A study of mobility in WLAN

Fengping Li
Helsinki University of Technology
Telecommunication Software and Multimedia Laboratory
fli@cc.hut.fi

Abstract

This paper studies mobility in wireless LAN (WLAN, IEEE802.11). Mobility is an instinct nature of WLAN as being wireless. But this mobility is limited by radio propagation and the IEEE802.11 specifications. For seamless mobility in a larger area, e.g., between two extended service sets (ESSs), some sorts of mobility management mechanisms are needed. Mobile IP has been widely considered as a solution for realizing mobility in IP networks. However, in WLAN, the speed of handover is a big challenge, because it is critical for real-time applications such as voice over IP. To counteract these challenges, many approaches for enhancing the performance of Mobile IP have been proposed. This paper will mainly study two mechanisms, namely Hierarchical Mobile IP and link-layer information (or Hint). By combining these enhancement based on Mobile IP, seamless mobility in WLAN should be achieved.

[KEYWORD: WLAN, seamless mobility, speed of handover, Mobile IP, hierarchical Mobile IP, link-layer information]

1 Introduction

WLAN has become more and more popular recently. It is used in enterprises, public places, and individual homes. One of the key motivations of deploying WLAN is mobility. That is, we can leave the cables behind. However, as new applications emerge, such as voice over IP (VoIP), challenges arise facing WLAN's capability of mobility. Solutions are needed to solve these problems.

Before we try to find appropriate solutions, we should first study mobility of WLAN. Mobility can be categorized as micro-mobility and macro-mobility. In WLAN, micro mobility can be defined as mobile nodes (MN) move within an Extended Service Set (ESS) [2][10] while keep connecting with the network during an ongoing session. Macro mobility can be defined as MNs move between ESSs while keep connecting with the network during an ongoing session. Compared to macro mobility, micro mobility can be more easily realized because there is some kind of support by 802.11 specifications [2]. Macro mobility faces more challenges because 802.11 specifications do not support mobility between two ESSs [2]. This paper will mainly focus on how to counteract these challenges. Mobile IP [1] is one solution for macro mobility. However for real-time applications such as VoIP, the speed of handover is critical. Mobile IP alone, as we will see, has difficulties to handle fast handover between different IP subnets (e.g., two ESSs). So we still need some enhancement based on Mobile IP.

The following chapters of this paper are organized as follows: chapter 2 introduces the system architecture of WLAN and studies mobility in WLAN, chapter 3 presents solutions to macro mobility, including Mobile IP, Hierarchical Mobile IPv6 and Link-layer information (or Hint), and chapter 4 conclusion sums up the ideas.

2 Overview of mobility in WLAN

In this chapter we will introduce the concept of mobility in WLAN, as well as problems related to satisfied performance of mobility. Before we start to study mobility of WLAN, we will first have an overview of IEEE802.11system.

2.1 802.11 WLAN overview

WLAN provides a wireless way to access the Internet, as LAN does through a wired way. WLAN standards are specified by the Institute of Electrical and Electronics Engineers (IEEE). Currently there are three standard versions for WLAN, 802.11a, 802.11b and 802.11g. Among them the most mature one is 802.11b which operates at the frequency band of 2.4GHz with a maximum throughput of 11 Mbps. 802.11g is an upgraded version of 802.11b. It operates at the same frequency band as 802.11b with an improved throughput. 802.11a operates at 5.4 GHz with a maximum throughput of 54 Mbps. Today the WLANs employed in the market are mostly 802.11b because it is cheaper and more technically sophisticated.

In this paper, the discussion of mobility does not distinguish between these versions. We use 802.11 to refer to all these variant standards mentioned above. And WLAN refers to IEEE802.11 WLAN.

2.1.1 WLAN architecture and network components

In 802.11 specification[2] and Achieving Seamless IPv6 Handover in 802.11 Network[10], the WLAN network components are defined as follows and are used in this paper.

- Access Point (AP) A key component of WLAN, which is a layer-2 entity that has station functionality and provides access to the distribution services, via the wireless medium for associated nodes.

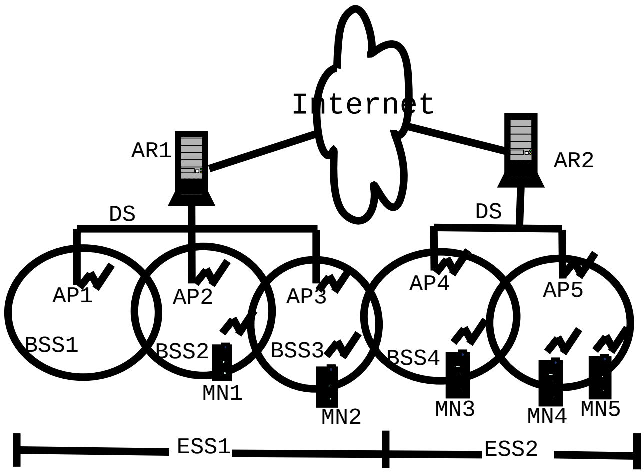

Figure 1: WLAN architecture and components

- Basic Service Set (BSS) The building block of WLAN, which is a set of nodes controlled by a single coordination function, where the coordination function may be centralized (e.g., in a single AP) or distributed (e.g., for an ad-hoc network). The BSS can be thought of as the coverage area of a single AP. Here a node is a machine (e.g., a note-book computer, a PDA, an IP phone, etc) served by the AP. When the node is moving, we call it a mobile station or a mobile node (MN).
- Distribution System (DS) A system that interconnects a set of basic service sets (BSSs) and integrates local area networks (LANs) to create an extended service set (ESS).
- Extended Service Set (ESS) A set of one or more interconnected basic service sets (BSSs) and integrated local area networks (LANs). An ESS appears as a single BSS to the logical link control layer at any node associated with one of those BSSs. The ESS can be thought of as the coverage area provided by a collection of APs all interconnected by the Distribution System. It may consist of one or more IP subnets.
- Access Router (AR) An Access Network Router resides on the edge of an access network and offers IP connectivity to mobile nodes.

Figure 1 shows a typical architecture of WLAN and its components. It consists of two ESSs which belong to different IP subnets.

2.2 WLAN mobility category

The mobility discussed in this paper differs from roaming where a session is simply terminated and restarted in a new WLAN cell. Handover is not required in roaming but is required in mobility to keep an ongoing session (e.g., VoIP) alive when a MN crosses cell boarders. Here, handover is defined as a MN moves from a cell to the next when kepping an ongoing session alive. A cell is defined as the coverage area of an AP. Mobility can be categorized as micro mobility and macro mobility as defined in chapter 1.

2.2.1 Mobility within a basic service set (BSS)

This is the basic mobility of WLAN as being wireless, which is shown in Figure1 as MN1 moving within BSS2. An AP covers a certain service area. The size of the area depends on the transmission power, the radio propagation environment, antenna sensitivity, etc. After association with an AP, a MN is capable of moving within the service area of the AP while remains connected with the network. This kind of mobility is the basic benefit WLAN offers to its users. By leaving the cable behind, users gain more freedom. In this case, no transition occurs because there is no handover in any layer. We can refer to this situation as 'static'.

2.2.2 Mobility between BSSs within an extended service set (ESS)

This is what we have defined as micro mobility, which is shown in Figure1 as MN1 or MN2 moving between BSS1, BSS2 and BSS3. Usually there are overlaps of AP's service areas. A MN continuously monitors the signal strength and quality from all APs to which it can communiate within an ESS. And each AP in the ESS is also aware of the location of the MN. When a MN moves out of an AP's service area and into another AP's, a transition occurs. For example, as shown in Figure1, MN3 leaves BSS4 and enters BSS5. The MN uses the re-association service to associate with the new AP without losing connection with the network.

The transition between APs requires the cooperation of APs to exchange information necessary for a successful handover. For example, as shown in Figure 1, if MN1 moves into BSS3, AP3 will inform AP2 that the MN1 is now associated with AP3, and then AP2 will deliver all the delayed packets to AP3 if any. Thus essentially micro-mobility involves intra-domain handovers. There is no need for external coordination. Issues such as timing, call control and handover control can be set (or bounded) by network design.

However, 802.11 does not specify the details of the communications between APs during the transition[2]. That would be a problem if two APs are manufactured by different vendors. Inter Access Point Protocol (IAPP) [3] should be able to solve this problem, which is under standardization by IEEE802.11 work group. Moreover, efforts are being taken by the IEEE 802.11i (security) and 802.11e (QoS) task groups in order to improve both AP-to-AP handovers and authentication [4].

2.2.3 Mobility between two ESSs

This is what we have defined as macro mobility, which is shown in Figure1 as MN2 moves form BSS3 to BSS4, which belong to ESS1 and ESS2 respectively. The two ESSs belong to diferent IP subnets which may be adminstrated by different organizations like two companies, two hot-spot operators, etc.

2.2.4 Problems of macro mobility

Because there is no direct link-layer communcation between the two ESSs, to realize seamless IP-layer mobility will have some problems, for example, handover speed, which is critical for real-time applications. Thus some mechanisms are needed to solve these problems in order to achieve seamless IP-layer mobility. Macro mobility is the main subject of this paper. The following solutions are used to solve problems related to macro mobility.

3 Solutions to macro mobility

To solve the problems of macro mobility, a MN needs to quickly acquire a new IP address when moving from one ESS to another without disrupting an ongoing session,e.g., a VoIP conversation. One solution is Mobile IP that helps realize IP-layer handover. However, in order to achieve fast handover, Mobile IP alone is not sufficient so enhancement mechanisms are needed. This seesion we will discuss these mechanisms, namely hierarchical network architecture and link-layer information.

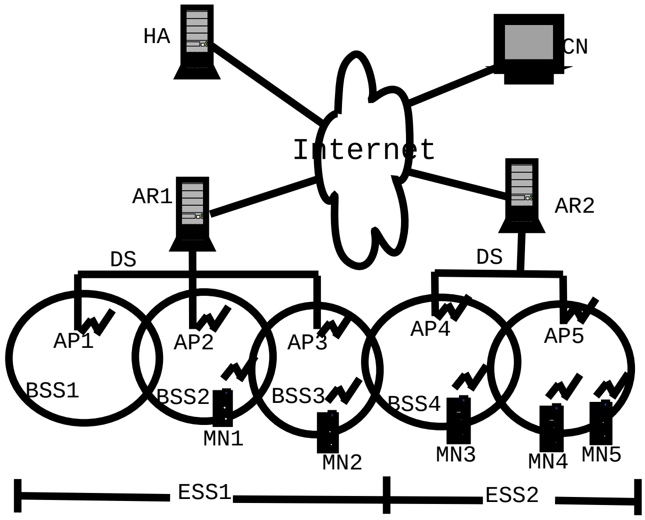

Figure 2: Mobile IPv6 network architecture and operation

3.1 Mobile IP

3.1.1 Mobile IP overview

The Internet Protocol (IP) was originally designed without any mobility support. When a MN changes its point of attachment to an IP network and moves to a new location, it would be unable to send or receive traffic.

The Mobile IP is an extension to the Internet protocol and is designed for Internet mobile node support. Mobile IP provides a function similar to the post-office forwarding scheme in order to provide network connectivity to moving mobile nodes. Mobile IP introduces two new network entities, namely the Home Agent (HA) and Foreign Agent (FA). There are two versions of Mobile IP, Mobile IPv4 and Mobile IPv6. Mobile IPv6 introduces some differences and improvement, for instance, it omits the FAs by distributing their functionality amongst the MNs and the network infrastructure. Moreover, in Mobile IPv6, support for route optimization is a fundamental part of the protocol, rather than a nonstandard set of extensions as in Mobile IPv4, so that the problem of 'tri-angel routing' is solved. In this paper,we mainly discuss Mobile IPv6 and use Mobile IP to refer to.

3.1.2 Mobile IP operation

In Mobile IP, a MN is always identified and reachable using its home address, whenever it is at its home network or at a foreign network. The “home address” is an IP address assigned to the MN within its home subnet prefix on its home link. When a MN is at home,packets addressed to its home address are routed to the MN's home link, using conventional Internet routing mechanisms.

A Mobile IPv6 network architecture is shown in Figure 2, and the following descrption of Mobile IP operation is based on Figure 2.

When a MN is attached to some foreign network away from home, for instance, MN2, as shown in Figure 2,(assuming MN2 is away from home) is visiting ESS1, it acquires a care-of address (CoA) from the foreign network (ESS1) through some conventional IP mechanisms.After that the MN2 registers this CoA with its HA by sending a “Binding Update” message to the HA. The HA then replies to MN2 by returning a “Binding Acknowledgement” message. A binding is the association between a MN's home address and CoA. MNs can also provide information about their current location to correspondent nodes (CN). In this example as shown in Figure 2, MN2 informs CN about its CoA. This happens through the correspondent registration. As a part of this procedure, a return routability test is performed in order to authorize the establishment of the binding.

There are two modes of communication between a MN and a CN. The first mode is bi-directional tunneling. In this mode, as shown in Figure 2, the HA operates as a proxy. Packets from the CN are routed to the HA and tunneled to MN2. Packets from MN2 are tunneled to the HA and then routed normally to the CN. This tunneling is performed using IP encapsulation. This mode needs neither Mobile IP support for CN nor CN registration. That is, CN does not know where MN2 is. It can see that the first mode introduces 'tri-angel routing' which may cause HA congestion and overload the home network. This problem can be solved by the second mode, 'route optimization', which requires the MN registers its current binding at the CN. Thus packets can be sent directly between the MN and the CN without the assist of HA. In this mode, as shown in Figure 2, when IPv6 is used, the CN sets the Destination Address in the IPv6 header to CoA of the MN2. A new type of IPv6 routing header is also added to the packet to carry the desired home address of MN2. Similarly, MN2 sets the Source Address in the packet's IPv6 header to its current CoA. MN2 adds a new IPv6 'Home Address' destination option to carry its home address. The inclusion of home addresses in these packets makes the use of the CoA transparent above the network layer (e.g., at the transport layer).

Above is a basic operation of Mobile IP. For more detailed information please refer to [1] and [6].

3.2 Mechanisms for enhanced performance

Mobile IP alone would not be sufficient to solve all the problems. We have known that one of the key requirement for seamless mobility is handover speed, specially for delay-sensitive applications like VoIP. Because if the MN is far from home or CN, the propagation time of binding messages between the MN and HA or CN would be too long thus make the handover delay unacceptable. Furthermore, for the cases of many MNs or frequent handovers, the huge amount of binding messages would consume too much network capacity. To solve these problems, there have been many proposals such as MIPv6 [6], Fast Handovers [7], HMIPv6[8], IDMP[5], Link-layer information [9] and Seamless IPv6 handover [10],etc.

Among them, there are two fundamental mechanisms we think would be efficient for WLAN to achieve seamless handover between two ESSs (macro mobility). They are hierarchical routing scheme and Link-layer information, which are discussed mainly in [8] and in [9] respectively. However, some other proposals are more or less involved one or both of these two mechanisms or alike.

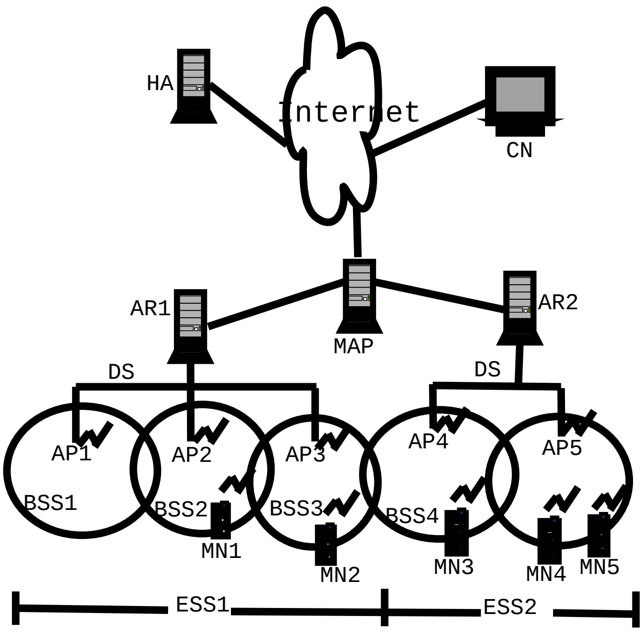

Figure 3: HMIPv6 system architecture

3.2.1 Hierarchical routing scheme

HMIPv6 Overview

This scheme is based on Mobile IPv6. So we may term this scheme as Hierarchical Mobile IPv6 (HMIPv6), which introduces a new function, the Mobility Anchor Point (MAP), and minor extensions to the mobile node operation. The correspondent node and home agent operation will not be affected. The MAP is a router located in a network visited by the MN. The MAP is used by the MN as a local HA. One or more MAPs can exist within a visited network.

The introduction of MAP solves the problems described above in the following ways:

- The MN sends Binding Updates to the local MAP rather than the HA and CNs;
- Only one Binding Update message needs to be transmitted by the MN before traffic from the HA and all CNs is re-routed to its new location. This is independent of the number of CNs that the MN is communicating with.

Thus a MAP is essentially a local Home Agent.

HMIPv6 operation

A HMIPv6 architecture is shown in Figure 3, and the description of HMIPv6 is based on this Figure 3.

Before we study the operation of HMIPv6, two concepts should be introduced first. One is Regional Care-of Address (RCoA), which is obtained by the MN from the visited network. An RCoA is an address on the MAP's subnet. It is auto-configured by the MN when receiving the MAP option. Another is On-link CoA (LCoA). The LCoA is the on-link CoA configured on a MN's interface based on the prefix advertised by its default router. In Mobility support in IPv6[6] this is simply referred to as the Care-of-address. However, in this paper, LCoA is used to distinguish it from the RCoA.

A MN entering a MAP domain will receive Router Advertisements containing information on one or more local MAPs. For example, as shown in Figure 3, MN2 receives advertisements from AR1, and the advertisement contains information about MAP. The information includes the MAP's global IP address(RCoA). This procedure is called MAP discovery. The discovery phase will also inform the MN2 of the distance of the MAP from the MN2. It should be noted that the MAP function could be implemented in the MAP, as shown in Figure 3, and at the same time also in AR1 and AR2. In this case a MN can choose the first hop MAP or one further up in the hierarchy of routers.

MN2 can bind its current location (on-link CoA) with an address on the MAP's subnet (RCoA) and registers the RCoA with its HA and CNs. Now the MAP acts as a local HA, packets sent to or from the MN are forwarded by the MAP, other parts of the communication is carried out normally as described in Mobile IP. If the MN changes its current address (LCoA) within a local MAP domain, for example, MN2 moves among BSS3, BSS2 and BSS1, it only needs to register the new address with the MAP. So, take MN2 as an example, only one Binding Update is needed between MN2 and its HA and CN. As a result, as long as MN2 moves within the domain of the MAP, the movement is transparent to MN2's HA and CN to which it is communicating with. Here a MAP domain's boundaries are defined by the Access Routers (ARs) advertising the MAP information to the attached MNs. As shown in Figure 3, the MAP domain consists of AR1 and AR2, which advertise MAP's information in both ESS1 and ESS2.

The process of MAP discovery continues as the MN from one subnet to the next. The MN will keep detecting whether it is still in the same MAP domain. The router advertisement used to detect movement will also inform the MN, through the MAP option, whether it is still in the same MAP domain. Whenever a MN detects a change in the advertised MAP's address (that means it has moved to a new MAP domain), the MN must act on the change by sending Binding Updates to its HA and CNs.

As a result, taking MN2 as an example, when MN2 first enters BSS3, it needs to perform a Binding Update once with its HA and CN as long as it moves within ESS1 and ESS2. That is, when MN2 moves from BSS3 (belongs to ESS1) to BSS4 (belongs to ESS2), it does not need to register with HA or CN. While in normal Mobile IP operation as shown before, this kind of registration is required.

Therefore, this hierarchical Mobile IPv6 scheme achieves faster handover by reducing Binding Update between a MN and its HA and CNs.

3.2.2 Link-layer information

In chapter 2 we have studied the mobility in WLAN. Where we know that one of the difficulties for mobility between ESSs is the lack of link-layer support. In this chapter we will study a mechanism of speeding up the Mobile IP handover with the support of link-layer information. Before introduce the link-layer information,summarize the Mobile IP process is necessary.

Summary of Mobile IP process

From the above study of Mobile IP and hierarchical routing scheme, we can summarize the processes of Mobile IP handover. Firstly, whenever a MN leaves a network and enters another, a handover is required because there is a transition (location switch) between cells. This transition is an 802.11 link-layer (L2) handover. During this transition a Mobile IP handover takes place if the two cells belong to different IP networks (subnets) so that the MN can acquire a new IP address and keep the IP connection alive. Mobile Agents (HA or FA) assist the Mobile IP handover. However no all L2 handovers are followed by Mobile IP handover. For instance, if one ESS only consists of one IP subnet then movement of a MN within this ESS will not require Mobile IP handover. Secondly, whenever a MN moves it is required to detect its movement by discovering a change in its environment, Thirdly, Whenever a MN enters a new cell and acquires a new IP address (CoA), it must register this new CoA with its HA and CNs (if route optimization is used). From this summary we can determine that a Mobile IP handover is the sum of the following individual processes [9]:

- Link-layer handover;
- Mobile IP movement detection;
- Mobile IP registration.

In previous chapters we have known that a Mobile IP registration involves the exchanges of a Binding Update message and a Binding Ack message between the MN and its HA and CNs (if route optimization is used). This duration is equivalent to the round-trip propagation time between the aforementioned entities including processing delay. We have proposed a hierarchical Mobile IP scheme in chapter 3.2.1 to reduce the binding messages between a MN and HA and CNs.

Before we study the link-layer information scheme, we will investigate how mobile IP movement detection and agent discovery are performed.

Movement detection and agent selection

Mobile IP handovers occurs as a consequence of link-layer handovers between different IP networks (e.g., two ESSs). Due to layer independence, the event of a link-layer handover cannot be communicated to Mobile IP so alternative means must be used to determine this information. For this, in Mobile IP Mobility Agents (HA, FA or AR) must periodically broadcast advertisements to indicate their existence, so that a MN can discover these agents by receiving their advertisements. Moreover, a MN determines its location by evaluating these advertisements. That is, the receipt of a new agent advertisement from an "undiscovered" mobility agent is perceived as an indication of movement into a new network. Similarly, the loss of contact with a "discovered" agent is perceived as an indication of movement out of a network

After evaluating the received advertisements, the MN will perform Agent Selection to select an appropriate agent before it carries out registration.

Both actions, movement detection and agent selection, may consume significant time [9]. For example, the speed of movement detection directly depends on the advertisement interval. A smaller advertisement interval leads to optimum movement detection. However a smaller interval will also result in less network efficiency. So a trade-off between performance of movement detection and network capacity utility is observed.

With the link-layer information, Mobile IP movement detection and agent discovery will not be required. Moreover the existence of link layer information makes the periodic mobility agent broadcast advertisements become unnecessary thus it enables a more effective utilization of the network capacity.

Use link-layer information for movement detection

This link-layer information used for movement detection is termed hint [11]. A hint is used to indicate the occurrence of a link-layer handover. In Hinted Cell Switch (HCS) algorithm [11],a MN, instead of waiting for the Mobile Agent advertisement, requests for an advertisement upon receiving a hint. This is done through the broadcast of an agent solicitation that forces all neighboring Mobility Agents to respond with an immediate unicast advertisement. However, in environments with a large number of roaming nodes that are required to broadcast solicitations with every location change, the effect of less capacity efficiency illustrated before may again be encountered.

To solve this problem, we can extend the amount of information communicated from the link layer to Mobile IP to include information about the environment such as the identity of the local Mobility Agent. Through this, the need for movement detection and agent selection is not required. This enhanced algorithm is call Fast HCS (FHCS) [9][11]. FHCS determines the identity of the local mobility agents through link-layer hints. A mobility agent is identified by the pair of the agent's IP and hardware addresses. However, the way the link-layer determines this information is bearer specific. In 802.11, implementation of HSCS requires that the mobility agent's information be transported within the network's Service Set IDentifier (SSID) field. As such, any link-layer handovers between two IP networks would be associated with a change in the SSID which would also reveal the identity of the local mobility agents. In that case, a MN may directly proceed to registration without going through broadcast solicitation as well as the Agent Selection required by HCS.

4 Conclusion

In this paper different categories of mobility in WLAN are introduced. Bigger challenges arise when a MN move from an ESS to the next. To achieve the goal of seamless mobility in WLAN, a fast Mobile IP handover is the key. Mobile IP serves as a base to realize mobility. In this paper we have studied the basic operation of Mobile IP. However, Mobile IP alone is insufficient to handle fast handover. For this reason, we have also studied some mechanisms such as Hierarchical Mobile IP and link-layer information or Hint. In the hierarchical scheme, a MAP is introduced to reduce the binding messages between MN and HA or CNs; In the FHCS scheme, local Mobile Agents' identifiers (e.g., IP and MAC addresses) are included in SSID field so that a mobile node can directly go to registration phase after receiving these hints.

As a result, the combination of these two mechanisms may achieve faster Mobile IP based handover. Thus performance of delay-sensitive applications like Voice over IP would not be compromised even if a MN is crossing a boarder between two WLAN networks.

References

[1] C. E. Perkins. IP Mobility Support. RFC2002, IETF Network Working Group, Oct. 1996.

[2] Matthew S. Gast. 802.11 Wireless Networks: The Definitive Guide. O'Reilly & Associates,Inc,2002.

[3] IEEE Std 802.11F, 2003 Edition. IEEE Trial-Use Recommended Practice for Multi-Vendor Access Point Interoperability via an Inter- Access Point Protocol Across Distribution Systems Supporting IEEE 802.11 Operation.

[4] Including VoIP over WLAN in a Seamless Next-Generation Wireless Environment. Texas Instruments, International Engineering Consortium, 2003. http://www.iec.org/online/tutorials/ti_voip_wlan/

[5] Subir Das, Anthony J McAuley, and Ashutosh Dutta, Archan Misra,Kaushik Chakraborty, Sajal K.Das. IDMP:An Intra-Domain Mobility Management Protocol for Next Geneation Wireless Networks. Telcordia Technologies Inc. http://www1.cs.columbia.edu/ dutta/research/idmp.pdf

[6] D.Johnson, C. Perkins, J. Arkko. Mobility Support in IPv6. Draft-ietf-mobileip-ipv6-24.txt, IETF Mobile IP Working Group, Dec.2003.

[7] K. El. Malki. Low Latency Handoffs in Mobile IPv4. IETF Network Working Group, Jan. 2004.

[8] Hesham Soliman, Claude Castelluccia, Karim El-Malki, Ludovic Bellier. Hierarchical Mobile IPv6 mobility management (HMIPv6), draft-ietf-mipshop-hmipv6-00.txt. IETF Mobile IP Working Group, June 2003.

[9] N.A.Fikouras, A.J.Könsgen, C.Görg. Accelerating Mobile IP hand-offs through Link-layer Information, An Experimental Investigation with 802.11b and Internet Audio. Department of Communication Networks, University of Bremen. http://www.comnets.uni-bremen.de/ niko/publications/mmb01.pdf

[10] Paul Tan. Recommendations for Achieving Seamless IPv6 Handover in IEEE 802.11 Networks. draft-paultan-seamless-ipv6-handoff-802-00.txt. Internet Draft, IETF, Feb 2003.

[11] Nikolaus A. Fikouras and Carmelita Görg. A Complete Comparison of Algorithms for Mobile IP Hand-offs with Complex Movement Patterns and Internet Audio. Department of Communication Networks, University of Bremen. http://www.comnets.uni-bremen.de/ niko/publications/wpmc01-mip.pdf