A SPEED ADAPTIVE MOBILE INTERNET PROTOCOL
OVER WIRELESS LOCAL AREA NETWORK

By

JUN TIAN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2005
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>List</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ix</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2 RELATED WORK</td>
<td>5</td>
</tr>
<tr>
<td>Network Layer Handoff Management</td>
<td>5</td>
</tr>
<tr>
<td>Mobile IP</td>
<td>6</td>
</tr>
<tr>
<td>Agent Discovery</td>
<td>8</td>
</tr>
<tr>
<td>Registration</td>
<td>10</td>
</tr>
<tr>
<td>Tunneling</td>
<td>12</td>
</tr>
<tr>
<td>Hierarchical MIP</td>
<td>15</td>
</tr>
<tr>
<td>Cellular IP</td>
<td>16</td>
</tr>
<tr>
<td>Routing</td>
<td>17</td>
</tr>
<tr>
<td>Handoff</td>
<td>17</td>
</tr>
<tr>
<td>Paging</td>
<td>19</td>
</tr>
<tr>
<td>HAWAII</td>
<td>20</td>
</tr>
<tr>
<td>Wireless LAN</td>
<td>24</td>
</tr>
<tr>
<td>Technology Overview</td>
<td>24</td>
</tr>
<tr>
<td>The IEEE 802.11 Established Standards</td>
<td>25</td>
</tr>
<tr>
<td>Standard 802.11</td>
<td>26</td>
</tr>
<tr>
<td>Standard 802.11b</td>
<td>28</td>
</tr>
<tr>
<td>Standard 802.11a</td>
<td>29</td>
</tr>
<tr>
<td>Standard 802.11g</td>
<td>31</td>
</tr>
<tr>
<td>Pending Specifications Within the 802.11 Suite</td>
<td>33</td>
</tr>
<tr>
<td>The IEEE 802.11 Wireless LAN Architecture</td>
<td>33</td>
</tr>
<tr>
<td>Wireless LAN Station</td>
<td>34</td>
</tr>
<tr>
<td>Basic Service Set (BSS)</td>
<td>34</td>
</tr>
<tr>
<td>Independent Basic Service Set (IBSS)</td>
<td>34</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Implementation of Speed Adaptive MIP</td>
<td>82</td>
</tr>
<tr>
<td>Home Agent</td>
<td>82</td>
</tr>
<tr>
<td>Mobile Node</td>
<td>84</td>
</tr>
<tr>
<td>Foreign Agent</td>
<td>86</td>
</tr>
<tr>
<td>Evaluation of Speed Adaptive Extension for MIP</td>
<td>87</td>
</tr>
<tr>
<td>6 SUMMARY AND FUTURE WORKS</td>
<td>90</td>
</tr>
<tr>
<td>LIST OF REFERENCES</td>
<td>91</td>
</tr>
<tr>
<td>BIOGRAPHICAL SKETCH</td>
<td>97</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1 Characteristics specified by the 802.11 standard</td>
<td>28</td>
</tr>
<tr>
<td>2-2 Characteristics specified by the 802.11b standard</td>
<td>29</td>
</tr>
<tr>
<td>2-3 Characteristics specified by the 802.11a standard</td>
<td>31</td>
</tr>
<tr>
<td>2-4 Characteristics specified by the 802.11g standard</td>
<td>32</td>
</tr>
<tr>
<td>2-5 Comparison of characteristics specified within the IEEE 802.11 suite</td>
<td>33</td>
</tr>
<tr>
<td>3-1 Average throughput at different speeds and AP distances</td>
<td>58</td>
</tr>
<tr>
<td>4-1 Handoff latency distribution of MIP over WLAN</td>
<td>71</td>
</tr>
<tr>
<td>5-1 Average throughput for speed-adaptive MIP</td>
<td>89</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>Macro and Micro mobility</td>
<td>6</td>
</tr>
<tr>
<td>2-2</td>
<td>Three functional entities of MIP</td>
<td>8</td>
</tr>
<tr>
<td>2-3</td>
<td>IP in IP encapsulation and minimal encapsulation</td>
<td>13</td>
</tr>
<tr>
<td>2-6</td>
<td>802.11 wireless LAN handoff procedure</td>
<td>39</td>
</tr>
<tr>
<td>3-1</td>
<td>RAMON testbed architecture</td>
<td>51</td>
</tr>
<tr>
<td>3-2</td>
<td>Dynamic MIP sample scenario</td>
<td>53</td>
</tr>
<tr>
<td>3-3</td>
<td>Time-sequence graph at speed 20m/s and AP distance 1000m</td>
<td>54</td>
</tr>
<tr>
<td>3-4</td>
<td>Throughput graph at speed 20m/s and AP distance 1000m</td>
<td>54</td>
</tr>
<tr>
<td>3-5</td>
<td>Time-sequence graph at speed 80m/s and AP distance 1000m</td>
<td>55</td>
</tr>
<tr>
<td>3-6</td>
<td>Throughput graph at speed 80m/s and AP distance 1000m</td>
<td>55</td>
</tr>
<tr>
<td>3-7</td>
<td>Time-sequence graph at speed 10m/s and AP distance 500m</td>
<td>56</td>
</tr>
<tr>
<td>3-8</td>
<td>Throughput graph at speed 10m/s and AP distance 500m</td>
<td>56</td>
</tr>
<tr>
<td>3-9</td>
<td>Time-sequence graph at speed 40m/s and AP distance 500m</td>
<td>57</td>
</tr>
<tr>
<td>3-10</td>
<td>Throughput graph at speed 40m/s and AP distance 500m</td>
<td>57</td>
</tr>
<tr>
<td>3-11</td>
<td>Average throughputs vs speeds</td>
<td>59</td>
</tr>
<tr>
<td>3-12</td>
<td>Time-sequence graph at AP distance 1000m with speed 20m/s without handoff</td>
<td>60</td>
</tr>
<tr>
<td>3-13</td>
<td>Time-sequence graph at AP distance 1000m with speed 80m/s without handoff</td>
<td>60</td>
</tr>
<tr>
<td>3-14</td>
<td>Average throughput vs handoff rate</td>
<td>62</td>
</tr>
</tbody>
</table>
4-1  Handoff procedure with message exchange..............................................................67
4-2  LCS handoff latency for MIP...................................................................................67
4-3  ECS handoff latency for MIP....................................................................................68
4-4  Handoff procedure with handoff latency distribution..............................................72
5-1  Traditional MIP Handoff Procedure........................................................................74
5-2  FA Set size vs handoff rate....................................................................................79
5-3  Normal vendor/organization specific extension......................................................80
5-4  SA-MIP handoff procedure.....................................................................................81
5-5  Function flowchart of registration in HA.................................................................84
5-6  Registration request message format....................................................................85
5-7  Registration request message extension format.....................................................85
5-8  Function flowchart of sending registration request...............................................86
5-9  Function flowchart for FA handling registration request........................................87
5-10 Time-sequence graph at speed 60m/s and AP distance 1000m...............................88
5-11 Time-sequence graph at speed 80m/s and AP distance 1000m...............................88
5-12 Average throughput vs. handoff rate......................................................................89
Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

A SPEED ADAPTIVE MOBILE INTERNET PROTOCOL OVER WIRELESS LOCAL AREA NETWORK

By

Jun Tian

December 2005

Chair: Abdelsalam (Sumi) Helal
Major Department: Computer and Information Sciences and Engineering.

This dissertation presents two novel contributions in the area of mobile network communication. The first is the performance/moving speed relationship of Mobile Internet Protocol (MIP) over Wireless Local Area Network (LAN). In this dissertation, the rapid mobility of MIP over Wireless LAN is emulated on a testbed. The performance of MIP over Wireless LAN at different moving speeds is evaluated. The result shows that current MIP protocol is not suitable for rapid moving environments. This dissertation analyzes the emulation results and depicts the relationship between the performance and the moving speed of the mobile devices. This relationship is used in a novel protocol, which is the second contribution, to improve the performance of MIP over Wireless LAN in rapid moving environments. The second contribution is the Speed Adaptive Mobile IP. In the Speed Adaptive Mobile IP, Mobile Node's registration message is extended by speed
extension. With the speed information popularized in the mobile IP network, the behavior of the Speed Adaptive Mobile IP will automatically adapt to the speed of the Mobile Node so that the performance of the Speed Adaptive Mobile IP won’t decline dramatically in a rapid moving environment. At the same time, the Speed Adaptive Mobile IP only uses reasonable resources that are enough for seamless handoff. The emulation result shows that the Speed Adaptive MIP greatly improves the performance of MIP over Wireless LAN in rapid-moving environments.
CHAPTER 1
INTRODUCTION

The population living on the world wide internet is exploding. According to the analysis of Internet usage across more than 50 countries, the latest report from Computer Industry Almanac(CIA) Inc.'s shows that as of the end of March 2004, there are 945 millions of internet users world wide. The report also indicates 1.12 billion Internet users projected for the end of 2005, and 1.46 billion for 2007. A significant number will be using wireless devices such as Web-enabled cell phones and PDAs to go online. In America, 27.9% of 193 millions of internet users are using wireless internet. At the end of 2007, 46.3% of 263 millions will be wireless internet users.

Throughout history, the economic wealth of people or a nation has been closely tied to efficient methods of transportation. The transportation speed is becoming faster and faster. A person can drive a car on high way at speed of 70mph. Some high speed trains such as France TGV, Japanese bullet, German maglev can travel at speeds of over 300km/hour(186mph). Could those people surf the internet, communicate with families and enjoy an online movie while traveling at high speeds? Could the current network infrastructure support rapid mobility?

While TCP/IP successfully overcomes the barriers of time and distance in a wired network, mobile IP is a promising technology to eliminate the barrier of location for the
increasing wireless internet usage. Third generation (3G) services combine high speed mobile access with IP-based services. With access to any service anywhere, anytime, from one terminal, the old boundaries between communication, information sharing, media distribution will disappear. 3G enables users to transmit voice, data, and even moving images whenever and wherever. But, 3G networks are not based on only one standard, but a set of radio technology standards such as cdma2000, EDGE and WCDMA. Mobile IP [Perk02] can be the common macro mobility management framework to merge all these technologies in order to allow mobile users to roam between different access networks. These radio technologies only need to handle Micro mobility issues such as radio specific mobility enhancements. Mobile IP is different from other efforts for doing mobility management in the sense that it is independent to any specific access technology[Mobi03].

Wireless local area networks (WLAN) have experienced incredible growth over recent years. WLANs provide wireless users with an always-on, wireless connection to each other, to local area networks (LAN), to wide area networks (WAN), and to the Internet. The major benefit of WLANs over wired network is its flexibility and mobility [Kapp02]. There are currently two major WLAN standards, and both operate using radio frequency (RF) technology. The two standards have heretofore been colloquially referred to as 802.11b and 802.11a. 802.11b operates in the radio frequency (RF) band between 2.4 and 2.485GHz while 802.11a operates between 5.15-5.35GHz and 5.725-5.825GHz. The performance of both 802.11b and 802.11a decreases as your distance from the antenna increases. This degradation is neither linear nor granular. Instead, each wireless
specification has a handful of pre-defined bandwidth levels at which it can operate (802.11b has four, while 802.11a has seven). Take 802.11b as an example. Within a closed office, the bandwidth will drop from 11, 5.5, 2 to 1mbps when the distance increases from 25, 35, 40 to 50 meters. For outdoors, the bandwidth will drop from 11, 5.5, 2 to 1mbps when the distance increases from 160, 270, 400 to 550 meters. So if you want to keep a high throughput, you have to reduce the distance between access points. For example, to keep 5.5mbps when outdoors, the distance between two access points should be no more than 500 meters. The smaller the cell the higher the bandwidth you get.

The use of current cellular/PCS high data rate services for data networking is not economically feasible due to high usage costs. The success of WLAN lies in the following factors. First, WLAN uses license-free band. 802.11b and 802.11g use Industrial, Scientific, and Medical (ISM) 2.4GHz radio band while 802.11a operates in the 5 GHz National Information Infrastructure (UNII) radio band. Second, WLAN offers reasonably high available data rates. 802.11b can transmit data up to 11 Mbps while 802.11g and 802.11a can provide data rate up to 54Mbps. Finally, there are lots of commercially available WLAN products around the world. Even though WLAN has been designed and used for mostly indoor applications, the possible use of WLAN technologies for high mobility outdoor applications, such as, telemetry, traffic surveillance, rescue operations, and outdoor data networking can provide reasonably high data rates at minimal operational costs. For outdoor applications WLANs provide support for link-layer handoff, which is used to switch a mobile node (MN) from one access point (AP) to another. For WLANs
connected by an IP backbone, Mobile IP[Perk02] is the protocol for location management and network-layer handoff. These attractions led us to investigate the performance of MIP over WLAN in outdoor rapid moving environments.

In this dissertation, Chapter 2 introduces related research in the area of mobile network protocols, wireless LAN standards, layer 3 and layer 2 handoff mechanisms and location tracking technologies. Chapter 3 introduces a protocol evaluation testbed, RAMON. The performance of MIP over wireless LAN and its relationship to speed are shown in Chapter 3 as well. Chapter 4 breaks down the handoff procedure of MIP over wireless LAN and presents a quantitative analysis of the handoff latency. A speed adaptive MIP protocol is proposed in Chapter 5 and the performance for this protocol is evaluated. Chapter 6 summarizes the dissertation and presents future works.
CHAPTER 2
RELATED WORK

Mobile computing and networking try to provide users confident accesses to the Internet anytime, anywhere. One big challenge for mobile computing and networking is how to manage global and seamless roaming among various access technologies. Mobility management contains two components: location management and handoff management [Akyi99]. In wireless network, there are two kinds of roaming, interdomain and intradomain roaming. Interdomain roaming, also called macromobility, refers to roaming among different domain of systems. Intradomain roaming, also called micromobility, refers to roaming among different cells in the same domain or system. In this chapter we will introduce network layer handoff management of macro/micro mobility, wireless LAN protocol standards and technologies to reduce handoff latency for wireless LAN and Mobile IP network. At the end of this chapter, some research works on location tracking will be introduced.

Network Layer Handoff Management

Macro Mobility protocols aim to handle global moving of users. An example is mobile IP[RFC3344]. Micro-mobility protocols are used to handle local moving (e.g., within a domain) of mobile hosts without interaction with the Mobile IP enabled internet.
Hierarchical MIP, Cellular IP, IntraDomain Mobility Protocol (IDMP), HAWAII are examples of micro mobility protocols. Figure 2-1 shows the macro and micro mobility.

**Figure 2-1 Macro and Micro mobility**

**Mobile IP**

IP mobility support for IPv4 is specified in RFC3344. The Mobile IP protocols support transparency above the IP layer, including maintenance of active TCP connections and UDP port bindings. It allows a node to continue using its 'permanent' home address no matter where the node physically attached to. Therefore, ongoing network connections to the node can be maintained even as the mobile host is moving around the internet.

Mobile IP defines three functional entities where its mobility protocols must be implemented: Mobile Node (MN), Home Agent (HA) and Foreign Agent (FA).

MN is a movable device whose software enables network roaming capabilities.

FA is a router that may function as the point of attachment for the MN when it roams to a foreign network, delivering packets from the HA to the MN. Mobile IP works by allowing the MN to be associated with two IP addresses: a home address and a dynamic,
Care-of Address (CoA). Home address is fixed IP address the MN gets from its home network. The CoA is the termination point of the tunnel toward the MN when it is on a foreign network. CoA changes at each new point of attachment to the Internet.

HA is a router on the home network serving as the anchor point for communication with the MN; it tunnels packets from a device on the Internet, called a Correspondent Node (CN), to the roaming MN. (A tunnel is established between the HA and a reachable point for the MN in the foreign network.). The HA maintains an association between the home IP address of the MN and its CoA, which is the current location of the MN on the foreign or visited network. The MN’s movement is invisible to the CN.

Figure 2-2 shows the three functional entities and routing of datagrams transmitted from a MN away from home. When a MN moves, it finds an agent on its local network by the Agent Discovery process. It listens for Agent Advertisement messages sent out by FAs or HAs. If it doesn't hear these messages it can sent Agent Solicitation message to ask for it. From the Agent Advertisement message, the MN determines whether it is on its home network or a foreign one. The MN works like any fixed node when it’s on its home network. When the MN moves away from its home network, it obtains a CoA on the foreign network. The MN registers each new CoA with its HA while away from home. This may be done either directly between the MN and the HA, or indirectly using the FA as a conduit. The packets from CN are tunnelled by HA to FA then to the CoA. The packets from MN to CN are either directly routed to the CN or reverse-tunneled from FA to HA then to the CN.
Figure 2-2 Three functional entities of MIP

MIP has three main processes, Agent discovery, registration and tunneling.

**Agent discovery**

The Mobile IP agent discovery process makes use of ICMP Router Advertisement Protocol (RFC 1256) and add one or more MIP extensions. HAs and FAs periodically broadcast a router advertisement ICMP messages with an advertisement extension. The router advertisement portion of the message includes the IP address of the router. The advertisement extension includes additional information such as lift time, care-of-address, etc. A MN listens for these agent advertisement messages. If a MN needs to get a care-of address and does not want to wait for that long time, the MN can broadcast or multicast an agent solicitation (also an ICMP message) and then listens for the agent advertisement messages. Another important rule of agent discovery process is movement detection. This can be done in two ways. One way is to make use of Lifetime field in the agent advertisement message. When a MN receives an agent advertisement from a FA that it is currently using or that it is now going to register to, it records the lifetime field as a timer. If the timer expires before the agent receives another advertisement from the agent, then
the node assumes that it has lost contact with that agent. In this situation, the MN may choose to wait for another advertisement or to send an agent solicitation. Another way is to use network prefix. The MN checks whether any newly received agent advertisement is on the same network as the current care-of address of the node. If it is not, the MN assumes that it has moved and uses the new advertisement.

The MN can also get a collocated care-of-address acquired from a Dynamic Host Configuration Protocol (DHCP) server. In this case, the MN acts as its own FA.

The agent advertisement extension consists of the following fields:

- **Type**: 16, indicates that this is an agent advertisement.
- **Length**: \((6 + 4N)\), where \(N\) is the number of care-of addresses advertised.
- **Sequence number**: The count of agent advertisement messages sent since the agent was initialized.
- **Lifetime**: The longest lifetime, in seconds, that this agent is willing to accept a registration request from a mobile node.
- **R**: Registration required. Registration with this foreign agent (or another foreign agent on this link) is required even when using a co-located care-of address.
- **B**: Busy. The foreign agent will not accept registrations from additional mobile nodes.
- **H**: This agent offers services as a home agent on this network.
- **F**: This agent offers services as a foreign agent on this network.
- **M**: This agent can receive tunneled IP datagrams that use minimal encapsulation.
- **G**: This agent can receive tunneled IP datagrams that use Generic Routing Encapsulation (GRE).
- **r**: Set as zero; ignored on reception.
- **T**: Foreign agent supports reverse tunneling.
- **Care-of Address(es)**: The care-of address or addresses supported by this agent.
Registration

When a MN realizes that it is on a foreign network and has acquired a care-of-address, it needs to notify the HA by sending a registration request message so that the HA can forward IP packets between MN and CN. There are two kinds of registration messages, registration request and registration reply, both sent to User Datagram Protocol (UDP) port 434. The MN sends the request to the FA, which then relays the request to the home agent. If the MN is using a collocated care-of-address, the MN sends its request directly to the HA, using collocated care-of-address as the source IP address of the request.

The registration request message consists of the following fields:

- **Type**: 1, indicates that this is a registration request.
- **S**: Simultaneous bindings. When set, the mobile node is requesting that the home agent retain its prior mobility bindings. The home agent will forward multiple copies of the IP datagram, one to each care-of address currently registered for this mobile node.
- **B**: Broadcast datagrams. Indicates that the mobile node would like to receive copies of broadcast datagrams that it receives if it were attached to its home network.
- **D**: Decapsulation by mobile node. The mobile node is using a collocated care-of address and will decapsulate its own tunneled IP datagrams.
- **M**: Indicates that the home agent should use minimal encapsulation.
- **G**: Indicates that the home agent should use GRE encapsulation.
• **R**: Sent as zero; ignored on reception.
• **T**: Reverse Tunneling requested.
• **X**: Set as zero; ignored on reception.
• **Lifetime**: The number of seconds before the registration is considered expired. A value of zero is a request for deregistration.
• **Home address**: The home IP address of the mobile node.
• **Home agent**: The IP address of the mobile node home agent.
• **Care-of address**: The IP address for the end of the tunnel. The home agent should forward IP datagrams that it receives with the mobile node home address to this destination address.
• **Identification**: A 64-bit number generated by the mobile node, used for matching registration requests to registration replies and for security purposes.
• **Extensions**: authentication extension must be included, and other optional extensions.

The registration reply message consists of the following fields:

```
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---------+---------+---------+---------+
<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Lifetime</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
+---------+---------+---------+
+---------+---------+---------+---------+
<table>
<thead>
<tr>
<th>Home Address</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
+---------+---------+---------+
+---------+---------+---------+---------+
<table>
<thead>
<tr>
<th>Home Agent</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
+---------+---------+---------+
+---------+---------+---------+
<table>
<thead>
<tr>
<th>Identification</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
+---------+------+
+---------+---------+---------+
<table>
<thead>
<tr>
<th>Extensions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
+---------+------+
```

• **Type**: 3, indicates that this is a registration reply.
• **Code**: Indicates result of the registration request. 0 for registration accepted, 77 for invalid care-of address, etc.
• **Lifetime**: If the code field indicates that the registration was accepted, the number of seconds before the registration is considered expired. A value of zero indicates that the mobile node has been deregistered.
• **Home address**: The home IP address of the mobile node.
• **Home agent**: The IP address of the mobile node home agent.
• **Identification**: A 64-bit number used for matching registration requests to registration replies.
• **Extensions**: authentication extension must be included, and other optional extensions.

The identification field of the registration request and reply messages and the authentication extension are used to protect replay attack. The Identification value enables
the mobile node to match a reply to a request. Two methods are described in RFC 3344: timestamps mandatory) and "nonces" (optional).

An authentication extension consists the following fields:

- **Type**: Used to designate the type of this authentication extension. 32 for MN-HA, 33 for MN-FA, 34 for FA-HA.
- **Length**: 4 plus the number of bytes in the authenticator.
- **Security parameter index (SPI)**: An index that identifies a security context between a pair of nodes. This security context is configured so that the two nodes share a secret key and parameters relevant to this association (for example, authentication algorithm).
- **Authenticator**: The value used to authenticate the message. The default authentication algorithm uses HMAC-MD5[RFC2104] to compute a 128-bit "message digest" of the registration message.

**Tunneling**

After a successful registration, the home agent must be able to intercept datagrams destined to the mobile node and tunnel them to the mobile node’s care-of-address. The tunneling can be done by one of several different encapsulation algorithms, IP in IP encapsulation [RFC2003], Minimal encapsulation [RFC2004] and GRE encapsulation [RFC1701]. By default, home agents and foreign agents must support tunneling datagrams using IP in IP encapsulation. Any mobile node that uses a collocated care-of address must support IP in IP encapsulation. In IP-within-IP encapsulation, the original entire IP datagram becomes the payload in a new IP datagram. The original IP header is unchanged except to reduce Time To Live (TTL) by 1. The outer IP header is a full IP header. Two fields are copied from the inner IP header: The version number, 4, which is the protocol identifier for IPv4, and the type of service field. Figure 2-3 is the IP in IP encapsulation and minimal encapsulation format.
Minimal encapsulation results in less overhead but is little complicated than IP in IP encapsulation. It can only be used if the MN, HA, and FA all agree to use it. With minimal encapsulation, a minimal forwarding IP header is inserted between the original IP header and the original IP payload. The original IP header is modified to form a new outer IP header. The minimal forwarding IP header includes the following fields:

- **Protocol**: Copied from the protocol field in the original IP header. It identifies the protocol type of the original IP payload.
- **S**: If 0, the original source address is not present, and the length of this header is 8 octets. If 1, the original source address is present, and the length of this header is 12 octets.
- **Header checksum**: Computed over all the fields of this header.
- **Original destination address**: Copied from the Destination Address field in the original IP header.
- **Original source address**: Copied from the Source Address field in the original IP header. This field is present only if the S bit is 1. The field is not present if the encapsulator is the source of the datagram.
The new outer IP header is modified from the original IP header. The modified field are as following.

- **Total length**: Incremented by the size of the minimal forwarding header (8 or 12).
- **Protocol**: 55, indicts the following header is minimal IP encapsulation header.
- **Header checksum**: recomputed over all the fields of this header.
- **Source address**: The IP address of the encapsulator, typically the home agent.
- **Destination address**: The IP address of the end of the tunnel, the care-of address.

Mobile IP is a macro mobility management protocol. MIP-based mechanisms use a flat hierarchy, whereby every change in the MN’s point of attachment requires a global binding update. Frequent global binding updates can not only incur high latency, thereby making rapid handoffs impossible, but also significantly increase the overall signaling overhead, especially when the number of MNs increases. Various solutions have been proposed to solve this problem. All these solutions implicitly or explicitly use a concept of micro-mobility regions where registrations with the home agent are not necessary if the MN is moving within these regions. Only if the MN moves between micro-mobility regions, registrations with the HA would be required. Micro-mobility management protocols are designed to reduce the high handoff latency of Mobile IP by handling mobility within micro-mobility regions.

The micro-mobility protocols can be categorized in two types: Hierarchical Tunneling and Mobile-Specific Routing [Camp02]. Hierarchical tunneling schemes rely on a tree-like structure of FAs. In Hierarchical tunneling schemes, HA delivers encapsulated traffic to the root FA. Each FA on the tree decapsulates and then re-encapsulates data packets while they forward the data down the FA tree towards the
MN’s point of attachment. As the MN moves between two FAs, location updates are made at the optimal point in the tree, which is the common root of the two FAs. Hierarchical Mobile IP [Soli02] is an example of Hierarchical tunneling scheme.

Mobile-Specific Routing schemes avoid the overhead introduced by decapsulation and re-encapsulation in hierarchical tunneling schemes. These proposals use mobile specific routes to forward packets toward a MN’s point of attachment. Examples of micro-mobility protocols that use mobile-specific routing include Cellular IP and HAWAII.

**Hierarchical MIP**

The Hierarchical Mobile IP (HMIP) employs a hierarchy of FAs to locally handle Mobile IP registration. In this protocol MNs send mobile IP registration request messages to update their respective location information. The Registration messages establish tunnels between neighboring FAs along the path from the mobile host to a gateway foreign agent (GFA). Packets addressed to mobile hosts travel through these tunnels from the GFA to MN. Figure 3-4 illustrates the operation of Hierarchical Mobile IP. The red dash arrow is a regional registration, which only need to reach a local entity, GFA. The blue real arrow is a normal
registration, which have to traverse the whole network to the HA. For the purposes of managing hierarchical tunneling the location register is maintained in a distributed form by a set of Mobility Agents (MA), i.e. GFAs. Each MA reads the original destination address of the incoming packets and searches its visitor list for a corresponding entry. The entry contains the address of the next MA one level lower in the hierarchy. Such entries are created and maintained by registration messages transmitted by MNs. [Soli02]

**Cellular IP**

The Cellular IP (CIP) protocol[Cam99] from Columbia University and Ericsson supports fast handoff and paging techniques. Cellular IP inherits features found in cellular networks, such as, seamless mobility, passive connectivity and paging, for mobile IP hosts. It uses Mobile IP to provide interconnectivity between a set of Cellular IP access networks, which in turn provide a cellular internetworking environment. The Cellular IP access networks will be connected to the Internet via gateway routers. In that case, host mobility between gateways(i.e., Cellular IP access networks) will be managed by Mobile IP, while mobility within access networks will be handled by Cellular IP. MNs attached to the network use the IP address of the gateway as their Mobile IP care-of address. The data packets from CN to MN will be first routed to MN's HA and then tunneled to the gateway. The gateway "detunnels" packets and forwards them toward base stations. Inside the Cellular IP network, data packets are routed directly to the MN. Data packets from MN to CN are first routed in the cellular IP network to the gateway and from there on to the
HA[Camp00]. The following presents an overview of the Cellular IP routing, handoff and paging algorithms

**Routing**

In Cellular IP, location management and handoff support are integrated with routing. To minimize control messaging, regular data packets transmitted by mobile hosts are used to refresh host location information. Uplink packets are routed from MN to the gateway on a hop-by-hop basis. The path taken by these packets is cached in base stations, which is call route cache. Cellular IP uses mobile originated data packets to maintain reverse path. This path is used to route downlink packets addressed to a mobile host. When the mobile host has no data to transmit then it periodically sends empty IP packets to the gateway to maintain its downlink routing state. The loss of downlink packets when a mobile host moves between access points is reduced by customized handoff procedures. Cellular IP supports two types of handoff scheme, hard handoff and semi-soft handoff.

**Handoff**

The Cellular IP hard handoff algorithm is based on simple approach that trades off some packet loss in exchange for minimizing handoff signaling. Hard handoff causes packet losses proportional to the round-trip time and to the downlink packet rate. Mobile hosts listen to beacons transmitted by base stations and initiate handoff based on signal strength measurements. To perform a handoff, a mobile host tunes its radio to a new base station and sends a route-update packet. The route-update message creates routing cache
mappings on route to the gateway hence configures the downlink route to the new base station.

Cellular IP semi-soft handoff exploits the notion that some mobile hosts can simultaneously receive packets from the new and old base stations during handoff. During semi-soft handoff a mobile host may be in contact with either the old and new Base Stations and receives packets from them. Packets intended to the mobile host are sent to both Base Stations, so when the mobile host eventually moves to the new location it can continue to receive packets without interruption. To initiate semi-soft handoff, the moving mobile host transmits a route-update packet to the new Base Station and continues to listen to the old one. The S flag is set in this route-update packet to indicate semi-soft handoff. Semi-soft route-update packets create new mappings in the Route and Paging Cache similarly to regular route-update packets. When the semi-soft route-update packet reaches the crossover node where the old and new path meet, the new mapping is added to the cache instead of replacing the old one. Packets sent to the mobile host are transmitted to both Downlink neighbors. When the mobile host eventually makes the move then the packets will already be underway to the new Base Station and the handoff can be performed with minimal packet loss. After migration the mobile host sends a route-update packet to the new Base Station with the S bit cleared. This route-update packet will remove all mappings in the Route Cache except for the ones pointing to the new Base Station. The semi-soft handoff is then complete. If the path to the new Base Station is longer than that to the old Base Station or if it takes non-negligible time to switch to the new Base Station,
then some packets may not reach the mobile host. To overcome the problem, packets sent
to the new Base Station can be delayed during semi-soft handoff. This way a few packets
may be delivered twice to the mobile host, but in many cases this results in better
performance than a few packets lost. Introduction of packet delay can be best performed in
the Cellular IP node that has multiple mappings for the mobile host as a result of a
semi-soft route-update packet. Packets that belong to flows that require low delay, but can
tolerate occasional losses, should not be delayed.

Semi-soft handoff minimizes packet loss providing improved TCP and UDP
performances over hard handoff. Distinguishing idle and active mobile hosts reduces
power consumption at the terminal side. The location of idle hosts is tracked only
approximately by Cellular IP. Therefore, mobile hosts do not have to update their location
after each handoff. This extends battery life and reduces air interface traffic. When packets
need to be sent to an idle mobile host, the host is paged using a limited scope broadcast. A
mobile host becomes active upon reception of a paging packet and starts updating its
location until it moves to an idle state again.

Paging

If a mobile host has not received data packets for a system specific time
active-state-timeout, it becomes idle. The idle mobile hosts allow their soft-state routing
cache mappings to be time out. Idle hosts transmit empty IP packets(paging-update
packets) at regular intervals(paging-update-time) to the gateway. Paging-update packets
are sent to the base station that offers the best signal quality. Paging-update packets are
also routed on a hop-by-hop basis to the gateway. Base stations may optionally maintain paging cache. A paging cache has the same format and operation as a routing cache except for two differences. First, paging cache mappings have a longer timeout period called paging-timeout. Second, paging cache mappings are updated by any packet sent by mobile hosts including route-update packets and paging-update packets. This results in idle mobile hosts having mappings in paging caches but not in routing caches. If the base station has no paging cache, it will forward the packet to all its interfaces except for the one the packet came through. Paging cache is used to avoid broadcast search procedures found in cellular systems. Base stations that have paging cache will only forward the paging packet if the destination has a valid paging cache mapping and only to the mapped interface(s)[Camp00].

HAWAII

The Handoff Aware Wireless Internet Infrastructure (HAWAII) protocol [Ramj99][Ramj02] proposes a separate routing protocol to handle intra-domain mobility. All issues related to mobility management within one domain are handled by a gateway called a domain root router. A MN entering a new foreign agent domain is assigned a collocated care-of address. The MN retains its care-off address unchanged while moving within the foreign domain, thus the HAs does not need to be involved unless the MN moves to a new domain. In this case, packets for the MN are intercepted by its HA first. The HA tunnels the packets to the domain root router serving the MN. The domain root router routes the packets to the MN using the host-based routing entries. When the
MN moves between different subnets of the same domain, only the route from the domain root router to the BS serving the MN is modified, and the remaining path remains the same. Thus, during an intra-domain handoff, the global signaling message load and handoff latency are reduced.

**HAWAII path setup messages**

There are three types of HAWAII path setup messages: powerup, path refresh, and path update. On power up a mobile host sends a Mobile IP registration request message to the corresponding base station. The base station then sends a HAWAII path setup power-up message to the domain root router which is processed in a hop-by-hop manner. This has the effect of establishing host specific routes for that mobile host in the domain root router and any intermediate routers on the path towards the mobile host. The domain root router finally acknowledges this path setup power-up message to the base station which finally notifies the mobile host with a Mobile IP registration reply.

If a router knows multiple paths to the domain root router, it can use any of them but it always has to use the same route for a specific host. The routing entries in the routers are soft-state, i.e. they have to be refreshed periodically by path setup refresh messages, which are sent independently by each network node and which can be aggregated. This increases the robustness of the protocol to router and link failures. The mobile host infrequently sends periodic path refresh messages to its base station to maintain the host based entries. The base station and the intermediate routers, in turn, send periodic aggregate hop-by-hop refresh messages towards the domain root router. Path setup messages are sent to only
selected routers in the domain, resulting in very little overhead associated with maintaining soft-state.

While the mobile host moves within a domain, maintaining end-to-end connectivity to the mobile host requires special techniques for managing user mobility. HAWAII uses path setup update messages to establish and update host-based routing entries for the mobile hosts in selective routers in the domain so that packets arriving at the domain root router can reach the mobile host with limited disruption. The choice of when, how, and which routers are updated constitutes a particular path setup scheme.

**HAWAII Path Setup Schemes**

The HAWAII handoff procedures are only activated when the mobile host’s next hop IP node is changed during the handoff. [Ramj02] assumes base stations have IP routing functionality and uses a tree-based topology for clarity, but the schemes will also provide for non-tree-based topologies.

[Ramj99] defines two schemes for implementing Handoff procedures within the domain, the forwarding and the non-forwarding scheme. The cross-over router is defined as the router closest to the mobile host that is at the intersection of two paths, one between the domain root router and the old base station, and the second between the old base station and the new base station.

**Non-forwarding path update scheme**

The non-forwarding path set-up is a two way Update handshake process. It is initiated by the Mobile Station sending an Update to the new Base Station. The path setup
Update message consists of the Mobile Station IP address, the old and new Base Station address, and some other informations. The following is the algorithm:

Step 1: Update the Cache with the combination of the IP address of the Mobile Station and the port on which the Update was received. This builds an element in a “reverse” chain in the direction from the current node towards the new Base Station. If the current node is the old Base Station, it sends an acknowledgement to the Mobile Station directly via the air interface. This completes the procedure and the old Base station will not receive further datagrams for the Mobile Station. The path from the gateway to the new Base Station will be refreshed, the rest (from the Crossover router to the old Base Station) will not and times out shortly.

Step 2: (recipient is not the old Base Station) The node extracts the forwarding port for the old Base Station from the routing table, and forwards the Update. Step 1 is then revisited.

**Forwarding path update scheme**

The forwarding path set-up is initiated by the Mobile Station. Its also a two way Update handshake process. The Mobile Station sends the old Base Station an Update message, which consists of the Mobile Station IP address, the old and new Base Station addresses, and some other informations. The following is the algorithm:

Step 1: If the node receiving the Update is the new Base Station, it sends an acknowledgement to the Mobile Station directly via the air interface, and updates the Cache with the IP address and port number for the Mobile Station. This completes the
procedure, leaving two Soft-state paths. One leading from the old to the new Base Station, and the other from the gateway, via the Crossover router, to the new Base Station. The second path will be refreshed while the first will time out shortly.

Step 2: (recipient is not the new Base Station) The node extracts the forwarding port for the new Base Station from the routing table, and updates the Cache with the IP address of the Mobile Station and this port number. Step 1 is then revisited. The Forwarding path set-up scheme packets in transit towards the old router are then forwarded from the Old Base Station to the new Base Station until the flow is diverted at the Crossover router.

The non-forwarding scheme is optimized for networks where the Mobile Station can listen/transmit to multiple base stations simultaneously, as in the case of Code Division Multiple Access (CDMA) networks. The forwarding scheme is optimized for networks where the Mobile Station can listen/transmit to only one base station, as in the case of a Time Division Multiple Access (TDMA) network. Both schemes ensure no BSS internal loss of in transit datagrams during handoff.

**Wireless LAN**

**Technology Overview**

Over recent years, the market for wireless communications has experienced incredible growth. Wireless technologies have quickly found a significant place and popularity in business and the computer industry. Their major motivation and benefit is increased flexibility and mobility. Unlike a wired LAN, which requires a wire to access the network, a Wireless LAN connects computers and other components to the network via an
Access Point (AP). Wireless LANs offer several fundamental benefits including user mobility, rapid installation, flexibility and scalability. However, there are some primary limitations [Gast02].

- The speed of wireless networks is constrained by the available bandwidth.
- Radio waves can suffer from a number of propagation problems that may interrupt the radio link, such as multi-path interference and shadows.
- On Wire LAN, sniffing is much easier because the radio transmissions are designed to be processed by any receiver within range. Security is still a prime concern.

The IEEE 802.11 Working Group was formed in September of 1990. Their goal was to create a wireless LAN specification that will operate in one of the Industrial, Scientific, and Medical frequency (ISM) ranges. The first 802.11 standard was released in 1997. The latest version is the 1999 edition. The official name of 802.11 is IEEE Standards for Information Technology -- Telecommunications and Information Exchange between Systems -- Local and Metropolitan Area Network -- Specific Requirements -- Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.

The 802.11 protocols address the Medium Access Control (MAC) and Physical (PHY) layers independently. The MAC layer handles moving data between the link layer and the physical medium. Figure 2-5 shows how the OSI model matches up to the 802.11 standards.

**The 802.11 Established Standards**

The 802.11 suite has the four established standards: 802.11, 802.11b, 802.11a and 802.11g. The IEEE is continuing to work on new standards that will extend the physical layer options, improve security, and add quality of service (QoS) features. In the following several sections, we will brief introduce these four standards [80211].
**Standard 802.11**

802.11 was the first IEEE standard used for wireless data networking applications with maximum data transfer rates at 2 Mbps in the 2.4 GHz radio band. Within 802.11, two different modulation schemes are supported that can be used to transmit data signals.

The first modulation scheme is frequency-hopping spread spectrum (FHSS). This transmission technique is used in WLAN transmissions where the data signal is modulated with a narrowband carrier signal that “hops” in a random sequence from frequency to frequency as a function of time over a wide band of frequencies. This technique reduces the chances of interference.

The other modulation scheme is direct-sequence spread spectrum (DSSS). In this method of transmission, the signal does not hop from one frequency to another but is passed through a spreading function and distributed over the entire band at once. DSSS
usually provides slightly higher data rates and shorter delays than FHSS, because the
transmitter and receiver don't have to spend time retuning. DSSS avoids interference by
configuring the spreading function in the receiver to concentrate the desired signal but
spread out and dilutes any interfering signal. A data signal at the sending station is
combined with a higher data rate bit sequence, or chipping code, that divides the user data
according to a spreading ratio. The chipping code, a redundant bit pattern for each bit that
is transmitted, increases the signal’s resistance to interference. If one or more bits in the
pattern are damaged during transmission, the original data can be recovered due to the
redundancy of the transmission.

Although the 802.11 standard supports both modulation schemes, the two types of
spread spectrum technologies are not compatible. The number of channels used by 802.11
compliant products depends on the modulation scheme used. More specifically,
FHSS-based products use 79 channels of the Unlicensed National Information
Infrastructure (UNII) band, whereas DSSS-based products use either 3 non-overlapping
channels or 6 overlapping channels of the Industrial, Scientific, and Medical (ISM) radio
band. Some of the common characteristics specified by the 802.11 standard are listed in
Table 2-1.
Table 2-1 Characteristics specified by the 802.11 standard

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>802.11 Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>Wireless data networking</td>
</tr>
<tr>
<td>Data Rate</td>
<td>1–2 Mbps</td>
</tr>
<tr>
<td>Typical Operating Frequency Band</td>
<td>ISM band: 2.4 to 2.4835 GHz.</td>
</tr>
<tr>
<td>Modulation Mechanism</td>
<td>FHSS or DSSS, CRC-16 in header</td>
</tr>
<tr>
<td>Channels available</td>
<td>79 channels with FHSS; 3 or 6 channels with DSSS</td>
</tr>
<tr>
<td>Coverage</td>
<td>40m to 400m</td>
</tr>
<tr>
<td>Mobility</td>
<td>Roaming between APs by mobile IP</td>
</tr>
<tr>
<td>Security</td>
<td>128-bit WEP</td>
</tr>
<tr>
<td>Link Layer</td>
<td>Carrier Sense Multiple Access With Collision Avoidance (CSMA/CA) with request to send (RTS)/clear to send (CTS)</td>
</tr>
</tbody>
</table>

**Standard 802.11b**

IEEE 802.11b[80211b] is the first enhancement 802.11 standard to be ratified in 1999. 802.11b uses the same radio signaling frequency(2.4GHz) as the original 802.11 standard. The 802.11b standard specifies operation on three channels in the 2.4–2.4835 GHz spectrum. 802.11b can transmit data up to 11 Mbps but will scale down to 1 Mbps based on conditions.

802.11b uses DSSS modulation scheme to transmit data signals through the 11 available channels(3 non-overlapping). This unlicensed portion of the radio band shares space with many low-power signals from home electronics, including microwave ovens, cordless telephones, Bluetooth-enabled devices, and garage-door openers. 802.11b compliant products have a range of up to 400 meters in ideal conditions and will be compatible with the products that meet the new 802.11g standard when it is finalized.

Some of the key characteristics specified by the 802.11b standard are shown in Table 2-2.
Table 2-2 Characteristics specified by the 802.11b standard

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>802.11b Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>Wireless data networking</td>
</tr>
<tr>
<td>Data Rate (Mbps)</td>
<td>1, 2, 5.5, 11</td>
</tr>
<tr>
<td>Typical Operating Frequency Band</td>
<td>ISM band: 2.4 to 2.4835 GHz</td>
</tr>
<tr>
<td>Channels available</td>
<td>11 (3 non-overlapping)</td>
</tr>
<tr>
<td>Modulation Mechanism</td>
<td>DSSS</td>
</tr>
<tr>
<td>Coverage (m)</td>
<td>40 to 400</td>
</tr>
<tr>
<td>Mobility</td>
<td>Roaming between APs by mobile IP devices</td>
</tr>
<tr>
<td>Security</td>
<td>128 bit WEP</td>
</tr>
<tr>
<td>Link Layer</td>
<td>CSMA/CA with RTS/CTS</td>
</tr>
</tbody>
</table>

Pros of 802.11b - lowest cost; signal range is best and is not easily obstructed.

Cons of 802.11b - Speed and channel restriction are significant limitations of 802.11b compliant networks. Interference within one’s own 802.11b network becomes more likely as the number of users and APs increase. Similarly, interference is more likely as 802.11b compliant networks are deployed near each other. 802.11b products share the bandwidth with other low-power signals, and thus, problems may arise when the technology is used near some electronic devices such as microwave ovens, Bluetooth-enabled devices, and cordless telephones.

**Standard 802.11a**

802.11a[80211a], a High-speed Physical Layer in the 5 GHz band standard for WLANs, was completed in September 1999. It is offered in the 5 GHz radio (UNII) band, and operates on 8 channels; however, the available radio spectrum in some countries permits the use of 12 channels. The additional number of channels used in the higher spectrum yields less interference from neighboring APs. The Federal Communications Commission (FCC) has divided the total of 300 megahertz (MHz) frequencies used by
802.11a WLANs into 3 distinct 100 MHz domains, each with a different legal maximum
power output. The “low” band operates in the 5.15–5.25 GHz range and has a maximum
output power of 50 milliwatts (mW). The “middle” band is located in the 5.25–5.35 GHz
range, with a maximum of 250 mW. The “high” band uses the 5.725–5.825 GHz range,
with a maximum of 1 Watt. Because of the high power output, most devices transmitting
in the high band are building-to-building bridge products. The low and medium bands are
more suited to in-building wireless products.

802.11a transfers data at rates of up to 54 Mbps in the available radio spectrum,
which is up to five times faster than 802.11b compliant networks. More commonly,
however, 802.11a compliant networks communications are at the 6 Mbps, 12 Mbps, or
24 Mbps data rates. As the distance between the user and the AP increases, the data rate
decreases.

802.11a compliant networks use Orthogonal Frequency Division Multiplexing
(OFDM) modulation to provide these data rates. OFDM is a type of digital modulation in
which a signal is divided into separate channels at different frequencies. Table 2-3 show
the major characteristics of 802.11a standard.

**Pros of 802.11a** - speed as 5 times as 802.11b; supports more simultaneous users;
regulated frequencies prevent signal interference from other devices

**Cons of 802.11a** - shorter range signal that is more easily obstructed; shorter range
costs more APs to cover the same area as an 802.11b network; consume more power than
802.11b products.
Table 2-3 Characteristics specified by the 802.11a standard

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>802.11a Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>Wireless data Networking</td>
</tr>
<tr>
<td>Data Rate (Mbps)</td>
<td>6, 9, 12, 18, 24, 36, 48, and 54 Mbps. Rates of 6, 12, and</td>
</tr>
<tr>
<td></td>
<td>24 Mbps are mandatory for all products.</td>
</tr>
<tr>
<td>Typical Operating Frequency Band</td>
<td>UNII band: 5.15-5.25 GHz, 5.25-5.35 GHz, and 5.725-5.825 GHz</td>
</tr>
<tr>
<td>Channels Available</td>
<td>12 non-overlapping</td>
</tr>
<tr>
<td>Modulation Mechanism</td>
<td>OFDM- Orthogonal Frequency Division Multiplexing</td>
</tr>
<tr>
<td>Coverage (m)</td>
<td>≤ 100</td>
</tr>
<tr>
<td>Mobility</td>
<td>Roaming between APs by mobile IP devices</td>
</tr>
<tr>
<td>Security</td>
<td>128-bit WEP, 64-bit WEP, 152-bit WEP</td>
</tr>
<tr>
<td>Link Layer</td>
<td>CSMA/CA with RTS/CTS</td>
</tr>
</tbody>
</table>

802.11a was ratified after 802.11b was already penetrating the market, so even though it offers higher speed and frequency, it may not be worth the switch for users who have already invested in 802.11b technology. Because 802.11a and 802.11b utilize different frequencies, the two technologies are incompatible with each other. Some vendors offer hybrid 802.11a/b network gear, but these products simply implement the two standards side by side.

**802.11g**

IEEE 802.11g was ratified as a standard in Jun. 2003. It operates in the same 2.4 GHz range as 802.11b but offers the same speed up to 54 Mbps as 802.11a does. This standard features increased data transmission rates while maintaining interoperability with 802.11b compliant products. The standard uses the same modulation scheme OFDM as 802.11a to achieve data rates from 22 Mbps to up to 54 Mbps; however, 802.11g products will be backward compatible with 802.11b products that use the modulation scheme DSSS. The backward compatibility feature allows an 802.11b compliant client adapter
card to interact directly with an 802.11g compliant AP. Communications between 802.11g and 802.11b devices are limited to data rates up to 11 Mbps. The common characteristics specified by the 802.11g standard are shown in Table 2-4.

Table 2-4 Characteristics specified by the 802.11g standard

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>802.11g Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>Broadband Wireless LAN Access</td>
</tr>
<tr>
<td>Data Rate (Mbps)</td>
<td>6, 9, 12, 18, 24, 36, 48, 54</td>
</tr>
<tr>
<td>Typical Operating Frequency Band</td>
<td>ISM band: 2.4 to 2.4835 GHz</td>
</tr>
<tr>
<td>Channels available</td>
<td>3 non-overlapping</td>
</tr>
<tr>
<td>Modulation Mechanism</td>
<td>OFDM/DSSS</td>
</tr>
<tr>
<td>Coverage (m)</td>
<td>20 to 400</td>
</tr>
<tr>
<td>Mobility</td>
<td>Roaming between APs by mobile IP devices</td>
</tr>
<tr>
<td>Security</td>
<td>128 bit WEP</td>
</tr>
<tr>
<td>Link Layer</td>
<td>CSMA/CA with RTS/CTS</td>
</tr>
</tbody>
</table>

**Pros of 802.11g** - fast speed as up to 54mbps; supports more simultaneous users; signal range is better than 802.11a and is not easily obstructed

**Cons of 802.11g** - costs more than 802.11b; just like 802.11a, appliances may interfere on the unregulated signal frequency when the technology is used near some electronic devices such as microwave ovens, Bluetooth-enabled devices, and cordless telephones.

Table 2-5 provides a comparison of the primary 802.11 standards.
Table 2-5 Comparison of characteristics specified within the IEEE 802.11 suite

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>802.11</th>
<th>802.11a</th>
<th>802.11b</th>
<th>802.11g</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectrum Band</td>
<td>ISM: 2.4 to 2.4835 GHz</td>
<td>UNII: 5.15-5.25 GHz, 5.25-5.35 GHz, and 5.725-5.825 GHz</td>
<td>ISM: 2.4 to 2.4835 GHz</td>
<td>ISM: 2.4 to 2.4835 GHz</td>
</tr>
<tr>
<td>Modulation Scheme</td>
<td>FHSS or DSSS</td>
<td>OFDM</td>
<td>DSSS</td>
<td>OFDM or DSSS</td>
</tr>
<tr>
<td>Number of Channels</td>
<td>79 channels with FHSS; 3 or 6 channels with DSSS</td>
<td>12</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Optimum Data Rates (Mbps)</td>
<td>2</td>
<td>54</td>
<td>11</td>
<td>54</td>
</tr>
<tr>
<td>Range (meters)</td>
<td>400</td>
<td>100</td>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td>Date established</td>
<td>July 1997</td>
<td>September 1999</td>
<td>July 1999</td>
<td>June 2003</td>
</tr>
<tr>
<td>Compatibility</td>
<td>802.11 only</td>
<td>802.11a</td>
<td>802.11b</td>
<td>802.11b/g</td>
</tr>
<tr>
<td>Operability</td>
<td>North America, Europe, Asia</td>
<td>North America, Europe, Asia</td>
<td>North America, Europe, Asia</td>
<td>North America, Europe, Asia</td>
</tr>
</tbody>
</table>

Pending Specifications Within the 802.11 Suite

IEEE 802.11a, 11b, 11g are major standard of wireless networking. There are various other standards which were developed to improve the transmission of data and promote the effective communication. The following are current standards which enhance and expand the functionality of the overall 802.11 protocol.[STD802]

- **IEEE 802.11c**: Defines wireless bridge operations
- **IEEE 802.11d**: Defines standards for companies developing wireless products in different countries.
- **IEEE 802.11e**: Defines enhancements to the 802.11MAC for QoS.
- **IEEE 802.11f**: Defines Inter Access Point Protocol (IAPP)
- **IEEE 802.11i**: Improved encryption
- **IEEE 802.11j**: 802.11 extension used in Japan.
- **IEEE 802.11n**: New standard expected to be completed in 2005 that is expected to support up to 100Mbps.
The IEEE 802.11 Wireless LAN Architecture

The 802.11 architecture is comprised of several components and services that interact to provide station mobility transparent to the higher layers of the network stack. The major components and services in Wireless LAN are as followings [Jain03].

Wireless LAN Station

The wireless LAN station (STA) is the most basic component of the wireless network. A station is any device that implements the MAC and PHY functionality of the 802.11 protocol. Typically the 802.11 functions are implemented in the hardware and software of a network interface card (NIC). A station could be a laptop PC, PDA, or an Access Point. Stations may be mobile, portable, or stationary and all stations support the 802.11 station services of authentication, de-authentication, privacy, and data delivery.

Basic Service Set (BSS)

802.11 defines the Basic Service Set (BSS) as the basic building block of an 802.11 wireless LAN. The BSS consists of a group of stations.

The Topologies could be Independent Basic Service Set (IBSS), Infrastructure Basic Service Set(BSS) or Extended Service Set (ESS)

Independent Basic Service Set (IBSS)

The most basic wireless LAN topology is a set of stations, which have recognized each other and are connected via the wireless media in a peer-to-peer fashion. This form of network topology is referred to as an Independent Basic Service Set (IBSS) or an Ad-hoc network. In an IBSS, the mobile stations communicate directly with each other. Every
mobile station may not be able to communicate with every other station due to the range limitations. There are no relay functions in an IBSS therefore all stations need to be within range of each other and communicate directly.

**Infrastructure Basic Service Set (BSS)**

An Infrastructure Basic Service Set is a BSS with a component called an Access Point (AP). The access point provides a local relay function for the BSS. All stations in the BSS communicate with the access point and no longer communicate directly. All frames are relayed between stations by the access point. This local relay function effectively doubles the range of the IBSS.

**Extended Service Set (ESS)**

An extended service set is a set of infrastructure BSS’s, where the access points communicate among themselves to forward traffic from one BSS to another to facilitate movement of stations between BSSs.

**Wireless LAN Handoff Management**

**Wireless LAN Handoff Management Frames**

The 802.11 standard defines various frame types that stations (NICs and APs) use for communications, as well as managing and controlling the wireless link. Every frame has a control field that depicts the 802.11 protocol version, frame type, and various indicators for WEP is on/off, power management is on/off, etc. In addition all frames contain MAC addresses of the source and destination station, a frame sequence number, frame body and frame check sequence (for error detection). 802.11 control frames assist in
the delivery of data frames between stations. Data frames carry protocols and data from higher layers within the frame body such as RTS, CTS, ACK. Management frames enable stations to establish and maintain communications. Here we only introduce the management frames which relative directly to handoff management. [Jim01]

- **Authentication frame:** 802.11 authentication is a process whereby the access point either accepts or rejects the identity of a radio NIC. The NIC begins the process by sending an authentication frame containing its identity to the access point. With open system authentication (the default), the radio NIC sends only one authentication frame, and the access point responds with an authentication frame as a response indicating acceptance (or rejection). With the optional shared key authentication, the radio NIC sends an initial authentication frame, and the access point responds with an authentication frame containing challenge text. The radio NIC must send an encrypted version of the challenge text, using its wired equivalent privacy (WEP) key, in an authentication frame back to the access point. The access point ensures that the radio NIC has the correct WEP key (which is the basis for authentication) by seeing whether the challenge text recovered after decryption is the same that was sent previously. Based on the results of this comparison, the access point replies to the radio NIC with an authentication frame signifying the result of authentication.

- **Deauthentication frame:** A station sends a deauthentication frame to another station if it wishes to terminate secure communications.

- **Association request frame:** 802.11 association enables the access point to allocate resources for and synchronize with a radio NIC. A NIC begins the association process by sending an association request to an access point. This frame carries information about the NIC (e.g., supported data rates) and the SSID of the network it wishes to associate with. After receiving the association request, the access point considers associating with the NIC, and (if accepted) reserves memory space and establishes an association ID for the NIC.

- **Association response frame:** An access point sends an association response frame containing an acceptance or rejection notice to the radio NIC requesting association. If the access point accepts the radio NIC, the frame includes information regarding the association, such as association ID and supported data rates. If the outcome of the association is positive, the radio NIC can utilize the access point to communicate with other NICs on the network and systems on the distribution (i.e., Ethernet) side of the access point.

- **Reassociation request frame:** If a radio NIC roams away from the currently associated access point and finds another access point having a stronger beacon
signal, the radio NIC will send a reassociation frame to the new access point. The new access point then coordinates the forwarding of data frames that may still be in the buffer of the previous access point waiting for transmission to the radio NIC.

- **Reassociation response frame**: An access point sends a reassociation response frame containing an acceptance or rejection notice to the radio NIC requesting reassociation. Similar to the association process, the frame includes information regarding the association, such as association ID and supported data rates.

- **Disassociation frame**: A station sends a disassociation frame to another station if it wishes to terminate the association. For example, a radio NIC that is shut down gracefully can send a disassociation frame to alert the access point that the NIC is powering off. The access point can then relinquish memory allocations and remove the radio NIC from the association table.

- **Beacon frame**: The access point periodically sends a beacon frame to announce its presence and relay information, such as timestamp, SSID, and other parameters regarding the access point to radio NICs that are within range. Radio NICs continually scan all 802.11 radio channels and listen to beacons as the basis for choosing which access point is best to associate with.

- **Probe request frame**: A station sends a probe request frame when it needs to obtain information from another station. For example, a radio NIC would send a probe request to determine which access points are within range.

- **Probe response frame**: A station will respond with a probe response frame, containing capability information, supported data rates, etc., when after it receives a probe request frame.

**IEEE 802.11 Handoff Procedure**

An IEEE 802.11 Handoff occurs when a STA moves out of the range of one AP, and enters another BSS. During the handoff, management frames are exchanged between the station (STA) and the AP. Also the APs involved may exchange certain context information (credentials) related to that STA via Inter Access Point Protocol (IAPP).

The handoff procedure can be divided into two steps [Mish03] [Shin04], discovery and reauthentication.

Discovery: this step involves the handoff initiation phase and the scanning phase.

When the STA is moving away from the current AP, the signal strength and the
signal-to-noise ratio of the signal may degrade and initiate the scanning phase. Scanning is to try to find a new available AP to associate with. There are two can of scanning mode: passive or active. In passive scanning mode, the STA listens to each channel of the wireless medium for beacon frames broadcasted by AP. Using the information obtained from beacon frames the STA can elect to join an AP. In active scanning, apart from listening to the beacon frames, the STA send probe request frames on each channel and listens to probe responses from the APs. The basic procedure of the active scanning includes the following steps [80211], as summarize by[Shin04] :

- Using the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) channel access mechanism gain control of wireless medium.
- broadcast a probe request frame.
- Start a probe timer.
- Listen to the channel for probe responses.
- If no response has been received by minChannelTime, scan next channel.
- If one or more responses are received by minChannelTime, stop accepting probe responses at maxChannelTime and process all received responses.
- Repeat the above steps to scan next channel. After all channels have been scanned, all information received from probe responses are processed so that the STA can select one AP to associate.

Reauthentication: The reauthentication process involves authentication and reassociation to the new AP. The STA sends a authentication request to the new AP, informing the AP of its identity. The new AP sends back an authentication response, indicating acceptance or rejection. After successful authentication, the STA sends a reassociation request to the new AP and waits for a reassociation response containing an acceptance or rejection notice.

Figure 2-6, taken from [Shin04], shows the IEEE 802.11 handoff process.
Techniques to Reduce IEEE 802.11 Handoff Time

A lot of researches have been done to analyze and reduce the handoff latency of wireless LAN. [Mish03] conducts experiments to accurately measure the handoff latency in an in-building wireless network. The measurements are done on two co-existing wireless networks, and using three wireless NICs from different vendors. It analyzes the handoff latencies by breaking down the whole process into discovery and reauthentication phases to assess the contribution of each phase to the handoff latency. The experiment results show that the discovery phase (scanning time) is the most time consuming part of the handoff process, taking over 90% of the total handoff latency. The variations in the
probe-wait time account for the large variations in the overall handoff latency. The reauthentication phase contributes only a few milliseconds.

[Mish04] use of an efficient data structure, neighbor graphs, which dynamically captures the mobility topology of a wireless network as a means for pre-positioning the station’s context ensuring that the station’s context always remains one hop ahead. This caching mechanism is based on the IAPP protocol in order to exchange the client context information between neighboring APs. The cache in the AP is built using the information contained in an IAPP Move-Notify message or in the reassociation request sent to the AP by the client. By exchanging the client context information with the old AP, the new AP does not require the client to send its context information in order to reassociate, hence reducing the reassociation delay. Its experimental and simulation results show that the use of neighbor graphs cache reduces the layer 2 handoff latency due to reassociation by an order of magnitude from 15.37ms to 1.69ms.

[Kim04] propose a selective scanning algorithm which depends on the use of neighbor graphs. This approach requires changes in the network infrastructure and use of IAPP. The scanning delay is defined as the duration taken from the first Probe Request message to the last Probe Response message. This definition does not take into consideration the time needed by the client to process the received probe responses. [Shin04] also propose a selective scanning algorithm and a caching mechanism. This caching data structure is maintained at the client side and no changes are required in the existing network infrastructure or the IEEE 802.11 standard. All the required changes are
done on the client side wireless card driver. And [Shin04] considers the time required for processing the probe responses received by the client. This processing time represents a significant part of the scanning delay especially when the number of probe responses received increased significantly.

Sangheon Pack and Yanghee Choi in [Par02] and [Park02] proposed a fast handoff scheme using the pre-authentication method based on IEEE 802.1x model. In their proposal, when a mobile host handoff, it performs authentication procedures not only for the current AP but for a set of multiple APs. Multiple APs are selected using a Frequent Handoff Region (FHR) selection algorithm considering users' mobility patterns and their service classes. The FHR is a set of adjacent APs. It is determined by the APs' locations and users' movement patterns. Namely, the FHR consists of APs with which mobile hosts are likely to communicate to in the near future. Since a mobile host is authenticated for FHR in advance, the handoff latency due to the reauthentication can be minimized.

**Low Latency Handoff Mechanisms for MIP over 802.11 Network.**

The HMIP, Cellular IP (CIP)[Cam99] and (HAWAII) [Ramj02] protocol we talked in Section 2 of this chapter are handoff management protocols without considering underlying layers. This clean separation between Layer 2 and Layer 3 protocol stack allows those protocols to run on most layer 2 technologies. The disadvantage of this clean separation is lower performance. In MIP over wireless LAN network, the MN may only keep connectivity with one AP, hereby one FA. So the MN can only start the registration process after completion of the L2 handoff. [Malk02] proposed two mobility protocols,
pre-registration and post-registration, that aim at low latency Layer 3 handoff based on Layer 2 information or called Layer2 trigger. In pre-registration, MN may communicate with the new FA while still being connected with the old FA. In post-registration, the data can be delivered to the MN at the new FA before the registration process has completed.

Here we briefly depict these two method summarize by [Blon04].

L2 Triggers

A L2 trigger is a signal related to the L2 handoff process. There are there kind of L2 triggers mentioned in [Malk02]: anticipation trigger- an early notice of an upcoming change in the L2 point of attachment of the MN. Line Down trigger (L2-LD)- indicates that the L2 link between the MN and the old AP is lost. Line Up trigger (L2-LU)- indicates that the L2 link between the MN and the old AP is established. A trigger initiated at the old FA is referred as a source trigger and a trigger initiated at the new FA is referred as a target trigger.

Pre-Registration

Pre-Registration allows the old FA and new FA to utilize information from layer 2 (the L2 "trigger") to set up a kind of "pre-registration" prior to receiving a formal Registration Request from the Mobile Node. The network assists the MN in performing an L3 handoff before the L2 handoff is completed. Both the MN (mobile-initiated) and the FAs (network-initiated) can initiate a handoff.
A mobile-initiated handoff occurs when the L2 anticipation trigger is received at the MN informing it that it will shortly move to the nFA. The L2 trigger contains information such as the nFA’s IP address identifier.

A network-initiated handoff can be initiated by a source trigger at the oFA (source-initiated handoff) or by a target trigger at the nFA (target-initiated handoff). A source-initiated handoff is initiated at the oFA by a received L2 trigger that informs the oFA of a MN’s upcoming movement from oFA to nFA. A target-initiated handoff is initiated at the nFA by a received L2 trigger that informs the nFA of a MN’s upcoming movement from oFA to nFA.

**Post-Registration**

The Post-Registration handoff method is based on a network-initiated model of a handoff. The Post-Registration occurs after the L2 handoff has been completed. This approach uses a bi-directional edge tunnel (BET) to perform a low latency change in the L2 point of attachment of the MN without requiring any involvement of it.

A handoff occurs when the MN moves from the oFA, where the MN performed a Mobile IP registration, to the nFA. The MN delays its registration with the nFA, while maintaining connectivity using the BET between the oFA and nFA. There are two different Post Registration handoff schemes, Source and Target Trigger Post Registration, depends on what kind of L2 is using. An FA becomes aware that a handoff is about to occur at L2 through the use of an L2 trigger. Two types of triggers can be received: a source trigger at the oFA and a target trigger at the nFA.
The FA receiving the trigger sends a Handoff Request (HRqst) to the other FA. The FA receiving the HRqst sends a Handoff Reply (HRply) to the first FA. This establishes a BET. The L2-LD (Link Down) trigger at the oFA and at the MN signals that the MN is not connected anymore with the oFA. When the oFA receives the L2-LD trigger, it begins forwarding the MN packets through the forwarding tunnel to the nFA. When the nFA receives the L2-LU (Link Up) trigger, it begins delivering packets tunneled from the oFA to the MN and forwards packets from the MN. When the MN receives the L2-LU, it decides to initiate the Mobile IP Registration process with the nFA by soliciting an Agent Advertisement or continues using the BET. Once the Registration process is complete (through the exchange of a Regional Registration Request and a Regional Registration Reply with the GFA), the nFA replaces the role of oFA.

**Location Tracking**

The ability to determine a user’s location in an existing 802.11 wireless network can provide many useful services for wireless users. Such services include: location sensitive content delivery, such as being able to send documents to a vicinal printer; creation of real-time roadmap, asset tracking (locating a valuable device), etc. Some location mechanisms use additional devices such as GPS, some not.

The Global Positioning System (GPS) is a worldwide radio-navigation system consists of a constellation of 24 satellites and their ground stations. GPS uses these "man-made stars" as reference points to calculate positions accurate to a matter of meters. In fact, with advanced forms of GPS the accuracy can be better than a centimeter[Trim04].
A GPS device, through triangulation of multiple signals received and determination of propagation (how long it took the signal to go from the satellite to the GPS device), is able to accurately determine a user’s location to within a meter. The problem with GPS is that the device must have a clear line of sight between itself and the satellite. This means the technology is unusable in heavily forested areas, urban environments with tall buildings and indoor environments.

Some works has been done to use the popular 802.11 network infrastructure to determine the user location without using any extra hardware. Generally, suck kind of system needs to measure the signal quantity as a function of distance and one or more reference point such as the APs in the wireless LAN. The signal strength decays logarithmically with distance in an open space. But in indoors, the wireless channel is very noisy and the radio frequency (RF) signal can suffer from reflection, diffraction, and multipath effect [Yous03], which makes the signal strength a complex function of distance. To overcome this problem, WLAN location determination systems may constructs radio-maps during offline by sampling the signal at selected locations in the area of interest and tabulate the complex function. When the system need to determine the location, the vector of samples received from each access point is compared to the radio-map and the “nearest” match is returned as the estimated user’s location.

[Yous04] divided the radio map-based techniques into two broad categories: deterministic techniques and probabilistic techniques. Deterministic techniques, such as RADAR system in [Bahl00] and Location Information Privacy Model in [Smai01],
represent the signal strength of an access point at a location by a scalar value, for example, the mean value, to estimate the user location. Probabilistic techniques measure the signal quantity as a function of distance from the APs and store information them into a radio map and use probabilistic techniques to estimate users location. [Cast01] [Ladd02] [Roos02] [You04] [You03] are all using probabilistic techniques.

RADAR, An In-Building RF-based User Location and Tracking System, was developed in Microsoft Research. In RADAR, the signal strength is measured when transmitting beacon packet between the mobile host and AP. They take sample of radio signals and build up a radio map for the area interested during offline phase. RADAR uses 3 APs as reference point of its location, which is called triangulation. During location phase, it matches the real time signal strength with the radio map and determines the user’s location. The match is done by linear search.

Horus system from the University of Maryland is an RF-based location determination system [You04] [You03]. It is implemented in the context of 802.11 wireless LANs. The system uses the stored radio map to find the location that has the maximum probability given the received signal strength vector. In [Yous04], they also proved formally that probabilistic techniques give more accuracy than deterministic techniques.

Other Related Work

IEEE802.11 standard was originally devised to replicate in a wireless fashion the structures of the wired LANs. Only recently the idea of utilizing IEEE802.11 technology
for high mobility scenarios has been taken into account and the range of WLAN based applications has been enriched. In [Mani03], Pierpaolo Bergamo from UCLA and Don Whiteman from NASA, experimentally studied the behavior of an IEEE802.11 wireless network when the nodes are characterized by mobility up to the speed of 240 km/h. The authors studied the survivability and the performance of a connection under various aggressive mobility conditions. These studies may be adapted for data telemetry from mobile airborne nodes to fixed networks or between airborne nodes. In [Sing02], authors assessed the performance of WLANs in different vehicular traffic and mobility scenarios. The network throughput and the quality of the wireless communication channel, measured on IEEE 802.11b compliant equipment, are observed to degrade with increasingly stressful communication scenarios. [Amic02] presents a project using a WiFi-like network for military telemetry applications. For military telemetry, aircrafts and/or cars equipped with IEEE802.11 enabled devices will communicate with a fixed backbone infrastructure. The authors of [Amic] focused on aspects like frequency selection and network security. In [Thor], authors developed their own frequency hopping transceiver working at 900 MHz for telemetry purposes. In [Bamb], authors assured through analytical considerations that these kinds of transceivers can guarantee an impressive tolerance to rapid moving environments.

A review on recent research on MIP shows a great amount of efforts contributed to reducing MIP handoff latency. Malki [Malk02] proposed two mobility protocols, pre- and post-registration, using L2 trigger. In pre-registration, MN may communicate with both oFA and nFA. In post-registration, data are cached in nFA before the registration is completed. Fast-handover [Kood02] for Mobile IPv6 network combines the about two methods. But they all depend on L2 information. S-MIP[Hsie03], uses MN location and
movement patterns to ‘instruct’ the MN when and how handoff should be carried out. [Wijn04] also uses MN’s movement model to predict handoff. But all these efforts didn’t consider the speed factor of MN, which may cause problems when the MN moving rapidly.
CHAPTER 3
PERFORMANCE OF MIP OVER WLAN AT DIFFERENT SPEEDS

MIP over Wireless LAN Handoff Procedure

MIP over wireless LAN provides more flexibility and mobility to mobile IP network. Unlike a traditional wired mobile IP network, which requires a wire to connect a computer to the network, wireless LAN users can access IP network from nearly anywhere without losing connectivity.

Mobile IP is designed independently for all Layer 2 technologies, so it can run on any layer 2 infrastructures. But such kind of independency also costs more overhead. Figure 3-1 is the handoff procedure of MIP over two wireless LAN. When a MN moves from wireless LAN1 to wireless LAN2, it performs a layer2 802.11b handoff between Access Point 1 (AP1) and Access Point 2(AP2). After the layer2 handoff, the MN begins a layer3 handoff, which is MIP handoff. Suppose there is a communication, for example a TCP stream, between MN and CN. After the layer2 and layer3 handoff, it will require a significant time interval to recover the communication. This time internal is called layer4 handoff latency, which is also a part of the whole handoff cost. Equation 1 gives the life-cycle of MIP over wireless LAN handoff procedure:
\[ t_{\text{handoff}} = t_{L2\text{handoff}} + t_{L3\text{handoff}} + t_{L4\text{handoff}} \]  

(Equation 1)

Where \( t_{\text{handoff}} \) is the total handoff latency of MIP over wireless LAN, \( t_{L2\text{handoff}} \), \( t_{L3\text{handoff}} \), \( t_{L4\text{handoff}} \) are the handoff cost of Layer2, Layer3 and Layer4 separately.

In the following section, we introduce an emulation testbed, RAMON, which is used to evaluate the performance of MIP over WLAN and to analyze the handoff latency of the MIP handoff procedure.

**RAMON Testbed**

In order to evaluate the performance of MIP over WLAN, we build up a MIP emulator RAMON\[Hern02\]. RAMON is a Rapid Mobile Network emulator. It’s a testbed combining software and hardware components to produce a realistic experimentation environment that can test the behavior and performance of actual mobile systems. The testbed provides the wireless and wired infrastructure to allow experimental testing of wireless and wired mobile network protocols. Figure 3-2 is the architecture of RAMON.

RAMON consists of a Pentium II pc as Emulator, a circuit board as Controller, three JFW Industries Attenuators with Antennas, three Cisco 350 Access Points, three FAs, a HA and one or more MNs. All the FAs, HA and MN, which are the major entities of MIP, are running Linux kernel 2.4.20 and are installed with HUT dynamic MIP implementation version 0.8.1. The Attenuators are program controllable devices. The Emulator manipulates the Attenuators by the Controller to control the signal strength coming out from the Access Points. By increasing or decreasing the signal strength of one AP, we can emulate the MN moving towards to or away from the AP. By varying the increasing or
decreasing speed of the signal strength, we can emulate the speed change of the MN. The emulation program running on the emulator can dynamically change the IP addresses for each AP and FA so that every physical AP(and FA) in Figure. 3-2 can emulate multiple logical AP(and FA) in Figure 3-3.

Figure 3-1 RAMON testbed architecture

**Hardware Architecture**

The hardware architecture of RAMON includes

- two PCs—one is emulator, one is home agent
  - The emulator has four Ethernet cards. IP addresses are
    - Eth0: 192.168.1.2 mask 255.255.255.0
    - Eth1: 192.168.2.2 mask 255.255.255.0
    - Eth2: 192.168.3.2 mask 255.255.255.0
    - Eth3: 192.168.4.2 mask 255.255.255.0
The HA has two Ethernet cards. IP address are
- Eth0: 10.3.3.14 mask 255.255.255.0
- Eth1: 192.168.4.2 mask 255.255.255.0

- 3 IBM ThinkPad laptops—as 3 foreign agents
  - FA1 : eth0: 192.168.1.1 mask 255.255.255.0
  - FA2 : eth0: 192.168.2.1 mask 255.255.255.0
  - FA3 : eth0: 192.168.3.1 mask 255.255.255.0

- 3 CISCO AIRONET 350 Aps, IP addresses are
  - AP1: 192.168.1.3
  - AP2: 192.168.2.3
  - AP3: 192.168.3.3
  - The backup configuration files of these 3 APs are saved in the emulator.

- 3 Omnidirectional 3dbi Cushcraft Antennas
- one control board—control the attenuator to emulate the signal fading
- 3 JFW Industries – 50p-1230 Attenuators
- one Laptop-- as mobile host
  - MN eth0: 192.168.4.5

Software Architecture

- Emulator:
  - Linux Kernel 2.4.7-10.
  - Modules IPIP
  - Script emulator to create Virtual interfaces and routing table
  - Emulation object file to run the emulation

- HA:
  - Dynamics HUT mobile IP home agent package: dynhad.
  - NAT
  - Modules IPIP

- FA:
  - Dynamics HUT mobile IP foreign agent package: dynfad.
  - Modules IPIP
  - Script FA1, FA2, FA3. simulate the action of foreign agents.
  - DynX.1 – dynfad.conf configure files.

- MN:
  - Dynamics HUT mobile IP foreign agent package: dynmnd.
  - Dynmndconf1 – dynmnd.conf configure file.
  - Tcpdump to capture data
  - Ethereal tool for analysis.
Performance Evaluation

Emulation Scenario and Result

Using RAMON, we emulate HUT dynamic MIP in the following scenario in Figure 3-3:

![Figure 3-2 Dynamic MIP sample scenario](image)

In this scenario, a rapid moving MN will travel through 8 APs. Each AP is wired to a FA. The distance between every two APs is \( d = 250m, 500m \) or \( 1000m \). The moving speed of MN is \( V \), varying from 10m/s to 80m/s. In our experiments, we used ftp to transfer a large file from the CN to the MN. During the ftp transfer, we tracked down TCP sequence numbers by using the tool tcpdump. We analyze the tcpdump data by using ethereal. Here we only give the experiment results for \( d = 500m \) and \( 1000m \), \( v = 10m/s \) to 80m/s. Figure 3-4 and figure 3-5 are the time-sequence graph and throughput graph at speed 20m/s and AP distance 1000m. Figure 3-6 and 3-7 shows the time-sequence graph and throughput
graph at speed 80m/s and AP distance 1000m. Figure 3-8 to figure 3-11 are those graphs at speed 10m/s, AP distance 500m and speed 40m/s, AP distance 500m.

Figure 3-3 Time-sequence graph at speed 20m/s and AP distance 1000m

Figure 3-4 Throughput graph at speed 20m/s and AP distance 1000m
Figure 3-5 Time-sequence graph at speed 80m/s and AP distance 1000m.

Figure 3-6 Throughput graph at speed 80m/s and AP distance 1000m.
Figure 3-7 Time-sequence graph at speed 10m/s and AP distance 500m.

Figure 3-8 Throughput graph at speed 10m/s and AP distance 500m
Experimental Result Analysis

To compare the performance of MIP/ WLAN at different speeds and different AP distances, we list the experiment data in table 3-1. In the table, the bytes transferred are the total bytes transferred from when the MN enters the first cell to when it moves out of the last cell. The average throughput is calculated by dividing bytes transferred by travel time.
The total handoff time is the summary of the handoff latency of 7 times handoffs. The effective time is the time for effectively transferring data, which equals to the travel time minus the total handoff time.

Table 3-1 shows the average throughput drops when the MN’s speed goes up. At the same speed of 20m/s, the average throughputs are 196.97kB/s for d=1000m and 167.172kB/s for d=500m. At the speed of 40m/s, the average throughputs are 167.512kB/s for d=1000m and 93.877kB/s for d=500m. The table shows that if we double the speed and at the same time double the AP distance, the average throughput shows no suggestive difference. For example, at the speed of 40m/s and AP distance 1000m the average throughputs is 167.512kB/s. At the speed of 20m/s and AP distance 500m the average throughputs is 167.172kB/s.

Table 3-1 Average Throughput at Different Speeds and AP Distances.

<table>
<thead>
<tr>
<th>Speed (m/s)</th>
<th>AP distance (m)</th>
<th>Bytes transferred (kB)</th>
<th>Travel Time (s)</th>
<th>Average throughput (kB/s)</th>
<th>Total handoff time(s)</th>
<th>Effective time(s)</th>
<th>PMaxavg (kB/s)</th>
<th>Handoff Rate (FAs/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>1000</td>
<td>78000</td>
<td>396</td>
<td>196.970</td>
<td>58</td>
<td>338</td>
<td>232.5</td>
<td>0.02</td>
</tr>
<tr>
<td>40</td>
<td>1000</td>
<td>33000</td>
<td>197</td>
<td>167.512</td>
<td>57</td>
<td>140</td>
<td>234.31</td>
<td>0.04</td>
</tr>
<tr>
<td>60</td>
<td>1000</td>
<td>16700</td>
<td>130.5</td>
<td>127.969</td>
<td>56</td>
<td>74.5</td>
<td>234.07</td>
<td>0.06</td>
</tr>
<tr>
<td>80</td>
<td>1000</td>
<td>9200</td>
<td>98.5</td>
<td>94.359</td>
<td>57</td>
<td>41.5</td>
<td>232.673</td>
<td>0.08</td>
</tr>
<tr>
<td>10</td>
<td>500</td>
<td>78500</td>
<td>397</td>
<td>197.733</td>
<td>58</td>
<td>339</td>
<td>233.01</td>
<td>0.02</td>
</tr>
<tr>
<td>20</td>
<td>500</td>
<td>33100</td>
<td>198</td>
<td>167.172</td>
<td>56</td>
<td>142</td>
<td>234.4</td>
<td>0.04</td>
</tr>
<tr>
<td>30</td>
<td>500</td>
<td>16600</td>
<td>129</td>
<td>128.682</td>
<td>56</td>
<td>73</td>
<td>232.86</td>
<td>0.06</td>
</tr>
<tr>
<td>40</td>
<td>500</td>
<td>9200</td>
<td>98</td>
<td>93.877</td>
<td>58</td>
<td>40</td>
<td>232.8</td>
<td>0.08</td>
</tr>
</tbody>
</table>
The analysis of table 1 also shows: (1). The total handoff time doesn’t change with speed. (2). Effective-time/total-travel-time ratio drops when the speed goes up. This is the reason why higher speed has lower throughput.

Figure 3-12, the average throughput vs. speed graph, gives a more obvious view of this conclusion.

![Figure 3-11 Average throughputs vs speeds.](image)

In order to figure out the relationship between the performance of MIP over wireless LAN and the moving speed, we measured the throughputs of MIP over wireless LAN at different moving speeds and AP distances when there are no handoffs. We call this throughput, $P_{\text{Maxavg}}$, the maximum average throughput without handoff. Here we only give the time-sequence graph at AP distance 1000m with speed 20m/s(left) and 80m/s(right).

From figure 3-13, we get $P_{\text{Maxavg}} = 93000\text{kB} / 400\text{s} = 232.5 \text{ kB/s}$. From the right graph of figure 3-14, we get $P_{\text{Maxavg}} = 23500\text{kB} / 101\text{s} = 232.673 \text{ kB/s}$. The $P_{\text{Maxavg}}$ at different moving speeds and AP distances are listed in table 1.
Figure 3-12 Time-sequence graph at AP distance 1000m with speed 20m/s without handoff

Figure 3-13 Time-sequence graph at AP distance 1000m with speed 80m/s without handoff

Let $P_{avg}$ – Average throughput

$P_{maxavg}$ – Average throughput without handoff

$T_{travel}$ – Total travel time
**Te**ffective – Total effective time for ftp transmission

**Th**andoff – Total handoff time while traveling

**K**handoff – The number of handoffs while traveling

**th**andoff – Average handoff time among 7 times of handoff

Then, 

\[
P_{\text{avg}} = \left( \frac{P_{\text{maxavg}}}{T_{\text{travel}}} \right) \times T_{\text{effective}}
\]

\[
= P_{\text{maxavg}} \left( \frac{T_{\text{travel}} - T_{\text{handoff}}}{T_{\text{travel}}} \right)
\]

\[
= P_{\text{maxavg}} \left( 1 - \frac{T_{\text{handoff}}}{T_{\text{travel}}} \right)
\]

\[
= P_{\text{maxavg}} \left( 1 - K_{\text{handoff}} \times \frac{\text{thandoff}}{T_{\text{travel}}} \right)
\]

\[
= P_{\text{maxavg}} \left( 1 - \left( \frac{K_{\text{handoff}}}{T_{\text{travel}}} \times \text{thandoff} \right) \right)
\]

Since thandoff doesn’t change, The change of Pavg is caused by Khandoff/T_{travel} ratio.

We define MN handoff rate as \( r_h = \frac{v}{d} \), which is the ratio of the MN’s speed and the cell size(AP distance). It means that how many APs or FAs the MN hands over in one second. \( r_h \) is also equal to \( K_{\text{handoff}} / T_{\text{travel}} \).

The relationship between the performance of MIP/WLAN and the moving speed is presented in Equation 2:

\[
P_{\text{avg}} = P_{\text{maxavg}} \left( 1 - r_h \times \text{thandoff} \right)
\]

Equation 2

Where Pavg is the average throughput for the MN; \( P_{\text{maxavg}} \) is the average throughput without handoff. \( \text{thandoff} \) is the average handoff time for each handoff procedure.
Since $t_{\text{handoff}}$ doesn’t change, the change of $P_{\text{avg}}$ is caused by handoff rate $r_h$. At handoff rate $0.02$ FAs/s, the average throughput is $197.35$ kB/s. When the handoff rate goes up to $0.08$ FAs/s, the average throughput drops to $94.118$ kB/s. The graphs in Figure 3-12 can be combined into graph in Figure 3-15.

![Figure 3-14 Average throughput vs handoff rate](image)

This chapter shows that the performance of MIP over WLAN is depends on the MN’s handoff rate. In Chapter 5, we will propose an idea of how to make use of this throughput/handoff-rate relationship to improve the performance of MIP over wireless LAN in rapid moving environment. In the following chapter, we will take a deep view of the handoff latency by breaking down the handoff procedure of MIP over wireless LAN.
CHAPTER 4
QUANTITATIVE ANALYSIS OF THE MIP OVER WIRELESS LAN HANDOFF LATENCY

Equation 1 in Chapter 3 shows that the life-cycle of MIP over wireless LAN handoff is the summary of Layer2, Layer3 and Layer4 handoff latency. In the following sections, we analyze the handoff characters of each layer and provide a quantitative analysis of the MIP over wireless LAN handoff latency.

Layer 2 Handoff Latency

In the case of IEEE 802.11b WLAN, Layer2 handoff is the change of APs. It causes an interruption of data frame transmission. Buffering and routing update make the handoff time for uplink and downlink traffic different. Some researches have been done to even this difference[El-Ho00][Ren99]. In our experiments, we only concern the downlink handoff time. In [Vela04], Hector Velayos splitting the Layer2 handoff time into three sequential phases: detection, search and execution. In our experiment, we also split it into three parts and name them as: movement detection, AP searching and reassociation.

The Layer2 handoff involves three participating entities, the station(here is the MN), an old AP(oAP) and a new AP(nAP). The oAP is the access point which the station had layer2 connectivity prior to the handoff, while the nAP is the access point to which the station gets layer2 connectivity after the handoff. The handoff process among 2 APs also includes information exchanges. This information typically consists of the station’s
credentials and accounting information. The message exchange between APs can be done by Inter Access Point Prototcol(IAPP)[11F03] or via a proprietary protocol. The following is a detail analysis of three phases of Layer 2 handoff.

**Layer 2 Movement Detection Phase**

In oAP’s coverage, the station keeps frame transmission. There are three reasons for frames lose: collision, radio signal fading, or oAP is out of range. The station first assumes the lost frame is cause by collision. In 802.11b standard, collision is handled by Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol. CSMA/CA is a basic protocol used to avoid signal collision and canceling. It works by requesting authorization to transmit for a specific amount of time prior to sending information. When collision happens, the sending device broadcasts a Request To Send (RTS) frame with information on the length of its signal. If the receiving device permits it at that moment, it broadcasts a Clear To Send (CTS) frame. Once the CTS is transmitted, the sending machine transmits its information. Any other sending device in the area that “hears” the CTS realize another device will be transmitting and allow that signal to go out uncontested. If the station tried to retransmit several times and still unsuccessful, then it assumes signal fading. This time the station sends out probe requests to probe the link. After several probe requests and without any response, the station assumes oFA is out of range and begin AP searching phase. In figure 10, from TCP point of view, when MN receives the last TCP package, it responses with TCP ACKnowledgement. After several
unsuccessful transmission of TCP ACK, the MN assumes the oAP is out of range and starts a new AP searching phase.

**Layer2 AP Searching Phase**

After the station assumes oAP is out of range, it tries to find new potential APs to associate to. This is done by 802.11b MAC layer function: SCAN. There are two methods of scanning, active and passive. In passive scanning, the station listen to each channel for beacon frames (broadcasted periodically by APs every 10ms). The station takes note of the corresponding signal strengths while scanning. The beacons contain information about the AP, including service set identifier (SSID), supported data rates, etc. The station can use this information along with the signal strength to compare APs and decide upon which one to chose. In active scanning, the station broadcasts a probe request frame and waits for response. The time to wait for responses depends on the channel status. If the channel is idle during MinChannelTime, the station can receive prove response form the AP on that channel. If there is any traffic during this time, the station will wait for MaxChannelTime to allow the data in the channel be transmitted and wait for AP’s response. After gathers several response from APs in range, the station will compare and choose one to associate to. Active scanning enables a station to receive immediate response from APs, without waiting for beacon frames. However, it imposes additional overhead on the network. In our experiment, only after got 3 probe responses from an AP, the station regards that AP is stably in range. This is a default configuration of Orinoco Wireless card.
Layer2 Reassociation Phase

After choose one AP in phase 2, the station sends out a reassociation request to nAP. If the nAP can get the credentials and other state information from oAP through IAPP, there is no Authentication message exchange between the station and nAP. Or else, the station will send out authentication request to the nAP and wait for response. After authentication, nAP reassociates the station and sends reassociation response back.

The above three phases complete Layer2 handoff. The layer2 handoff latency can be expressed in Equation 3.

\[ t_{L2handoff} = t_{L2detection} + t_{L2searching} + t_{L2reassociation} \]  

(Equation 3)

Where \( t_{L2detection} \), \( t_{L2searching} \) and \( t_{L2reassociation} \) are the time costs for Layer2 movement detection, Layer2 AP searching and Layer2 reassociation. Figure 4-1 shows these three phases in green arrows and are indexed as L2.

Layer 3 Handoff Latency

Only after the layer 2 link has been established, could the Layer 3 handoff starts, because the MN can only communicate with the FA on the same link. The Layer 3 handoff involves 2 phases, agent discovery and registration.

Agent Discovery

The well know agent discovery algorithms are Lazy Cell Switching(LCS) and Eager Cell Switching(ECS)[Perk98].
The LCS method is a reactive handoff initiation strategy. In LCS the MN keeps receiving Agent Advertisement messages from the oFA and refreshes the lifetime of the CoA and stays in the original network until it moves and loses contact with oFA for the duration of three advertisement (FA broadcast Agent Advertisement message every 1
second), which means oFA becomes unreachable. A handoff will be initiated if a nFA is discovered after this moment. If the nFA hasn’t been discovered before the oFA becomes unreachable, the handoff latency will be much higher. An advantage of the LCS is to reduce the frequency of handoff when the MN hangs around among several FA. As to MIP over WLAN, because the MN can only keep physical link with one FA, the new agent can’t be discovered before the old agent becomes out of range. Figure 4-2 is the LCS handoff latency plot for MIP.

ECS is a proactive initiation strategy. It dictates an immediate MIP handoff as soon as a new agent is discovered. ECS is effective for the moving patterns that the MN rarely change its moving direction. Figure 4-3 is the ECS handoff latency plot for MIP.

![Figure 4-3 ECS handoff latency for MIP](image)

**Registration**

When a MN realizes that it is on a foreign network and has acquired a care-of-address from the nFA, it needs to notify the HA so that the HA can forward IP
packets between MN and CN. This is done by registration. The registration process involves four steps.

- The MN sends a registration request to nFA.
- The nFA relays this request to the GFA or HA.
- The HA either accepts or denies the request and sends a registration reply to nFA. If it accepts the request, it will build a tunnel downward to nFA (if FA decapsulation is used).
- The nFA relays this reply to the MN. If the registration reply is positive, it will build a tunnel upward to HA or GFA.

If the MN is using a collocated care-of-address, it will register directly with the HA, which is not the case in this paper.

The layer3 handoff latency can be splitted into Equation 4 [Fiko01]. Figure 4-1 shows these two phases in red arrows and are indexed as L3.

\[
 t_{L3\text{handoff}} = t_{\text{mipagentdicovery}} + t_{\text{mipregistration}} \\
\text{ (Equation 4)}
\]

**Layer 4 Handoff Latency**

TCP is a connection-oriented, end-to-end reliable protocol designed to support error recovery and flow control. Reliability is insured by a sliding-window acknowledgement and retransmission mechanism. All data sent by TCP must be acknowledged by the receiver. TCP maintains a variable-sized window of data that is unacknowledged for a given time. If the window is full, no data will be sent until an acknowledgement is received. TCP maintains a Retransmission Time Out (RTO) timer. If no ACK has been received when the RTO timer expired, TCP assumes that the data has lost and retransmits all of the data in the window. The retransmission follows the exponential back-off algorithm. According to this algorithm TCP doubles the timeout value on unsuccessful
successive retransmissions [Hsie03]. In our case, during the Layer2 and layer3 handoff, the TCP doubles the retransmission timeout value several times. So even after the layer2 and layer3 handoff is over, TCP still have to wait for RTO to timeout to recover the retransmission. In figure 4-1, the dash blue arrows depict the TCP retransmission interval has been doubled. This latency is cost by TCP exponential back-off algorithm. So we call it TCP back-off delay: $t_{\text{tcp-back-off}}$.

We define $t_{\text{L4handoff}} = t_{\text{tcp-back-off}}$ (Equation 5)

**Quantitative Analysis of the Handoff Latency**

According Equation 1, 2, 3 and 4, the handoff latency distribution for MIP over WLAN is show in Equation 6.

$$t_{\text{handoff}} = t_{\text{L2detection}} + t_{\text{L2searching}} + t_{\text{L2reassociation}} + t_{\text{mipagentdicovery}} + t_{\text{mipregistration}} + t_{\text{tcp-back-off}}$$ (Equation 6)

We used RAMON introduced in Section 3 to emulate the same scenario as in Section 3. We did 20 times experiments to get the average handoff latency. The experimental result of the handoff latencies of MIP over wireless LANs listed in table 4-1. Table 4-1 gives 20 times of experiment data. Each row is one experiment. Each column is the time latency for that handoff phase. The data in the last column are the total handoff latencies for every experiment. The number in the bottom right cell is the average handoff latency.
Table 4-1 Handoff latency distribution of MIP over WLAN

<table>
<thead>
<tr>
<th>Latency Exp#</th>
<th>L2 movement detection</th>
<th>L2 AP searching</th>
<th>L2 reassociation</th>
<th>MIP agent discovery</th>
<th>MIP registration</th>
<th>TCP backoff</th>
<th>Handoff latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.033</td>
<td>0.061</td>
<td>0.005</td>
<td>2.996</td>
<td>0.073</td>
<td>5.058</td>
<td>9.226</td>
</tr>
<tr>
<td>2</td>
<td>1.064</td>
<td>0.044</td>
<td>0.009</td>
<td>1.945</td>
<td>0.042</td>
<td>6.01</td>
<td>9.511</td>
</tr>
<tr>
<td>3</td>
<td>1.133</td>
<td>0.063</td>
<td>0.006</td>
<td>3.023</td>
<td>0.052</td>
<td>5.345</td>
<td>9.622</td>
</tr>
<tr>
<td>4</td>
<td>1.032</td>
<td>0.100</td>
<td>0.008</td>
<td>2.563</td>
<td>0.050</td>
<td>5.323</td>
<td>9.076</td>
</tr>
<tr>
<td>5</td>
<td>1.044</td>
<td>0.065</td>
<td>0.003</td>
<td>2.756</td>
<td>0.052</td>
<td>5.125</td>
<td>9.045</td>
</tr>
<tr>
<td>6</td>
<td>1.131</td>
<td>0.057</td>
<td>0.004</td>
<td>2.578</td>
<td>0.043</td>
<td>5.004</td>
<td>8.817</td>
</tr>
<tr>
<td>7</td>
<td>1.009</td>
<td>0.056</td>
<td>0.010</td>
<td>2.436</td>
<td>0.060</td>
<td>5.625</td>
<td>9.196</td>
</tr>
<tr>
<td>8</td>
<td>1.120</td>
<td>0.060</td>
<td>0.006</td>
<td>3.001</td>
<td>0.704</td>
<td>5.002</td>
<td>9.893</td>
</tr>
<tr>
<td>9</td>
<td>1.023</td>
<td>0.059</td>
<td>0.026</td>
<td>2.213</td>
<td>0.054</td>
<td>4.998</td>
<td>8.373</td>
</tr>
<tr>
<td>10</td>
<td>1.039</td>
<td>0.076</td>
<td>0.005</td>
<td>3.008</td>
<td>0.053</td>
<td>5.006</td>
<td>9.187</td>
</tr>
<tr>
<td>11</td>
<td>1.100</td>
<td>0.045</td>
<td>0.030</td>
<td>2.770</td>
<td>0.041</td>
<td>5.728</td>
<td>9.714</td>
</tr>
<tr>
<td>12</td>
<td>1.013</td>
<td>0.049</td>
<td>0.010</td>
<td>2.545</td>
<td>0.042</td>
<td>4.768</td>
<td>8.427</td>
</tr>
<tr>
<td>13</td>
<td>1.021</td>
<td>0.051</td>
<td>0.009</td>
<td>3.001</td>
<td>0.065</td>
<td>5.202</td>
<td>8.896</td>
</tr>
<tr>
<td>14</td>
<td>1.006</td>
<td>0.043</td>
<td>0.017</td>
<td>2.600</td>
<td>0.046</td>
<td>5.312</td>
<td>9.024</td>
</tr>
<tr>
<td>15</td>
<td>1.104</td>
<td>0.069</td>
<td>0.006</td>
<td>2.598</td>
<td>0.047</td>
<td>4.544</td>
<td>8.368</td>
</tr>
<tr>
<td>16</td>
<td>1.003</td>
<td>0.064</td>
<td>0.013</td>
<td>2.674</td>
<td>0.062</td>
<td>4.806</td>
<td>8.622</td>
</tr>
<tr>
<td>17</td>
<td>1.110</td>
<td>0.054</td>
<td>0.010</td>
<td>2.783</td>
<td>0.054</td>
<td>5.705</td>
<td>9.716</td>
</tr>
<tr>
<td>18</td>
<td>1.100</td>
<td>0.064</td>
<td>0.006</td>
<td>3.012</td>
<td>0.057</td>
<td>5.602</td>
<td>9.841</td>
</tr>
<tr>
<td>19</td>
<td>1.302</td>
<td>0.056</td>
<td>0.009</td>
<td>2.349</td>
<td>0.070</td>
<td>5.71</td>
<td>9.496</td>
</tr>
<tr>
<td>20</td>
<td>1.098</td>
<td>0.044</td>
<td>0.004</td>
<td>2.404</td>
<td>0.062</td>
<td>5.172</td>
<td>8.784</td>
</tr>
<tr>
<td>Avg</td>
<td>1.074</td>
<td>0.059</td>
<td>0.010</td>
<td>2.660</td>
<td>0.086</td>
<td>5.253</td>
<td>9.142</td>
</tr>
</tbody>
</table>

We redraw figure 4-1 with handoff latency distribution in figure 4-4.
Figure 4-4 Handoff procedure with handoff latency distribution
CHAPTER 5
SPEED ADAPTIVE MIP AND ITS PERFORMANCE EVALUATION

From above analysis of handoff latency distribution, we can see the largest part is TCP back-off latency $t_{tcp-back-off}$. Because of TCP exponential back-off algorithm, if we reduce the L2 and L3 latency, $t_{tcp-back-off}$ will be reduced exponentially. In this chapter, we deal with L3 latency first. L2 and L4 latency will be considered in future works.

Traditional MIP over WLAN Handoff Procedure

The physical coverage of an IEEE 802.11-base wireless LAN is limited. To increase the coverage of a wireless network, one can deploy multiple wireless LAN cells or segments in an overlapped fashion where each cell is associated with an AP. AP serves as a layer-2 bridge between the high-speed wired network and the wireless LAN. As MNs move in and out of these overlapped cells, they can associate with the corresponding APs according to beacon signal strengths. In IEEE 802.11b-based networks, the intelligence to measure signal strength and switch among network segments is built into the wireless LAN NIC (Network Interface Card), which exposes various status and control information to the software device driver. To enable cellular-like networking structure, wireless LAN NIC need to be configured to run in the access point mode, which is also known as the infrastructure mode. Mobile IP provides MNs the ability to roam across wireless IP subnets without loss of network-layer connectivity. Any network application executing on
a mobile host with mobile IP support can continue to run regardless of any change in the mobile node’s point of attachment. With mobile IP, mobile nodes do not need to reconfigure their IP addresses while migrating from home subnets to foreign subnets. A generic wired and wireless network topology with which mobile IP operates is shown in Fig. 5-1 [Srik04].

![Figure 5-1 Traditional MIP Handoff Procedure](image)

In this topology, there are one HA and several FAs running on the wired network. The MN is communicating with CN through the wireless link with AP1. The FAs periodically broadcast mobile IP advertisements on the wireless LANs (message 1, 2, 3 and 4 in figure 5-1). Because there no wireless link between the MN and AP2, AP3 and AP4, the mobile IP advertisements messages 2, 3 and 4 can not be transferred to the MN. The mobile IP advertisements messages 1 can reach the MN. Since MN already registered on FA1, message 1 will be discarded by the MN. Whenever the MN migrates from one subnet
to another (foreign) subnet, it first needs to establish wireless connection with the corresponding AP then starts receiving mobile IP advertisements from the corresponding FA.

When an IEEE 802.11b-based wireless network is configured in infrastructure mode, the MN is associated with the AP, which is AP1 in figure 5-1, of the wireless LAN cell in which it currently resides. Each AP periodically broadcast beacon frames every 10ms in passive scanning mode (message 5, 6, 7 and 8 in figure 5-1). The beacons contain information about the AP, including service set identifier (SSID), supported data rates, etc. The station can use this information along with the signal strength to compare APs and decide upon which one to chose. If the MN chooses AP2, it initiates a link-layer handoff from AP1 to AP2. The MN sends a reassociation request message to AP2 (message 9 in figure 5-1). If the nAP can get the credentials and other state information of the MN from AP1 through IAPP, there is no Authentication message exchange between the MN and AP2. Or else, AP2 will send out authentication request to the MN and wait for response. After authentication, AP2 reassociates the MN and response with a reassociation response message (message 10). In all known IEEE 802.11b cards, this link-layer handoff logic is built into the firmware of the NIC, and does not generate any interrupts to notify the higher-layer software. If the new wireless LAN cell belongs to the same IP subnet as the old wireless LAN cell (like AP3 and AP4 belongs to the same subnet to FA3), then to the IP layer and above on the mobile node there is no change in connectivity and the network applications continue without any disruptions. However, if the new wireless LAN cell
belongs to a different IP subnet, then the MN can no longer communicate with CN until a
network layer handoff is completed. In this case, the MN would eventually receive an
advertisement from the FA2 through AP2 (message 2 in figure 5-1). The mobile IP
software running on the MN intercepts these advertisements and sends a registration
request to FA2 (message 11). This registration request is forwarded by FA2 to the
HA (message 12). After the authentication (not show in figure 5-1) a registration reply is
sent to the FA2 (message 13) and is relayed to the MN (message 14). The mobile IP handoff
is over and an IP-over-IP tunnel is established between the HA and FA2. From this point
onwards, the HA, acts as a proxy for the MN, forwards all packets to FA2 over the tunnel.
FA2 de-encapsulates the packets and forwards them to the MN. Similarly, all packets that
the MN transmits to the CN are first received by FA2 and are tunneled over to the HA,
which further routes them to the CN. This process is known as bidirectional tunneling.

The above process of switching from FA1 to FA2 as the MN moves across adjacent
wireless cells is called mobile IP handoff. After the moves to a new wireless LAN cell but
before the associated mobile IP handoff completes, the mobile node is essentially cut off
from the wired network. For a rapid moving MN, this mobile IP handoff latency greatly
duces the network performance. In extreme cases, the MN may even not be able to
accomplish mobile IP handoff. For example, assume a rapid moving MN moves at speed
V (m/s), the wireless LAN cell size is D (m) and the mobile IP handoff latency is T (s). If V
x T > D, the MN can never register to the wired network. Therefore, it is critical to reduce
the mobile IP handoff latency in rapid moving environments.
Algorithm of Speed Adaptive MIP

In Chapter 3, we define MN handoff rate as \( r_h = \frac{v}{d} \). It means MN move through how many APs or FAs per second. Chapter 3 also shows that the performance of MIP over WLAN is depends on the MN handoff rate among FAs. Figure 3-13 shows when the handoff rate is 0.02 FA/s, the average throughput is above 90kBytes/s. When the handoff rate rises to 0.08 FA/s, the average throughput drops to around 50kBytes/s. This means lower handoff rate has higher throughput. \( r_h \) is also equal to the ratio of \( K_{\text{handoff}}/T_{\text{travel}} \). We rewrite the handoff rate \( r_h = \frac{v}{d} \) in Equation 7.

\[
 r_h = \frac{K_{\text{handoff}}}{T_{\text{travel}}} \quad \text{(Equation 7)}
\]

Where \( K_{\text{handoff}} \) is the number of handoffs occurred during the MN traveling. \( T_{\text{travel}} \) is MN’s total travel time. In order to reduce handoff rate without changing total travel time, we can reduce the number of handoffs. The optimal is \( K_{\text{handoff}} = 0 \).

Let \( N \) be total FA numbers on the way MN traveling. Let’s assume somehow \( M \) is the number of FAs with whom the MN can communicate without L3 latency. The optimal is \( M = N \). But it costs too many resources, especially when the number of active MNs is large. Also we don’t know how long will the MN travel at the beginning.

We call \( M \) the size of the FA Set with whom the MN can communicate without L3 handoff latency. From IP level of view, \( M \) is the number of FAs that MN has registered to and can communicate with at that moment.
Now the question is:

- How to decide FA set size $M$
- How to guarantee MN can communicate with a FA set almost like to do with a single FA.

The first problem SA-MIP needs to deal with is to decide FA set size $M$. In SA-MIP algorithm, $M$ is decided by the following Equation.

$$M = \left| t_{\text{handoff}} \times r_h \right| + 1$$  \hspace{1cm} (Equation 8)

where $t_{\text{handoff}}$ is the handoff time for every handoff procedure, and $r_h$ is the handoff rate. Here we use the experimental average handoff time 9.142s for $t_{\text{handoff}}$. $r_h$ is dynamic.

For example, at speed 40m/s, AP distance 500m, $M = \left| 9.142 \times 40/500 \right| + 1 = 2$. At speed 80m/s, AP distance 500m, $M = 3$.

The second problem is how to guarantee MN can communicate with a FA set just like it can do with one FA. Our solution is to let MN pre-register $M$ potential FAs along the way MN traveling, at the same time multicast IP packets to those FAs in this FA set. So MN won’t feel any handoff latency from the IP level of view.

In Speed Apative MIP (SA-MIP), the set of FAs that MN can talk to without L3 latency is extended from one point at low moving speed to a line at high moving speed. The length of the line dynamically changes with the MN handoff rate as in figure 5-2. The behavior of SA-MIP will automatically adapt to the handoff rate of the MN so that the performance of SA-MIP won’t decline dramatically in rapid moving environments. At the same time SA-MIP only cost reasonable resource that is as much as enough for seamless handoff.
M = 1  \quad M = 2  \quad M = 3  \quad M = 4
\begin{align*}
r_h &= 0 & 0 \leq r_h &\leq 0.109 & 0.109 \leq r_h &\leq 0.218 & 0.218 \leq r_h &\leq 0.328
\end{align*}

Figure 5-2 FA Set size vs handoff rate

Speed detection and location tracking is an interesting topic on mobile computing. [Bahl00] [Yous03] are all making use of signal strength information to locate and track wireless users. [Erge02] uses GPS to inform mobile users about the prospective future location and to improve performance of the ad hoc routing. In this paper, we assume the MN has GPS system to detect its location. When the MN moves at speed $v$, if 
$30$ m/s ($67.10$ miles/h), it performs a normal registration. If $30$ m/s $\leq v \leq 40$ m/s ($89.4$ miles/h), it initializes registration after receiving two successive agent advertisements. If $v \geq 40$ m/s ($89.4$ miles/h), we assume the MN won’t change its direction largely in a short distance. It initializes registration once it gets a new agent advertisement.

MN’s registration message is extended by speed extension. According to Mobile IP Vendor/Organization-Specific Extensions [RFC3115]. Two Vendor/Organization Specific Extensions are allowed for MIP, Critical (CVSE) and Normal (NVSE) Vendor/Organization Specific Extensions. The basic difference is when the CVSE is encountered but not recognized, the message containing the extension must be silently discarded, whereas when a NVSE is encountered but not recognized, the extension should be ignored, but the rest of the Extensions and message data must still be processed. We use the NVSE extension.
The following is the NVSE format.

![Figure 5-3 Normal vendor/organization specific extension](image)

In figure 5-3, the type here is 134 for NVSE extension. Length is the size in bytes of the extension, not including the type and length bytes. The vendor/org-ID is assigned in RFC 1700. We pick up a large unassigned number 5205. Vendor-NVSE-Type Indicates the particular type of Vendor-NVSE-Extension. The administration of the Vendor-NVSE-Types is done by the Vendor. Vendor-NVSE-Value here is a floating point number for handoff rate.

Figure 5-4 shows the SA-MIP handoff procedure and message exchange.

Whenever the MN needs to handoff to a new FA set, after it gets that many times of agent advertisements which is determined by speed(step 1 in figure 5-4), it sends a registration request with up-to-date handoff rate information to the very first FA in a new FA set(step 2). The first FA relays the registration request to upper FA or HA(step 3). Meanwhile, it decapsulates the speed extension, refill the MIP header and authentication extension and then forward it to other FAs(M-1 FAs) in this FA set(step 4). Assume the handoff rate is below 0.109. The FA set size at this time is 2. These other FAs relay the registration request to upper FA or HA as well, just like the request comes from the
MN (step 5). When the GFA or HA received these registration requests, it builds up tunnels downwards to each FA and responses with registration reply (step 6 and 7). When the FA received the registration reply, it builds up tunnel upwards to the GFA or HA.

![Diagram of SA-MIP handoff procedure](image)

**Figure 5-4 SA-MIP handoff procedure**

Whenever the MN setups the Link-layer contact with the FA, the later forwards the registration reply to the former (step 9 or 10). The MN gets the care-of-address from agent advertisement message (step 10 or 9) or registration reply message (step 9 or 10), and begins data communication. At the same time, it sends registration request to the new FA with up-to-date speed information (step 11). This new FA decapsulates the registration request message and sets up a new FA set. Assume the handoff rate is between 0.109 and 0.218. The FA set size is 3 at this time. The new FA (FA2) refill the MIP header and authentication extension and then forward it to other FAs (FA3 and FA4 in the figure) in
this FA set and repeats the above process. In Figure 5-4, the FA set size $M$ changes from 2 to 3 when the MN handoff rate changes from 0.08 to 0.11.

**Implementation of Speed Adaptive MIP**

Mobile IP has three main entities, HA, FA and MN. HUT dynamic MIP implementation version 0.8.1, originally developed at Helsinki University of Technology (HUT), is a scalable, dynamical, and hierarchical Mobile IP software for Linux operating system. The SA-MIP is developed on HUT dynamic MIP implementation version 0.8.1.

**Home Agent**

The HA implementation of SA-MIP is almost the same as HUT dynamic MIP except the Registration Request validation check function. The following describes the basic functionalities of HA.

The HA is responsible for encapsulating and forwarding packets to its MNs when they are away from their Home Network. It also decapsulates and forwards tunneled packets originating from its Mobile Nodes. The HA communicates with FAs and MNs using Berkeley IP sockets. The HA listens to ICMP agent solicitation messages from MNs on a "packet" socket. ICMP agent advertisement messages are sent in reply to these messages on the same socket. The HA also listens to Registration Requests on a UDP socket (port 434 by default) originating from FAs or MNs. If Registration Requests is validate a mobility binding for the requested Mobile Node will be established or, if one already exists, updated. The request is then answered with an corresponding Registration Reply.
When received of a Registration Request Message the HA performs a Registration Request validation check process. It first looks up the shared secret for the corresponding MN. The shared secret is used to check the MAC of the request message. If a Mobility Binding for the MN exists, then the timestamp in the request is checked to be greater than the one in the Mobility Binding. If either of these checks fails the HA responds to the sender with a Registration Reply indicating registration failure. If the checks succeed the HA determines the smaller lifetime value of the one in the request and the HA's pre-configured maximum value. It then generates a Session Key and creates a Mobility Binding consisting of the MN's address, its highest FA, the identification timestamp and the Session Key. The HA then responds with a Registration Reply indicating registration success. The message includes the same timestamp as the request, the lifetime value, a MAC, the Session Key encrypted with the shared secret and the Session Key encrypted with the highest FA's public key. The HA configures a tunnel between itself and the highest FA and works as a proxy for the registered MN. If the lifetime in the request is set to zero, the HA interprets this as a deregistration from the MN. On deregistration the HA purges the tunnel configuration and stops the proxy ARP functionality for the MN’s address. If the FA differs in a reregistration, a Registration Reply with a lifetime set to zero is sent to the previous FA to indicate that the old tunnel should be torn down.

In order to focus on performance issues of mobile IP, we ignore the security check part. When the HA checks the validation of the Registration Requests, the MN-HA
authentication check is comment out. Figure 5-5 is the function flowchart of Registration in HA.

Mobile Node

In addition to the basic function of HUT dynamic MIP’s MN, the MN of SA-MIP needs to transfer moving speed information to FAs. This is done by extending the Registration Request message with speed extension. The Registration function in HUT dynamic MIP implementation is the method by which MN requests forwarding services when visiting a foreign network, informs their HA of their current CoA, renews a registration which is due to expire, and/or deregisters when they return home.

The Registration Request message has the following format.
The Registration Request message header consists of the fields from Type to Identification. The send_registration() function in the MN implementation first fills out the Registration Request header with corresponding data then fills out the extension.

Figure 5-7 is the Registration Request Message extension format.

The speed extension is as following.
struct speed_ext * mn_speed; // speed_ext struct for SA-MIP

if(speedChanged) //compare the handoff rate send last time with current one.
{
    mn_speed = (struct speed_ext *) pos;
    if (left < sizeof(struct speed_ext)) //left is the message size after Registration
        //header
        return -1;
    mn_speed ->type = VENDOR_EXT_TYPE2;
    mn_speed ->length = sizeof(struct speed_ext) - 2;
    mn_speed ->reserved = 0;
    mn_speed ->vendor_id = htonl(VENDOR_ID_DYNAMICS);
    mn_speed ->sub_type = 25
    mn_speed -> mn_spd = handoff_Rate;
    pos += sizeof(struct speed_ext);
    left -= sizeof(struct speed_ext);
}

Figure 5-8 show the function flowchart of sending Registration Request

Figure 5-8 Function flowchart of sending registration request

**Foreign Agent**

Whenever the FA received a Registration Request from the MN, it decapsulates the message, checks the speed extension. If the handoff rate is non-zero, this FA calculates the FA set size M. It fills out the Registration Request header with new CoA and new MD5
MN-HA authentication. This new Registration Request message is sent to next M-1 FAs, which in turn forward the Registration Request one level up or the HA.

Figure 5-9 is the function flowchart for FA handling Registration Request.

**Figure 5-9 Function flowchart for FA handling registration request.**

**Evaluation of Speed Adaptive Extension for MIP**

We evaluate the performance of SA-MIP over WLAN under the same scenario as in Section 3. Figure 5-10 and 5-11 are the time-sequence graph at speed 60m/s($r_h = 0.06$) and
80m/s (\(r_h = 0.08\)) and AP distance 1000m. The average throughput at different speed is listed in table 5-1.

Figure 5-10 Time-sequence graph at speed 60m/s and AP distance 1000m

Figure 5-11 Time-sequence graph at speed 80m/s and AP distance 1000m
Table 5-1 Average throughput for speed-adaptive MIP

<table>
<thead>
<tr>
<th>Speed (m/s)</th>
<th>AP distance (m)</th>
<th>Bytes transferred (kB)</th>
<th>Travel Time (s)</th>
<th>Avg throughput (kB/s)</th>
<th>Handoff Rate (FAs/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>1000</td>
<td>85000</td>
<td>399</td>
<td>213.03</td>
<td>0.02</td>
</tr>
<tr>
<td>40</td>
<td>1000</td>
<td>37500</td>
<td>198</td>
<td>189.39</td>
<td>0.04</td>
</tr>
<tr>
<td>60</td>
<td>1000</td>
<td>19400</td>
<td>130</td>
<td>149.23</td>
<td>0.06</td>
</tr>
<tr>
<td>80</td>
<td>1000</td>
<td>11600</td>
<td>99</td>
<td>117.17</td>
<td>0.08</td>
</tr>
<tr>
<td>10</td>
<td>500</td>
<td>84400</td>
<td>398</td>
<td>212.06</td>
<td>0.02</td>
</tr>
<tr>
<td>20</td>
<td>500</td>
<td>37400</td>
<td>198</td>
<td>188.89</td>
<td>0.04</td>
</tr>
<tr>
<td>30</td>
<td>500</td>
<td>19500</td>
<td>131</td>
<td>148.55</td>
<td>0.06</td>
</tr>
<tr>
<td>40</td>
<td>500</td>
<td>11500</td>
<td>98</td>
<td>117.34</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Figure 5-12 is the average throughput vs. handoff rate before and after the speed adaptive MIP is installed. After installing SA-MIP, at handoff rate 0.02 FA/s, the average throughput is improved by \((212.54 - 197.35)/197.35 = 7.69\%\). At handoff rate 0.04, 0.06 and 0.08 FA/s, the average throughput is improved by 13.02%, 15.97% and 24.73% respectively.

![Figure 5-12 Average throughput vs. handoff rate](image-url)
CHAPTER 6
SUMMARY AND FUTURE WORKS

In this dissertation, in order to evaluate the rapid mobility of MIP in a laboratory environment, we build up the performance evaluation testbed on Wireless LAN. The emulation experiments showed that MIP is not suitable for rapid moving environments. We depicted the relationship between the performance and the handoff rate of MN and quantitatively analyzed the handoff latencies of the MIP over wireless LAN. A Speed Adaptive MIP is proposed and evaluated. The emulation showed that the SA-MIP can improve the performance from 8% to 25% when the handoff rate changes from 0.02 FA/s to 0.08 FA/s. Compared to the mechanisms of Malki[Malk02] and Koodli’s mechnism[Kood02], SA-MIP combines the pre- and post-registration methods, but keeps indenpendency from L2 infrastructure. Compared to Hsieh[Hsi03] and Wijngaert’s mechnism[Wijn04], SA-MIP not only predicts its next move but also involves next M number of FAs according to MN’s moving speed.

In our work so far, SA-MIP only deal with L3 handoff latency. But there is still physical link break from the Layer 2 handoff. And also we noticed that even in SA-MIP, the biggest part of handoff latency was still the layer4 TCP back-off-latency. In future works, the speed adaptive scheme should be applied to layer 2 and layer 4 handoff latencies.
LIST OF REFERENCES


[80211g] IEEE 802.11g-2003 IEEE Standard for Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications—Amendment 4: Further Higher-Speed Physical Layer Extension in the 2.4 GHz Band
BIOGRAPHICAL SKETCH

Jun Tian got his B.S. and M.S. in electrical engineering from Shandong University, China. In 2002, he was awarded an alumni fellowship to pursue his Ph.D degree in computer sciences at the University of Florida. He got his Ph.D degree and joined Motorola Corporation in May 2005, and works on mobile device local connectivity technologies.