervasive Computing and Communications (PerCom)*, (Florida, USA, March 2004), 155-164.
- [3] W. Floroiu, R. Ruppelt, D. Sisalem, J.V. Stephanopoli. Seamless Handover in Terrestrial Radio Access Networks: A Case Study. *IEEE Communication Magazine*, 41,11, (November 2003), 110-116.
- [4] T. Pfeifer. Redundancy vs. Imperfect Positioning for Context-dependent Services. In *Proceeding of 1st International Workshop on Advanced Context Modeling, Reasoning, and Management (UbiComp 2004)*. (Nottingham, England, September 2004).
- [5] K. Fukuda, N. Wakamiya, M. Murata, H. Miyahara. QoS Mapping between User's Preference and Bandwidth Control for Video Transport. In *Proceedings of 5th International Workshop on Quality of Service (IWQoS 97)*, (May 1997), 291 - 302.
- [6] C. Prehofer, N. Nafisi, Q. Wei, A Framework for Context-aware Handover Decisions. In *Proceeding of PIMRC 2003*, (Beijing, China, September 7-10, 2003).
- [7] D. Garlan, D. P. Siewiorek, A. Smalagic, P. Steenkiste. Project Aura: Toward Distraction-Free Pervasive Computing. *IEEE Pervasive Computing, special issue on "Integrated Pervasive Computing Environments"*, Vol. 21, 2, (April-June, 2002), 22-31.
- [8] S. Balasubramaniam, J. Indulska, Vertical Handover supporting pervasive computing in future wireless networks, *Computer Communications* 27, (2004), 708-719.
- [9] M. Stemm, R.H. Katz, Vertical handoffs in wireless overlay networks. *Mobile Networks and Applications*, 3, (1998), 335-350.
- [10] K. Pahlavan, P. Krishnamurthy, A. Hatami, M. Ylianttila, J.P. Makela, R. Pichna, J. Vallstron. Handoff in Hybrid Mobile Data Networks. *IEEE Personal Communications*, 7, 2, (April 2000), 34-47.
- [11] S. Helal, C. Lee. An Architecture for Wireless LAN/WAN Integration. In *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC) 2000*. (Illinois, USA, September 2000).
- [12] M. Roman, C. Hess, R. Cerqueira, A. Ranganathan, R. H. Campbell, K. Nahrstedt. A Middleware Infrastructure for Active Spaces. *IEEE Pervasive Computing*, Vol. 1, 4, (October-December, 2002), 74-83.