Computing and Informatics, Vol. 24, 2005, 481–494

EFFICIENT VERTICAL HANDOFFS IN WIRELESS OVERLAY NETWORKS

Hakim Badis, Khaldoun Al Agha

LRI Laboratoty
University of Paris-Sud XI
91405 Orsay, France
e-mail: badis@lri.fr, alagha@lri.fr

Manuscript received 24 Januray 2005; revised 24 June 2005
Communicated by Ladislav Hluchý

Abstract. Mobile IP is used to keep track of location information and make the data available to the mobile device anytime, anywhere. It is designed to address the macro-mobility management, it does not address micro-level mobility issues such as handoff latency and packet loss. In this paper, we propose a mobility management scheme to handle the movements of mobile nodes among different wireless network technologies. Our scheme combines: (a) A hierarchical mobility management architecture to hide mobility of mobile nodes within the foreign domain from the home agent; (b) A handoff protocol to reduce packet loss during the transition from one cell to another; (c) The use of our proposed virtual cells in order to reduce the upward vertical handoff latency and disruption as much as possible. Our design is based on the Internet Protocol (IP) and is compatible with the Mobile IP standard (MIP). We also present simulation results showing that our handoff scheme is very fast to meet the requirements of an interactive communication session such as Internet telephony and avoiding packet loss.

Keywords: Overlay networks, micro-mobility, macro-mobility, mobile IP, vertical handoffs

1 INTRODUCTION

With increasing technological developments in digital wireless transmission and location devices, cell sizes are becoming smaller and smaller (*macrocells*, *microcells*, *picocells*, ...), increasing the available bandwidth per cell. Thus, the handoff latency between two cells is becoming an important aspect to minimize in order to reduce data loss and maintain uniform connectivity. Most existing wireless network technologies can be divided into two categories: those that provide a low-bandwidth service over a wide geographic area and those that provide a high bandwidth service over a narrow geographic area. While it would be desirable to provide a high bandwidth service to mobile users at all times, this is unlikely. It is necessary to use a combination of wireless networks to provide the best possible coverage over a range of geographic areas. This combination of wireless network interfaces fits into a hierarchy of network interfaces, which we call a wireless overlay network structure [1]. By overlaying a group of picocells with a microcell and a group of microcells with a macrocell, the number of handoff is reduced significantly. In this paper, we propose a scalable mobility management scheme for wireless internetworks.

Mobility management is often divided into two parts: macro-mobility and micro-mobility. Macro-mobility concerns the management of mobile nodes moving on large scale, between different domains (or wide wireless access network), while micro-mobility covers the management of users movements at a local level, inside a particular wireless network. The IETF Mobile IP standard [2, 3] is designed to address the macro-mobility management. When the Mobile Host (MH) moves away from its current network point-of-attachment, handoff is invoked to choose another suitable point of attachment. In such an environment, handoff latency and mobility dynamics pose a challenge for the provision of efficient handoff. Our handoff scheme applies to the Internet. It is based on IP and integrates seamlessly with Mobile IP. Our Base Stations (BSs) are network-layer routers and are capable of buffering the last few IP packets sent to a mobile, as before.

We also present a scalable simulation model close to real operations in wireless overlay networks. We have implemented a vertical handoff scheme [1, 4] that allows 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. For example, when the user leaves his/her office, where his/her Personal Digital Assistant (PDA) is connected via an in-room infrared network, to elsewhere in the building, where it is connected via a building-wide radio frequency (RF) network, his/her PDA performs a vertical handoff. We also use retransmission buffers in base stations and mobile nodes to recover from packet losses incurred during the transition between cells. The simulations show that our handoff scheme is very suitable for roaming mobile nodes that may encounter numerous handoffs while they are in the midst of an interactive voice communication session such as Internet telephony.

The rest of this paper is organized as follows. In Section 2, we describe in more detail the concept of wireless overlay networks, the vertical handoff and the virtual cells concept. Section 3 describes the mobility management. Section 4 presents the simulation model and performance results, and Section 5 concludes the paper.

2 OVERLAY NETWORKS AND VERTICAL HANDOFFS

In this section, we describe the wireless overlay network concept, and why wireless overlay networks present new challenges compared to existing handoff systems.

2.1 The Wireless Overlay Network Structure

A key concept is overlay networking, which is the unification of several heterogeneous networks, of varying coverage and performance, into a single logical network that provides the best of all coverage areas with performance corresponding to the best network in range. Figure 1 shows an example of a wireless overlay network. The lower levels consist of high-bandwidth wireless cells that cover a relatively small area. The higher levels in the hierarchy provide a lower bandwidth per unit area connection over a larger geographic area. Some devices already include multiple wireless link technologies to benefit from wireless diversity. Our goal is to offer the best service by executing a fast handoff without disrupting communications.

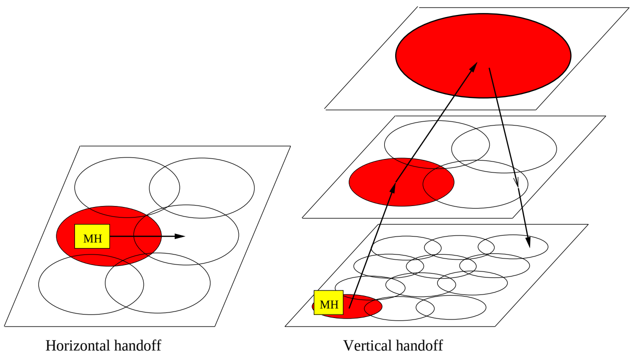

Fig. 1. Horizontal vs. vertical handoffs

A standard handoff (horizontal handoff) is simply migrating within the same cellular architecture, i.e., between base stations that are using the same type of wireless network interface. A vertical handoff (Figure 1) is moving from a cell in one cellular architecture to a cell in a different architecture, i.e., between base stations that are using different wireless network technologies (for example, GSM [5] and IEEE 802.11 [6]). There is a distinction between up and down in vertical handoffs: an upward vertical handoff is a handoff to an overlay with a larger cell size and lower bandwidth, and a downward vertical handoff is a handoff to an overlay with a smaller cell size and higher bandwidth. Downward vertical handoffs are less time-critical, because a mobile can always stay connected to the upper overlay during handoff.

The primary technical objectives in the design of a seamless vertical handoff system are the balance among handoff latency to minimize data loss and interruption, power consumption to minimize the power drain due to multiple simultaneously active network interfaces and wasted bandwidth to reduce the amount of additional network traffic used to implement handoffs.

2.2 Triggering Handoffs

In a network of homogeneous BSs, the relative signal strength of beacons is compared and the BS with the highest is chosen as the forwarding BS (Figure 2).

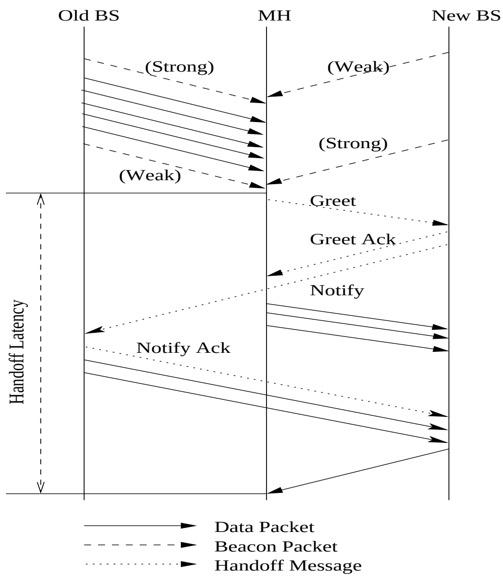

Fig. 2. Breakdown of horizontal handoff

BSs send out periodic beacons and forward data packets from the old BS. Each beacon carries the address of the BS that sent it. If the signal strength of the new BS is greater than that the old BS, the MH may decide to initiate a handoff from its current BS to the new BS. The ensuing message exchange and related processing are as follows:

- The MH sends a *Greet* message to the new BS, conveying its own address as well as the old BS and the list of *IDs* of the last few packets it has received. It also makes the new BS its default BS.
- The new BS creates a routing table entry for the MH so that it can forward packets to the MH. It also responds with a *Greet Ack* message. When the MH receives this message, if it has packets recently sent to the old BS in its retransmission buffer, it sends them to new BS.

- The new BS sends a *Notify* message over the wired link to the old BS to inform the old BS that the MH has moved. This message conveys the address of the new BS and the list of *IDs* sent by the MH.
- The old BS deletes its routing table entry for the MH. It resends only those buffered packets that do not appear in the list of *IDs*. It also returns a *Notify Ack* message to the new BS.

The handoff latency time in the homogeneous system is measured from the time the MH decides that the new BS has larger signal strength until the first data packet arrives from it.

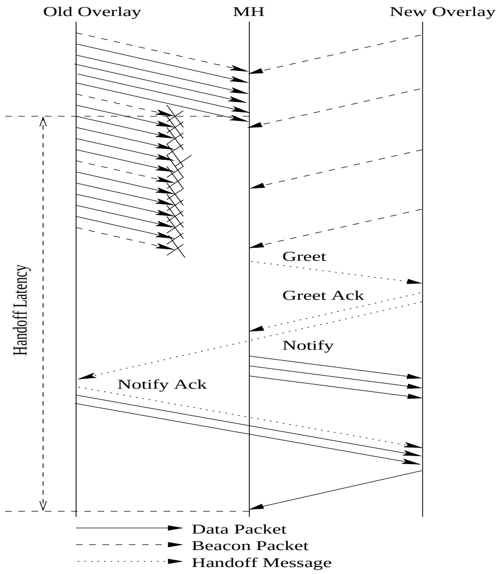

Fig. 3. Breakdown of upward vertical handoff

The vertical handoff starts when the lower overlay becomes reachable or unreachable, and ends when the first data packet forwarded from the new overlay network arrives at the MH. As shown in Figure 3, an upward vertical handoff is initiated when several beacons from the current overlay network are not received. The MH decides that the current network is unreachable and hands over to the next higher network. The arrival of beacons on a lower overlay initiates a downward vertical handoff.

2.3 Virtual Cell

The BSs broadcast beacons periodically (about every 100 ms). The beacon signal contains information about the BS such as timestamp, beacon interval and BS identifier. When a MH powers up it associates itself with the BS with the strongest beacon as a forwarding BS. The MH keeps track of the Received Signal Strength (RSS) from its current BS. When the RSS becomes weak, it starts to scan for stronger beacons from neighboring BSs.

In the horizontal handoff, when the signal strength received from any BS is greater than that received from the current BS, the MH instructs the new BS to start forwarding and the old BS to stop forwarding packets. An upward vertical handoff is initiated when T_B beacons (T_B: threshold number of beacon packets heard or not heard before initiating a handoff) from the current overlay network are not received. The time of T_B beacons is the component of latency during which the mobile discovers that it must handoff to a new wireless overlay. This time becomes critical with a larger beacons frequency. In order to eliminate this problem, we propose the virtual cell concept.

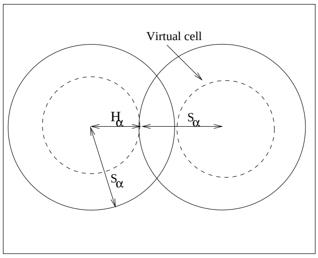

Fig. 4. The concept of the virtual cell

We define two parameters – S_α and H_α. The threshold of an acceptable Signal Strength (S_α) is the lower bound of RSS from the forwarding BS below reliable communication is not possible. The threshold of homogeneous cell Signal Strength (H_α) is the lower bound of the RSS from the forwarding cell such that the RSS from a homogeneous cell is greater than S_α. As we know the average of the cell diameters and user-visible bandwidths, we can fix S_α and H_α for each wireless network interface of the MH.

As shown in Figure 4, in each cell there is a virtual cell. For a MH out of this cell, the RSS from the BS of this cell is less than H_α. When the RSS from the current forwarding BS falls where H_α and there is no signal strength with acceptable level S_α heard from a homogeneous cell or a lower overlay network interface, the MH initiates an upward vertical handoff when the RSS from the current forwarding BS falls below S_α.

2.4 Analysis

We define the vertical handoff latency L as the amount of time from when the mobile is disconnected from the old BS to when the mobile receives the first packet from the new BS. L is given by the following expression:

$$L = L_D + L_{Greet} + L_{Notify} + L_T + L_N$$

where:

- L_D is the time during which the mobile discovers that it must handoff to a new wireless overlay. The MH must wait for approximately T_B beacons to determine that the current overlay is no longer reachable. Let N_B be the spacing between beacon packets (in seconds). The upper bound of L_D is given by $N_B \times T_B$, the lower bound is given by $N_B \times (T_B - 1)$ and the average is given by $N_B \times (T_B - 1) + N_B/2$. *With the use of virtual cell concept, this component of the latency is zero for the upward handoffs.*
- L_{Greet} is the latency needed for the *Greet* message from the MH to the new BS. L_{Greet} takes $L_U + S_{Greet}/Bw_U$ seconds for the upward handoff and $L_L + S_{Greet}/Bw_L$ seconds for downward handoffs. L_U and L_L represent the latency of the upper and lower network interface (in seconds), respectively. The size of the *Greet* message is S_{Greet}. Bw_U and Bw_L represent the bandwidth of the upper and lower network interface (in bits/s), respectively.
- L_{Notify} is the latency of sending the *Notify* message from the new BS to the old BS. L_{Notify} takes into account the number of hops between the old and the new BSs (H), the latency of a link in the wired backbone L_W and the bandwidth of the wired backbone network (Bw_W). $L_{Notify} = [(S_{Notify}/Bw_W) + L_W] \times H$.
- L_T is the latency of sending the first buffered packet that does not appear in the list of *IDs*. $L_T = [(S_{Data}/Bw_W) + L_W] \times H$.
- L_N is the latency for the new BS to send the first data packet across the new network to the mobile. If there is no outstanding data to send to the MH, then this component of the latency is zero. L_N takes $L_U + S_{Data}/Bw_U$ seconds for the upward handoff and $L_L + S_{Data}/Bw_L$ seconds for downward handoffs.

3 AN HIERARCHICAL STRUCTURE FOR MOBILITY MANAGEMENT

The proposed architecture is depicted in Figure 5. It is composed of several functional entities:

- Home Agent (HA).
- Correspondent Host (CH).
- Gateway Foreign Agent router (GWFA router): a router implementing the role of a Foreign Agent. It can also implement the role of a HA if its domain is the home network. The GWFA router periodically broadcasts Agent's advertisement messages containing its IP address. The GWFA router manages macro and micro cells. It uses an overlay network.
- Base Station (BS): a network-layer router having two interfaces: wireless and wired. Mobiles access the Internet via the BSs over the wireless links. The BSs can buffer the last few IP packets sent to a mobile.

- Router (R).
- Mobile Host (MH).

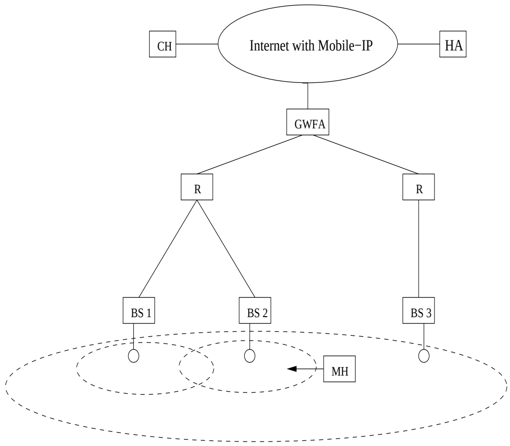

Fig. 5. Hierarchical mobility management

When the MH is away from its home network, the HA intercepts packets addressed to it and sends these packets through a tunnel up to the current mobile node's attachment. The MH would previously have registered its GWFA router care-of-address with its HA. The GWFA router decapsulates packets sent by the HA and forwards those packets meant the MH.

BSs send out periodic beacons similar to Mobile IP foreign-agent advertisements. The MH listens to these packets and determines which BS should forward packets for the mobile. Mobiles keep track of the *IDs* of the last few packets it has received. During a handoff, mobiles send to the new BS that list of *IDs* as part of the *Greet* message. The new BS forwards the list to the old BS as part of the *Notify* message. The old BS then resends only those buffered packets that do not appear in the list of *IDs*.

4 PERFORMANCE EVALUATION

In this section, we evaluate the performance of our handoff system. We have carried out simulations to analyze the vertical handoff in different configurations and scenarios. We have used the OPNET simulator.

4.1 Topology Generation Model and Traffic

In our simulation model, we have two overlay levels. The lowest level comprises a collection of cells of ring's format with high bandwidth networks. The second level consists of a collection of a disjoint of a square's cells which provide a lower bandwidth connection over a much wider geographic area (Figure 6).

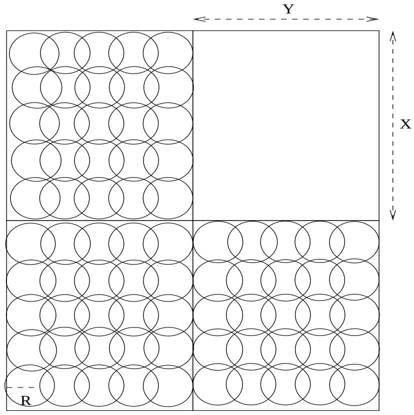

Fig. 6. Topology model

For each square's cell, we select its type (contains or not ring's cells) by randomly distributing in a way that $T\%$ (T is a parameter threshold) of the square's cells contain cells. The number, diameter and bandwidth of cells are given as parameters. Each BS is represented by a subqueue to queue up the new arriving packets, and placed in the center of its corresponding cell by selecting its x and y co-ordinates. Each node is represented by a subqueue and placed in the region by randomly selecting its x and y co-ordinates.

4.2 Buffer Size

To eliminate the packet loss due to the handoff, the amount of buffers needed at the BSs should be equivalent to the maximum possible amount of packet loss due to the handoff. As shown in Figure 7, we define the detection time as the latency during which the mobile discovers that it must handoff to a new wireless overlay.

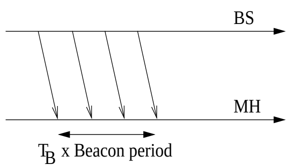

Fig. 7. Detection handoff time

The upper bound detection handoff time is equal to $T_B \times$ beacon_period. The maximum possible number of packet loss during a handoff (without any buffering scheme) is given by [(detection_time)/(packet_inter-arrival_time)] + 1. The size of buffer needed at the base station to support handoff can be extrapolated easily. A smaller beacon frequency decreases the detection handoff time.

4.3 Mobility Model

The random mobility model [7] proposed in this section is a continuous-time stochastic process. Each node's movement consists of a sequence of random length intervals, during which a node moves in a constant direction at a constant speed (Figure 8). To calculate the co-ordinate of node n at t during an interval i of duration T_n^i, angle θ_n^i and speed V_n^i, we calculate at the first time the distance D covered by n, $D = V_n^i T_n^i$. Then, we calculate the x, y local co-ordinates, $x = D\sin(\theta_n^i)$, $y = D\cos(\theta_n^i)$. At the end, we calculate the global co-ordinates by changing scale.

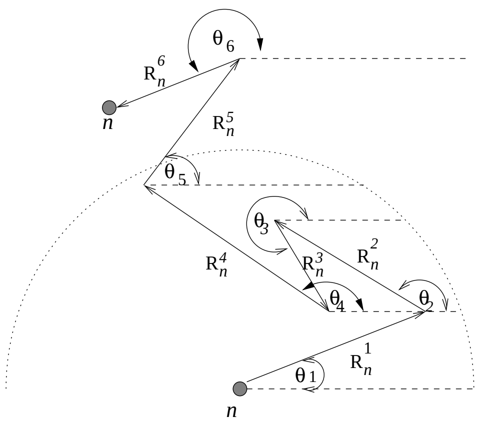

Fig. 8. Interval random mobility vectors

4.4 Evaluation Results

The objective of these simulations is to validate our protocol handoff and virtual cell as an enhancement to reduce handoff latency and packet loss. We have taken two overlay levels. The lowest level is the AT & T WaveLAN [8]. It has a bandwidth of 1.6 Mb/sec, 100 meters of cell diameter and 2 ms of latency. The second level consists of 4 cells of Metricom Ricochet Network [9]. It has a bandwidth of 60 Kb/sec, 1 km of cell diameter and 100 ms of latency. The GWFA sends out UDP unicast packets according to the Poisson distribution at an interval of 60 ms to simulate the forwarding of encapsulated packets from the GW to the MH. We have taken the mean packet size as 1 kb and the maximum packet size as 64 kb. We have fixed T_B (number of beacons to determine that the current overlay is no longer reachable) at 3, T at 80 %, number of nodes at 50 and the speed-max at 5 km/h. By using virtual cells, all packets reach destinations in both upward and downward vertical handoff.

In Figure 9, we compared the basic horizontal handoff system without virtual cell to our horizontal handoff (Wavelan to Wavelan, Ricochet to Ricochet). The Wavelan to Wavelan handoffs achieve lower handoff latency than Ricochet to Ricochet handoffs (4.25 ms in Wavelan and 201.06 ms in Ricochet).

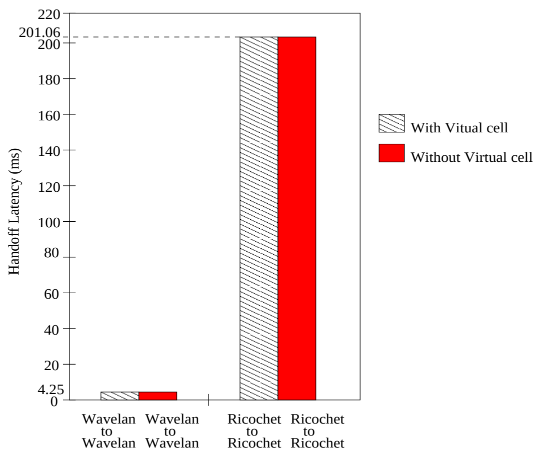

Fig. 9. Horizontal handoff latency

In Figure 10, we have taken beacon period of 200 ms. The upward vertical handoff from Wavelan to Ricochet with virtual cells presents the high performance. We note a variation of 75 % (more than 600 ms). This variation is caused by the use of the virtual cells which do not take in consideration the parameter T_B. The downward vertical handoff latency is the same in both methods.

The next simulation is to find out whether our scheme can minimize the effects of a handoff while the MH is in the midst of an Internet telephony session during handoff. We have taken two overlay levels. The lowest level has a bandwidth of 1.6 Mb/s, 100 meters of cell diameter and 2 ms of latency. The second level has a bandwidth of 1 Mb/sec, 400 meters of cell diameter and 10 ms of latency. The GWFA sends a stream of packets of 200 bytes every 20 ms to simulate the forwarding of a real time Internet telephony stream to the MH, while the MH moves between BSs. The playout delay for an interactive conversation should not be more than 200 ms (human factors studies have shown that the maximum tolerable delay for interactive conversations is approximately 200 ms).

Figures 11 and 12 show the packet inter-arrival times. The normal packet spacing is 20 ms. The first packet after the handoff, sequence number 13, is delayed by approximately 4.2 ms while the horizontal handoff in the lowest level completes. The jitter introduced by the horizontal handoff is the same in both methods (with or without using the virtual cell concept). The longest packet inter-arrival time

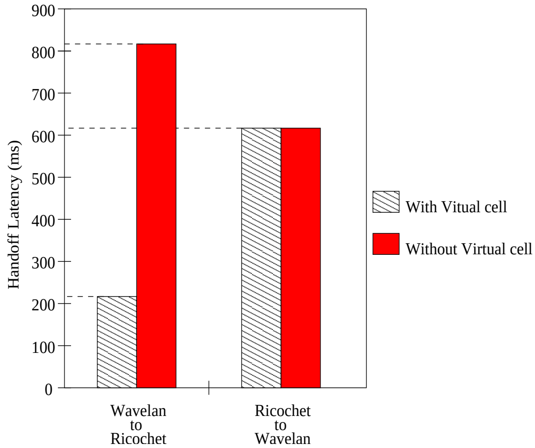

Fig. 10. Vertical handoff latency with beacon period of 200 ms

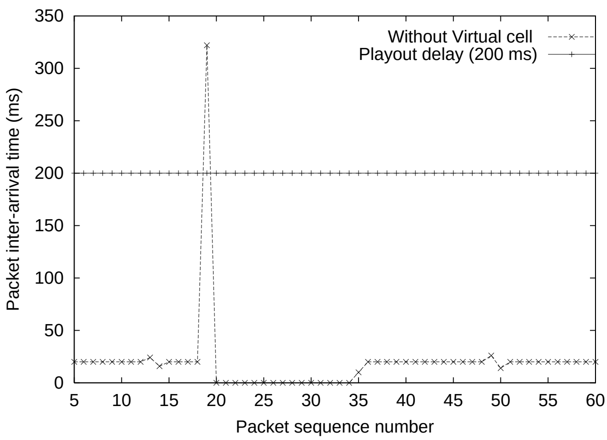

Fig. 11. Packet inter-arrival time before, during and after a basic handoff, from a stream of UDP packets that emulates an audio stream without virtual cell

is 322 ms, of which 300 ms is contributed by the $T_B \times 100$ for the upward vertical handoff without using the virtual cell and 22 ms with using the virtual cell concept. The jitter introduced by the upward vertical handoff without using the virtual cell is over the limit imposed by the playout buffer. Packets with sequence number from 20 to 35 are transmitted back-to-back after the vertical handoff from the new BS and have very short packet inter-arrival time between them. However, the jitter introduced by the vertical handoff with the using of the virtual cell concept is well below the limit (200 ms).

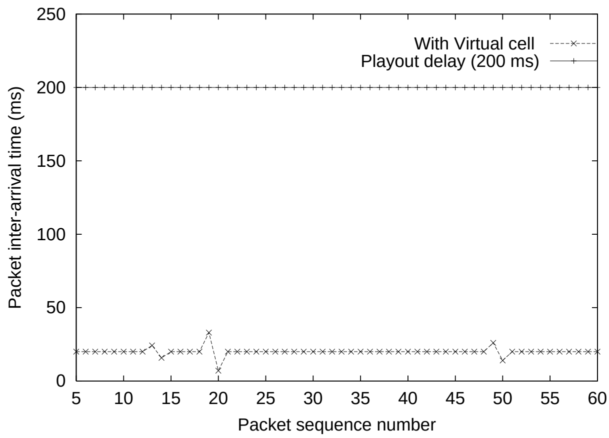

Fig. 12. Packet inter-arrival time before, during and after a basic handoff, from a stream of UDP packets that emulates an audio stream with virtual cell

From the sequence number, we can see that no packets are lost, duplicated or have arrived in the wrong sequence. We have managed this situation by setting the buffer size equal to the maximum possible amount of packet loss due to the handoff.

5 CONCLUSIONS

In this paper, we have described additions to a horizontal handoff system to support the simultaneous operation of multiple wireless network interfaces. This vertical handoff allows a mobile user to roam among multiple wireless networks. We have presented a handoff scheme with hierarchical mobility management architecture to handle the frequent handoffs. In order to reduce the upward vertical handoff latency time and loss packets, we have proposed the virtual cells concept which do not take in consideration the parameter T_B. Simulation results show the performance of our virtual cells concept to meet the requirements of an interactive communication session such as Internet telephony. The emergence of Internet telephony and teleconferencing has demonstrated that the Internet can provide acceptable quality of service to an important class of real-time applications.

REFERENCES

[1] Stemm, M.—Katz, R. H.: Vertical Handoffs in Wireless Overlay Networks. Mobile Networks and Applications, ACM Press, Vol. 3, 1999, No. 4, pp. 70–80.

[2] Perkins, C.: IP Mobility Support. RFC, 2002.

[3] Perkins, C. E.—Johnson, D. B.: Route Optimization in Mobile IP. Mobile IP Working Group, Internet Draft, Work in progress, November 1997.

[4] Tourrilhes, J.—Carter, C.: P-Handoff: A Protocol for Fine Grained Peer-to-Peer Vertical Handoff. PIMRC, 2002.

[5] Rahnema, M.: Overview of the GSM System and Protocol Architecture. Electronic Industry Association, April 1993.

[6] IEEE 802.11 standard. Wireless LAN Medium Acess Control (MAC) and Physical Layer (PHY) specifications, June 1997.

[7] Badis, H.—Al Agha, K.: Scalable Model for the Simulation of OLSR and Fast-OLSR protocols. Med-Hoc-Net '03, Mahdia, Tunisia, June 2003.

[8] AT & T WaveLAN: PC/AT Card Installation and Operation. AT & T manual, 1994.

[9] IBM Infrared Wireless LAN Adapter Technical Reference. IBM Microelectronics, Toronto Lab, 1995.

Hakim Badis is a Post-Doc researcher in computer science at the University of Computer Science Laboratory at École polytechnique, France. He received his Diplom-Engineer from the Algeria Institute of Computer Science, Algiers, in 2001. He obtained the M. Sc. and Ph. D. degrees from the University of Paris-Sud XI, France, in 2002 and 2005, respectively. His research interests include computer networking and mobile computing, random graph theory and distributed algorithms. He is a member of QOLSR team at LRI laboratory in Paris.

Khaldoun Al Agha received the State Engineering Degree from École Supérieure d'Électricité (Supelec), Paris, France, in 1993. He obtained the Master's Degree and the Ph. D. from the Versailles University, France, in 1995 and 1998, respectively. In 2002, he received the HDR (Habilitation à diriger des recherches) from Paris XI University. He was Assistant Professor in Versailles University in 1998. In 1999, he joined INRIA (Institut National de la Recherche en Informatique et en Automatique) for one year where he worked on the European Project BRAIN on full IP networks. He is Full Professor at Paris XI University. He creates and conducts the Networking Group in the LRI laboratory (Laboratoire de Recherche en Informatique). He participates in different projects on mobility (SAMU, Arcade and SAFARI). His research interests includes resource allocation and IP mobility for cellular (GSM, UMTS) and ad-hoc networks.